



D1.5 – Evaluation of Flexgrid Technologies

Status and Version:	Final version	
Date of issue:	30.10.2014	
Distribution:	Public Report	
Author(s):	Name	Partner
	Luis Velasco (Editor)	UPC
	Andrew Lord	BT
	Annalisa Morea	Alcatel Lucent
	Matteo Dallaglio	CNIT
	João Pedro	Coriant PT
	António Eira	Coriant PT
	Nelson Costa	Coriant PT
	Matthias Gunkel	Deutsche Telekom
	Antonio D'Errico	Ericsson
	Juan Fernandez-Palacios	Telefonica
	Marco Quagliotti	TI
	Laura Serra	TI
	Emilio Riccardi	TI
	Kostas Christodouloupoulos	University of Patras
	Manos Varvarigos	University of Patras
	Chris Matrakidis	UoP
	Alexandros Stavdas	UoP
	Marc Ruiz	UPC
	Alberto Castro	UPC
	Adrià Asensio	UPC
	Lluís Gifre	UPC
	Gabriel Junyent	UPC
	Jaume Comellas	UPC
	Mateusz Żotkiewicz,	WUT



IST IP IDEALIST
(Industry-Driven Elastic and Adaptive Lambda
Infrastructure for Service and Transport
Networks)

**Evaluation of Flexgrid
Technologies**

	Artur Tomaszewski	WUT
	Mirosław Klinkowski	WUT
	Michał Pióro	WUT
Checked by:	Juan Fernandez-Palacios	



Abstract

Optimisation of various IDEALIST flexgrid technologies on reference networks and comparison with fixed-grid solutions. CAPEX and energy efficiency comparisons for all the flexgrid technologies under study. Discussion of the impact of an adaptive network manager and an elastic-enabled control plane proposed in WP3 in terms of OPEX reduction (including reduced power consumption) in a network operator. Main algorithm results, including offline and online, defragmentation and energy profile optimisation. First working planning tool prototype. Techno-economic benchmarking of key data plane technologies identified in WP2 during the first year and identification of the most cost effective data plane technologies and node architectures to be implemented in WP4. Techno-economic analysis of the different routing and restoration algorithms proposed in Task 1.2 during the first year and identification of the most cost effective algorithms to be implemented in the IDEALIST control plane during the third year.



Contents

1	EXECUTIVE SUMMARY	7
1.1	<i>SLICEABLE BIT RATE VARIABLE TRANSPONDERS (S-BVTs).....</i>	<i>7</i>
1.1.1	<i>Transponder design parameters</i>	<i>7</i>
1.1.2	<i>Optical node architecture</i>	<i>9</i>
1.1.3	<i>Network impact of different S-BVT designs</i>	<i>11</i>
1.2	<i>USE CASES</i>	<i>12</i>
1.2.1	<i>Multi-Layer Shared Backup Router (MLSBR) Use Case</i>	<i>12</i>
1.2.2	<i>Elastic connectivity for Cost Optimization in Federated Datacenters</i>	<i>12</i>
1.2.3	<i>Multicast</i>	<i>13</i>
1.2.4	<i>BRAS centralization in metro networks.....</i>	<i>14</i>
1.2.5	<i>Benefits of flexgrid in Long haul networks.....</i>	<i>14</i>
1.3	<i>TECHNO-ECONOMIC BENEFITS</i>	<i>16</i>
1.3.1	<i>Cost model including flexible client mapping</i>	<i>16</i>
1.3.2	<i>OPEX benefits of S-BVTs - sparing</i>	<i>17</i>
1.3.3	<i>Migration.....</i>	<i>18</i>
1.4	<i>POWER SAVINGS</i>	<i>18</i>
1.5	<i>ALGORITHM DEVELOPMENT.....</i>	<i>19</i>
1.6	<i>CONCLUSIONS.....</i>	<i>20</i>
2	INTRODUCTION.....	20
2.1	<i>PURPOSE AND SCOPE.....</i>	<i>20</i>
2.2	<i>REFERENCE MATERIAL</i>	<i>20</i>
2.2.1	<i>Acronyms.....</i>	<i>20</i>
2.3	<i>DOCUMENT HISTORY</i>	<i>22</i>
3	IMPACT OF ADAPTIVE NETWORK MANAGER	23
3.1	<i>CAPEX OPTIMIZATION BY MULTILAYER SERVICE PROVISIONING AND RESTORATION</i>	<i>23</i>
3.1.1	<i>Multi-Layer Shared Backup Router Use Case</i>	<i>23</i>
3.1.2	<i>MLSBR Availability Analysis.....</i>	<i>24</i>
3.1.3	<i>Impact on CAPEX Reduction</i>	<i>26</i>
3.2	<i>IMPACT ON OPEX REDUCTION.....</i>	<i>28</i>
3.2.1	<i>SLA relaxation by Multi-layer Restoration</i>	<i>29</i>
3.2.2	<i>Reparation WorkForce Model</i>	<i>30</i>
3.3	<i>THE ROLE OF THE ADAPTIVE NETWORK MANAGER.....</i>	<i>33</i>
4	TECHNO-ECONOMIC AND ENERGY ANALYSIS.....	35
4.1	<i>IDEALIST BUILDING BLOCKS</i>	<i>35</i>
4.2	<i>TECHNO-ECONOMIC ANALYSIS OF IDEALIST BUILDING BLOCKS</i>	<i>40</i>
4.2.1	<i>Fixed vs Flex-grid evolution models</i>	<i>40</i>
4.2.2	<i>SBVT OPEX analysis</i>	<i>41</i>
4.2.3	<i>Conclusions from the techno-economic analysis</i>	<i>45</i>
4.3	<i>NODE AND NETWORK ARCHITECTURE ANALYSIS</i>	<i>45</i>
4.3.1	<i>BRAS centralization in metro networks.....</i>	<i>46</i>



4.3.2	<i>Viability of Flexgrid in long distance networks.....</i>	<i>54</i>
4.3.3	<i>Evaluation of SBVT architecture based on multi-wavelength sources.....</i>	<i>62</i>
4.3.4	<i>Traffic grooming under fixed and elastic spectrum allocation</i>	<i>66</i>
4.3.5	<i>SERANO approach</i>	<i>72</i>
4.3.6	<i>Architecture on Demand (AoD) Approach.....</i>	<i>75</i>
4.4	POWER CONSUMPTION	82
4.4.1	<i>Power Consumption Reduction in Survivable Networks.....</i>	<i>83</i>
4.4.2	<i>Elastic Optical Networks to leverage Energy Savings</i>	<i>88</i>
5	MAIN ALGORITHM RESULTS.....	94
5.1	NETWORK OPTIMIZATION AND OFF-LINE ALGORITHMS	96
5.1.1	<i>Valid inequalities and a column-and-cut generation for RSA-based problems.....</i>	<i>96</i>
5.1.2	<i>Gradual Network Design</i>	<i>100</i>
5.1.3	<i>Performance evaluation of light-tree schemes</i>	<i>103</i>
5.1.4	<i>Multi-layer planning of flexible optical networks</i>	<i>106</i>
5.2	ON-LINE ALGORITHMS FOR REAL TIME OPERATION	114
5.2.1	<i>Spectrum defragmentation</i>	<i>114</i>
5.2.2	<i>Reoptimization of Dynamic Flexgrid Networks After Link Failure Repairs.....</i>	<i>118</i>
5.3	SELECTED ALGORITHMS TO BE IMPLEMENTED.....	127
5.3.1	<i>Spectrum Defragmentation</i>	<i>127</i>
5.3.2	<i>Re-optimization After Link Failure Repairs.....</i>	<i>128</i>
5.3.3	<i>Inter-datacenter Multicast Connectivity</i>	<i>129</i>
5.3.4	<i>Distance Adaptive RSA</i>	<i>131</i>
5.3.5	<i>Algorithms to be implemented in PLATON.....</i>	<i>132</i>
6	PLANNING TOOL PROTOTYPE	133
6.1	MANTIS	135
6.2	PLANNING TOOL FOR OPTICAL NETWORKS (PLATON)	135
6.2.1	<i>Test-bed set-up</i>	<i>137</i>
6.2.2	<i>Summary of results.....</i>	<i>138</i>
6.2.3	<i>Performance evaluation.....</i>	<i>138</i>
7	CONCLUSIONS	139
8	APPENDIX: SUPPLEMENTAL MATERIAL.....	140
8.1	IMPACT OF ADAPTIVE NETWORK MANAGER	140
8.1.1	<i>Modeling the Field operations and Repair.....</i>	<i>140</i>
8.2	HARDWARE SUPPORT MODEL	141
8.3	SERVICE LEVEL AGREEMENT (SLA) FOR THE REPAIR & RETURN MODEL.....	141
8.4	TECHNICAL SUPPORT SERVICE LEVELS	142
8.4.1	<i>Elastic connectivity for Cost Optimization in Federated Datacenters</i>	<i>144</i>
8.5	COST OPTIMIZATION IN FEDERATED DATACENTERS.....	144
8.5.1	<i>Dynamic connection and transfer mode requests.....</i>	<i>145</i>
8.5.2	<i>Minimizing energy expenditures.....</i>	<i>146</i>
8.5.3	<i>Performance evaluation.....</i>	<i>147</i>
8.5.4	<i>Conclusions</i>	<i>150</i>
8.6	TECHNO-ECONOMIC AND ENERGY ANALYSIS	151
8.6.1	<i>Idealist Building Blocks.....</i>	<i>151</i>
8.6.2	<i>CAPEX analysis of a survivable multi-layer IP-over-Flexgrid network.....</i>	<i>154</i>
8.6.3	<i>BRAS Centralization network design problem</i>	<i>160</i>
8.6.4	<i>Complete version of the study on viability of flexgrid in long distance networks.....</i>	<i>164</i>



8.6.5	<i>Traffic grooming under fixed and elastic spectrum allocation</i>	188
8.6.6	<i>SERANO approach</i>	194
8.7	MAIN ALGORITHM RESULTS.....	202
8.7.1	<i>Model and solving methods for the GRANDE Problem</i>	202
8.7.2	<i>Mathematical model for the P2MP-RSA Problem</i>	206
8.7.3	<i>Multi-Layer Network Planning Algorithm.....</i>	208
8.7.4	<i>Mathematical model for the SPRING problem</i>	215
8.7.5	<i>Mathematical models for the AFRO problem</i>	217
9	REFERENCES	221



1 Executive summary

This is the fifth deliverable from Work Package 1 of Idealist and represents work completed in the second year of this three-year project. The main results from the work are summarised in this Executive Summary, but the main report contains many more details and discussion, including a large section on algorithmic development, not adequately covered in this summary.

The objective of Idealist, and therefore of work Package 1 is to decide on and evaluate the benefit of Elastic Optical Networking in general, and to identify how to make the best use of it. As shown in this Executive Summary, in all cases there is benefit in these new elastic technologies – both from flexible transponder resources and from a more flexible use of optical spectrum.

The summary focuses on the specific Sliceable Bit Rate Variable Transponder technology first, as this is currently receiving a great deal of attention within Idealist and in the community in general. Then we focus on specific network Use Cases, followed by a review of some of the techno-economic results, power savings and developments in the algorithms used.

1.1 *Sliceable Bit Rate Variable Transponders (S-BVTs)*

1.1.1 Transponder design parameters

S-BVTs are a generalised version of a BVT: the BVT has a single wavelength carrier and modulation function, producing a single optical sub-channel such as DP-QPSK giving 100Gb/s, or DP-16QAM giving 200Gb/s. S-BVTs contain more than one carrier / modulation function, allowing their overall bandwidth to be used more flexibly. Fig.ES1 shows an example of how super-channels can be created, highlighting also the key role of OTN in this process, providing the containers to carry the traffic. In the figure, a 1Tb/s S-BVT contributes to two media channels: the first consisting of 6 sub-carriers making a total of 600 Gb/s. This 1Tb/s S-BVT also contributes 4 sub-carriers to a second media channel, which also makes use of 1 fifth sub-carrier from a second S-BVT. This second S-BVT only has a 200 Gb/s capability, which is split into two as shown in the figure.

Additionally there needs to be provision for multiple client inputs and switching to ensure any client can be carried on any sub-carrier. In the medium term it is expected that the dimensions of the overall unit (client + switching + line slide) will be of order 1 – 1.2 Tb/s. The S-BVT or Multiflow Module will also be limited, for example to the following:

- 2 sub-carriers module, which support up to two independent flows: 6 modules are necessary to reach the total flexibility of 12 carriers (for 1.2Tb/s)
- 4 sub-carriers module, which supports up to four independent flows: 3 module are necessary to reach the total flexibility of 12 carriers

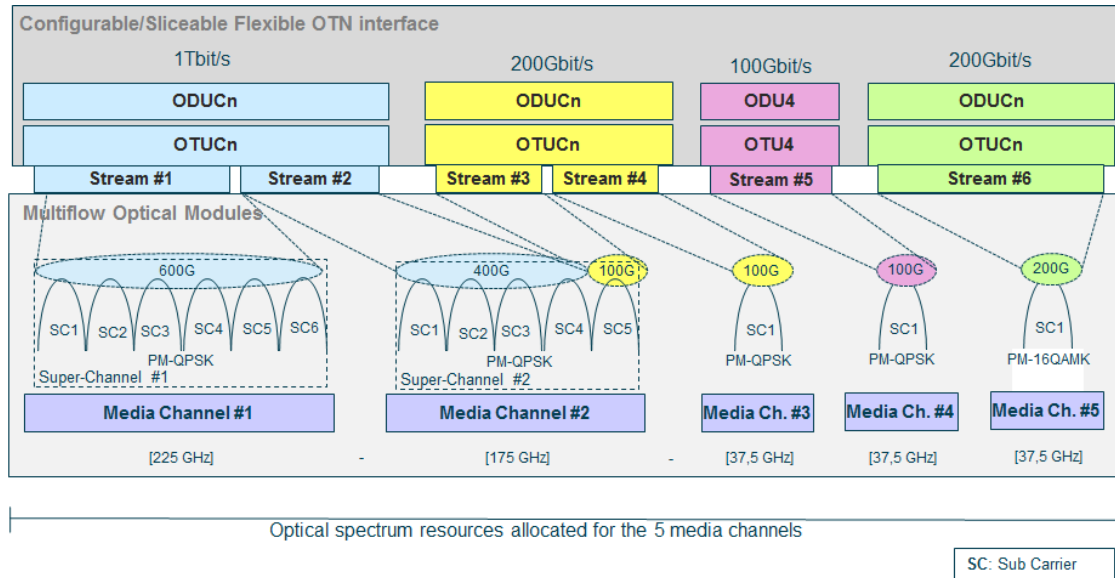


Fig.ES1. Example of ways in which S-BVTs can be used to create optical channels

One important question for the community is to quantify the additional value of S-BVT technology beyond non-sliceable BVTs (i.e. single modulator devices). Much of this value relates to the assumed need for dynamic bandwidth, and fig.ES2 below shows one result of this, taking into account the kind of additional flexibility shown in Fig.ES1. Here it is apparent that fewer S-BVTs can do the job due to their ability to cater for a wider range of traffic demands.

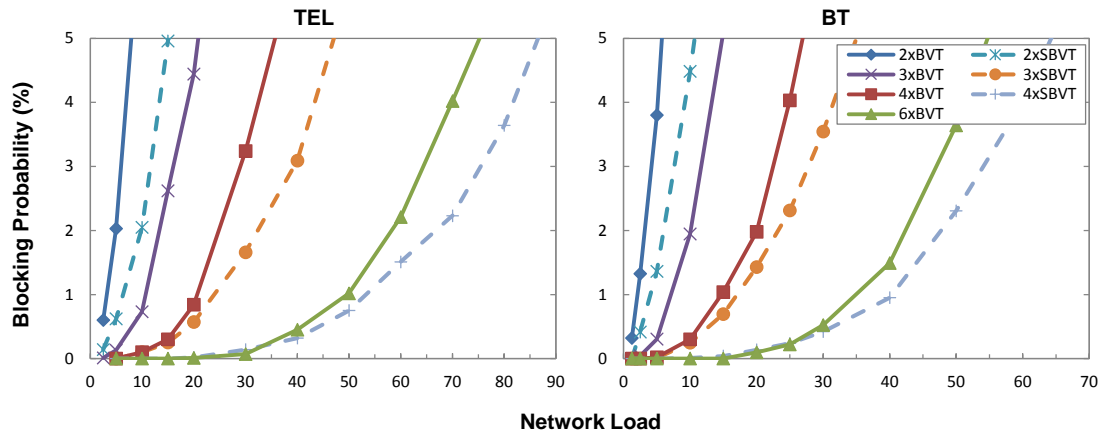


Fig.ES2. Impact of numbers of BVTs or S-BVTs on blocking probability for dynamic traffic

A further analysis to compare BVTs and S-BVTs relates to an IP over DWDM architecture, effectively involving the router client traffic demands in the consideration. Here, inclusion of

edge and aggregation routers makes the scenario more complex. Some results for the Telefonica network are shown in Fig. ES3. The benefits of S-BVTs are again evident, with diminishing returns going beyond about $f = 4$.

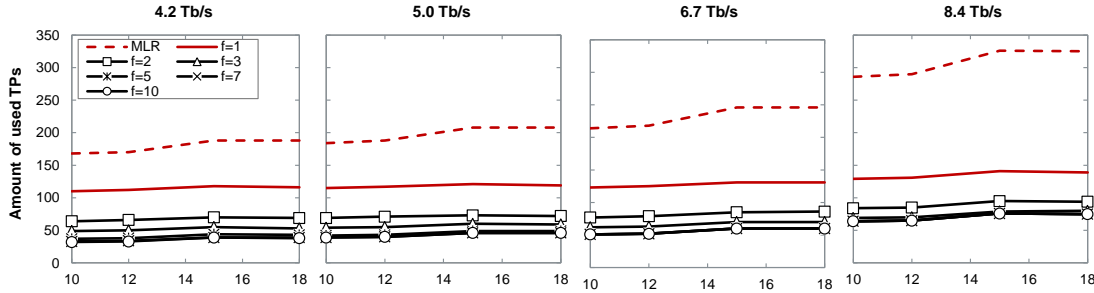


Fig.ES3. Amount of transponders needed for different traffic dimensions and for S-BVTs with different numbers of modulators (f).

Additional flexibility in need of further study in the final year of Idealist concerns the net and gross baud rate combinations. Net baud rates in the medium term could range between 25 and 33.3Gbaud/s, and a range of FEC options added to this (e.g. 7, 11 or 25% depending on reach).

In table ES1 the various superchannel configurations to give multiple 100Gb/s client rates are shown, including number of number of carriers and net baud rates (in brackets) for the different modulation formats. Some client rates can be created in a wider combination of ways – for example 1Tb/s can be created in two ways for all modulation formats, giving many transmission options.

1.1.2 Optical node architecture

The development of flexgrid towards future S-BVT super-channel transmission will have a direct impact on the design of the optical node which will perform the various switch and bypass functions required. Current architectures are based around present-day 1x9 WSS filters and also multicast switches, including functions such as colourless, directionless and contentionless add/drop. Additional requirements now include scope for slicing and add/dropping to S-BVTs as well as a more dynamic operating mode.

Fig.ES4 describes the IDEALIST approach towards developing the proposed new Flex-OXC node architecture, allowing different configurations based on the internal organization of a number of functional components. An Architecture on Demand (AoD) based Flex-OXC node is proposed together with a Switchless Elastic Rate Node (SERANO) module, which is an alternative to the wavelength switching architectures.



Table ES1: Number of carriers and Net Baud rate (in brackets, expressed in Gbaud/s) required in a S-BVT for different client bit rates multiple of 100G

Client bit rate (Gb/s)	PM QPSK		PM 8QAM		PM 16QAM		PM 32QAM	
	max#carr.	min #carr.	max#carr.	min#carr.	max#carr.	min#carr.	max#carr.	min#carr.
100	1 (25.000)	-	-	-	-	-	-	-
200	2 (25.000)	-	1 (33.333)	-	1 (25.000)	-	-	-
300	3 (25.000)	-	2 (25.000)	-	-	-	1 (30.000)	-
400	4 (25.000)	3 (33.333)	2 (33.333)	-	2 (25.000)	-	-	-
500	5 (25.000)	4 (31.250)	3 (27.778)	-	2 (31.250)	-	2 (25.000)	-
600	6 (25.000)	5 (30.000)	4 (25.000)	3 (33.333)	3 (25.000)	-	2 (30.000)	-
700	7 (25.000)	6 (29.167)	4 (29.167)	-	3 (29.167)	-	-	-
800	8 (25.000)	6 (33.333)	5 (26.667)	4 (33.333)	4 (25.000)	3 (33.333)	3 (26.667)	-
900	9 (25.000)	7 (32.143)	6 (25.000)	5 (30.000)	4 (28.125)	-	3 (30.000)	-
1000	10 (25.000)	8 (31.250)	6 (27.778)	5 (33.333)	5 (25.000)	4 (31.250)	4 (25.000)	3 (33.333)
1100	11 (25.000)	9 (30.556)	7 (26.190)	6 (30.556)	5 (27.500)	-	4 (27.500)	-
1200	12 (25.000)	9 (33.333)	8 (25.000)	6 (33.333)	6 (25.000)	5 (30.000)	4 (30.000)	-

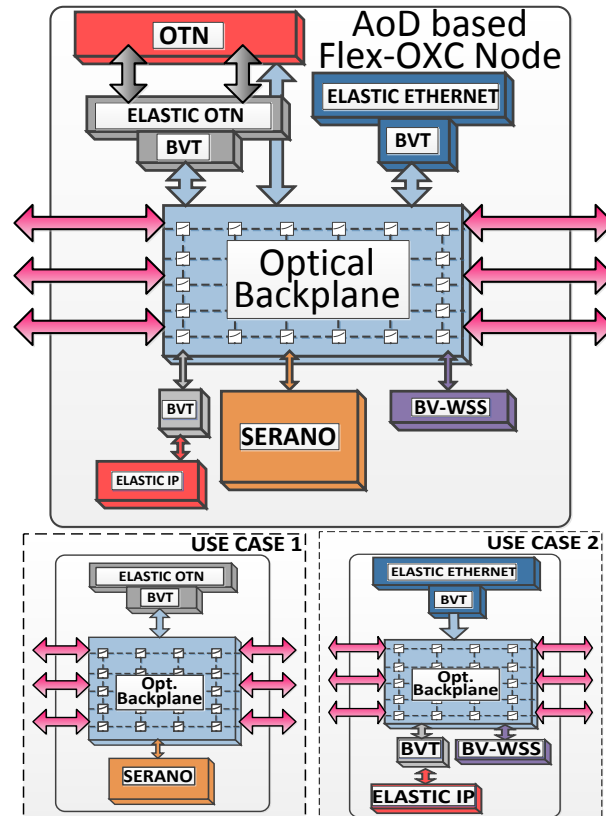


Fig.ES2 Idealist future optical node architecture

The heart of the architecture is an optical backplane (e.g. a large portcount 3D-Micro-Electro-Mechanical (MEMS) cross-connect) allowing 'Architecture on Demand' (AoD), in which functions can be switched in and out dynamically – examples shown in the figure include OTN or ethernet flexibility.

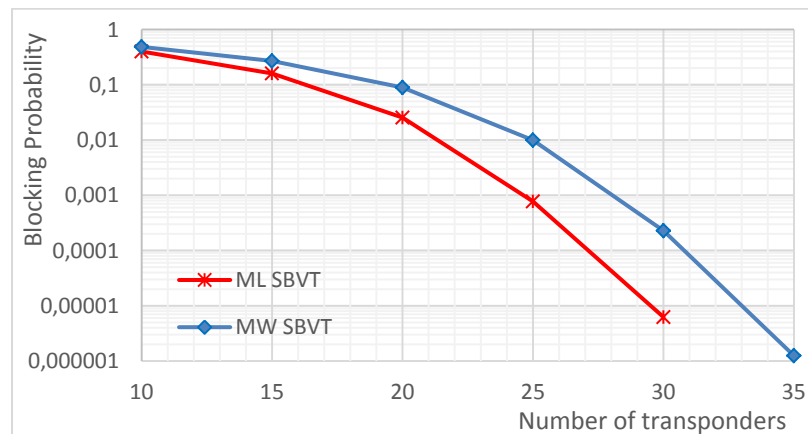
One feature shown in Fig.ES4 is the SERANO node function. It is included as an option in a broader architecture to allow functions that optics can't do alone. Examples of SERANO functions include (i) 3R regeneration (re-time, re-transmit, re-shape a selected input signal), (ii) Modulation format translation (using S-BVTs to translate the modulation format of a flow), (iii) Change of carrier central frequency (to avoid spectrum conflict). These can have a wide range of applications including complex network procedures such as defragmentation. Other ways to utilise SERANO include optimising the optical reach and unblocking by relaxation of frequency continuity.

1.1.3 Network impact of different S-BVT designs

Different technologies and solutions can be considered to design and implement SBVT architectures. The laser source in an SBVT can be developed using either a multi-laser (ML) source, i.e. an array of N tuneable lasers, or a multi-wavelength (MW) source that consists in a single laser able to generate various carriers.

The ML source does not introduce any constraint on the routing and spectrum assignment (RSA), whereas the MW source provides greater stability among the carriers thus better spectrum compactness. Moreover, MW allows to reduce the cost of the SBVT by decreasing the number of lasers and therefore the power consumption. As drawback, the MW source imposes new constraints on the tuning and relative spacing among the carriers.

Results for network blocking probability based on the two designs are shown in fig. ES5:



ES5 Blocking Probability as a function of numbers of transponders per node for the two SBVT designs.

Here it is clear that the additional spectrum restrictions of the MW source cause a higher blocking probability. This device would therefore need to be significantly cheaper than the ML device for it to make sense.



1.2 Use Cases

1.2.1 Multi-Layer Shared Backup Router (MLSBR) Use Case

Coordination between the router and optical transport layer provides many benefits, one of which is the potential for orchestrated protection in case of failures. Rather than have independent protection in both layers, the MLSBR scheme allows back-up router resources to be shared, saving significant cost. Modelling the Telefonica network, just having two back-up routers saved up to 24% on router ports (as well as supporting chassis). Sharing back-up routers also improves OPEX too – this stems from the fact that the sharing relaxes the requirement for fast repair time. For example, in the modelling presented, a 5-9s availability leads to a 3-4 day repair requirement for the full protection case whereas just two back-up routers increase this to 14 days.

This relaxation in repair time (see fig.ES6 for a plot of this benefit) has a direct impact on the SLAs required for the repair contracts and the impact of this is a reduction in the requirement for the number of repair teams. Whilst the details of this depend on the exact nature of the repair contract, nevertheless in all cases this reduction is significant, leading to worthwhile OPEX savings.

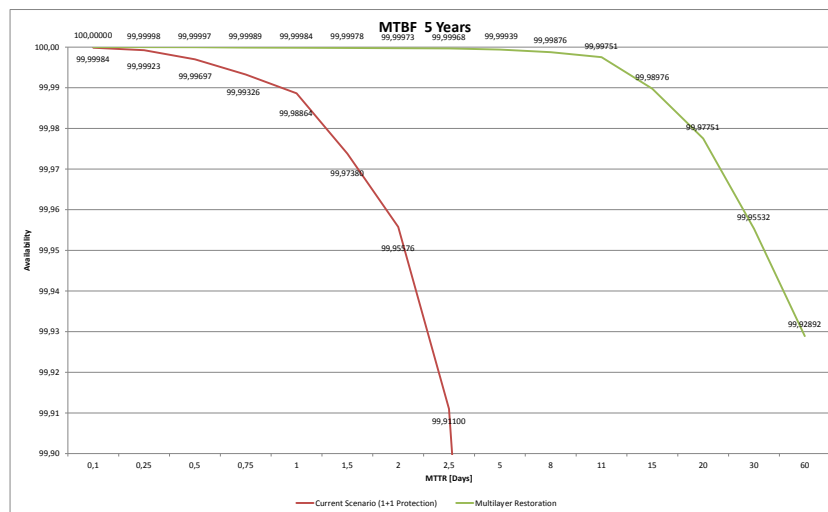


Fig.ES6 – Availability of current 1+1 and multilayer protection schemes as a function of MTTR in days

1.2.2 Elastic connectivity for Cost Optimization in Federated Datacenters

Current inter-datacenter (DC) operation assumes large bandwidth static connections, over-provisioned to handle the peak traffic demands. This traffic is dynamic in nature, and is dominated by the migration of Virtual Machines (VM) and database synchronisation between the datacenters. The potential of flexgrid to dynamically make fine spectrum granularity adjustments, allows it to be highly useful in these dynamic traffic environments.

One way to provide a more flexible service to meet these demands is to use SDN to open up the control of the underlying network resources thus allowing these to be manipulated in real time. Both energy and cost advantages of doing this depend on the amount of traffic that can be scheduled, compared to the background traffic, and some example results are shown in fig.ES7:

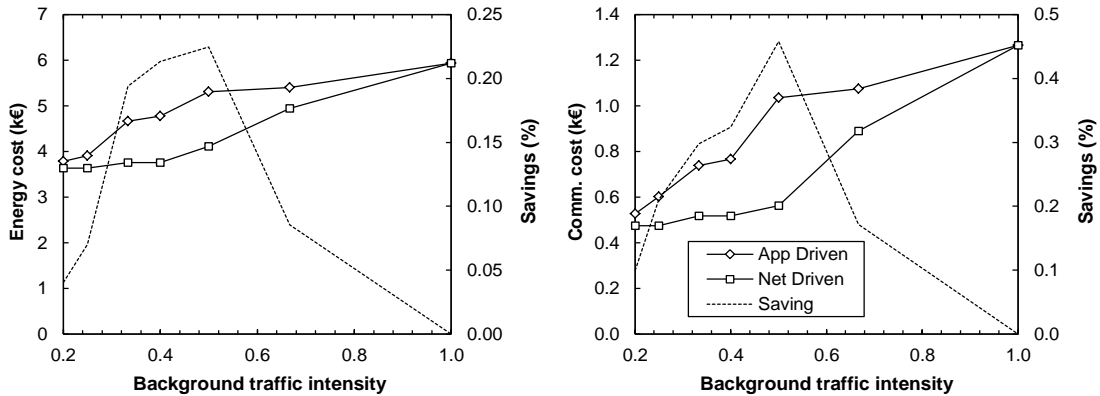


Fig.ES7. Daily savings in energy and communications, if the resources required for inter data centre traffic can be set up on demand

From the results, we observed that when the network operates under low and medium traffic load costs savings as high as 20% and 40% in energy and communications, respectively can be obtained.

1.2.3 Multicast

Content distribution with high bandwidths such as 4k uncompressed TV will have a large impact in future networks. One application for flexgrid and SBVTs relates to this application and some studies in Idealist have looked at this. Some results are shown in fig.ES8. The advantage of S-BVTs for multicast is that each slice can be spectrum-tuned giving much more scope for the Routing and Spectrum Assignment algorithm to find paths (compared to passive splitters for example). This reasoning explains the benefits of higher numbers of modulators for this particular Use Case.

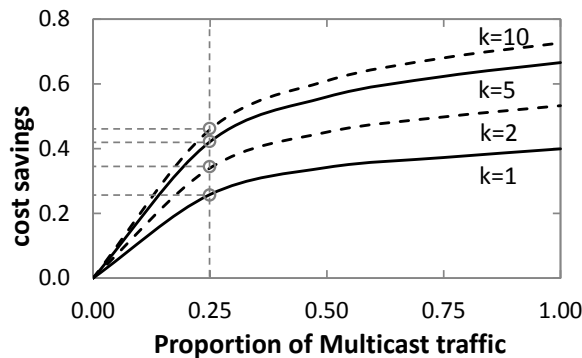


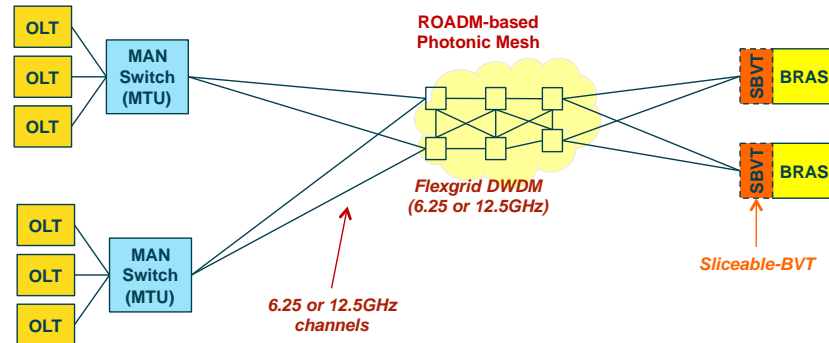
Fig.ES8. Cost savings for SBVTs as multicast traffic increases

1.2.4 BRAS centralization in metro networks

Typically, several BRAS servers are installed in the network and are placed close to end-users. To reduce the cost, network operators are considering centralizing the BRAS's within a region and transferring the traffic destined to them over a regional optical mesh network over a WDM network which is the current practice for building regional mesh networks. Since flexgrid technology is considered an alternative to WDM, Idealist is evaluating the applicability of this technology in this regional/MAN network scenario.

The transmission requirements of a realistic network scenario are evaluated, and this suggested the use of appropriate flexgrid transponders that are based on direct detection (DD) that is less complicated and much cheaper than flexgrid transponders that would be used in core-long haul networks.

The network architecture proposed is shown in fig.ES9:



ES9. Architecture utilising flexgrid and S-BVTs to interconnect BRAS with Access

The study highlights that the transmission distances involved don't necessarily require high performance coherent solutions and so an improved S-BVT architecture based on multiband OFDM (MB-OFDM) and 10Gb/s direct detection (DD) is proposed. A techno-economic comparison between the WDM and flexgrid solutions shows very reasonable costs for the Direct Detection S-BVT to give cost benefits of 30% - and more would be available with reducing flexgrid WSS costs (which dominate the solution).

1.2.5 Benefits of flexgrid in Long haul networks

To properly assess the benefits of elastic networking and flexgrid to the longer distances, we have focused on a European network shown in fig.ES10.

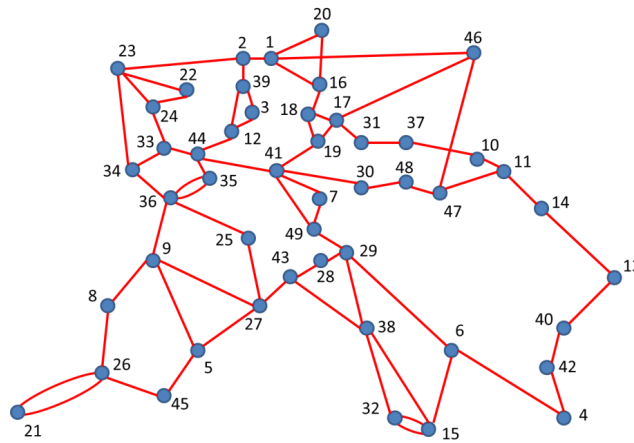
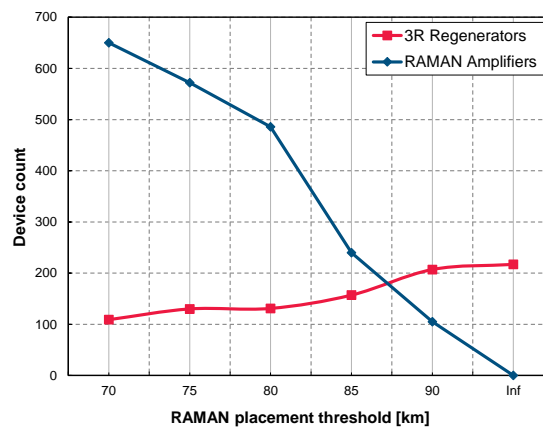


Fig.ES10: Topology of Pan European Network (Telecom Italia Sparkle Network).

The network has 49 nodes and 72 links, with node degree ranging from 2 to 5. Links have lengths from 2 to 1251 km, with about 390 km on average. Type of fiber is a mix of G655 and G652. Working paths have an average length of about 1800 km and a maximum length of 5000 km. Back up paths have an average length of about 3000 km and a maximum length a little less than 7000 km. This means that to avoid signal regenerations, a cost effective network design requires very long reach systems and every possible care to limit the signal degradation.

One key technology here is Raman amplification, with its potential for very low noise figure gain. It can be used to significantly reduce the amount of 3R regeneration in long haul systems, and fig.ES11 shows the impact on the Pan EU network:



ES11. Impact of Raman gain on number of 3R regenerations.

Even with maximal Raman, the network still requires 100 3R regens, but these climb to more than 200 if Raman isn't used. There is clearly a benefit for using Raman gain in this

type of network. The cost analysis showed that Raman gives a 15% reduction considering only the Raman and Transponder costs.

The study continued to have a preliminary look at the benefits of S-BVTs compared to their fixed rate counterparts. It was found that the cost of an S-BVT to give a comparable overall network cost was challengingly low, suggesting no obvious advantage for S-BVTs in this regime.

1.3 Techno-economic benefits

1.3.1 Cost model including flexible client mapping

The flexibility being made available through new line side developments requires additional flexibility between client and line side for full benefits to be appreciated. This is well summarised in fig.ES12 which shows three different possibilities, involving flexibility both from the line cards and the transceivers that plug into them.

The first (fixed+fixed) assumes fixed line card slots (e.g. 100G or 150G etc) into which fixed transceivers of the same rate are plugged. There needs to be a switching function from client to line side to allow clients to be mapped onto multiple line transceivers if necessary.

In the second example (flex+fixed), the line cards are no longer specific but can handle the whole range of line rates. However the transceivers are still fixed rate. This is a more flexible solution, requiring increased switching between clients and line side.

In the third example both the line cards and transceivers are flexible, providing the most versatile option but with the most client-line switching required.

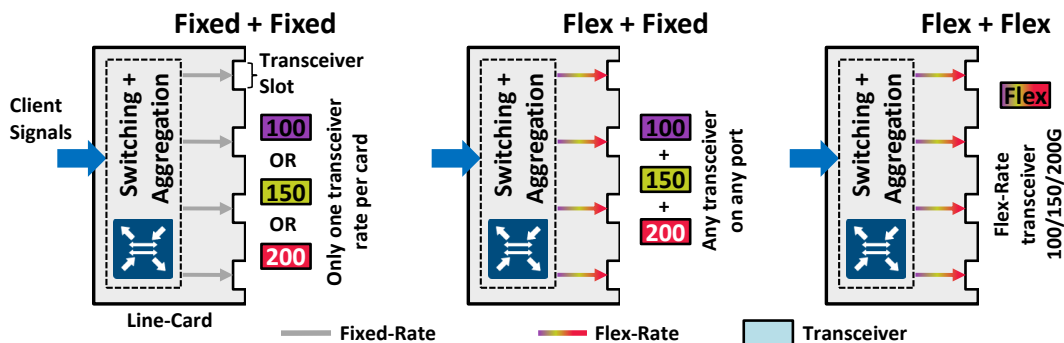


Fig.ES12. Different client + line architectures to take advantage of flexible transceivers

The underlying trade-off behind these line-card and transceiver designs is that the possibility of having equipment more adaptable to changing traffic conditions comes at the price of more complex hardware that will entail higher investments in the initial network roll-outs. Given the early stage of the development of flex-rate technology, it is hard to pinpoint accurate cost values for the different elements in this modular architecture. The

approach adopted in Idealist is to attempt to evaluate the relative differences for several target cost values. Issues such as the relative costs of line cards and transceivers will dictate which of the solutions is optimal, as will the cost difference between the use of different bit rates. The last factor of cost uncertainty is the premium for having flex-rate line-cards. The flex-rate line-card must have additional embedded functions in the switching module in order to output multiple rates for the transceiver ports. In addition, the digital signal processing in the ports must be adaptive to receive signals with different modulation formats. On the other hand, having a single line-card in the entire network brings economies of scale and facilitates development and component management, potentially bringing the cost per unit down. All these factors are hard to account for simultaneously without having the equipment in production, so the evaluation in the main report also includes an analysis of various cost ratios between a flex-rate line-card and a 200 Gb/s line-card (the highest capacity in the fixed-rate system).

The results are complex and require sensitivity analyses to fully explore the assumptions. However a sample result is shown in Fig.ES13:

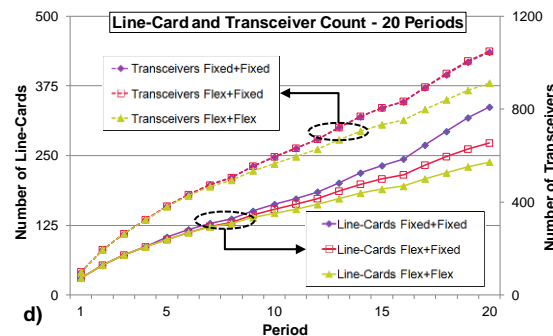


Fig.ES13. Numbers of line-cards and transceivers for the various solutions as a function of multiple period traffic growth.

1.3.2 OPEX benefits of S-BVTs - sparing

One operational benefit of a transponder capable of a wide range of modulation formats and bit rates, as well as multiple slices of output spectrum, relates to the operational advantages of sparing. A study has been completed here on an IP over DWDM scenario on the Spanish network with results shown in fig.ES14.

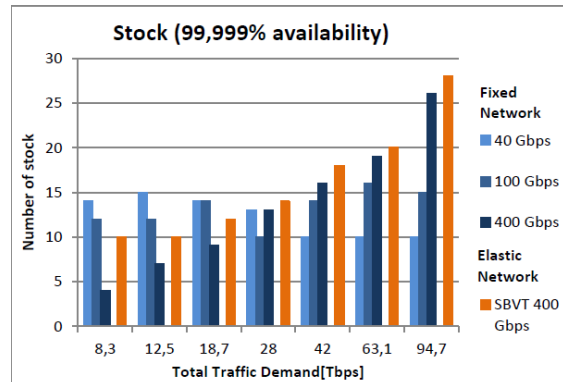


Fig.ES14. Stock required for fixed and S-BVTs vs traffic demand

The combined fixed transponder spares portfolio (including 40Gb/s, 100Gb/s, 400Gb/s) clearly requires higher sparing than the SBVT solution. It is interesting to add to this analysis some estimates on possible costs for the various transponders. Although these are obviously estimated for the higher bit rates, there is nevertheless scope for an S-BVT to cost up to 36% higher than a fixed rate 400Gb/s transponder and STILL deliver 30% network savings. This shows that S-BVTs have clear benefits even considering the OPEX sparing issue.

1.3.3 Migration

A key topic relates to the question of if and when to migrate to flexgrid operation, and in particular when to deploy flexgrid resources. A great deal of modelling continues to address this question. Fig.ES15 gives a strong impression of the interplay between the various effects. The black curve relates to a non-flexgrid solution which becomes inefficient when demands start to exceed 100 Gb/s. The blue curve implies building two networks – a non flex-ready one first and then a second one when required, thus explaining the large jump in devices required. The clear winner on this measure is a flex-ready EON which is able to efficiently switch super-channel demands when they are required by the network. There may of course be an up-front cost the make the network flex-ready.

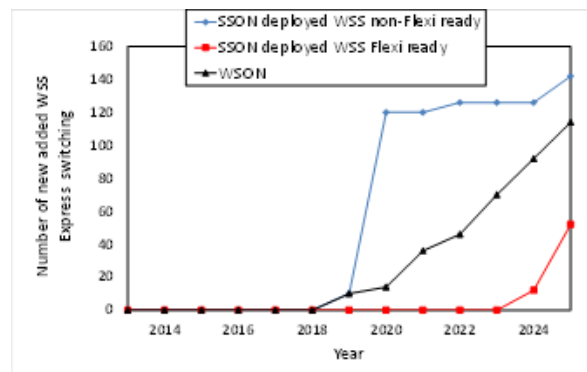


Fig.ES15 Numbers of additional WSS over time dependent on build strategy.



1.4 Power Savings

Several studies have focused more on the opportunities for elastic optical networks to save power (as distinct from cost).

One of these focused on the issue of shared protection in IP over DWDM networks. The scope for flexible programming of SBVTs leads to fewer devices provisioned for protection purposes which in turn leads to lower power. Indicative results are shown in Table ES2 where the Elastic solution consumes less power than the single and mixed line rate (SLR, MLR) solutions for all traffic loads.

A second study looks more generally at the E2E network and seeks to optimise the metro clusters feeding into a core network based on the transmission technology. In this way, the enhanced flexibility of EON, with appropriately optimised grooming, can lead to significant energy savings. The study made use of the BT 1113 node network in order to fully model the interplay between metro and core boundaries. The work shows that efficient hierarchical topologies can be constructed exploiting the transponder bandwidth granularity, reducing overall network cost and power consumption.

Table ES2. Power consumption for different technology solutions in IP over DWDM networks

		Load = 25%		Load = 50%		Load = 75%		Load = 100%	
	DWDM layer	UN	SURV	UN	SURV	UN	SURV	UN	SURV
Optical layer	SLR	27.3	49.7	27.3	49.7	27.3	49.7	27.3	49.7
	MLR	23.7	41.1	23.7	41.1	23.7	41.1	23.7	41.1
	Elastic	14.4	26.9	15.6	28	18.9	31.3	23.8	36.3
IP layer (router ports)	SLR	34.8	50.5	34.8	50.5	34.8	50.5	34.8	50.5
	MLR	37.7	54.8	37.7	54.8	37.7	54.8	37.7	54.8
	Elastic	35.6	50.5	35.6	50.5	35.6	50.5	35.6	50.5
IP layer (processing)	SLR	29.9	29.9	59.8	59.8	89.7	89.7	119.6	119.6
	MLR	31.2	31.2	62.4	62.4	93.6	93.6	124.8	124.8
	Elastic	30.3	30.3	60.5	60.5	90.8	90.8	121	121
Total power consumption (IP + optical)	SLR	92	130.1	121.9	160	151.8	189.9	181.7	219.8
	MLR	92.6	127.1	123.8	158.3	155	189.5	186.2	220.7
	Elastic	80.3	107.7	111.7	139	145.3	172.6	180.4	207.8

One key aspect of an optimisation on this scale is grooming. Metro demands are groomed to create larger pipes that match the transponder technology on the core. Achieving optimal metro clusters, such that the groomed core pipes are the right size, is a complex optimisation procedure.

Results examine two traffic matrices representing a network growth of roughly 10 years. In the earlier traffic matrix, without SBVT availability, power could still be reduced by means of electronic grooming and optical bypassing, which can achieve power consumption gains of up to 45%. The introduction of SBVTs to the later traffic matrix can improve power performance by a further 10%, assuming the network is already utilising optimal grooming



and bypass benefits. Optical bypassing itself becomes more significant with the larger traffic matrix – up to 60% compared to the 45% benefits for the smaller matrix.

1.5 Algorithm Development

A great deal of effort in Idealist has been committed to developing algorithms to apply to the various scenarios above. There is a major chapter in the main deliverable outlining the developments in these, both for offline planning and also real time operations.

Offline algorithmic developments over the past year include a new algorithm for gradual network design called **GRANDE** which finds the additional least cost upgrade in terms of equipment (new nodes, part of nodes and system on links) when an updated traffic matrix is assigned. Also there is a new version of the multi-layer planning problem for a flexible optical network.

Online algorithm developments include a new *reactive algorithm* for spectrum defragmentation called **SPRING**, which applies the concept of spectrum shifting for defragmentation.

1.6 Conclusions

The case for moving to elastic transponders and flexible grid has been made comprehensively over the course of the Idealist project. The detailed studies here will be highly useful to operators as they evaluate the details of the elasticity required for different use cases. One key technology – the S-BVT – is receiving significant attention and this too appears to have many applications where it could offer cost advantages to non-sliceable BVTs. Ultimately this will depend on the motivation of manufacturers to make them in sufficient volume for the cost benefits to be realised.

2 Introduction

2.1 Purpose and Scope

This is the fourth deliverable from Work Package 1 of Idealist. The motivation and an overview of this deliverable have been provided in the Executive Summary. The main document is organized into four sections:

Section 3 analyzes the impact of the adaptive network manager (proposed in WP3) in terms of CAPEX and OPEX reduction in a network operator.

Section 4 focuses on techno-economic and energy analysis.

Section 5 reports the main algorithm results including offline and online. The algorithms to be implemented in the planning tool are eventually selected.

Section 6 summarizes the results performed to the planning tool.

Extended material is presented in the appendix.



2.2 Reference Material

2.2.1 Acronyms

ABNO	Application-Based Network Operations
AoD	Architecture On Demand
BRAS	Broadband Remote Access Server
BVT	Bandwidth Variable Transponder
CAPEX	Capital Expenditures
CSO	Cross-Stratum Orchestrator
DAC	Digital-To-Analogue Converters
DC	Data Center
DD	Direct Detection
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbium-Doped Fiber Amplifiers
EON	Elastic Optical Networks
FRR	Fast Re-Route
GN	Gaussian Noise
LSP	Label Switched Path
MAN	Metropolitan Area Network
MDT	Mean Delivery Time
ML	Multi-Layer
MMTTR	Minimum Mean Time To Repair
MPLS	Multi-Protocol Label Switching
MTBF	Mean Time Between Failures
MTR	Maximum Time To Repair
MTTR	Mean Time To Repair
NOC	Network Operation Center
OB	Operational Business
OE	Opto-Electronic
OMFM	Optical Multi Flow Module
OPEX	Operational Expenditures
OTN	Optical Transport Network



OXC	Optical Cross Connect
RMSA	Routing, Modulation And Spectrum Allocation
ROADM	Reconfigurable Optical Add-Drop Multiplexer
RSA	Routing And Spectrum Allocation
RT	Repair Time
RW	Reparation Workforce
SBR	Shared Backup Router
SBVT	Sliceable Bandwidth Variable Transponder
SDN	Software Defined Networking
SERANO	Switchless Elastic Rate Node
SLA	Service Level Agreement
SPFF	Shortest Path With First Fit
TCO	Total Cost Of Ownership
UNI	User-To-Network Interface

2.3 Document History

Version	Date	Authors	Comment
Draft 1	30.06.14	Luis Velasco Andrew Lord	1 st draft. Contains placeholders for all contributions
Draft 2	15.07.14	Juan Fernandez-Palacios	2 nd draft after wp1 conf call
Draft 3	27.07.14	Luis Velasco	3 rd draft integrating proposed contributions
Draft 4	19.09.14	All partners	1 st integrated version
Draft 5	17.10.14	All partners	reviewed version
Final Version	30.10.14	All partners	



3 Impact of Adaptive Network Manager

This section analyzes the impact of an adaptive network manager and an elastic-enabled control plane proposed in WP3 in terms of CAPEX and OPEX reduction in a network operator.

TCO (Total Costs of Ownership) reduction is mainly due to more efficient network provisioning and resilience processes enabled by the control architecture proposed in IDEALIST.

- CAPEX Optimization by multilayer service provisioning and restoration
- OPEX reduction by IP-optical coordination

3.1 CAPEX optimization by multilayer service provisioning and restoration

Multi-layer network survivability is a topic of high interest for the research community. First multi-layer mechanisms were proposed in [2] for ATM over SDH/WDM architecture. However, as the network architecture has changed, new studies are done based on current IP/MPLS over WSON networks. The scope of the research in multi-layer networks is wide: new metrics to decide how to recover from failures are defined in [3], while authors in [4] focus on CAPEX reduction in an IP/MPLS-Over-WSON network. Authors in [5] present routing mechanisms suitable for multi-layer restoration. Multi-layer restoration has been an extensively researched topic in the past few years. In [7], [8] more cost-efficient architectures are presented as an alternative to current dual-plane protection. However, they have not been sufficiently explored. Building on advanced optical layer capabilities and multi-layer control, authors in [9] presented the MLSBR concept and compared the availability of MLSBR and traditional dual-plane approach.

The idea behind MLSBR consists on having extra shared backup routers to restore the traffic in case of a failure of an IP router. This technique is compared with the common design of today's networks, where two IP planes are created in order to deal with node failure. This paper presents comparative results to quantify the savings that can be obtained by applying this novel approach. The present document extends the techno-economic study of MLSBR carried out for the Spanish Telefónica core network in [10]. The impact of the number of shared backup routers on the overall savings is quantified.

This section describes the Multi-layer Shared Backup Router use case and explains its availability in comparison with 1+1 protection.

3.1.1 Multi-Layer Shared Backup Router Use Case

MLSBR use case consists on providing backup routers, which are available in case of a node failure. We assume that there is an optical mesh connection access, transit and interconnection routers. As previously described, the whole IP nodes must be duplicated in order to solve IP router failure. Let us assume a hierarchical architecture with three levels, as shown in Figure 1. This structure is typical to many IP networks. Let us call the lowest level in the hierarchy access routers, transit routers to the second level and interconnection routers to the higher level.

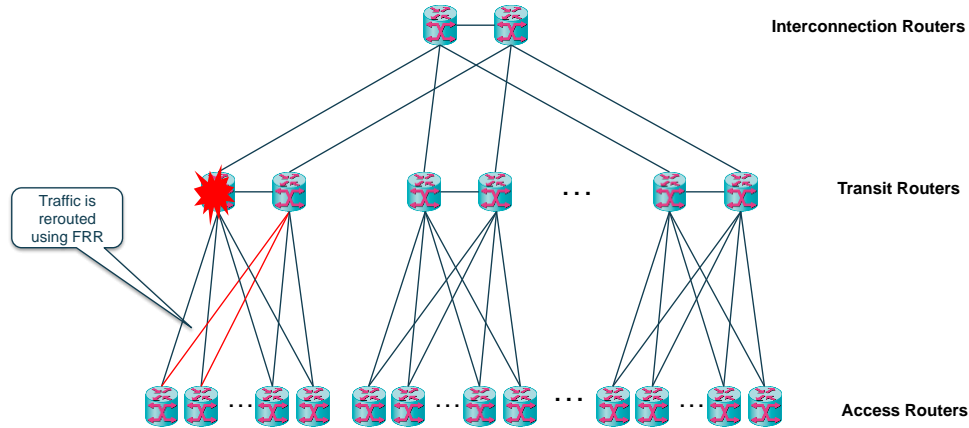


Figure 1: 1+1 Protection in a hierarchical network

In this example, the transit routers are duplicated to recover to a transit failure. When using MLSBR, a set of Shared Backup Routers (SBRs) are available so, when there is a failure in the transit routers, the failed transit router configuration is copied and new connections are created to the access and interconnection nodes. This scheme is presented in Figure 2.

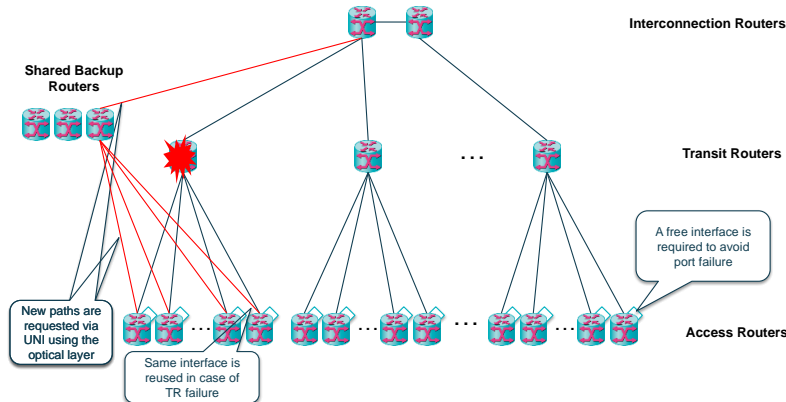


Figure 2: MLSBR scheme in a hierarchical network

Let us remark that the recovery time using dual-plane protection is faster than applying MLBSR, because MLBSR requires optical connection setup to the backup router (which will takes minutes to set up), as well as the time to configure the backup router with the configuration of the failure router – which again could take a few minutes. As previously mentioned, it is assumed that there is an optical mesh between access and transit nodes. However, the network availability when using MLSBR approach is better than traditional dual-plane protection.

3.1.2 MLSBR Availability Analysis

This section highlights the availability of this approach based on our previous work in [12]. To measure how “survivable” a network is, the availability concept is defined as how long a



user can access to the services provided by the network. Equation (1) presents the availability parameter definition.

$$Availability = 1 - \frac{Time\ unaccessible}{Total\ operating\ time} \quad (1)$$

Network status is defined in terms of the presence of failures in IP/MPLS nodes in the network. As there is an optical mesh, any failure in the optical layer is recovered by the optical mechanisms. The analytical study is done using Markov's chain model where the node reachability status is defined as follows:

Failure: 1

No failure: 0

The transitions between states of the Markov's model are defined based on the MTBF and MTTR parameters:

$$0\ to\ 1 \Rightarrow \lambda = 1 / MTBF \quad (2)$$

$$1\ to\ 0 \Rightarrow \mu = 1 / MTTR \quad (3)$$

3.1.2.1 One region case

First, the problem is defined for a one-region scenario. A region is defined as a group of access routers, which are connected by a pair of transit routers (Figure 1). In this scenario, the MLSBR and 1+1 protection schemes behave the same, because there are no more regions that can be reached via the optical mesh. As in the one-region scenario, there are two IP/MPLS transit nodes capable of driving the whole region traffic (1+1 protection 50% capacity dimensioning), the states can be defined as follows:

- 2 Active routers: No service affected
- 1 Active router: No service affected
- 0 Active routers: Services affected

The model with transition between states for the one region case is depicted in Figure 3.

Applying Markov's model, the resulting expression for availability is presented in eq. (4). The result in this case is the same for the MLSBR approach than the 1+1 protection scheme, because the availability resources for both cases are the same. Two transit routers at 50% capacity can drive the traffic in case of one failure, but in case of double failure, the service is affected.

$$A = 1 - \frac{(MTTR/MTBF)^2}{1 + 2 \frac{MTTR}{MTBF} + \left(\frac{MTTR}{MTBF}\right)^2} \quad (4)$$

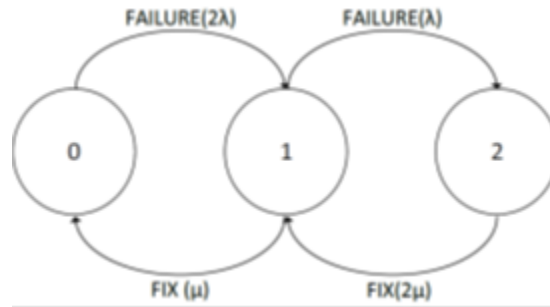


Figure 3: One region case

3.1.2.2 N regions case

The problem is generalized using equation (6), where N is the number of regions with duplicated transit routers and k is the relation between MTTR and MTBF (equation (5)). With this expression, multi-layer restoration network availability can be calculated.

$$k = MTTR/MTBF \quad (5)$$

$$A(N) = 1 - \Pi_0(N) \left(\sum_{i=n+1}^{2n} (k)^i \left[\frac{2N!}{(2N-i)! i!} \right] \right) \quad (6)$$

3.1.2.3 Availability comparison of the schemes

Table 1 shows the MTTR for protection and MLSBR schemes assuming a MTBF of 3 years in the IP routers for the scenario presented in Figure 1 with seven locations for transit routers. Based on the results, MLSBR allows increasing the MTTR for the same availability. This means that OPEX can be reduced using this protection scheme.

Table 1: Comparison between MLBSR and Protection in terms of MTTR (days)

Availability		99,99%	99,999%	99,9999%
Number of Backup Routers	2	33	14,3	6,7
	3	59,2	31,9	17,6
	4	86,6	51,6	31,9
	5	110,7	72,4	47,2
	6	132,6	91	63,6
Protection		11,1	3,4	0,1

3.1.3 Impact on CAPEX Reduction

The MLSBR concept is proved in the Core Telefónica Spanish Network (Figure 4). It shows the optical national mesh with the transit and interconnection nodes. The transit routers are shown co-located at the same location of some optical transit nodes.

This network has the structure the exposed in Figure 1. It is composed by 6 interconnection routers and 14 transit routers. The layer 3 connectivity, between transit routers, is a ring and star with the core in Madrid. It has been modeled 3 point of connection between transit and interconnection layers, two in Madrid and one in Barcelona. Interconnection and transit routers in Madrid and Barcelona are collapsed on two routers in Madrid and Barcelona.

Let us highlight that for this study only the two upper levels have been taken into account for the numerical results. One port is required in the transit and interconnection routers. The savings in the access level depend on the number of connections between the access and the transit routers (which in turn are dependent on the traffic volume and the capacity of the ports). With this technique just an extra port in the access routers is required to avoid a single point of failure (instead of one for each connection between access and transit routers in the 1+1 protection schema). However, these savings are independent of the number of SBRs.

This numerical study has used the traffic demands of 2012 and a traffic growth per year of 35%, in order to evaluate the same network in five years (2017).

The network dimensioning is done using the dual-plane protection approach using the 20 nodes network in Figure 4.

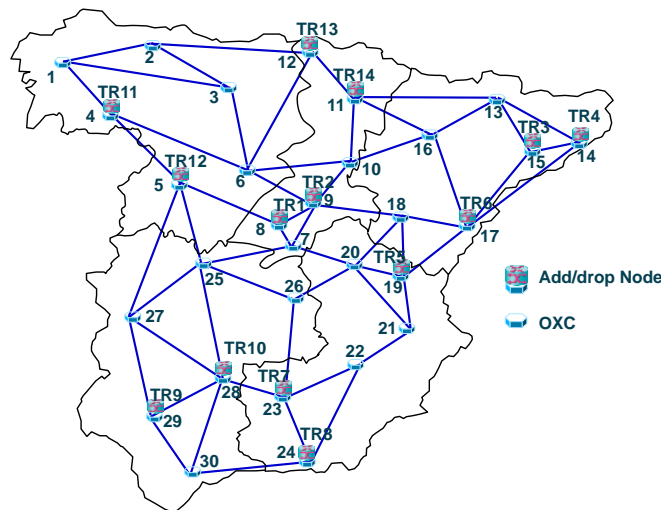


Figure 4: Telefónica Spanish optical mesh

For the MLSBR mechanism, the dimensioning process is done just for one plane (10 nodes in the network). IP layer is dimensioned with a maximum occupation of 80% in case of any failure in the network. The number of SBRs can vary based on how many node failures the network is protected. Depending on the number of SBRs, the number of IP ports is obtained using MLSBR. The results of the IP-ports savings of compare the dual-plane protection dimensioning approach versus the MLSBR approach are presented in Figure 5.

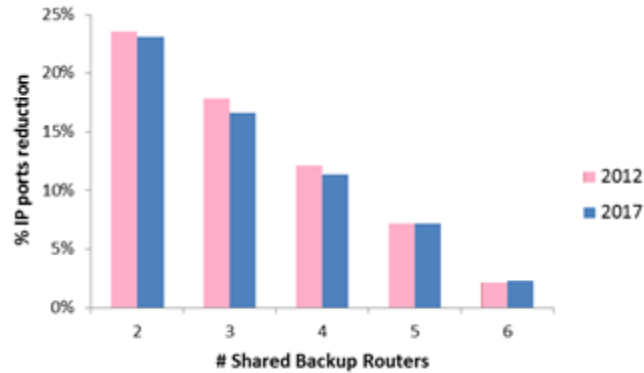


Figure 5: MLSBR Simulation results of network dimensioning

In light of the results, it is seen that the savings by introducing two SBRs it is obtained almost 24% of savings in the number of IP ports needed to deploy. The percentage decreases as the number of SBR grows, but the savings are conserved in 2017. If there were 7 SBRs, there would be the same number of IP routers than in the dual-plane protection case, as there are seven transit routers.

Finally, let us highlight that not only the ports, but also the chassis of the routers are reduced. In this scenario, there are 14 transit nodes and 6 interconnection routers. If we have 5 SBRs, MLSBR reduces from 14 transit routers to 12 (7 to carry the traffic and 5 for backup purposes). This means a reduction of 14.28% in routers. The maximum savings for this scenario are when 2 SBRs are using, which leads to 35.7% savings. Instead of using 14 transit routers, 9 routers are used (7 routers for normal operation and 2 backup routers).

3.2 Impact on OPEX reduction

The Operational Expenditures (OPEX) is the amount of money that network operators spend on an ongoing, day-to-day basis in order to run its business.

The operational expenditures for an operator can be divided in to seven general categories, network operation, interconnection and roaming, marketing and sales, customer service, charging and billing, IT and general support and service development. In particular, **the OPEX category where the IP over Optical Integration can play a significant role is the Network operation**, which includes OSS operation, maintenance and repair of the network elements, equipment and software licenses, rental of network resources, costs for site rental and electricity.

First, in order to have an insight on the order of magnitude of the Network operation related OPEX, Yankee Group estimates that for a fixed line operator, the network operation takes a significant part of expenses, accounting for 39% of the total OPEX.

The advantages in terms of OPEX savings of IP/optical coordination by SLA relaxation are analysed in this section. The repair model used in this analysis is described in detailed in section 8.1.1 .



3.2.1 SLA relaxation by Multi-layer Restoration

The survivability of the network is enhanced when the User-to-Network interface (UNI) can be used to request connectivity on demand. Nowadays, protection mechanisms at IP layer enter into action whenever there is a failure at the MPLS layer. This issue is now considered as a severity 2 issue, very urgent, because the redundancy is lost and there is a risk that, in case of a second failure, there are service losses affecting the business. That second failure would be considered severity 1 that is emergency, as there is a service loss. With the multi-layer re-route, it is possible to recover from an exceptional second failure. Thus, in case the second failure happened, the situation would be automatically solved, and the problem would be still of severity 2 (instead of 1), relaxing the Technical Support and hardware replacement requirements.

For the case study presented in 5.4.6, it has been shown that the minimum Time to Repair in order to have 99,999% availability would be less than 3 hours. There is no Technical support level available that guarantees such time to repair. However, just adding 2 additional backup routes, a Silver Minus contract is needed.

3.2.1.1 Objective MTTR for a desired availability

A set of simulations has been done for the Telefonica of Spain Reference Network, which includes about 2000 network elements (IP/MPLS routers, ROADMs, amplifiers, etc.). Two survivability mechanisms have been compared, in order to get an insight on the OPEX in terms of Technical Support Services needed to achieve a desired availability. The Mean Time Between failures was set to 5 years. The first mechanism is 1+1 protection at IP/MPLS layer. The second procedure considers the multi-layer control plane mechanism analyzed in this document, taking advantage of the UNI interface. In this case, after a second failure, it is possible to recover IP/MPLS traffic through new transport connections, increasing the availability.

The results are shown in Figure 6. In this figure the Minimum Mean Time to Repair for a given availability is presented. As seen before, a MMTTR can be translated to a requirement of Technical Support Level Needed and Hardware Support Service (damaged equipment repair/supply). A summary is presented in Table 2. It can be clearly seen that, for an availability of 5 9s, it would be required with current protection scheme a Gold Technical Support service (the second most expensive type of support). In comparison, using Multi-layer Restoration Schemes, for the same target availability (99,999%), the Minimum MTTR is extended and just a Bronze Technical Support is needed (Bronze is the second cheapest service level in Telefonica).

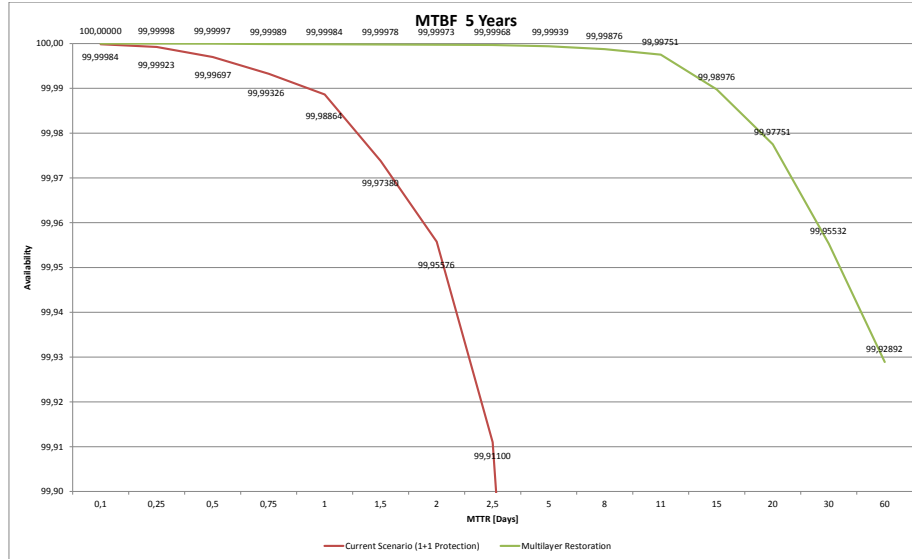


Figure 6 Availability vs MTTR (Spain Network Scenario)

Table 2 Required Technical Support Level Current vs MLR

Scenario	Availability	MTTR min	Required Technical Support Level	Required Spare Parts Repair Support
Current	99,9999%	Unrealistic time	N/A	-
Current	99,999%	12 hours	Gold	24x7x4 swap
Current	99,99%	18 hours	Gold minus	24x7x4 swap
MLR	99,9999%	12 hours	Gold	24x7x4 swap
MLR	99,999%	5 days	Bronze	NBD swap
MLR	99,99%	11 days	Iron	7 BD swap

With regards to the type of Hardware Support Service (spare parts repair) needed, it is possible to reduce the service, for example, in the case of 4 9s from a 24x7x4 swap to a Next Business Day swap contract (much cheaper).

3.2.2 Reparation WorkForce Model

In order to be able to quantify the Operational Expenses related to network failures reparation we propose a Reparation WorkForce (RW) model that obtains the number of repair teams needed, that is the work force devoted to reparation. The proposed RW

model is based on a centralized location of the repair teams. Each repair team is specialized in failures of a given network layer and can only repair one network element at a time. In sake of simplicity the model assumes a fixed repair time (RT) per network element, measured since the repair team is assigned to a failure and is independent on the failure location, that is, travel times to the failure location are not considered. It is important to highlight that the RT is measured from the moment that the repair team starts working on the failure (not the moment of the failure). Optionally, location and travel time can be considered. In this case, the RT can be calculated as a travelling time (depending on the location) plus actual repair time. The travelling time is the time that the repair team needs for going from their base location to the failure location. The actual repair time is the time that the repair team needs to work on the failure.

In order to guarantee a given Service Level Agreement or a given Network Availability, we introduce the notion of Maximum Time to Repair (MTR). This is the actual time by which the failure must be solved, measured since the time when the failure is notified. Based on this MTR value, a repair team could delay the start of the repair process provided that the MTR is satisfied. As is shown later, this optional delay can help to minimize the number of repair teams, as the same team could work on several nearby faults or close in time failures or operations.

With the intention of understanding better the relationship between RT, MTR and number of repair teams, Figure 7 below illustrate an example in which two failures happen in the network, and the second failure occurs before the repair team working in the failure can finish their task.

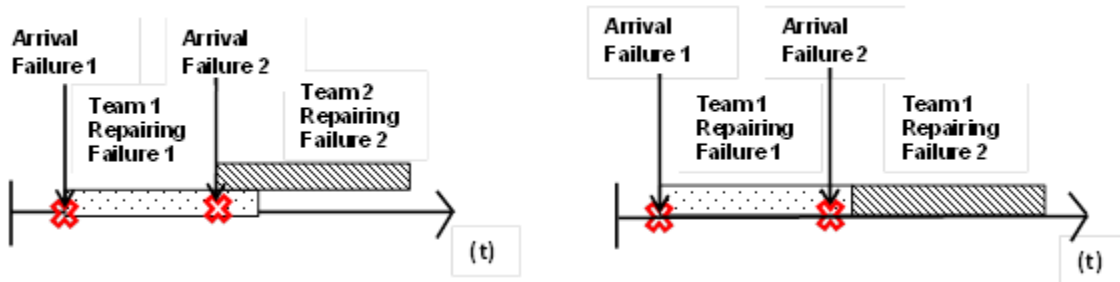


Figure 7 (a) Urgent Reparation

(b) Delay Allowed

Figure 7 (a) shows one example in which the reparations are very urgent and no delays are allowed (the maximum repair time equals the repair time). Thus, when the second failure arrives, a second repair team would be needed, as with only one the actual repair time would exceed the MTR. However, in the case of Figure 7 (b), the reparation is not so urgent, that is maximum repair time is bigger than the repair time. In this case it is observed that the beginning of the reparation of the second failure is delayed until the team repairing the first failure finishes. Therefore just one repair team would be necessary to repair the two failures, saving thus one repair team.

The examples in Figure 7 assume that repair teams are always available (at any time). However, each repair team has certain availability. The two most common models in the market are 8x5 (8 hours per day, 5 days per week) and 24x7 (24 hours a day, 7 days per



week). In the 8x5 it is assumed that the repair teams are only available from Monday to Friday, eight hours a day. In the 24/7 it is assumed 3 shifts on weekdays (Monday to Friday) and a Weekend shift. Based on this model, given a certain mean time between failures, a repair time, a maximum repair time, and the number of equipment subject to failure, the number of repair teams is obtained to guarantee the compliance of the repair time 100% of the time. Then, iteratively, the process is repeated but eliminating one repair team each time. For each number of team, the availability in terms of percentage of reparation of failures when the guarantee is assured

Let us assume a multi-layer network in which failures can happen in the one of the network layers. We take the results of [9] which obtains the maximum mean time to repair to guarantee a certain service availability with different survivability schemes, specially protection and multi-layer restoration. This paper follows the same network scenario, an IP/MPLS over Optical backbone topology, with a mean time between failures of 5 years. The repair work time has been fixed to 6 hours. Two maximum repair times are considered, 12 hours, that guarantees 99,999% service availability with single layer protection and 5 days, which guarantees the same service availability for multi-layer restoration. This case study has been performed for the two different Repair Team availabilities previously mentioned, 8/5 and 24/7.

The results obtained are, for each number of repair teams needed, the percentage of successful repair cases when the maximum time to repair is achieved. Simulations are performed for periods of time of 10 years generating random time between fail.

Figure 8 shows the number of reparation team versus availability for the four different shifts of the 24/7 availability model when the mean time to repair is 12 hours versus 6 hours needed to fix the failure. In this case the failure reparation can be delayed until 6 hours.

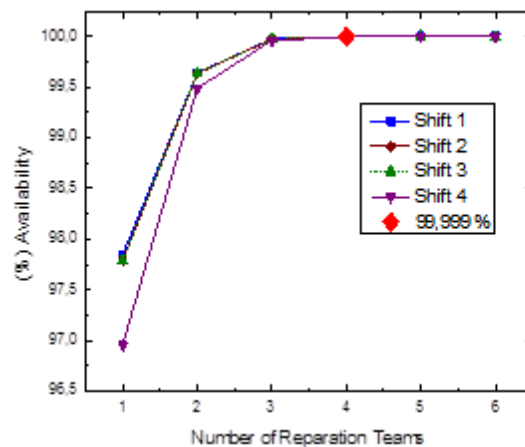


Figure 8 Single layer protection 24x7 model

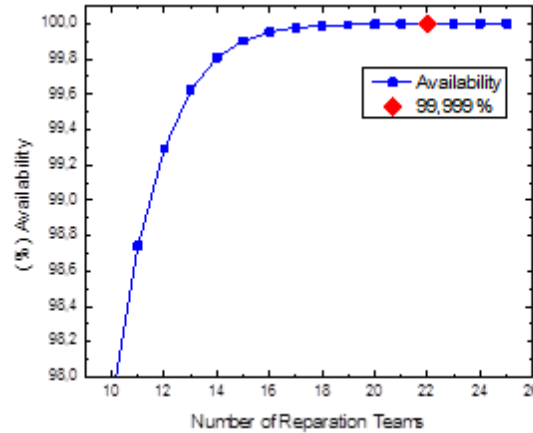


Figure 9 Multilayer restoration 8x5 model

In this model (24x7), if a failure is not completely fixed for the repair teams of a shift, the teams which are working in the next shift must be continue with this reparation.

The results show that around 16 reparation teams (4 teams of each shift) are needed to achieve the maximum repair time guarantee in 99,999% of the failures in case of single layer protection.

When Multilayer restoration is used the maximum repair time is relaxed to 5 days which allows to delay much more the reparation that in the case of single layer protection. According with the same procedure, in this case, it is obtained that just one reparation team is necessary for each shift to accomplish with 99,999% (not shown in the graphs) which sums 4 reparation teams.

On the other hand, the same study has been performance for the model 8x5 based on multilayer restoration (MTR=5days). The number of reparation teams has been calculated again in order to accomplish with certain availability. In Figure 9 it can be observed that in this case, at least 22 reparation teams are necessary to reach 99,999% availability.

Looking at the above results, definitively we can assume that is much better a model which covers every hour of the day divided in different shifts than a model which just covers 8 hour a day with just one shift.

Operators and service providers typically consider two main models for the repair costs, a flat fee model with a fixed monthly cost during the year, which includes the salary of the repair teams, and event base model, used when the repair teams are shared between operators.

3.3 The role of the Adaptive Network Manager

Even though from an analytical point of view this mechanism can offer savings to the network operators, there are some control plane requirements, which are fulfilled in IDEALIST control architecture that must be solved to see this solution deployed in real network scenarios. IDEALIST Adaptive Network Manager is an effective alternative to solve routing and path computation in a multilayer scenario composed by an IP/MPLS network over an elastic optical transport network. This element has been validated to operator in IP/MPLS services provisioning in [10] using IETF ABNO architecture. An SDN



controller, such as the Adaptive Network Manager proposed in IDEALIST, must program the back-up paths so any router in case of failure knows which UNI path to establish. In addition, the SDN controller may need to program the backup router differently depending on which transit router has failed (for example, the routing metric of the links should mimic the metric of the original links).

This means that network operators can reduce their CAPEX and OPEX using this multilayer restoration. However, there are some control plane requirements that have to be fulfilled. Following table summarizes the approaches and defines requirements to take into account for possible deployments.

	Advantages	Drawbacks	Requirements
Original planning	<ul style="list-style-type: none">• Simple operation• Traffic restoration in less than 50ms with FRR.	<ul style="list-style-type: none">• Resource duplication.	<ul style="list-style-type: none">• By-pass selection is required to reduce the cost of this approach.• FRR to minimize restoration time.
FRR+Optical Restoration	<ul style="list-style-type: none">• Reduction of IP layer investment (17% in 2012 or 32% in 2017).• Extension of MTTR after a link failure.	<ul style="list-style-type: none">• Optical restoration takes around 1 minute.	<ul style="list-style-type: none">• Optical mesh• GMPLS enabled in the optical mesh.• FRR to minimize restoration time.
FRR+Optical protection	<ul style="list-style-type: none">• Reduction of IP layer investment (17% in 2012 or 32% in 2017).• Extension of MTTR after a link failure.• No traffic cut after failure.	<ul style="list-style-type: none">• Transponders are duplicated to have protection.• Lightpaths are duplicated to have two separate paths for each IP demands.	<ul style="list-style-type: none">• Optical mesh• FRR to minimize restoration time.
MLRR+Optical Restoration	<ul style="list-style-type: none">• Minimize investment in IP ports (29% in 2012 or 40% in 2017).• Extension of MTTR after a link failure and a port failure.	<ul style="list-style-type: none">• MLRR takes around 1 minute. It is limited by optical restoration time.	<ul style="list-style-type: none">• Optical mesh• GMPLS enabled in the optical mesh.• UNI enabled.• Back-up routes pre-loaded in routers.• FRR to minimize restoration time.
MLRR+Optical Restoration +MLSBR	<ul style="list-style-type: none">• Minimize routers investment in chassis and ports.• Extension of MTTR after a link failure, a port	<ul style="list-style-type: none">• MLSBR takes around 1 minute. It is limited by optical restoration time.	<ul style="list-style-type: none">• Optical mesh• GMPLS enabled in the optical mesh.• UNI enabled.



	failure and node failure.		<ul style="list-style-type: none">• Adaptive Network Manager (Back-up routes calculation and transit back-up routers configuration).• FRR to minimize restoration time.
--	---------------------------	--	--

4 Techno-economic and energy analysis

4.1 IDEALIST building Blocks

The IDEALIST architecture building blocks used for the techno-economic analysis described in this deliverable are based on the mid-term solution (target 2018 for first deployment) proposed by WP2 for SBVT and Flex-OXC.

In particular, the SBVT is based on the model described in D2.2, with some more specification developed by WP2 partners during the plenary meeting in Bristol (May 2014) and further refined during last months. The general SBVT building blocks and multiplexing architectures are summarized in Figure 10 and Figure 11.

The SBVT comprises a configurable/sliceable flexible OTN interface module for client mapping from the OTN fabric into ODU_{Cn}/OTUC_n, with a net capacity that varies from the 100G OTUC₁¹ to the maximum of 1.2Tb/s OTUC₁₂ for each SBVT module, with steps of 100G. The 100G modularity has been chosen according to the standardization evolution, both for OTN and Ethernet layers, to 100 Gb/s and multiples of it (for example the 400G). The sliceability of the module manifest itself in that the total net capacity can be dynamically partitioned to serve different and independent clients at the same time, ranging from a minimum of 1 (with a maximum allowable net capacity of 1.2Tb/s) to a maximum of 12 (100Gb/s each) through all intermediate combinations.

The Configurable/Sliceable Flexible OTN interface module can flexibly fragment high-data rates (e.g., 1Tb/s) into lower-data streams (e.g., 600Gb/s plus 400Gb/s), if it is impossible to fit the whole client traffic into a single super-channel (for example due to unavailable contiguous spectrum along the path). An example is shown in Figure 11 the traffic from a client at 1Tb/s is divided by the OTN layer into two different streams, each one feeding two separate media channels (#1 and #2). Similarly, the traffic from an interface at 200 Gb/s is divided into two different streams, one complementing media channel #2 and the other one feeding media channel #3. Traffic could be also not fragmented as in the case of the other 200 Gb/s interface and the 100Gb/s interface in Figure 11.

¹ Actually standard OTU4 interfaces already cover 100G capacity, but in a “static” way; within IDEALIST project, we suppose that a flexible interface should also cover 100G traffic demands and not only higher bit rates ones as is proposing ITU-T, so we have extended the OTUC_n terminology to include the $n=1$ case not to exclude 100G single vendor proprietary intra domain interfaces.

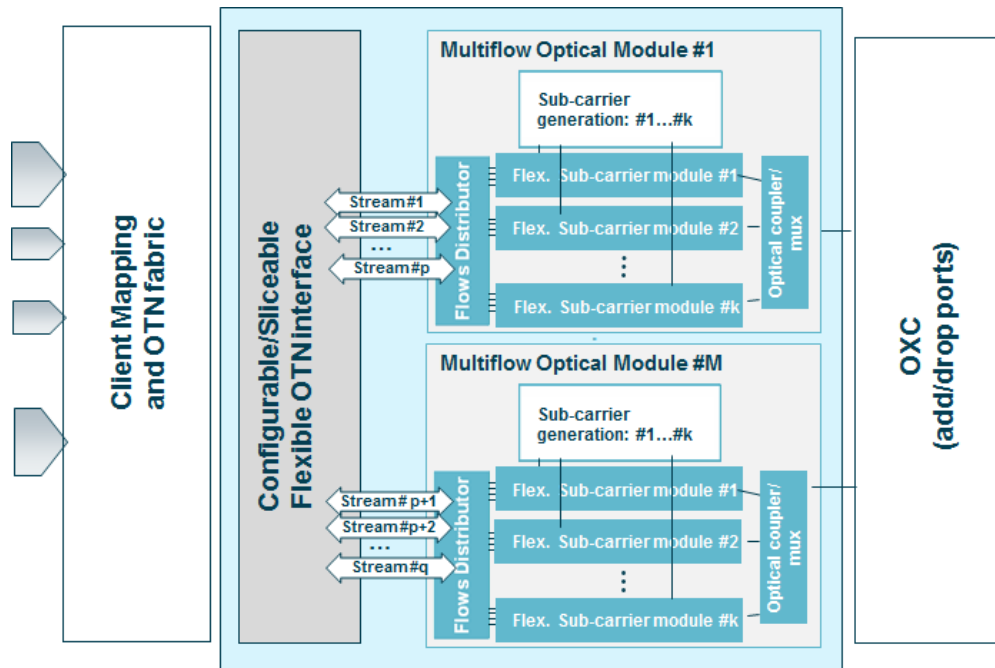


Figure 10 : SBVT general building blocks architecture

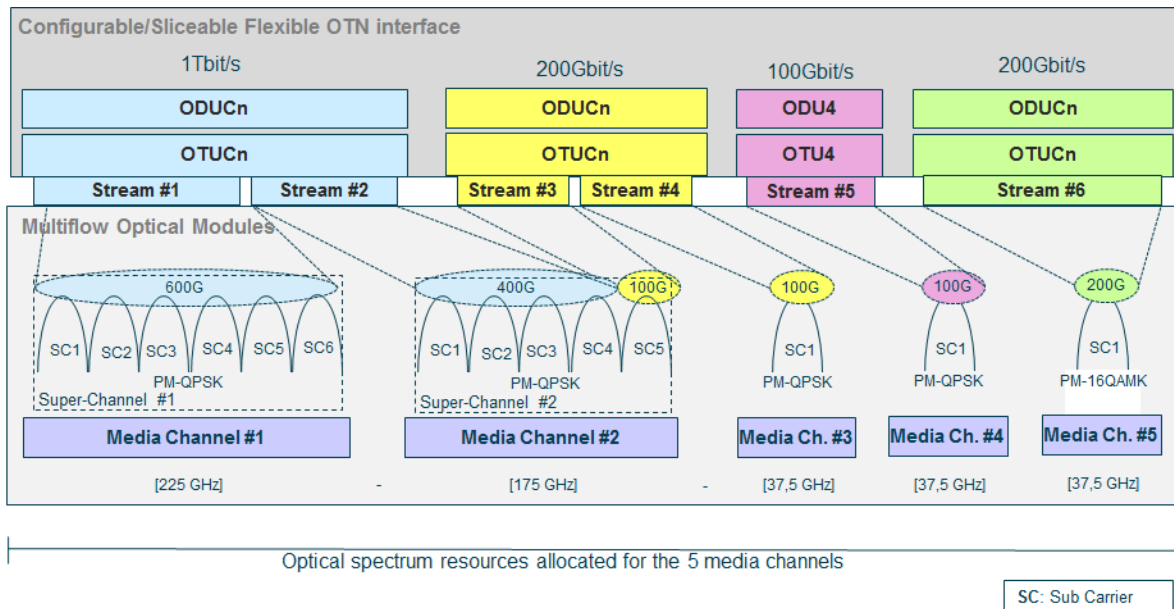


Figure 11 : General multiplexing architecture on SBVT

Then, the generation of super-channels is in charge of the Multi-flow Optical modules (Figure 10). Each Multi-flow Optical module generates multiple ($k = 2$ or 4 for medium-term SBVTs) sub-carriers that can be distributed among different super-channels. The p independent streams from the OTN layer are fed to a Flow Distributor module which



distributes each OTN stream into one or more Flex Sub-carrier modules depending on the stream capacity, reach and path along the network.

Each Flex Sub-carrier module modulates a sub-carrier wavelength with a specific portion of the traffic stream. It encodes data, shapes the spectrum and, possibly, pre-distorts the signal. Digital-to-analogue converters (DACs) are exploited to adapt the modulation format (e.g., providing multi-level signals for 16QAM) and IQ modulators to modulate the sub-carrier at a specific baud-rate with orthogonal polarizations. The bit rate and bandwidth can be tuned by changing the modulation format, the baud-rate, the spectral shaping and the coding. FEC with different gain and overhead is suitably added. At the receiver side each sub-carrier is independently demodulated and decoded by a coherent DSP based receiver. All generated sub-carriers are first logically grouped into super-channels feeding different media-channels and optically coupled together into a single add/drop port of the source Optical Cross Connect (OXC) node Figure 11.

Three modulation formats are considered: PM-QPSK (1.2Tb/s with 12 carriers), PM-8QAM (1.2Tb/s with 8 carriers), PM-16QAM (1.2Tb/s with 6 carriers) and a reference net baud-rate of 25 Gbaud/s (but a tunable baud-rate ranging between 25 and 33.3 Gbaud/s is also considered as an option). Additionally the PM-32QAM modulation format (1.2Tb/s with 4 carriers at 30 Gbaud/s) is to be considered too, although it is not yet clear its application context.

As already stated the maximum total net capacity supported by the SBVT amounts to 1.2Tb/s, and accordingly a modular solution is adopted in terms of Multiflow Optical Modules -with a smaller number of carriers than the expected total of 12 carriers needed for maximum sliceability and flexibility of the SBVT- to be equipped incrementally in a “pay as you grow” fashion. This in order to prevent as much as possible the waste of resources in case of combination of input clients demands that reach the total electronic capacity of 1.2Tb/s but are mapped in less than 12 optical carriers (e.g., 4 clients at 300Gb/s mapped each into a two sub-carriers PM-8QAM super-channel), or simply do not reach the total capacity of the SBVT. Likewise, modularity is needed at the client side to equip different and variable combination of client interfaces: this is particularly true if the SBVT is a muxponder and not a flexible line side interface of a big OTN DXC.

Referring to Figure 12 an SBVT is considered composed by three blades dedicated respectively to the client feeding, to the core electronic for organization fragmentation and distribution and to the housing of up to six Multiflow modules.

Two sizes for the Multiflow Optical Module are considered:

- 2 sub-carriers module, which support up to two independent flows: 6 modules are necessary to reach the total flexibility of 12 carriers;
- 4 sub-carriers module, which supports up to four independent flows: 3 module are necessary to reach the total flexibility of 12 carriers.

Mixing modules, several combinations are allowed in terms of flexibility, total traffic and simultaneous use of different modulation formats for demand needing different reaches.

However, the limitations in terms of a maximum number of equipable Multiflow Optical Modules (e.g. 6 for each SBVT) and the maximum number of Flexible Sub-carrier modules within each module (e.g. 4 for each Multiflow Optical Modules) implies that some client demands must be subdivided into more than one Multiflow Optical Module inside the SBVT. For example, if we consider in Figure 11 the 1Tb/s demand (divided into 600Gb/s Stream #1 and 400Gb/s Stream #2), together with the 200Gb/s demand (divided into

100Gb/s Stream #3 and 100Gb/s Stream #4), all based on 100G PM-QPSK sub-carriers, we need an SBVT equipped with three 4 sub-carriers Multiflow modules, the first one completely dedicated to Stream 1, the second one with other 4x100 sub-carriers (2 serving the last part of Stream #1 and the first part of Stream #2) and the third one with 2x100 sub-carriers serving the second part of Stream #2 and the last two sub-carriers dedicated respectively to Stream #3 and #4. Note that, in principle, the fragmentation of demands among media channels and the spreading among different Multiflow Optical Modules are completely independent of one another: the first is governed by the size and allocation of media channels in the optical spectrum, while the latter is correlated to technological choice in terms of modularity and size of the modules.

An important concern in SBVT design is the number of line ports that can assure in the future both proper traffic aggregation and cost optimization. In WP2 it has been proposed that each Multiflow Optical Module will have only one output port to collect, through an optical coupler/mux, the Flex Sub-carrier modules spectrums and interconnect them to the OXC. This choice derives by the expectation of future big single client demands (several Gb/s to 1Tb/s or more) necessarily requiring super-channels, i.e. many carriers following the same path along the network: consequently a single port should allow a significant saving in the OXC add/drop contentionless ports. The drawback of this solution is that it is not possible to reuse the same lambda for different demands within a single Multiflow Optical Module, as discussed later dealing with the Flex-OXC. A discussion on the implications of the baud rate values in the use of superchannels for nx100G client rate mapping, and assuming that the baud rate could take values within a bounded range, is given in the subsection of the appendix 8.6.1.1.

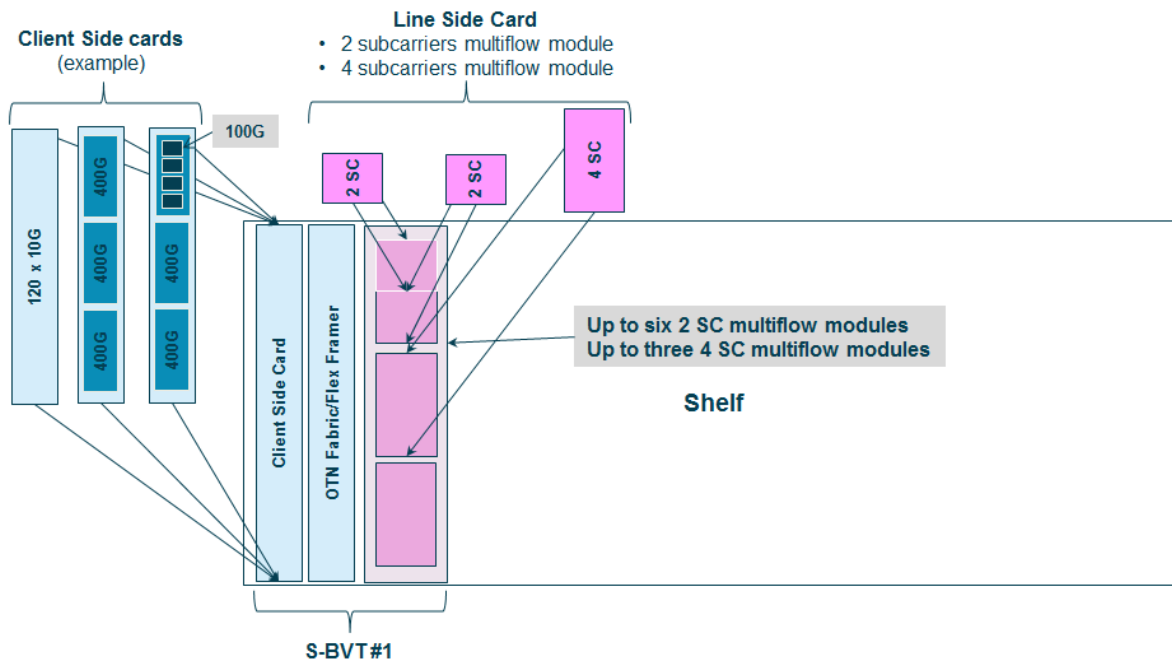


Figure 12 Maximum vs. minimum Multiflow Optical modules numbers w.r.t. maximum number of Flexible Sub-carrier modules for each SBVT

Regarding the Flex-OXC architecture, today available solutions are mostly considered for the techno-economic analysis. In the appendix (Sec. 8.6.4.2) a reference model for a state of the art flexgrid ROADM is reported.

Looking at mid and long term, IDEALIST project is developing a model for the optical node that goes over the concept of the state of the art OXC that is based on ROADM. The IDEALIST Node model assumes flexibility, scalability, resilience and adaptability as reference requirements.

Figure 13 describes the IDEALIST approach towards developing the proposed new Flex-OXC node architecture, allowing different configurations based on the internal organization of a number of functional components. An Architecture on Demand (AoD) based Flex-OXC node is proposed together with a Switchless Elastic Rate Node (SERANO) module, which is an alternative to the wavelength switching architectures.

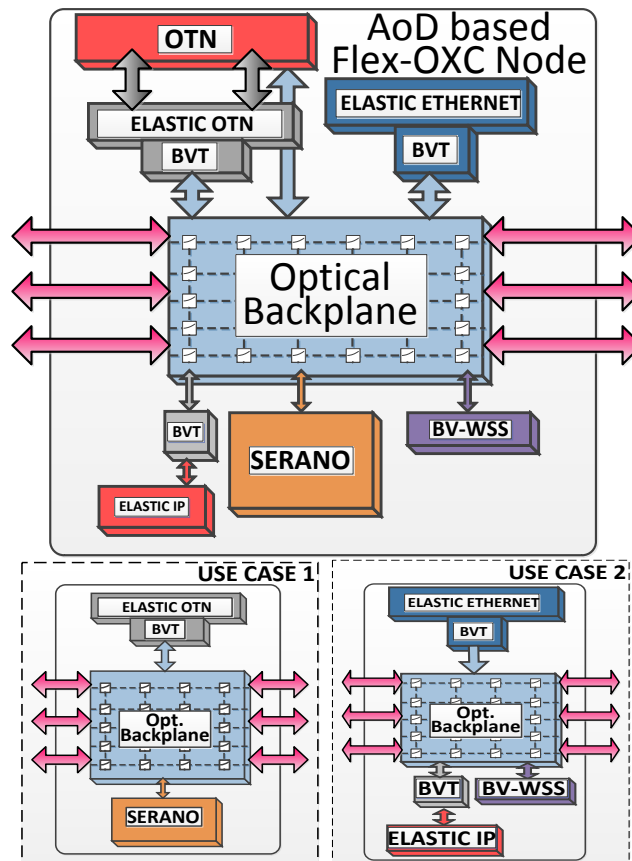


Figure 13 Node adaptability to some examples of different network applications

Two examples of adaptability to different network applications that this architecture can achieve are also shown in Figure 13. For instance, if the AoD based Flex-OXC node is considered within an optical transport network (OTN), network traffic demands are directed through a bandwidth variable transponder (BVT) to the optical node backplane, which in turn can lead this traffic to a SERANO module. On the other hand, if an Ethernet infrastructure is considered, an elastic BVT interface will interconnect demands coming



from Ethernet switches to an optical IP elastic layer with a router equipped with BVT optical interfaces through the optical backplane. Moreover, this elastic Ethernet and/or IP layer could be directly connected to a geographical Metropolitan Area Network (MAN) network by means of a Bandwidth Variable WSS (BV-WSS). Basically, the AoD based Flex-OXC node consists of an optical backplane (e.g. a large port count 3D-Micro-Electro-Mechanical (MEMS) cross-connect), connected to several signal processing modules, such as BV-WSS, Erbium-Doped Fiber Amplifiers (EDFAs), SERANO modules and the node's inputs/outputs. This extends the architectural framework that has been previously reported in [113], [114] to efficiently support AoD based Flex-OXC nodes and EON operations.

4.2 Techno-economic analysis of IDEALIST building blocks

The work presented in this section identifies the most cost effective data plane technologies and node architectures to be implemented in WP4. It deals with the techno-economic benchmarking of key data plane technologies identified in primarily in WP2 during the first year. This is the comparison of a fixed-grid versus a flex-grid evolution model quantifying the number of WSS needed. Then the benefit of SBVTs is studied in comparison to that of BVTs as a function of the traffic matrix and the aggregation level obtained by using different sizes of IP/MPLS transit network. Furthermore, this section represents an update of D1.3 performance analysis including costs and power consumption figures [51]. Finally, this section ends with an analysis of the CAPEX impact of fixed vs. flex-rate modular line interfaces in multi-period planning scenarios.

4.2.1 Fixed vs Flex-grid evolution models

This section quantifies the number of WSSs and transponders of different types following the fixed and the Flex-grid evolution models.

In order to compare which is the most appropriate evolutionary strategy, this section quantifies the equipment necessary in every year of the network. The performed network planning obtains the equipment necessary in the network in order to carry out the full traffic demands. The considered equipment is the total number of WSS required in each network evolution model and the number and type of transponders required to support the optical channels.

Figure 14 shows the number of WSSs used for the express switching (there is one WSS per ROADM degree) and the number of WSSs for the ADD/DROP chains. It can be seen that, the SSON evolution model that considered non flexi-grid capable ROADMs lead to a high number of WSSs for the express switching, as it is deployed a parallel network. Considering the number of WSSs, all the strategies follow the same pattern until 2019. At this point, the SSON non-flexi ready case needs to make a big investment due to the new parallel SSON network. However, in the long term, the WSON evolution strategy needs more WSSs in the add/drop chains than both SSON models, as it is needs to keep using a high number of transponders of 100 Gb/s, while the SSON models can use less transponders (but with higher capacity), as shown in Figure 15.

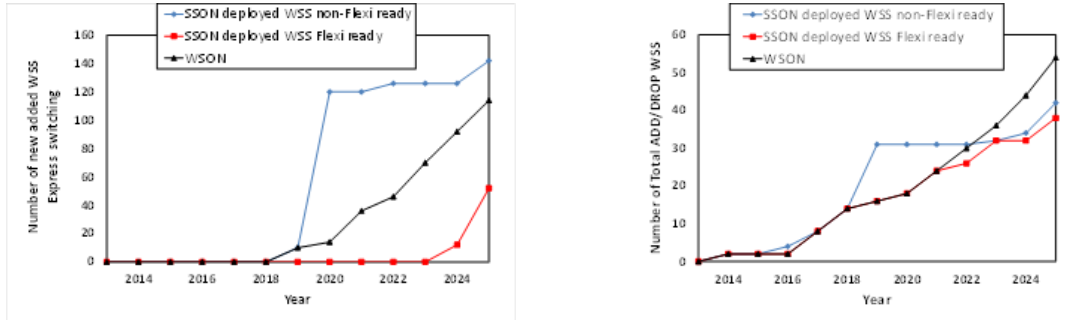


Figure 14 a) Number of 1x9 WSS in express switching (one WSS per ROADM degree); b) number of 1x20 WSS in the add/drop chain. Note: the ROADM configuration model has been obtained from [9].

The number of transponders is shown Figure 15.

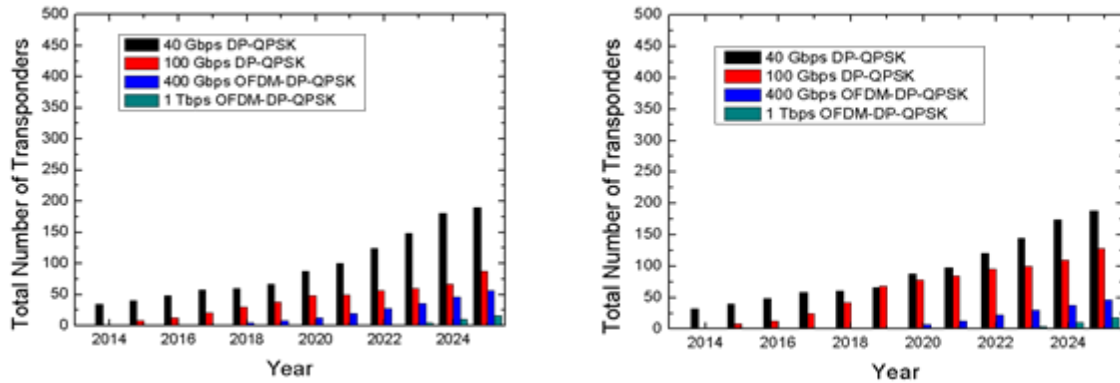


Figure 15 Total number of transponders: a) SSON evolution model with current deployed WSS Flexi-grid ready; b) SSON evolution model with current deployed WSS non-Flexi-grid ready.

If high-speed transponders achieve a high level of cost-efficiency, compared with the 100 Gb/s transponders, the cost to upgrade the network after 2020 using Flexi-grid will be lower, even considering the worst-case where new Flexi-grid equipment needs to be deployed.

4.2.2 SBVT OPEX analysis

Operators of telecommunication networks have to deal with an ever-increasing demand for new services. Competitiveness has grown in the last years, bringing many troubles to operators to maintain a sustainable business, and hence, cost reduction, both in Capital Expenses and Operational Expenses is a key issue. Emerging technologies can help



network operators to cope with the demand and keep both CAPEX and OPEX under control.

One of the emerging technologies which has received lot of attention from research community is the Elastic Optical Networks (EONs) which allow the use of optical spectrum in more flexible way, for instance, by changing the frequency slot granularity and central frequency granularity. In particular, one of the main novel blocks of EON is the Sliceable Bandwidth Variable Transponder (SBVT). This kind of transponder enables transmitting from one point to multiple destinations, changing the traffic rate to each destination and the number of destinations on demand. Different studies in the literature show that SBVTs can achieve CAPEX savings in network operators [32]. However, operational expenditures savings by SBVTs has received little attention. OPEX is the amount of money that network operators spend on an ongoing, day-to-day basis in order to run their business. In particular, the OPEX category where SBVTs can play a significant role is the network operation, which includes costs related to maintenance of the network.

Hence, the goal of this work is to quantify the reduction of network maintenance and reparation related OPEX by using SBVTs instead of a set of fixed rate transponders. Part of the cost of maintenance and reparation is related to keeping a stock of spare parts. Whenever there is a failure in a network element, there may be damaged parts that need to be replaced and a stock of spare parts for replacement needs to be maintained. Such stock can be maintained either by the network operator or a third party supplier and translates into a yearly cost.

In the case of optical transport networks, spare transponders need to be stocked. Following current model, transponders of different rates are needed, and thus a number of each rate needs to be stored. However, with the newly proposed Sliceable Bandwidth Variable Transponders, it is possible to reduce the number of transponders and reduce the variety of transponders. This study aims to analyze whether equipping a network with Sliceable Bandwidth Variable Transponders instead of fixed rate transponders of multiple rates reduces the needs in terms of stock maintenance.

4.2.2.1 Procedure to calculate the number of required stock of spare parts in a communication network

This study is based on a centralized stock model. It is assumed that a central warehouse has stored all the spare elements (e.g. transponders) and in case of a damaged part event, such part is shipped from the central location. The procedure described in this section can be used to calculate the number of required stock of any component subject to fail, but it has been focused on transponders for a nation-wide optical core network.

The procedure uses as a starting point the number of transponders required in the network for normal operation for two cases, one with different kinds of fixed rate transponders (40, 100 and 400 Gb/s are used, and a second case with sliceable BVT (400 Gb/s for the results). The Mean Time Between Failures (MTBF) of transponder is fixed and equal for all transponders. Then, the second step is performing a set of simulations where the transponders are set to fail according to an exponential distribution with mean MTBF. Every time an element fails, it is replaced and a new element is requested to factory. The mean delivery time (MDT) is the time to receive a new transponder from the vendor. Then, it has been obtained the number of stock that needs to be maintained to guarantee certain availability.



The steps of the procedure are summarized below:

1. Number of transponders needed for normal operation is obtained.
2. Failures in every transponder are distributed randomly in time based on an exponential distribution with mean MTBF.
3. At the same moment that failure happens, a spare transponder is taken from the stock and a replacement request to factory is made.
4. Stock accounting:
 - i. One is added to the stock counter of the given transponder if a failure happens.
 - ii. One is subtracted to the stock counter when the replacement happens (a new transponder arrives from factory to the stock).
5. From the previous steps, the peak value of the minimum stock number to be maintained at the warehouse is obtained.
6. Once the maximum stock is known, steps 2-4 are repeated, reducing the maximum stock in one each time, with one remark:
 - i. When a failure happens, if the stock is zero, one is added to the number of failed cases. Otherwise, one is added to the successful case.
7. The percentage of the successful cases is obtained for each stock value, until the stock is one. Given that the time between failures is a random variable, the previous steps are repeated several times in order to achieve a higher accuracy.

4.2.2.2 Results

This study has been performance for two different IP topologies of the Spanish IP/MPLS over Optical Core Network. The first scenario represents a full mesh topology of IP/MPLS nodes. The second scenario is based on current IP/MPLS topology and traffic is routed through the shortest path between IP/MPLS nodes.

TxP parameters	Cost
40Gb/s, 2500km, 50 GHz	6
100Gb/s, 2000km, 50 GHz	15
400Gb/s, 75GHz, 500km	22
1000Gb/s, 175GHz, 500km	25

Table 3 Non Sliceable Transponder Costs

At the end of the study the results obtained are the percentage of successful cases for each stock value, which are represented on the next Figure 16. These figure show the percentages for year 2014 using the first scenario and comparing the case based on fixed transponders and the case based on SBVT transponders.

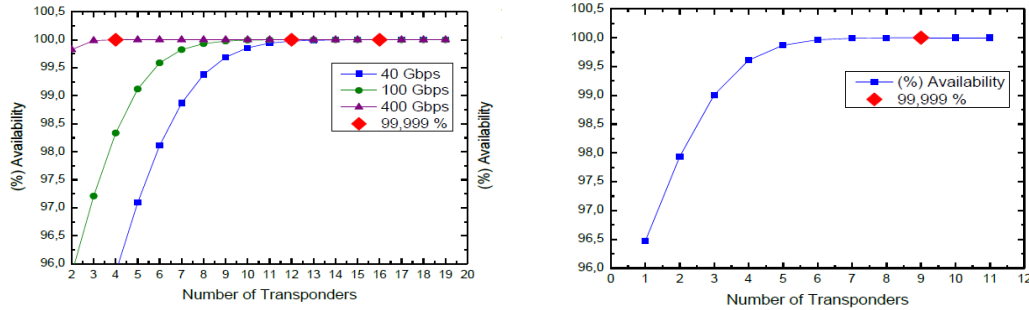


Figure 16 a) Stock number of fixed transponders; b) Stock number of SBVT transponders.

In light of these results, we can say that a less number of transponders are necessary in the case of using SBVT in a communication network to keep the service with certain availability. The previous results have been obtained from year 2014 to 2020 for the two earlier commented scenarios. Figure 17 show the number of stock needed to accomplish with 99,999% availability. The study of each year has been done based on the traffic of the Telefonica network in 2012. In order to compare the result in several years a 50% traffic increase has been assumed. It has to be taken into account the remaining stock of one year is not used the next year.

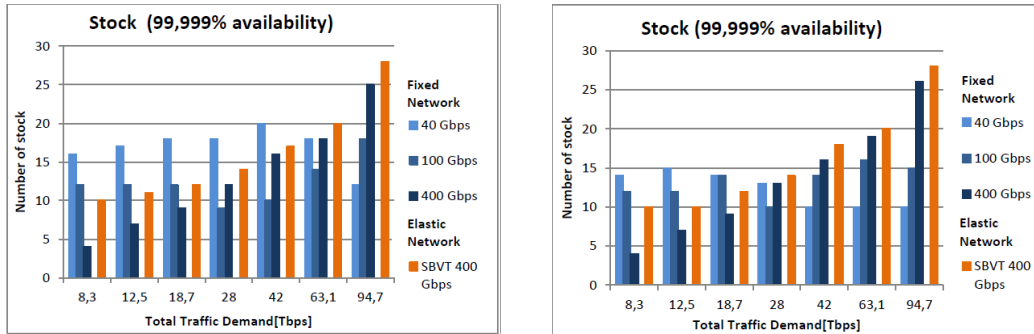


Figure 17 a) Stock for scenario 1 (full mesh) b) Stock for scenario 2 (IP topology shortest path).

It can be observed that a higher number of total fixed transponders are needed versus SBVT to keep 99,999% of service available. According with [4] the cost of the fixed transponder is 6, 15 and 22 for 40 Gb/s, 100Gb/s and 400Gb/s transponders respectively. Based on the previous information we have calculated what would be the cost of SBVT to save 30% operational expenditures.

Figure 6 shows that the possible cost of SBVT can vary between 22 which is the cost of fixed 400 Gb/s transponder and 30 which imply an increase to 36% in the cost of fixed 400Gb/s transponder. The peak target cost value is reached in 2015 (a total traffic demands of 12,5 Tb/s) then it steadily drops so that in 4-5 years the cost of the 400 Gb/s SBVT should be similar to the cost of a fixed transponder to achieve overall 30% savings in the operational expenditures of the network.

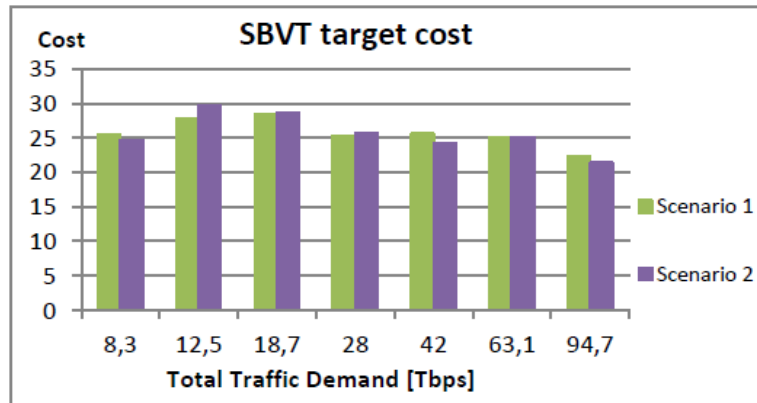


Figure 18 SBVT target cost.

This work shows how it is possible to save operational expenditures using SBVT versus fixed transponders despite the cost of the SBVT is greater than the fixed transponders. Another advantage of using SBVT is that it has to be stored just one kind of transponder. This study also presents the target cost of SBVT transponder in order to save 30% of operational expenditures, obtaining that this cost can vary from fixed 400 transponder cost (22) to an increase of 36% of this cost (30).

4.2.3 Conclusions from the techno-economic analysis

This section has investigated two models how to evolve from current fixed-grid network deployments to flexgrid networks. Benefits from sliceable bandwidth-variable transponders have been addressed under single and multi-layer network conditions. A special analysis of unidirectional SBVT interfaces has been done enabling point-to-multipoint demands by elastic light-trees over flexgrid networks. A new two-phase approach has been presented to optimizing total bandwidth-variable transponder costs in IP-over-Flexgrid networks. This could be easily solved to optimality for light traffic cases by a mixed-integer solver. Finally, the cost-effectiveness of fixed and flex-rate based modular SBVTs were evaluated for different network upgrade periods occurring every year or every trimester, as well as cost uncertainties regarding the deployment of modular flex-rate SBVT cards. While cumulative CAPEX is roughly comparable, the number of installed line-cards and transceivers is not. Hereby, flex-rate cards and transceivers allow 15% to 30% less hardware equipment and corresponding footprint which is due to a better line-card port utilization.

4.3 Node and Network architecture analysis

This section describes and evaluates potential application scenarios for the introduction of the flexgrid technology in Metro Area Networks (MAN) and in the core.

Regarding MAN, a key use case identified in D1.1 [37] was to support to Broadband Remote Access Server (BRAS) centralization. This support is evaluated in this section.

As for the core network, the viability of flexgrid in ultra-long haul (Pan-European) networks is first investigated, where Raman amplification and low cardinality formats (DP-B/SK) options are considered.



Next, and after SBVTs have been validated in the previous sections, the performance of 400Gb/s SBVT single multi-wavelength source architecture proposed in [66] is evaluated. Power consumption reduction when the flexgrid technology is used in survivable multi-layer optical networks is quantified. Several energy reduction strategies using power-adaptive interfaces adjusting their power to the transmitted traffic are proposed.

A network planning methodology efficiently exploiting traffic grooming over flexgrid and elastic rate systems is presented next. A power consumption model is built and used to show that efficient hierarchical topologies can be constructed exploiting the bandwidth granularity of the allocated transponders reducing overall network cost and power consumption.

Finally, algorithms for the Switchless Elastic Rate Node (SERANO) architecture proposed in D2.1 [65] are proposed.

4.3.1 BRAS centralization in metro networks

Typically, several BRAS servers are installed in the network and are placed close to end-users. To reduce the cost, network operators are considering centralizing the BRAS's within a region and transferring the traffic destined to them over a regional optical mesh network over a WDM network which is the current practice for building regional mesh networks. Since flexgrid technology is considered an alternative to WDM, we would like to evaluate the applicability of this technology in this regional/MAN network scenario.

In what follows we formally describe the BRAS centralization problem, present a number of algorithmic problem variations and outline algorithms to solve these variations. We use these algorithms to evaluate the transmission requirements of a realistic network scenario, which in turn drive the specification and design of appropriate flexgrid transponders that are based on direct detection (DD) that is less complicated and much cheaper than flexgrid transponders that would be used in core-long haul networks. We conclude with a techno-economic comparison between the WDM and flexgrid network to identify a target cost for the DD flexgrid transponders to achieve cost benefits of 30%, which seems quite low and reasonable.

4.3.1.1 Problem description

Typically, metro architectures are composed of two main levels of aggregation:

- i) First level of aggregation (named MTU level), in charge of collecting traffic from the end users (OLTs);
- ii) Second level of aggregation (named Access level), that aggregates the traffic from the MTUs mainly through direct fibre connections (i.e. dark fibre).

Nowadays, the IP functionality (i.e. traffic classification, routing, authentication, etc.) is implemented in the Broadband Remote Access Servers (BRAS) which are usually located at the second level of aggregation. Therefore, they are distributed in many sites along the regional area high CAPEX impact.

Reference WDM metro architecture

Since, in the last years, the main European network operators have been deploying / expanding their photonic mesh to the regional networks, it is proposed to move the

BRAS's (i.e. IP functionality) to the transit level, that is, in two sites for redundancy purposes. This way, the IP equipment needs (and its associated costs) would be reduced. A conveniently dimensioned pool of BRAS's (remote BRAS) could be co-located at each of these two transit sites in order to provide the required redundancy. So the regional photonic mesh provides the optical transport capacity from the second aggregation level (i.e. access level) towards the remote BRAS. This scenario is conceptually described in Figure 25.

A representative scenario has 200 MTU switches that generate traffic and access level with 62 nodes and an optical transport network comprising 30 ROADMS that are connected to 2 BRAS servers over that network.

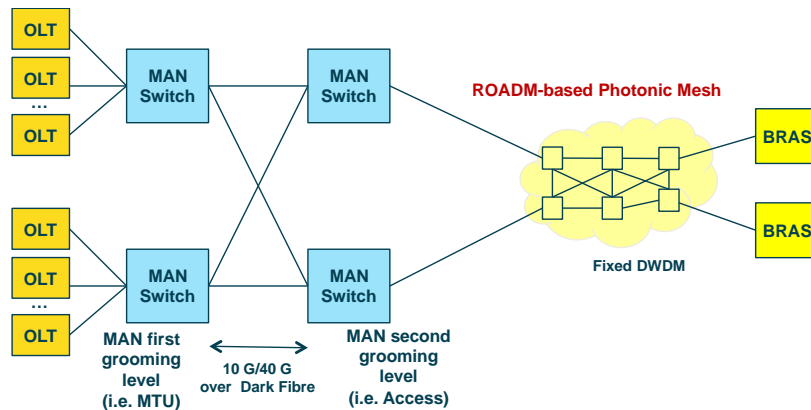


Figure 25: Scenario A – WDM evolutionary approach

This architecture does not provide too many savings in terms of CAPEX, since the first level of aggregation (i.e. MTU) is connected to the second level in the electronic domain, or in terms of bandwidth efficiency, since most lightpaths through the regional photonic mesh will require a capacity of 10Gb/s for the 50GHz-grid channels.

The flexgrid solution

As an alternative solution to the WDM network we propose a flexgrid solution to replace the WDM network. We assume that flexgrid transponders are used at the MTU switches and the Sliceable Bandwidth Variable Transponders (SBVT) are used at the BRAS servers. This scenario is presented in Figure 26. In particular we will examine two different variations for the flexgrid case. In the first variation we will assume the use of coherent BVT and SBVT transponders that are developed for core networks. The second variation will assume the use of customized BVT and SBVT transponders for metro network operation. We will drive the design of such customized transponders by examining a realistic metro network scenario and obtain requirements and specifications of the worst case optical path.

The routing problem is formally described in section 8.6.3.

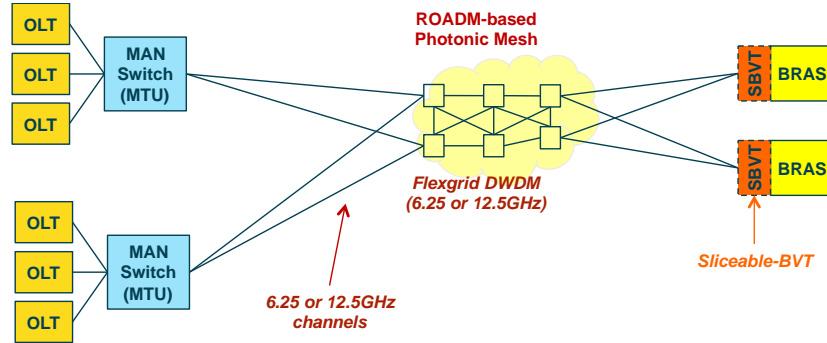


Figure 26: Flexgrid-based approach with SBVTs at the remote BRAS

4.3.1.2 Transmission Reach Requirements – Direct Detection flexgrid transponders for Metro networks

Routing case study

We now use the proposed algorithms to find transmission reach requirements for a realistic network topology. In particular the network topology that we use is the Regional-A topology of Telefonica network, which is shown in Figure 27.

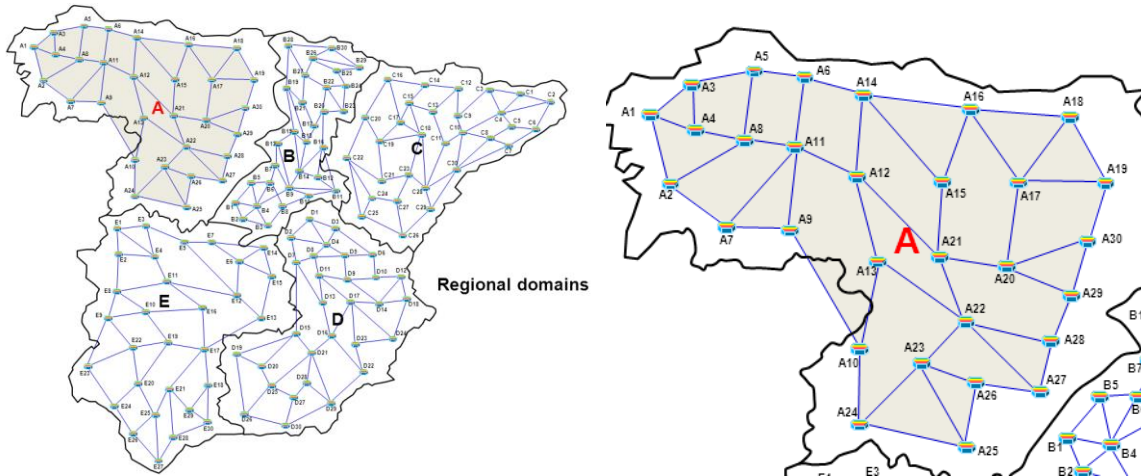


Figure 27: TID regional domains and focus on Regional-A topology.

We focus on problem variation 1, that is, we assume that we are given the placement or the number of BRAS nodes and we report on the worst (longest) path found.

In the results shown in Table 9 we assume that each MTU switch has one backup ROADM source ($|BS_n|=1$, for all $n \in V$, the backup source being the closest ROADM to the primary source n). We used the ILP algorithm that produced results for 2 BRAS nodes and the heuristic algorithm that searched among $k=4$ or 10 alternate primary and backup paths for each connection.

Table 9. Results when each MTU switch has one backup ROADM source.

	Heuristic ($k=4$)	Heuristic ($k=10$)	ILP
--	---------------------	----------------------	-----



	Worst path length	Worst path	Time (sec)	BRAS nodes	Worst path length	Worst path	Time (sec)	BRAS nodes	Worst path length	Worst path	Time (sec)	BRAS nodes
D = {21,22}	570	4 8 11 9 10 13 22	0.5	21,22	545	1 2 7 9 10 13 22	0.9	21,22	545	1 2 7 9 10 13	4.5	21,22
2 BRASes	490	5 6 14 15 21 22	62	13,22	420	30 29 28 22 13	150	13,14	420	30 29 28 22 13	76	13,14
3 BRASes	365	18 17 20 21 12	87	1,12,28	345	26 25 24 10 9	444	9, 12, 17				
4 BRASes	285	18 17 20 21	495	6, 9, 21,28	285	18 17 20 21	1300	6, 9, 21, 28				

We can observe that as expected, as we increase the number of BRAS nodes the worst path length reduces. We also observe that the heuristic has very good performance (at least for the cases that we were able to obtain the optimal solution with the ILP algorithm, the heuristic with $k=10$ was able to find the optimal solution). Thus, $k=10$ seems a good choice for networks of this size.

In the results shown in the following Table 10 we assume the backup ROADM sources for each MTU switch are the nodes within 120 km from the source ($|BS_m|$ is different for each $n \in V$). We can observe that the worst path lengths are reduced as opposed to the previous table where there was a single backup source. As in the previous case, the heuristic seems to have an outstanding performance, being able to find the optimal solution for the cases that we were able to track it.

Table 10. Results when backup ROADM sources for each MTU switch are the nodes within 120 km from the source.

	Heuristic (k=4)				Heuristic (k=10)				ILP			
	Worst path length	Worst path	Time (sec)	BRAS nodes	Worst path length	Worst path	Time (sec)	BRAS Nodes	Worst path length	Worst path	Time (sec)	BRAS nodes
D={21,22}	545	1 2 7 9 10 13 22	0.5	21,22	500	3 5 6 14 15 21	0.8	21,22	500	3 5 6 14 15	4	21,22
2 BRASes	370	1 4 8 11 12	48	12,15	370	1 4 8 11 12	83	12,15	370	1 4 8 11 12	51	12,15
3 BRASes	300	25 24 10 13 12	74	6,12,29	300	25 24 10 13 12	205	6,12,29				
4 BRASes	270	17 20 21 22	415	1,6,22, 29	270	17 20 21 22	1446	1,6,22, 29				

In the above presented results we used the proposed algorithms for finding the primary and backup connections for the centralized BRAS routing problem. These results give us a feeling of the expected path lengths in a realistic network scenario. The BRAS centralization use case requires the comparison of 2 network scenarios (as described in Idealist Deliverable D1.1), and in particular the cost comparison of the solutions so as to evaluate the applicability of the flexgrid technology in the Metro. Since the cost



calculations require the solving of the related routing problem and finding the maximum path lengths, the above results are used in the following to evaluate the different BRAS network scenarios.

Direct Detection flexgrid transponders

We proposed an improved SBVT architecture based on multiband OFDM (MB-OFDM) and DD for the BRAS network centralization scenario. We used the results of the previous section, which correspond to a representative network scenario considering a regional domain of the TID Spanish network, and transformed them into transmission requirements. A first cost-efficient OFDM-based SBVT solution using real-valued processing, double sideband modulation and DD, was presented in [70]. An improved and robust to physical layer impairments approach with multiband (MB) OFDM was then proposed in [71]. We numerically assessed the performance of the proposed MB-OFDM SBVT building block according to the transmission requirements. Furthermore, we experimentally validated our solution within the 4-node ADRENALINE photonic mesh network. Details of these comparisons can be found in the related references.

The cost-effective SBVT is able to serve $N \times M$ MTUs: the array of modulators generates N slices that can be directed towards different destinations nodes and each BVT building block (each of the N modulators) serves M MTUs with a single optoelectronic front-end by using only one laser source and simple/cost-effective DD. The number (M) of MTUs per MB-OFDM slice (corresponding to a single optical carrier) is limited by the bandwidth of digital-to-analog converter (DAC) and optoelectronic components (we consider 10GHz bandwidth devices). Considering current technology, $N=8$ and $M=5$ are realistic values.

4.3.1.3 Techno-economic Analysis

We will now techno-economically analyse two different network scenarios, the first one based on WDM technology and the second one based on flexgrid, with two variations one assuming coherent reception (CR) and components used in core networks and the second one assuming DD. The algorithms outlined in Section 2 were used for calculating the routing decision, and thus finding the length of the paths in Section 3.1. These in turn were used as specifications for the development of the DD flexgrid transponders in Section 3.2. So the second variation of the flexgrid network technology will assume the use of such transponders. The first scenario (WDM) corresponds to the reference scenario that is considered the current practice. We will calculate the cost of this scenario and use that as reference to calculate the targeted costs of components used in the flexgrid scenarios. For each scenario the cost calculation is going to be done based on the Idealist cost model presented in deliverable D1.1 [67] (which is an extension of the model presented in [69]). In all scenarios in the cost calculation we will neglect the cost of the MTU switches (base node and line cards of the switches) but not the cost of the transceivers used at these switches, which will be different in the different scenarios, and will also neglect the cost of the BRAS servers, since they are the same in all scenarios, but not the cost of the transceivers that are placed there. Also in the WDM network scenario we will neglect the cost of the WDM ROADMs, assuming that they are already in place.



Scenario 1: WDM solution

In this scenario the optical network consists of WDM ROADMs and the traffic from the MTU switches is aggregated at a second level of switches before it enters the WDM network (what we call aggregation switches). Taking the reference TID Regional-A domain, we have 200 MTU switches and 62 aggregation switches. Each MTU switch forwards $C=10$ Gb/s of traffic and is connected to two different aggregation switches for protection. Assuming equal distribution of load in the aggregation switches, each aggregation switch receives $200 \text{ (# MTU switches)} \cdot 2 \text{ (protection)} \cdot 10 \text{ Gb/s} / 62 \text{ (# aggregation switches)} = 6.45 \cdot 10 \text{ Gb/s} = 64.5 \text{ Gb/s}$ traffic, from the MTU switches, that is $\sim 7 \times 10 \text{ Gb/s}$ ports. Doubling this for the uplink direction (towards the BRAS) we find that each aggregation switch handles $\sim 140 \text{ Gb/s}$ of traffic, and needs in total $\sim 14 \times 10 \text{ Gb/s}$ ports. We also assume that the communication between the MTU switches and the aggregation switches is done with 10 Gb/s grey – short reach transceivers and that we use 10 Gb/s coloured transceivers for the communication of the aggregation switches and the BRAS servers (traffic router over the WDM regional mesh network. Figure 28 presents the aggregation network specifications of this scenario.

Cost of WDM solution = (i) cost of MTU switch transceivers (grey – short reach) + (ii) cost of MTU switches (140 Gb/s capacity, grey transceivers facing the MTU switches and coloured transceivers facing the WDM network)+ (iii) cost of transceivers (coloured) at the BRASes

Cost of (i) = $200 \cdot 2 \cdot \text{Cost of 10 Gb/s transceiver (grey – short reach)}$

Cost of (ii) = $200 \cdot 2 \cdot \text{Cost of 10 Gb/s transceiver (grey– short reach)} + 62 \cdot \text{Cost of aggregation switch (basic node + line cards)} + 200 \cdot 2 \cdot \text{Cost of 10 Gb/s transceiver (coloured - 50 GHz)}$

Cost of (iii) = $200 \cdot 2 \cdot \text{Cost of 10 Gb/s transceiver (coloured - 50 GHz)}$

According to the IDEALIST cost model, we have:

- Cost of metro switch (basic node 2Tb/s) = 0.42 ICU, cost of line card 20x10GE=1.99 ICU (we need just 1 such line-card to support the $14 \times 10 \text{ Gb/s}$ connections at each aggregation switch).
Since we need $\sim 140 \text{ Gb/s}$ metro switches, as an alternative we can assume a “lower end” aggregation switch than the 2Tb/s included in the IDEALIST cost model, with cost = 0.1 ICU.
- Cost of 10Gb/s transceiver (grey – short reach) = 0.008 ICU
- Cost of 10 Gb/s transceiver (coloured – 50 GHz) = 0.08 ICU

→ Cost of WDM solution = $[200 \cdot 2 \cdot 0.008] + [200 \cdot 2 \cdot 0.008 + 62 \cdot (0.42+1.99) + 200 \cdot 2 \cdot 0.08] + [200 \cdot 2 \cdot 0.08] = 3.2 + [3.2+ 149.42+ 32] + 32 = 219.82 \text{ ICU}$.

Alternative (assuming cost of switch chassis = 0.06): Cost of WDM= $3.2 + [3.2+ 129.58+ 32] + 32 = 199.98 \text{ ICU}$.

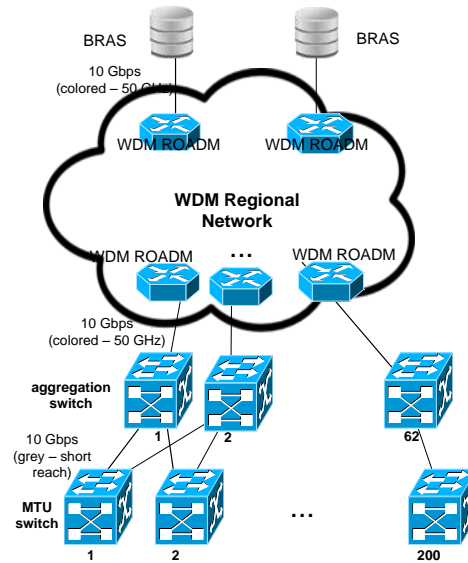


Figure 28: The aggregation network for the WDM network scenario.

Scenario 2: Flexgrid Network

In this scenario we assume that we replace the WDM network with a flexgrid network. So in this scenario we will include in the cost calculation the cost of the flexgrid WSSs (we will assume that all the other components of the ROADM nodes are the same). To count the number of WSSs we will assume that we need 2 WSSs per network interface (fibre link) and that each ROADM has one add-drop terminal, each add-drop using 2 WSSs (one for the add and one for the drop). For the TID region-A network we have $N_{nodes}=30$, $N_{links}=51$. We will examine two variations for the transceivers:

- (i) Using 10 Gb/s grey – short reach transceivers at the MTU switches and coherent 100Gb/s flexgrid BVTs (10x10Gb/s muxponders) at the add/drop terminals of the flexgrid regional network, and coherent 400 Gb/s SBVT at the BRAS servers (functioning as 4 x 100Gb/s). We assume that each ROADM node is equally loaded, so it serves $2 \cdot 200/30=14$ MTU switches (the two factor is for protection purposes). Thus, each ROADM gets 2 100Gb/s flexgrid BVTs (10x10Gb/s muxponders). Each 400 Gb/s SBVT (4 x 100Gb/s), serves 2 different ROADMs. Thus in total we need 15 400 Gb/s SBVTs.
- (ii) Using direct detected 10 Gb/s MB-OFDM transceivers at the MTU switches and direct detected MB-OFDM SBVT at the BRAS servers (each supporting $N=8$ different flows, each flow serving $M=5$ MTU switches, with a total capacity of 400 Gb/s). As above, we will assume that each ROADM node is equally loaded, so it serves 14 MTU switches. Thus, ROADM needs 3 flows of the DD MB-OFDM SBVT, and in total we need 10 DD MB-OFDM SBVTs.

Coherent case

Cost of the flexgrid network: (i) cost of 10 Gb/s transceivers (grey) at the MTU switches + (ii) cost of flexgrid BVT muxponders (100 Gb/s - 10x10Gb/s) + (iii) cost of flexgrid ROADMs + (iv) cost of SBVT transponders at the BRASes



Cost of (i) = $200 \cdot 2 \cdot \text{Cost of 10 Gb/s transceivers (grey – short reach)}$

Cost of (ii) = $30 \cdot 2 \cdot \text{Cost of 100 Gb/s (10x10G) BVT muxponders. Assuming equally loaded } N_{\text{nodes}} \text{ ROADMs}$

Cost of (iii) = $2 \cdot 30 \cdot \text{Cost of flexgrid-WSS (1x20)} + 2 \cdot 51 \cdot \text{Cost of flexgrid-WSS (1x9)}$

Cost of (iv) = $15 \cdot \text{Cost of 400 Gb/s SBVT (coherent – 4 x 100 Gb/s)}$.

According to the IDEALIST cost model, we have:

- Cost of 10Gb/s transceiver (grey – short reach) = 0.008 ICU
- Cost of coherent 100 Gb/s (10x10G) BVT muxponder = 1.28 ICU
- Cost of flexgrid-WSS = $1.20 \cdot \text{cost of WDM-WSS}$, Cost of WDM-WSS (1x9)= 0.32 ICU and Cost of WDM-WSS (1x20)=0.48 ICU

We will denote the cost of the 400 Gb/s SBVT (coherent – 4 x 100 Gb/s) as $X_{400G-SBVT}$.

→ Cost of coherent-flexgrid = $[200 \cdot 2 \cdot 0.008] + [30 \cdot 2 \cdot 1.28] + [1.2 \cdot (2 \cdot 30 \cdot 0.48 + 2 \cdot 51 \cdot 0.32)] + [15 \cdot X_{400G-SBVT}] = 3.2 + 76.8 + 73.728 + 15 \cdot X_{400G-SBVT} = 153.728 + 15 \cdot X_{400G-SBVT}$

As calculated in the previous subsection the total cost of the WDM solution is 219.82 ICU. To obtain 30% savings it means that the total cost of the flexgrid network should be $0.7 \cdot 219.82 = 153.874$

So this means that $X_{400G-SBVT} \leq 0.01$ ICU, which is quite unlikely to have (according to the IDEALIST cost model, a basic 400 Gb/s SBVT is expected to cost 3 ICU in 2015).

If we assume that the WDM network uses lower-end aggregation switches, which yields total cost of the WDM solution equal to 199.98, then the calculation is infeasible.

Direct Detection BVTs and SBVTs

Cost of the flexgrid network: (i) cost of 10 Gb/s DD MB-OFDM flexgrid transceivers at the MTU switches + (ii) cost of flexgrid ROADMs + (iii) cost of 400 Gb/s DD MB-OFDM SBVT transponders at the BRASes.

Cost of (i) = $200 \cdot 2 \cdot \text{Cost of 10 Gb/s DD MB-OFDM flexgrid transceivers}$

Cost of (ii) = $4 \cdot 30 \cdot \text{Cost of flexgrid-WSS (1x20)} + 2 \cdot 51 \cdot \text{Cost of flexgrid-WSS (1x9)}$

Cost of (iii) = $10 \cdot \text{Cost of 400 Gb/s DD MB-OFDM SBVT}$

The cost of the WSSs is as in the coherent case. We will denote the cost of 10 Gb/s DD MB-OFDM flexgrid transceivers as X_{10G-DD} and the cost of 400 Gb/s DD MB-OFDM SBVT as $X_{400G-SBVT-DD}$. To make the calculations easier we will assume that the X_{10G-DD} cost is one tenth of the cost of the coherent 100 Gb/s (10x10G) BVT muxponder used in the coherent flexgrid network case. In reality it would be even less, since X_{10G-DD} is for direct detection, as opposed to coherent. Thus, we will assume that $X_{10G-DD} = 0.128$ ICU. Just to give another reference, the cost of 50 GHz WDM 10 Gb/s transponder is 0.08 ICU, making the cost of X_{10G-DD} 60% higher than that.

→ Cost of the DD flexgrid network: $51.2 + 73.728 + 10 \cdot X_{400G-SBVT-DD} = 124.928 + 10 \cdot X_{400G-SBVT-DD}$



To achieve cost savings of 30% from the WDM network, the total cost of the flexgrid should be $0.7 \cdot 219.82 = 153.874$. Thus, $X_{400G-SBVT-DD} \leq 2.9$ ICU.

This result is quite promising, since the cost of a 400 Gb/s coherent SBVT is expected to be equal to 3 ICU in 2015, and the cost of the direct detection SBVT should be much lower.

If we assume that the WDM network uses lower-end aggregation switches, which yields total cost of the WDM solution equal to 199.98 ICU, then doing the same calculations we find $X_{400G-SBVT-DD} \leq 1.51$ ICU which again is within target.

Note that for calculating the above costs, when considering the flexgrid network we assumed that we change the WDM WSSs to flexgrid ones. This had a huge cost that was about 60% of the total cost. If this was neglected, then we would end up with much higher costs for the direct-detection transponders, while even the coherent network scenario would be feasible. Also note that the above costs for the direct detection flexgrid network were found so as to have a 30% cost saving in the total network.

4.3.1.4 Conclusions

We described and evaluated a potential application of the flexgrid technology in MANs. We outlined the related planning problem and applied algorithms to find the worst case path for a realistic network scenario. This in turn was used as specifications for the development of a cost effective DD flexgrid transponder. Finally, we performed a techno-economic comparison between a WDM and a flexgrid MAN solution. We found that coherent flexgrid transponders envisioned for core are rather expensive, while the cost of DD flexgrid transponders to achieve 30% cost savings from the WDM network seems quite realistic, making DD a viable solution for flexgrid in the MAN

4.3.2 Viability of Flexgrid in long distance networks

The introduction of Bandwidth Variable Transponders and Nodes in the Optical Networks able to elastically switch the spectrum (with a resolution higher than ITU-T fixed grid slices of 50 or 100 GHz) is undoubtedly very interesting and stimulating from the technical viewpoint. It remains to determine, which are the affordability conditions of such flexgrid technologies in the different network contexts.

This study analyses the introduction of flexgrid technologies in the specific context of a Pan European Network (one of the reference Networks defined within IDEALIST project, which is a long haul network with a diameter of about 6000 km) and in presence of static or semi-static traffic demand.

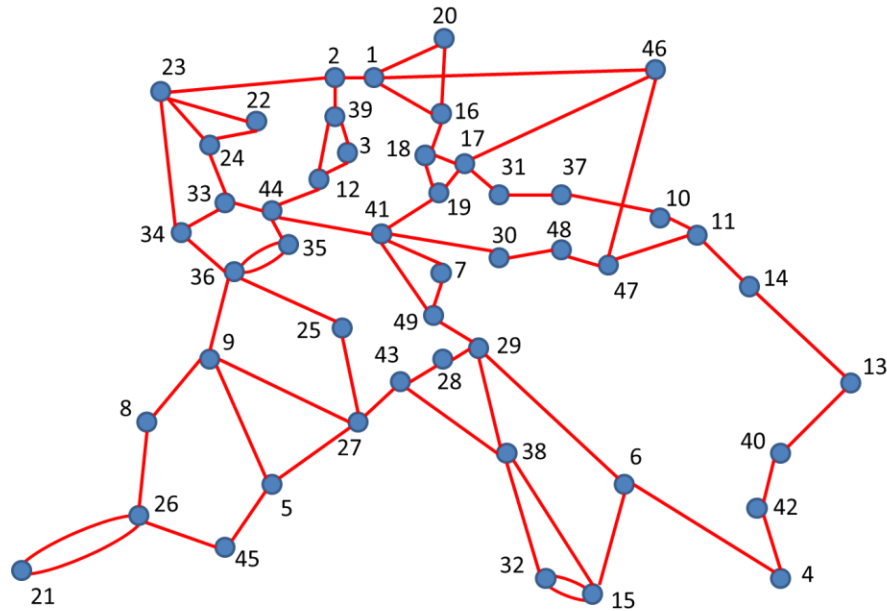


Figure 29: Topology of Pan European Network (Telecom Italia Sparkle Network).

The graph of the topology (not in scale) is shown in **Figure 29** while the main topological and routing characteristics are reported respectively in Table 41 and Table 42 in the Appendix. In the same appendix are also specified the characteristics of the static traffic demand used for the evaluations. See in particular Table 43 and Table 44 for the characteristics of the four traffic matrixes assumed, which differ each other for the number of active node pairs (node pairs that exchange traffic) and for the variance of the exchanged traffic in the set of active node pairs (from uniform to highly polarized on few directions).

By means of a comparison between a reference network solution which employs fixed grid technology (transponders which require a fixed and assigned optical bandwidth (50 GHz) and fixed grid node) and a solution which uses SBVT and flexgrid nodes, some indications are achieved in terms of the costs that would be required for flexgrid equipment to make this solution economically viable.

The study encompasses many aspects involved in the optical network planning, design and cost evaluation, with particular reference to typical characteristics of the long haul context.

Nodes architecture are based on existing or declared as soon commercially available equipment. A deep overview of such nodes and their cost model, in line with the IDEALIST cost model developed so far, is given in a specific subsection of the appendix (8.6.4.2).

The Transponders considered for fixed grid are the fixed bandwidth transponders and the fixgrid BVTs. The last ones can switch among different modulation format for adapting to different client bit rates; they require a fixed optical bandwidth (50 GHz) and show different reach performances depending on the operating modulation format. The transponder considered for flexgrid are the flexgrid BVT and the sliceable SBVT under specification in WP2 of the IDEALIST project. In particular the SBVT recently emerged is capable to



switch an aggregate client bandwidth up to 1.2 Tbit/s and to generate up to 12 separate and independent carriers modulated with a set of formats (QPSK, 8QAM, 16QAM) that can be freely merged in a set of separate superchannels. Additional flexibility in Baud Rates (ranging from 30 to 50 Gbaud) and in the FEC type (from Soft to Hard decision FEC, with overheads from 7% to 25%) allows to make available a large set of combination of superchannel types, each one characterized by its own client bandwidth (nx100 Gbit/s, OTN ODU4-Cn compliant) and OSNR tolerance (and consequently optical reach). Some more details with technical and economic features are given in 8.6.4.3.

The cost model applied is derived from results of previous projects (STRONGEST in particular, initially adapted for IDEALIST in D1.1, and now further developed with some extensions involving the CAPEX parts, also for taking into account some OPEX aspects. In particular, since the context is a transnational network in which the network provider does not own the greater part of the infrastructural facilities, the fiber rental and the equipment hosting are considered, together with energy cost, as a significant OPEX component. The CAPEX and OPEX are combined together assuming an amortization period for investment, and summarizing all the cost expenditures in the synthetic annual cost.

The overall cost model used in this study is described in 8.6.4.4 where an excel file including the complete set of network parts (at level of subsystem, board and equipment) with their costs and power consumption figures is also provided as embedded object in the document.

Due to the use of coherent reception in the transponders, dispersion unmanaged transmission is assumed. To evaluate the performance of the optical channels, the well known Gaussian Noise (GN) model is used. According to this model, a set of analytical formulas can be used to estimate under certain conditions (a constant spectral power density) the equivalent OSNR at the receiver with a good degree of accuracy, using a limited set of system parameters. Equivalent OSNR estimated by GN model takes into account the impairments due to both linear (ASE noise introduced by the EDFA amplifiers) and nonlinear (four wave mixing noise due to nonlinearity) effects. For a given modulation format the feasibility of a lightpath (i. e. the certainty that the demodulated signal has a post FEC BER under a given value, for instance 10^{-7}) is checked comparing the resulting equivalent OSNR with the sensitivity of the format, which depends on the format and on some other aspects (type of FEC and implementation penalty).

The use of RAMAN amplification in combination with EDFA in the so called hybrid EDFA Raman amplifier is also considered. The Raman amplification consists in transferring power from lower wavelengths to higher wavelengths [110]. Thus, if a high power (pump) signal is launched into an optical fiber in a smaller wavelength than the one of data channels, the data channels will experience amplification, thus mitigating the impact of fiber attenuation. Figure 30 shows the two possible Raman amplification approaches: the pump signals are launched at the fiber input (co-directional propagation) or the pump signals are launched at the fiber output (counter-directional propagation).

In these study the counter directional propagation with a single pump is considered. The complete physical and impairment propagation model is reported in 8.6.4.5.

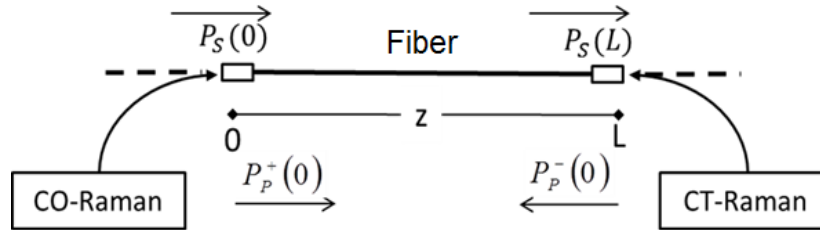


Figure 30 - Raman amplification approaches.

This analysis allows evaluating the trade-off between adding Raman amplification to improve the reach of some demands or increasing the number of regenerators. Two average span lengths (80 and 100 km) is considered for this analysis on the real network scenario which is at continental scale.

The results of this study are of two types. Firstly, in the context of a Continental Network, the conditions which make viable the use of RAMAN in addition to EDFA amplification are analysed. Secondly, the price conditions (target prices) of SBVT that make the flexgrid solution economically viable (if compared with the reference fixed grid network solution, considering the same network context and the same network requirements) are evaluated. This part of the work is still in a preliminary version that uses a simplified approach with approximate results and will be extended in a further version.

Network Design Framework

A customized network design framework, which to a large extent mimics the dimensioning functionalities of a commercial optical network planning tool, was augmented to enable performing the case study reported in the next section. Without going into detail (for an extended version on the design framework see 8.6.4.6), the main components of the network design framework exploited within the scope of this work are:

- Routing path computation functionality
- 3R regenerator placement functionality
- Wavelength assignment functionality
- Traffic demand routing ordering functionality
- Equipment inventory functionality

Importantly, the main modification realized in the already existing (custom-made) network design framework to allow performing the case study consisted of implementing the impairment model described above and integrating it within the appropriate steps of the framework.

An equipment inventory is generated after running the complete network design framework over the Sparkle Pan-European network with the assumed specific ROADMs architectures and link capacities and resorting to the line interface technology whose performance is modelled via the impairment model. This equipment inventory is used as input to the cost, footprint and power consumption model. As a result, an estimation of the overall cost associated to deploying and operating a DWDM network over the Sparkle topology is provided enabling the analysis, for example, the impact of adding Raman amplification to the Sparkle network.

Results obtained on the analyzed cases

Two different distributions of inline amplifiers along the links of the Sparkle topology were assumed for the case study. The first one considered a more typical amplifier spacing of around 80 km (in the following named S80 topology), whereas the second one assumed longer spacing values of around 100 km (named S100 topology).

For each network topology – Sparkle S80 and Sparkle S100 – the dimensioning was performed for different configurations in terms of RAMAN amplifiers placement, ranging from the case in which RAMAN is not deployed at all to the cases in which RAMAN is deployed for spans that are relatively short. The input parameter, used to decide which spans require RAMAN amplification or not, consists of the minimum span length above which RAMAN can be considered useful. A granularity of 5 km was used to obtain the results.

The traffic pattern utilized in the case study was the polarized pattern, whose main properties have been identified, assuming the traffic demands are unprotected. The baseline fiber topology of Sparkle was considered (no additional fiber deployment) and each link was capable of carrying 96 channels (in a 50GHz fixed grid).

Figure 31 shows the number of RAMAN amplifiers and 3R regenerators required, as a function of the span length threshold above which RAMAN is deployed, where plot (a) corresponds to Sparkle S80, plot (b) corresponds to Sparkle S100, and “Inf” denotes an infinite value of the threshold which implies not to install any RAMAN amplifier.

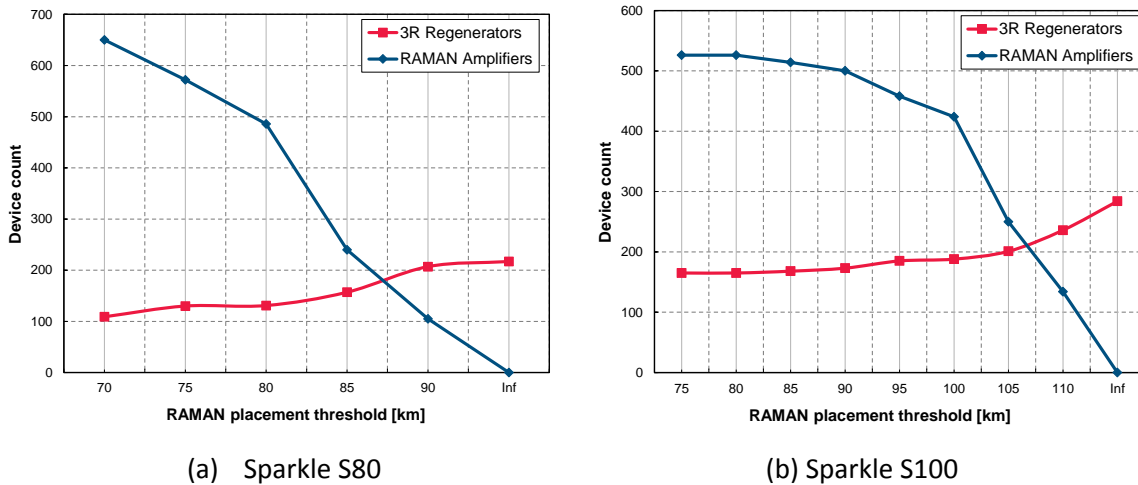


Figure 31: Number of RAMAN amplifiers and 3R regenerators required.

As can be seen from the results, and more notoriously in Figure 31 (b), increasing the threshold in the region of small span length values results at the beginning in the smooth decrease in the number of RAMAN amplifiers deployed and also a minor increase in the number of 3Rs needed. However, as the average span length value is attained, there is a steep drop in the number of RAMAN amplifiers deployed at the expense of a visible increase in the number of 3R regenerators. This trend highlights the trade-offs expected between RAMAN amplification and 3R regeneration.

Based on the cost model employed, Figure 32 depicts the cost contribution (in ICU) in terms of CAPEX only, associated to RAMAN amplifiers (e.g., additional cost of deploying an EDFA/RAMAN amplifier instead of just an EDFA amplifier) and to 3R regenerators (e.g.,

the cost of two 100G transponders used to perform the 3R functionality). Particularly, according to the enforced cost model, adding RAMAN amplification to an existing amplifier (for the two ways of transmission) costs 30% of deploying a single 100 G transponder (Raman for the two directions in a hybrid amplifier costs 0.3 ICU while a 100 G transponder costs 1 ICU).

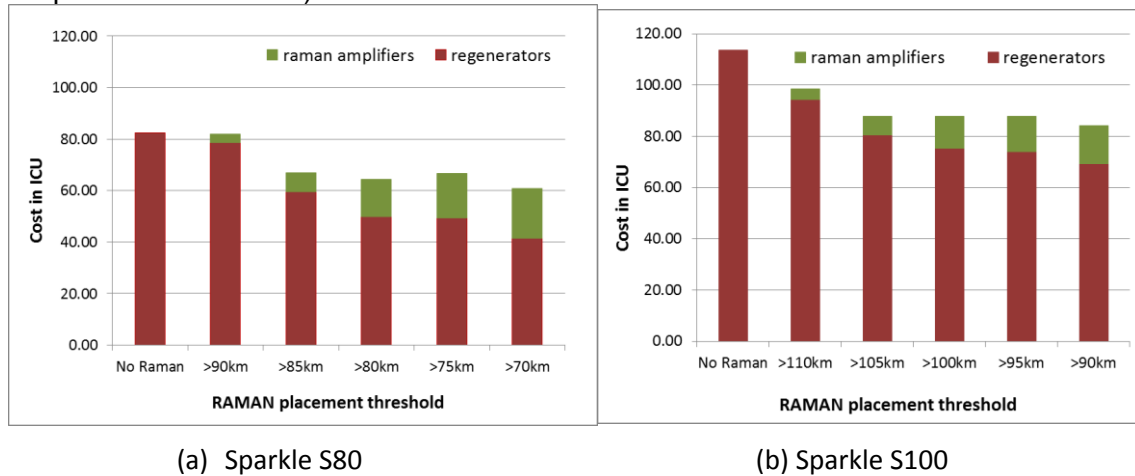


Figure 32: Cost in ICU of RAMAN amplifiers and 3R regenerators only.

Importantly, this is only a first and straightforward cost analysis made on CAPEX of Regenerators and RAMAN amplifiers only, that is not considering, for example, the impact of the power consumption and footprint of RAMAN vs. 3R regeneration and the indirect implications on other network co- factors..

Figure 33 shows the total cost diagram with the contribution of several components for some of the cases analyzed above, when the concept of Yearly cost is assumed is applied. The yearly cost includes the CAPEX components (Nodes, OLAs (EDFA), Transponders, Regenerators and RAMAN amplifiers) which are calculated with a 5 years amortization period, and the OPEX components (Fiber rental, Hosting, Energy). Diagram shows that the differences in terms of absolute yearly cost are not significant in all the analyzed cases. It also shows that the six cost components from “Fiber rental” to “Transponder” (the lower part of the bars in the diagrams) are very steady and do not depend significantly from the strategy used to place RAMAN amplifier in the network. This is due to the fact that the basic infrastructure of the network (nodes, fibers, OLAs) has to be placed regardless the traffic volume carried by the network. About the remaining two cost components, (related to the regenerators and RAMAN), their individual changes depend on the strategy applied to put RAMAN amplifiers (the threshold on span distance), but the sum of their costs is approximately constant and give in the best case an advantage of only a 6% on the total yearly cost w.r.t. the worst case.

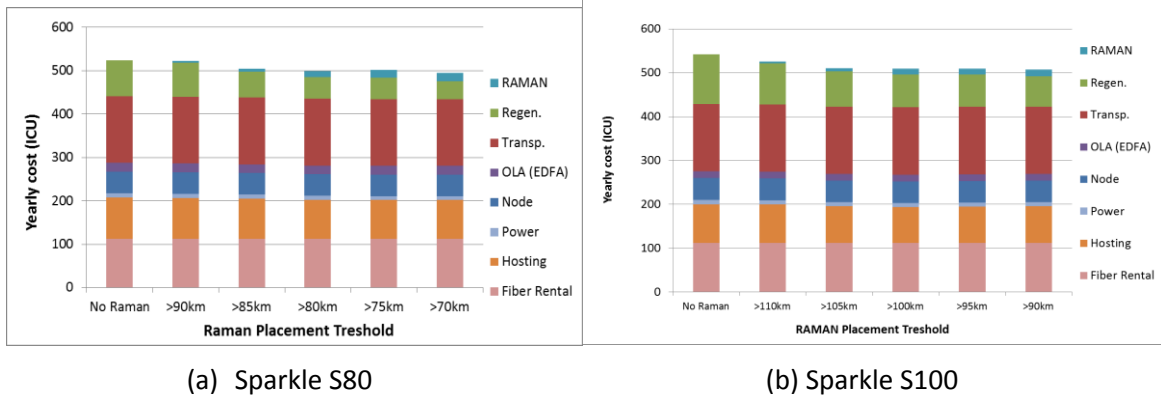


Figure 33: Yearly costs for Polarized reduced traffic.

Comparing the cases S80 and S100 in Figure 33, it can be observed that the total cost is approximately the same (all values differ each other at most of 6%, which is very little) but the values of the S80 case have a higher component of OLA (EDFA) and a lower component of Regenerators. This is a direct consequence of shorter spans, which implies higher number of OLA sites, but better performances in terms of impairment and consequently the need of less regenerations. The overall lowest cost case for S80 is for RAMAN put in spans longer than 70 km while for S100 the lowest cost is for RAMAN put in spans longer than 95 km.

The cost evolution shows that deploying RAMAN amplification leads to cost savings when compared to not deploying it. However, it also suggests that deploying RAMAN at most of the spans does not necessarily result in significant additional cost savings. Instead, it seems that there is a threshold value below which the sum of the costs of Raman amplifiers and 3R regenerators cost is almost constant. For example, for Sparkle S100 this threshold is 105 km, whereas for Sparkle S80 is 85 km. When correlating these findings with the results shown in Figure 31, it can be hinted that a good cost solution is the one where the number of RAMAN amplifiers is halfway between extensively deploying it in most of the spans and not deploying it at all along the network.

Preliminary extension to flexgrid networks

An extension of the dimensioning and cost study done for fix grid network has been done in case of use of SBVT and flexible grid network. The aim of this study is to find the target cost of an SBVT, i. e. the absolute cost of this component that equals the total cost of the fixed grid network using only transponders at 100 Gb/s in 50 GHz fixed grid.

The transponder type used as tributary for the flexgrid network traffic is always the SBVT, thus also for flows at the lowest bit rates (100 Gb/s) and also for regenerations which have been supposed implemented by SBVTs connecting back to back two tributary ports. These ports have to be at the same rate of the flow to be regenerated.

The same polarized demands used in the fixgrid dimensioning are considered, but assigning the node to node bandwidth requirement to different types of client port rates. The assortment of circuits at different client port rates for the whole network are listed in Table 54. Higher port rate circuits (1 Tb/s and 400 Gb/s) concern the most demanding nodes relationships. Circuits at 100 Gb/s are highly dominant in terms of number (196 of



250), carrying 52% of the total traffic volume. The higher rate circuits (from 200 Gb/s to 1 Tb/s) share almost equally the remaining traffic demand volume (48% of the total).

The details on the method used to calculate the target cost of SBVT is given in 8.6.4.8 while the final results are reported hereafter. Table 11 and Table 12 report the target cost in ICU for the SBVTs evaluated respectively for NO Raman and for Raman amplification on spans greater than 105 km. These costs are CAPEX absolute cost (not yearly cost as in previous tables) and are referred to a fully equipped SBVT as well to the SBVT without the Optical Multi Flow Modules (i.e. only with the system electronic part plus the client transceivers) and the single Optical Multi Flow Module (OMFM, which includes the optical line parts). In lack of predictions about the SBVT cost structure, the target cost is evaluated considering a set of very different conditions in terms of cost partitioning between SBVT electronic parts and optical line parts (OMFMs). For instance, if the OMFM cost percentage is 60% of the total cost and given that an SBVT has three OMFMs, it means that 40% of the total cost is imputed to the SBVT electronic part plus client transceivers and 20% of the total cost is imputed to each of the three OMFMs (60% for the three).

Table 11: Target cost of SBVT for S100 and without Raman.

S100 -No Raman				
SBVT part	% of cost of OMFM			
	80%	60%	40%	20%
SBVT without OMFM	2.31	4.51	6.62	8.61
Single Optivcal Multiflow Module	3.08	2.26	1.47	0.72
SBVT fully equipped with 3 OMFM	11.56	11.29	11.03	10.76

Table 12: Target cost of SBVT for S100 and with Raman used in spans longer than 105 km.

S100 – Raman > 105 km				
SBVT part	% of cost of OMFM			
	80%	60%	40%	20%
SBVT without OMFM	2.49	4.85	7.08	9.19
Single Optical Multiflow Module	3.32	2.42	1.57	0.77
SBVT fully equipped with 3 OMFM	12.44	12.12	11.81	11.49

Results from Table 11 and Table 12 show that the target cost for the complete SBVT is always very close to 12, which is 12 times the cost of a 100 Gb/s DP-QPSK transponder, with a slightly better solution for the configuration which use RAMAN (it has to be noticed that the higher the SBVT target cost, the more chance there is for the cost saving at network level because any cost reduction of SBVT has an impact on total cost reduction). This means that, at this first rough evaluation, it seems that there are no particular advantages (but neither penalties) using the flexgrid technologies on transnational networks, i.e. over networks requiring long-reach lightpaths that limit the use of high efficient modulation formats.



These results have to be confirmed with a more accurate design, that takes into account all the flexgrid model peculiarities (RSA, Add Drop chain of ROADM targeted to SBVT) and assuming different traffic conditions in term of global volume, distribution and client rates assortment.

Conclusions

This preliminary cost study, mainly focused on a fixed grid Pan European Network in presence of static or semi-static traffic demand, has been done in order to have a term of comparison between a reference network solution which employs fixed grid 50 GHz transponders and fixed grid nodes, and a solution which uses SBVTs and flexgrid nodes. The work has also been enriched by a focus on the analysis of the impact of using Raman amplification in this trans-national fixed grid network. It was shown that RAMAN amplification provides benefits on the yearly cost of about 25% when only RAMAN and Transponder cost are considered. If the overall yearly cost are analysed (including OPEX), Raman amplification gives benefits on cost of a factor of 6%. The best Raman amplification solution corresponds to inserting RAMAN amplification on the majority of the spans of the network (on spans longer than 70 km when average span lengths of 80 km are considered and on spans longer than 95 km when average span lengths of 100 km are considered), even though from significantly higher values of the placement threshold (85 km for S80 and 105 km for S100) very similar results in terms of total network cost are obtained. These lower thresholds imply that Raman amplification is required only on about half of the spans. The impact of traffic volume and traffic pattern on these conclusions still have to be assessed.

As regards flexgrid, some indications are achieved in terms of the costs that would be required for flexgrid equipment to make this solution economically viable. In particular a preliminary evaluation on target costs of SBVTs has shown that there are not particular advantages to apply the flexgrid technology to a trans-national network. The target cost of a 1.2 Tb/s SBVT is very close to the cost of 12 fixed transponders at 100 Gb/s. Therefore, there are margins to introduce SBVTs without penalties on the total cost if there is the chance that the 1.2 Tb/s SBVT will cost less than 12 100Gb/s fixed transponders.

Also in this case, a more accurate dimensioning work should be performed (in particular, RSA and appropriate A/D chains for SBVT) in order to confirm the results and an extension with other traffic demands should also be considered as for the case of fixed grid.

4.3.3 Evaluation of SBVT architecture based on multi-wavelength sources

Different technologies and solutions can be considered to design and implement SBVT architectures enabling efficient use of spectral bandwidth in EONs.

The laser source in an SBVT can be developed using one of the following two alternative solutions (proposed and detailed in Sect. 5.2.2 of D2.2 and [72]). The first solution uses a multi-laser (ML) source, i.e. an array of N tuneable lasers. The second solution uses a multi-wavelength (MW) source that consists in a single laser able to generate various carriers.

The ML source does not introduce any constraint on the routing and spectrum assignment (RSA), guaranteeing total freedom in the tuning of the carriers.

The MW source provides greater stability among the carriers thus better spectrum compactness. Moreover, MW allows to reduce the cost of the SBVT by decreasing the

number of lasers and therefore the power consumption. As drawback, the MW source imposes new constraints on the tuning and relative spacing among the carriers. These constraints reduce the solution space of the RSA.

In this section, the comparison among the two technological solutions is performed in terms of networking performance.

Extended Graph

In order to solve the RSA problem, the SBVT constraints have been introduced within the topology graph. The constraints are mapped into new virtual nodes and virtual links, forming an extended version of the topology graph. The RSA algorithm is then applied to the extended graph.

More specifically, to solve the path computation on the graph in Figure 34 for a source destination pair (S, D):

- The source node is split by creating a new virtual source node (SV).
- Each SBVT transponder (t) in the source node becomes a new node (TS_t).
- Each TS_t is connected with S through a new link (SL_t).
- Each TS_t is connected with SV through a new link (SVL_t).
- The two links SL_t and SVL_t have the same spectral description, which is based on:
 - Type of transponder TS_t
 - Transponder's state
 - Path computation parameters (number of carriers, carrier spacing, etc.)

Similarly for the destination node:

- The destination node is split by creating a new virtual destination node (DV).
- Each transponder (t) in the destination node becomes a new node (TD_t).
- Each TD_t is connected with S through a new link (DL_t).
- Each TD_t is connected with SV through a new link (DVL_t).
- The new links DL_t and DVL_t have the same spectral description based on the parameter mentioned before.

The path computation is performed considering the new graph (Figure 35) and the new virtual source/destination couple (SV, DV).

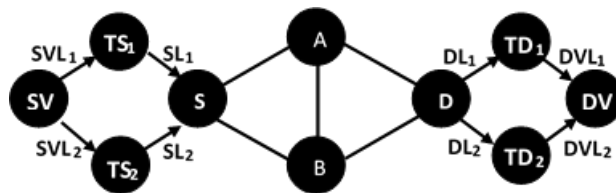
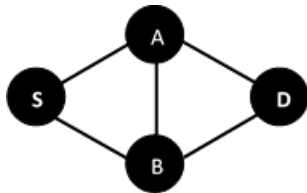


Figure 34: Real topology graph

Figure 35: Extended graph with virtual nodes and virtual links



Simulation scenario and results

Simulations are performed using OPNET Modeller. The developed model includes an accurate implementation of the RSVP-TE, PCEP and OSPF-TE protocols with the extensions required for flexible optical networks supporting the two SBVT solutions.

The considered network is a Spanish topology, with 30 nodes and 56 bidirectional links with 256 frequency slices of 12.5GHz per link. Spectrum assignment performed at the PCE is first-fit. The traffic is based on 100Gb/s and 400Gb/s connection demands and the network load is fixed to 800 Erlang, half of which is generated by the 100Gb/s connections and the other half by the 400Gb/s connections according to the equation (7). The traffic is uniformly distributed among node pairs and LSPs arrive following a Poisson process, the mean service time is fixed to 1 hour for all the connections in all the nodes.

The mean inter-arrival time per traffic demand is deduced according to the equation (7).

$$Network\ Load = \sum_n \left(\sum_d \left(\frac{S_{d,n}}{I_{d,n}} \right) K_d \right) \quad (7)$$

where:

n : Node index

d : Demand index (100, 400 in our case)

$I_{d,n}$: Mean inter-arrival time for demand d at node n

$S_{d,n}$: Mean service time for demand d at node n

$K_d = \frac{d [Gbps]}{100 [Gb/s]}$: Demand weight

In all scenarios, the considered SBVT can generate up to four carriers (using a single laser for the MW-SBVT, and four lasers when the ML-SBVT is used).

Three different scenarios are considered:

1. Nodes are equipped with ML SBVTs.
2. Nodes are equipped with MW SBVTs with a maximum relative carrier spacing of 50 GHz.
3. Nodes are equipped with MW SBVTs with a maximum relative carrier spacing of 100 GHz.

The aforementioned scenarios are compared in terms of achieved blocking probability (P_b) during provisioning. P_b is defined as the ratio between the blocked traffic and the overall traffic requests.

Figure 36 compares the blocking probabilities achieved by the first and the second scenario as a function of the number of transponders per node. The results show that given the same number of transponders per node, the ML solution achieves a lower blocking probability with respect to the MW. This gap is due to the new constraints

introduced by the MW SBVTs, in particular by the limited spectrum range that a MW source can cover.

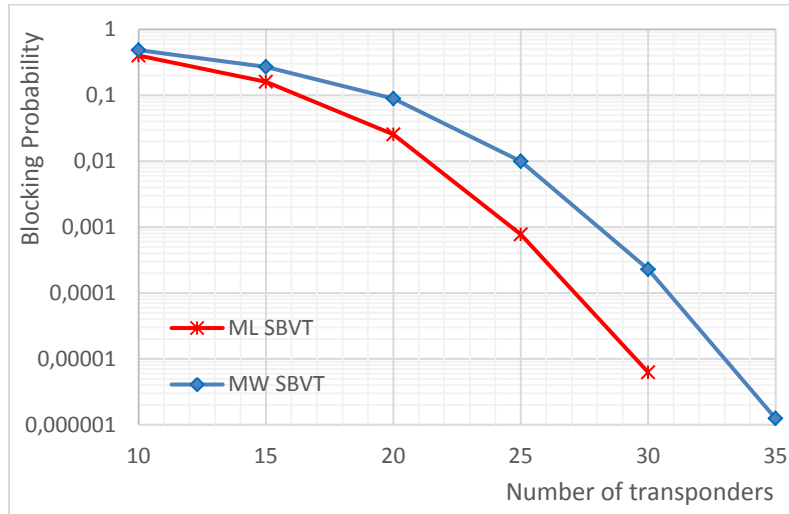


Figure 36: Blocking probability as a function of the number of transponders per node. Maximum carrier spacing 50GHz.

Figure 37 compares the same ML-SBVT of scenario 1 with the MW-SBVT supporting the larger carrier spacing considered in scenario 3. Results show that, increasing the maximum spacing between carriers allows the MW blocking probability to get closer to the ML curve.

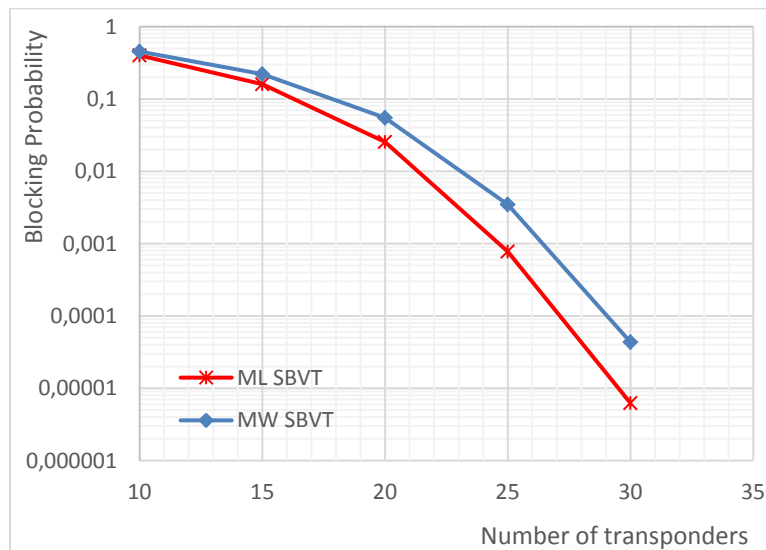


Figure 37: Blocking probability as a function of the number of transponders per node. Maximum carrier spacing 100GHz.

Assuming a service level agreement with a blocking probability less than 10^{-4} , results show that each network node has to be equipped adopting one of the following solutions:



1. 27 ML-SBVTs.
2. 31 MW-SBVTs with a maximum carrier spacing of 50GHz.
3. 29 MW-SBVTs with a maximum carrier spacing of 100GHz.

According to the considered scenario, in case the first solution is adopted, $27 \times 30 = 810$ transponders ($810 \times 4 = 3240$ lasers) are installed in the network. Using the second solution $4 \times 30 = 120$ additional transponders are needed but the total number of lasers is around 71% less with respect to the first solution (from 3240 to 930 lasers).

Considerations and future work

In these simulative results, the additional benefit provided by MW in terms of frequency stability among carriers (i.e., better spectrum compactness) has not been considered yet. Therefore, in the next studies, the blocking probability gap between the two architectures is expected to be further reduced.

Current MW technology enabling 50GHz relative spacing among carriers, if applied as in this study to a worst case scenario including many independent sub-carrier connections, results less attractive with respect to ML-based SBVTs. However, under different networking scenarios (results not reported here for space reasons), the two solutions become comparable from a networking perspective, particularly if, as feedback to WP2, the MW relative maximum channel spacing increases to 100GHz or above.

4.3.4 Traffic grooming under fixed and elastic spectrum allocation

Elastic spectrum allocation has been conceived to address the inefficient use of resources in case of traffic demands that do not adequately utilize the allocated resources at the optical layer (i.e. wavelengths). However, further improvements in resource allocation are expected by appropriately designing the network to accommodate the expected demands and by applying enhanced traffic engineering techniques in order to improve network efficiency. In this section we present a network planning methodology that exploits traffic grooming for improved CAPEX and energy efficiency under fixed and elastic spectrum allocation. Network planning is the process of link and node dimensioning under traffic demands with specific spatial and temporal properties. The objective is to reduce the total network CAPEX and OPEX, which in turn are strongly affected by the total power consumption. To meet this objective, a high degree of resource utilization should be achieved, while guaranteeing a certain QoS level to the established traffic flows. In wavelength switched optical networks several techniques have been proposed towards increasing resource utilization by means of traffic grooming [73][74]. Traffic grooming aims to maximize switch port and link capacity utilization by appropriately routing traffic flows to pre-selected points in the network and aggregating them so that they can share common paths. These approaches may assume flat, mesh-based topologies or hierarchal network organizations based on node clustering (Figure 38).

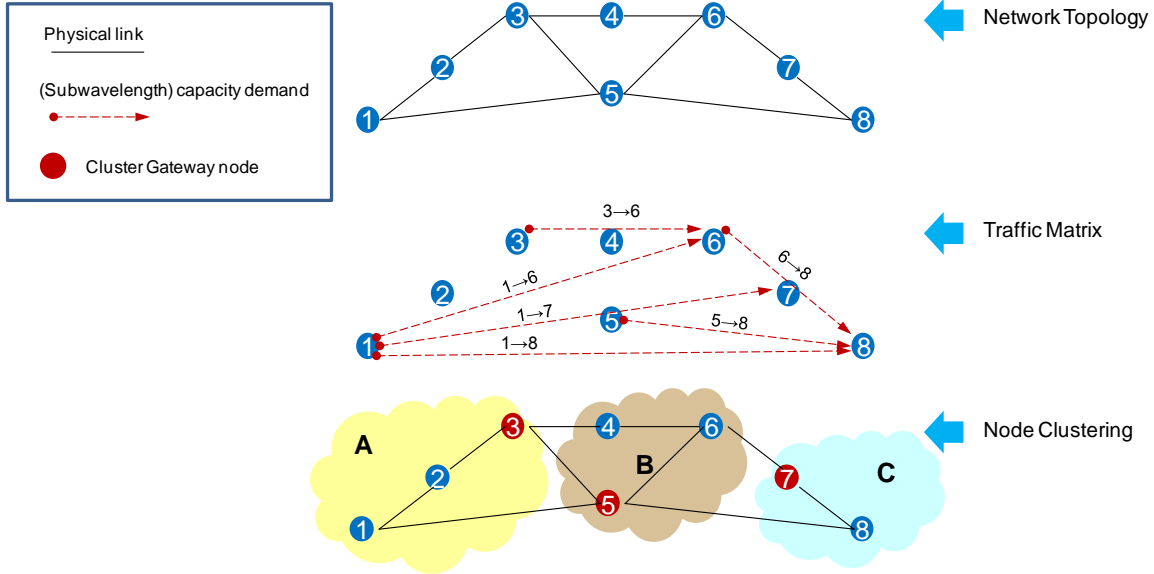


Figure 38: Example network topology, traffic matrix (demands) and hierarchical topology creation based on node clustering.

Recently, new advanced optical transmission technologies have emerged trying to address the same objective i.e. increased resource utilization. Flexgrid and elastic systems [75] aim to improve spectral efficiency and allow dynamic adaptation to line-rate changes exploiting SBVTs and flexible spectrum allocation, overcoming the limitations imposed by the fixed ITU grid. However, efficiency gains are strongly dependent on the traffic distribution across nodes, which in turn is determined by the routing and traffic aggregation techniques used. Hence, appropriate network planning techniques should be developed in order to maximize the benefits of flexgrid.

In this section we present the results of planning on the BT network with 1,113 nodes. We show that efficient hierarchical topologies can be constructed that can exploit the bandwidth granularity of the allocated transponders reducing overall network cost and power consumption. The details of the network planning methodology efficiently exploiting traffic grooming over flexgrid and elastic rate systems and the power consumption model we used are presented in section 8.6.5.

4.3.4.1 Performance evaluation results

The network consists of 1,113-nodes, partitioned into two categories: 103 are Metro/Core nodes forming a mesh topology using 164 links, while the rest operate as Aggregation nodes. A 7.8Pb/s total traffic matrix, having the year 2010 as a reference point, was generated consistent with the figure presented in [90]. This was obtained assuming that aggregation nodes groom and then forward incoming traffic from a given number of residential and business customers to their topologically nearest metro/core nodes using the shortest available path. 30% of this traffic is evenly distributed between metro/core



nodes. The remaining 70% is routed from each metro/core node directly to the topologically nearest of 7 Internet exchanges. With the same approach, a 312 Pb/s total traffic matrix was obtained based on projections for the year 2020.

For the 2010 case we assume fixed grid 2.5 Gb/s and 10 Gb/s transceivers and a CRS-1 router [78] with 40 Gb/s line cards, while for the year 2020 a 40-fold increase of traffic is expected and the full range of muxponders connected to a CRS-X router [78] with 400 Gb/s line cards, with the 400 Gb/s considered only on the flex-grid scenarios.

Using the two traffic matrices (2010 and 2020) we calculate the equipment required across all nodes of the network and obtain the total power consumption by accounting for the contribution of each sub-system (optical amplifiers, WXC's, line cards and SBVT's in the 2020 case). In Table 13 and Table 14 we report the main components that need to be deployed throughout the network.

For the 2010 traffic matrix, since SBVT's were not available the only option to reduce the number of active components is by means of electronic grooming in the opaque network (O-Fixed in Figure 39) and optical bypassing in cluster (BC-Fixed) or mesh based (BM-Fixed) topologies as discussed above. Evidently optical bypassing can achieve power consumption gains by 45% especially for hierarchical topologies, which prove more efficient. The introduction of SBVT's in 2020 can improve performance by 22% as observed in Figure 39 for the opaque network (O-Fixed vs. O-Flex), while it can also improve the performance of the clustered and mesh-based solutions employing optical bypassing by 14% and 10% respectively. Optical bypassing solutions are also shown for the 2020 traffic matrix to achieve significant gains compared to the opaque network up to about 60% i.e. even higher than the 2010 scenario.

Table 13: Network and node dimensioning for 2010 scenarios

	Line Cards	10 Gb/s transceivers	2.5 Gb/s transceivers	Max n.d.
O-fixed	2305	7160	32	8
BM-fixed	1205	2184	2054	14
BC-fixed	1189	2168	1516	9

Table 14: Network and node dimensioning for 2020 scenarios

	Line Cards	400 Gb/s SBVT	100 Gb/s MXP	40 Gb/s MXP	Max n.d.
O-fixed	8566		27738	52	18
O-flex	8648	7138	50	22	12
BM-fixed	3680		6992	3076	26
BC-fixed	3754		7880	1570	24
BM-flex	3839	1721	1024	2416	18
BC-flex	3824	1952	516	1246	15

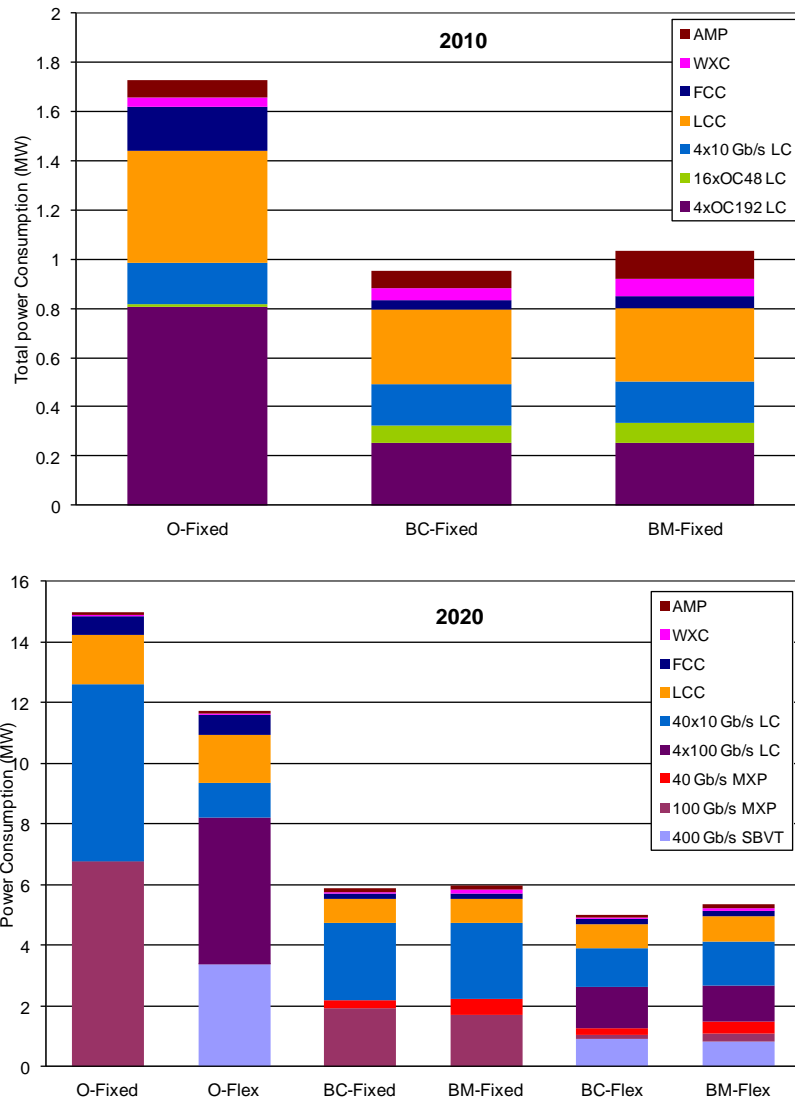


Figure 39: Total power consumption per scenario.

In order to be able to better identify the contributions of the various components to the total power consumption, we show in Figure 40 the percentage of each component for all 2020 scenarios.

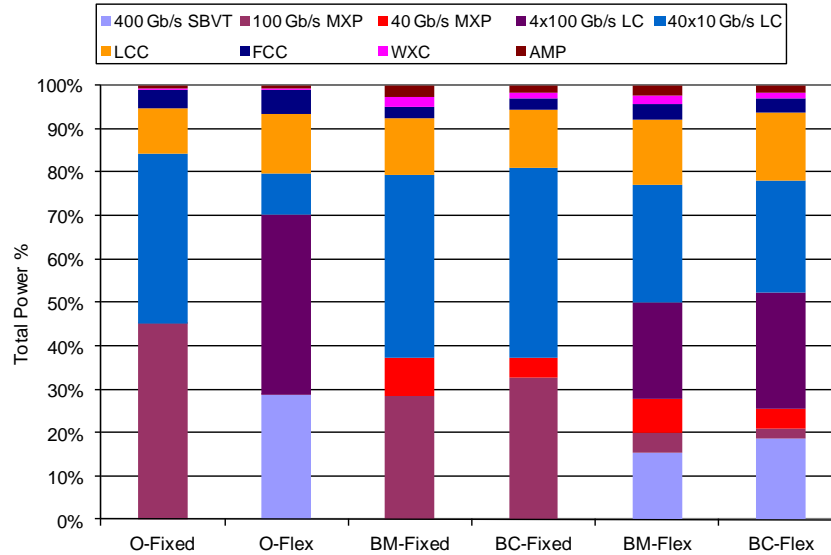
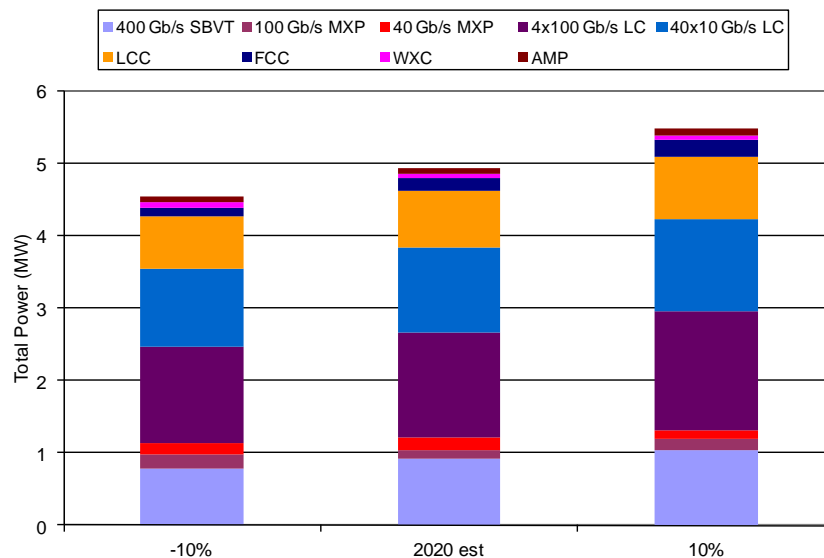
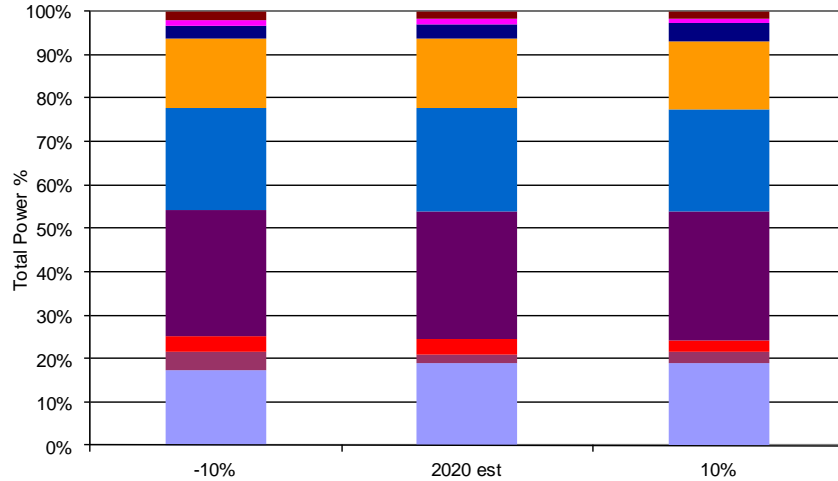


Figure 40: Percentage breakdown of total consumption per 2020 scenario.

We see that the major contributor to the power consumption is router chases and line-cards for all scenarios. All flexgrid configurations allow for better utilization of the available bandwidth and comparing Figure 39 and Figure 40 we find that the major difference between fixed grid and flexgrid scenarios for the opaque network and two grooming approaches is in the transceiver power consumption, with the contribution for router chases and line-cards remaining almost constant. To further see how this conclusion can be generalized, we show the results for the BC-flex scenario with a 10% variation of traffic in Figure 41 both as total power and percentage.



a)



b)

Figure 41: BC-flex 2020 scenario power consumption sensitivity to 10% variation of input traffic, as a) total power and b) percentage breakdown.

Figure 41a shows an almost linear increase in total power consumption in this traffic variation, even though the routing solutions differ considerably. Figure 41b shows that the contributions of the different components remain unchanged across this range.

In Figure 42 we also present the total cost across all nodes of the network and the contribution of each sub-system (optical amplifiers, WXC's, LCs and SBVTs) for the 2020 scenarios. The gains of grooming can be deduced by the maximization in the usage of the highest rate transceivers that helps to reduce both CAPEX and OPEX figures (Figure 39 and Figure 42) and can be maximized by exploiting flexgrid technologies. These gains are also demonstrated by the reduced average nodal degree (n.d.) of the network (Table 14), which can lead to efficient implementations within the limits of existing technologies (e.g. WSS size). Moreover, hierarchical topologies exploiting pre-planned aggregation points and a core network of hub nodes [74], [89] is shown to be advantageous compared to the distributed grooming approach of [73].

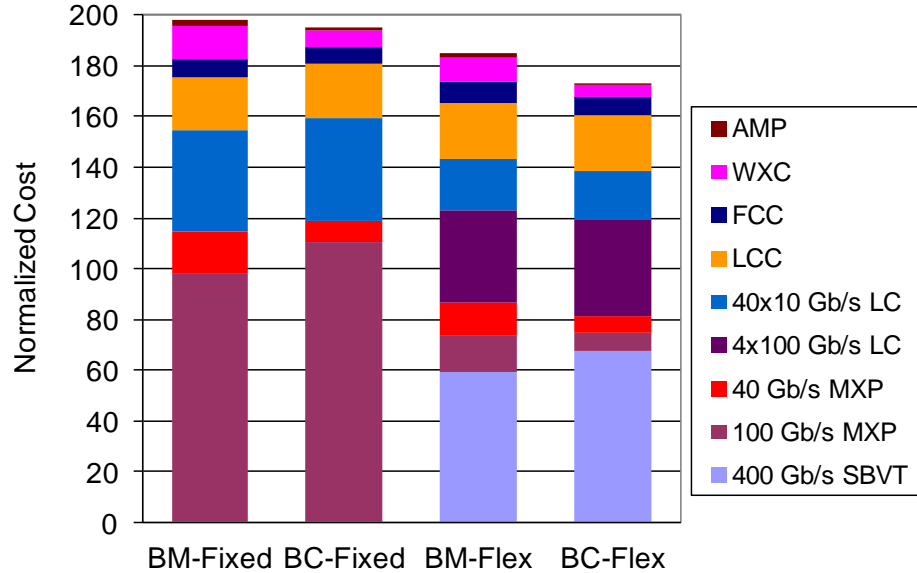


Figure 42: Total cost per 2020 scenario.

4.3.4.2 Conclusions

We proposed a planning and optimization methodology based on flexgrid extensions of traditional grooming techniques and a novel RWSA algorithm to exploit flexible grid and elastic spectrum allocation. The results show that using this methodology, efficient hierarchical topologies can be constructed that can exploit the bandwidth granularity of the allocated transponders reducing overall network cost and power consumption.

4.3.5 SERANO approach

EONs are aiming to improve spectral efficiency and allow dynamic adaptation to traffic volume changes. EONs today are employing WSSs to route spectral slots transparently in a multi-hop path. The WSS cannot provide for important on-the-fly adaptation functions like: change of carrier frequency; change of the modulation format to modify the optical bandwidth a flow may occupy and 3R regeneration, at the same time. These functions are only feasible via an electronic switch. To overcome these limitations, without resorting to electronic switching, the SERANO architecture has been proposed (D2.1 section 3.1.2) that provides for these operations as well as a number of networking functions like: i) spectral slot conversion; ii) fine tuning of the optical reach; iii) spectral defragmentation and iv) system vendor interoperability. The SERANO approach is explained in detail in section 8.6.6.

SERANO in the RMSA context

In striking contrast to typical WDM networks where traffic handling requires to consider both the grooming function as well as the RWA problem in EONs the latter function is reformulated as a Routing, Modulation and Spectrum Allocation (RMSA) problem. The



RMSA algorithms/mechanisms enhance the typical RSA algorithms/mechanisms with the capability of selecting different modulation format (MF) for each network flow. This capability improves the network spectral efficiency at the cost of the need for support of different formats and a more complex problem formulation. In this context, the emerge of SERANO can become beneficial for EON networks in terms of improving spectral efficiency and blocking probability, cost reduction etc., because SERANO -in contrast to other solutions- **can exploit the full potential of the RMSA algorithms/mechanisms**.

The four major functionalities of SERANO node are:

- 3R regeneration: SERANO node can 3R regenerate (re-time, re-transmit, re-shape) a selected input signal (flow) before forwarding it to the appropriate output.
- Modulation format translation: By using SBVTs as the main building blocks, SERANO has the capability to translate the modulation format of a flow pre segment basis and therefore can fine tune between smaller spectral occupancy and support of greater distances.
- Change of carrier central frequency: The SERANO node can modify the central carrier frequency of each flow entering the infrastructure. In this way, for example, if two or more flows compete to occupy the same frequency slot of an output fibre will not be collided, as it will happen otherwise, but they will be shifted to different non overlapping spectral zones. This functionality is vital in order to improve network utilization and accomplish spectral defragmentation.

Moreover, an important feature of SERANO is that all three aforementioned functionalities can be concurrently applied to the same flow. For example, a congested network may resort to SERANO in order to improve the network's spectral efficiency and capacity by concurrently regenerating, modifying the modulation format and changing the central frequency of suitably chosen flows.

The aforementioned capabilities of SERANO lead to a new generation of RMSA algorithms in which the modulation and spectrum allocation part is not assigned statically from source to destination, but can potentially selectively change at one (or more) intermediate network nodes. Therefore an efficient RMSA algorithm should take into consideration the idiosyncrasy of SERANO in order to capitalise on the expected gains. The main objectives of such an RMSA algorithm/mechanism are described in detail in section 8.6.6.1.

An RMSA reinforcing SERANO's features for an agile optical layer

In this section, the most important network functions offered from SERANO will be elaborated in association to the adjacent RMSA objectives. SERANO is providing the following functions: i) Optical reach optimization; ii) Relaxation of frequency continuity constraint; iii) 3R regeneration placement.

a) The optical-reach optimisation is implemented as follows: i) the RMSA algorithm selects the most efficient modulation format for the requested transparent length like typical RMSA algorithms; ii) our RMSA algorithm incorporates the features of SERANO, as they are presented allowing disintegrating the end-to-end path of a specific flow into a number of consecutive paths where different modulation formats are employed in each segment.

b) Relaxation of frequency continuity constraints: Another important feature of our RMSA



algorithm is that it relaxes any frequency continuity constraints typical in other legacy RMSA counterparts. Indeed, in typical RMSAs, the central carrier frequency of a flow should be such as the flow occupies the same spectral slots throughout the path from source to destination. This is not the case in SERANO which it allows for: i) arbitrary selection of carrier's central frequency and ii) modulation format translation. So in our RMSA, due to the former feature, some flows may occupy a group of successive spectral slot from source to an intermediate node and then, a group of different successive spectral slots from the intermediate node to the destination. In this case, the frequency continuity constrain is not requested for the entire path, but instead it is necessary only within the boundaries of the two consecutive paths. As such, new requests that will be blocked in typical RMSA mechanisms can now be granted under in SERANO. Due to the second feature, the use of different modulation formats per span allows to either expand or contract the number of successive spectral slots used in a path according to the RMSA objectives. In an EON employing SERANO the change of flow's central carrier frequency is implemented pre-emptively or on-demand. In the pre-emptive mode, change of carrier's frequency is implemented on carefully selected flows during the network operation in order to retain a defragmented network. In the on-demand mode, changes are realized when new flows are emerged or existent flows need to be rerouted of the capacity of the flow is expand/contract. Then, SERANO modules are activated in intermediate nodes in order to change carrier frequencies and accept new / existing flows.

c) 3R regeneration placement: An important optimisation problem in EONs is the optimum placement of regenerators. In general, the regeneration placement problem and the RMSA problem can be studied separately or in combination. In the second case, the objectives of RMSA problem are updated with the objectives of the regeneration placement problem, typically the minimization of the number of regenerations elements in the network or/and the minimization of the number of regeneration nodes. In our RMSA algorithm, the two problems are examined as a joint optimization problem in order to take advantage of the inherent regeneration capability of SERANO (SBVTs support 3R regeneration). Therefore the RMSA decides whether a flow requires 3R regeneration, splitting the transparent length into two (or more) parts and then it decides if different modulation formats/carrier frequencies may be used in order to optimise a specific performance metric (e.g., minimum CAPEX). The first step is typical for all regeneration placement problems when the use of regenerators is inevitable (e.g., when the QPSK modulation format cannot support the source/destination distance), while the second step is possible thanks to the unique features of SERANO. As such, the combined RMSA-regeneration placement mechanism chooses to regenerate a flow at an intermediate node and also to employ a different modulation format and therefore to alter the slot occupancy of some critical network links, if, of course, the gains from this action surpass the cost of the regeneration. In section 8.6.6.2 a set of examples realising our RMSA algorithm are presented.

Optimisation steps in all phases of network's lifecycle

The proposed RMSA-SERANO algorithm/mechanism may apply in three different problems/operational conditions:



- Off-line RMSA algorithm: RMSA algorithm applied during the network planning phase in order to optimise the network according to projected traffic and user requirements.
- On-line RMSA algorithm: RMSA algorithm applied during the network operation. The objectives of this mechanism are to retain an always optimised network. In this direction, the algorithm adapts its behaviour based on current network state.
- On-update RMSA algorithm: RMSA algorithm targeting on updating the previous network plan based on massive traffic changes and new/foreseeable requirements.

The three aforementioned algorithms are explained in detail in section 8.6.6.

Conclusions

We described the major functionalities of SERANO node that can improve spectral efficiency, blocking probability and cost reduction in EON networks. In addition, the optimisation steps of the proposed RMSA-SERANO algorithms/mechanisms were presented, which retain an always optimal network during all phases of its lifecycle.

4.3.6 Architecture on Demand (AoD) Approach

Considering the Architecture-on-Demand node design within a network topology, as described in D2.2 [66], in this deliverable, the design of cost-efficient architectures of networks with AoD OXCs is formalized as an ILP problem as well as with a heuristic approach for dynamic scenarios. The evaluation of realistic networks, such as the NSF and COST239 is also considered to assess the proposed strategies. The cost of the optimized AoD-based networks proposed is calculated considering the minimisation of the number of switching modules (e.g. SSSs), the number of cross-connections in the switching modules and the power consumption

The detailed design of the ILP model for the optimal network fabrication is described [115], where also the algorithms used for the network fabrication heuristic approach are illustrated. In this deliverable, the optimal design assessment is presented for various network topologies.

Optimal Design Assessment

The planning strategy is evaluated for both the optimal (i.e., ILP formulation) and suboptimal (i.e., heuristic) solutions for synthetic optical networks with AoD OXCs. The ILP formulation is solved by running a commercially available ILP solver, i.e., CPLEX. To ensure the optimality of the solution (for the ILP case), all the possible routes for demand (d) are computed by setting the value of π^d (set of candidate paths for d) to its maximum value (i.e., $|\pi^d| = \hat{K}_{max}$), whereas it is set to 5 for the heuristic strategy. Furthermore, the paths are computed by using the Yen algorithm [116].

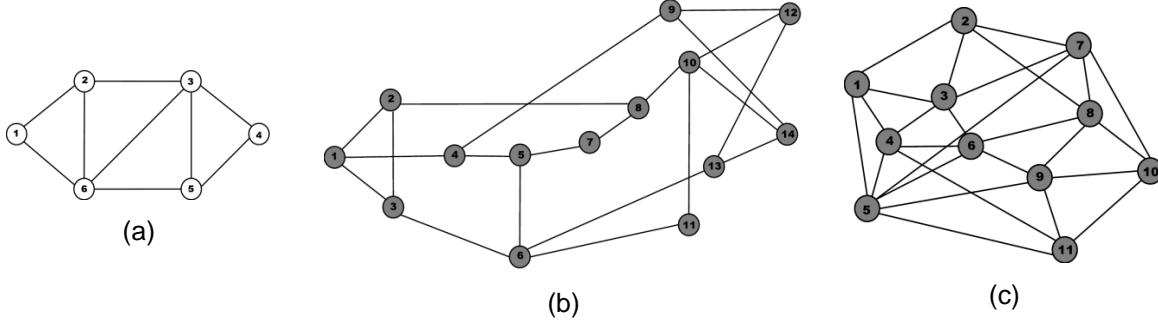


Figure 43: a) Sample topology, b) NSF topology with 14 nodes and 42 unidirectional links, and c) COST239 topology with 11 nodes and 52 unidirectional links.

For the assessment of the proposed strategy, a sample topology (Figure 43a) with six nodes and 18 unidirectional links, along with NSF (Figure 43b) and COST239 (Figure 43c) topologies are considered. For the sample topology, it is assumed that each link supports 40 slices, i.e., $F = 40 \forall (i,j) \in \mathcal{E}$, while for other topologies this value is set to 80. Different demand patterns (of the same quantity) may result in fabricated networks that need a distinct number of switching modules per nodes, with diverse location (i.e., input/output ports) within the node. Thus, for a given fixed number of traffic demands, a set of demands is generated by uniformly selecting the source and destination nodes, until a confidence interval of 5% or less is reached with a confidence level of 95%. The value of π^d for each demand is selected uniformly from the set $\{2,4,6,8\}$.

Sample Topology

This section presents results obtained for the sample topology using the CPLEX solver and the proposed heuristic algorithm. The planning strategy is evaluated for both non-contiguity and contiguity constrained networks with AoD OXCs. Figure 44a shows the number of components as a function of the number of demands along with total required slices (e.g., 10 demands need a total of 50 slices).

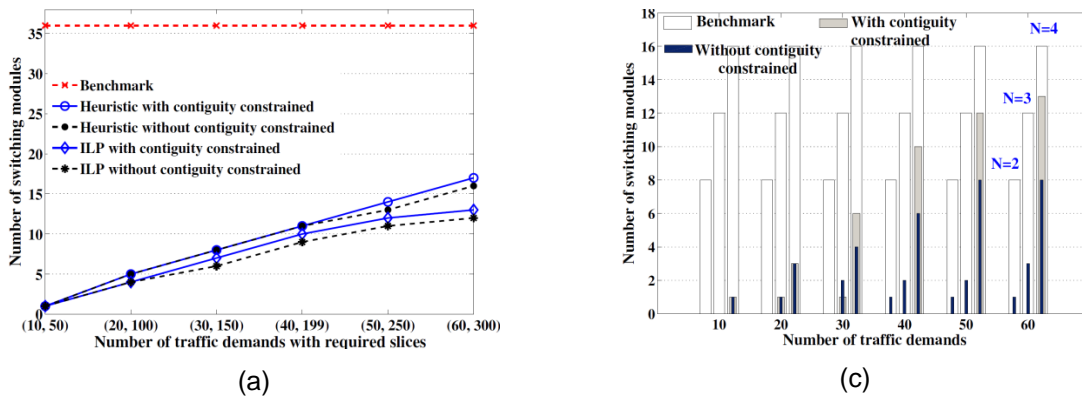


Figure 44: Sample topology: a) number of switching modules versus traffic demands and b) number of switching modules versus traffic demands, for nodes with different nodal degree.

The optimal solutions from the ILP formulation are compared with the results from the heuristic. Several insights can be gained from the results shown in Figure 44a. First, the proposed heuristic is able to well approximate the optimal solution (i.e., the maximum average error in terms of saving modules percentage is equal to 10% at high loads). Second, significant module savings is achievable by exploitation of the versatility and flexibility provided by AoD. Third, introducing flexibility through assigning non-contiguous spectrum slices further elevates this saving for higher loads.

Figure 44b exhibits the number of switching modules deployed in network nodes (with different N values) for the ILP results. For the contiguity constrained case, almost all of the modules are required for nodes with $N = 4$ (i.e., nodes 3 and 6) especially at medium and high loads. On the other hand, for the non-contiguity constrained case the planning strategy employs a small number of modules even at nodes with $N = 3$ (nodes 2 and 5) and $N = 2$ (nodes 1 and 4). By analysing the solutions (computed by the ILP solver) it is found that in the presence of the contiguity constraint, paths with diverse links are selected for setting up traffic demands. However, these paths mostly traverse through highly connected nodes and thus necessitate deploying maximum switching modules at these highly connected nodes. By mitigating the contiguity constraint, the link diversity among the selected paths decreases on $N = 4$ nodes for medium and high traffic, which leads to adding more modules on the other less connected nodes.

NSF and COST239 Topologies

Using the ILP formulation, the optimal solution for topologies with high nodal degree and/or number of nodes is intractable, because of its high computational complexity. Thus, for NSF and COST239 topologies only the results of the proposed heuristic algorithm are presented. Moreover, results for only the contiguity constrained network are presented here as for these topologies a modest improvement in reduction of network modules is observed by relaxing the contiguity constraint.

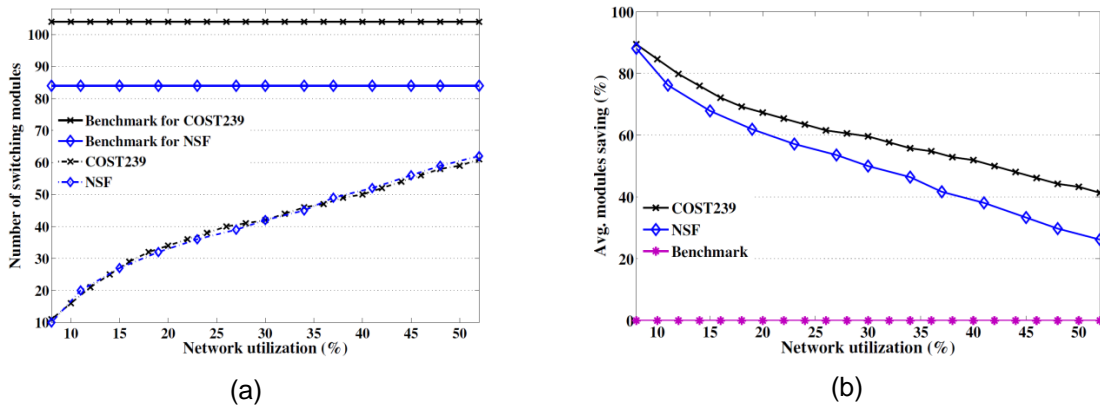


Figure 45: a) Number of switching modules versus network load, for NSF and COST239 topologies and b) number of switching modules versus traffic demands, for nodes with different nodal degree.

To compare the performance of the proposed heuristic strategy for these different topologies, the network utilization instead of number of demands is employed. The

network utilization is computed by measuring the percentage of the number of spectrum slices occupied by the established demands out of the total number of slices available on all the links of the network. Figure 45a displays the number of modules as a function of network resource utilization. It is worth mentioning that beyond 52% of network utilization some demands are not able to be set up in the network due to the resource unavailability constraint. Consequently, this utilization value is set as an upper limit on the traffic demands to be provisioned. Different observations can be made from the results shown in Figure 45a. First, the figure confirms the same trend witnessed for Figure 44a; i.e., the proposed heuristic strategy has the capability to significantly reduce the network components. Second, the strategy is more effective in terms of reducing the network modules for topology with high nodal degree (i.e., COST239). In fact for highly connected networks the solution space is larger, which facilitates the strategy to select the best paths for demands in terms of switching modules. Third, for a less connected topology the number of modules rapidly increases with load, whereas for the dense connected case the number of devices grows at a much lower rate. Figure 45b depicts the average module savings for different values of network utilization. At low loads, the heuristic saves up to 70% of modules with respect to the benchmark for both of the network topologies. However, for higher loads the gain drops to 40% and 25% for COST239 and NSF, respectively. However, this signifies a major cost saving especially for scalable and evolvable networks.

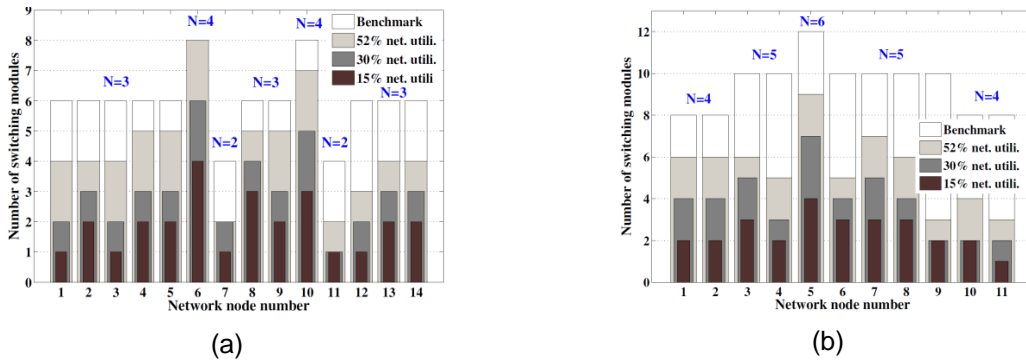


Figure 46: a) NSF topology: number of switching modules installed in the network nodes versus network node number and b) COST239 topology: number of switching modules installed in the network nodes versus network node number.

To further analyse the performance of the proposed heuristic strategy, Figure 46a and Figure 46b demonstrate the number of modules deployed at different network nodes for NSF and COST239, respectively, at different utilization levels. The figures manifest that the strategy works to reduce the number of modules, especially for sparsely connected nodes (compared to benchmark), even at higher load. Moreover, most of the modules are installed at the inner and higher connected nodes, which route most of the traffic demands. The network resource utilization per links for NSF and COST239 is assessed in Figure 47a and Figure 47b, respectively. Two important conclusions can be drawn from the figures. First, resource utilization for the links connected to nodes with more network modules is comparable to the other links. Second, the variance of the resource utilization per link for COST239 is high compared to the NSF topology. The reason is that, for a less connected

network (i.e., NSF), the paths for demands are relatively longer and more diverse in terms of links compared with a more connected one (i.e., COST239). Thus, even the links not connected to high degree nodes have a chance to route some of the demands in contrast to the densely connected network in which the links carry either most or very few of the demands.

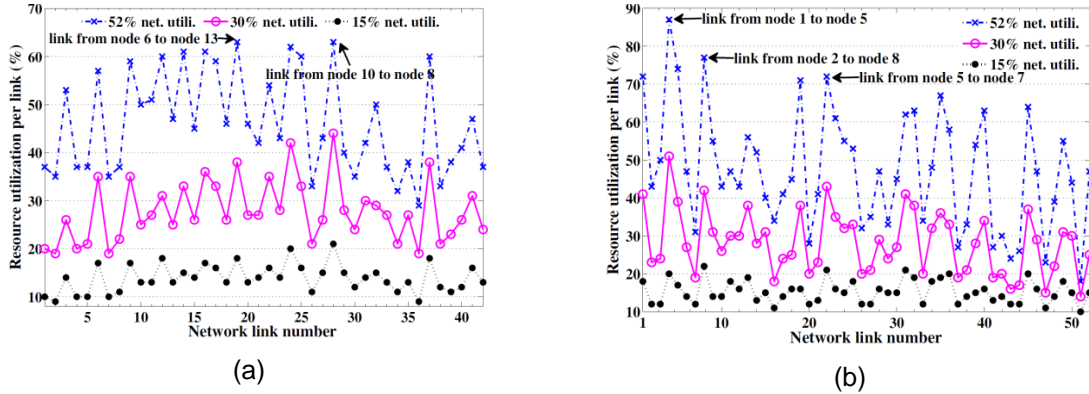


Figure 47: a) NSF topology: network resource utilization per link versus network link number and b) Cost239 topology: network resource utilization per link versus network link number.

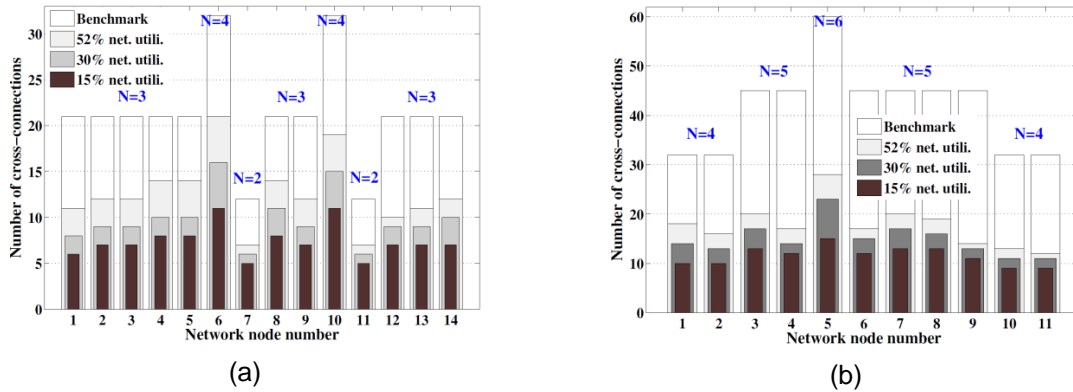


Figure 48: a) NSF topology: number of switching modules installed in the network nodes versus network node number and b) COST239 topology: number of switching modules installed in the network nodes versus network node number.

Finally, Figure 48a and Figure 48b evaluate the number of cross-connections per node required to implement AoD for NSF and COST239 topologies, respectively. Notice that the number of cross-connections indicates the optical backplane port count, and thus can be considered as a scalability parameter for AoD. The figures confirm the same trend witnessed so far; i.e., the heuristic strategy has the capability of reducing the number of cross-connections enormously compared to the benchmark. Even for high nodal degree node (i.e., $N = 6$) the strategy managed to scale down the number of cross-connections by half, through utilizing the versatility offered by AoD. Furthermore, the number of required

cross-connections (shown in the figures) indicates that the networks designed and studied are realizable with commercially available optical backplanes.

Power Consumption Analysis

For the evaluation of the power consumption for the network with AoD nodes, the parameters characterizing the power consumption are set according to the values highlighted in Table 15.

Table 15: Power Consumption Contributions

Device	Power [W]
Common equipment	100
SSS	40
3D-MEMS	150

The common equipment accounts for the controller, the cooling fans, and the power supply. Note that for the overall network power consumption, only the devices installed in the nodes are considered without taking into account the link devices (e.g., optical amplifiers). Figure 49a shows the network power consumption as a function of network utilization for NSF and COST239 topologies. The figure reveals that by exploiting the flexibility provided by the AoD node the power consumption of the network substantially decreases, compared to the network with rigid design nodes. Moreover, there is high potential for a densely connected network (COST239) when it comes to reduction of power consumption. The average power savings at different network loads is shown in Figure 49b. At low loads, the network with fabricated AoD nodes saves up to 50% and 40% for the COST239 and NSF topologies, respectively. These values drop to 30% and 20%, respectively, for higher loads.

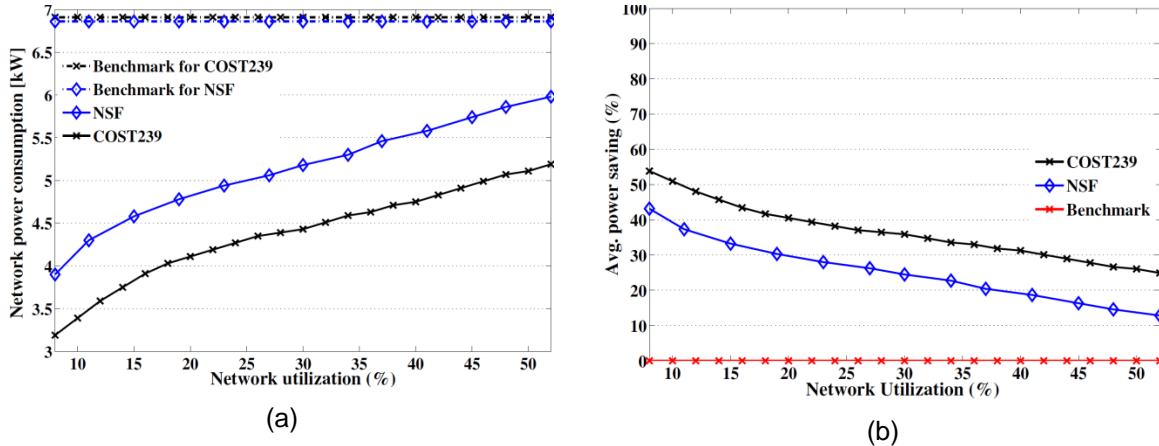


Figure 49: a) Network power consumption versus network load, for NSF and COST239 topologies and b) Average power saving versus network load, for NSF and COST239 topologies.

Dynamic Scenario Assessment

The performance of the proposed synthesis algorithm for the dynamic scenario is evaluated with a custom-built event-driven simulator. Demands are assumed to arrive in the network following a Poisson process. Established connections are assumed to have an exponentially distributed duration, whose average value is set to 1. It is also assumed that detailed link status information is promptly disseminated throughout the network so that each node controller can execute the proposed algorithm with updated information. Also, the latency for reserving the network resources is considered negligible. Demands are uniformly distributed among all node pairs, and the number of candidate paths is set to 5 ($|\pi_d| = 5, \forall d$).

The probability of rejecting demands due to an insufficient number of switching modules at AoD nodes is used to evaluate the performance of the proposed strategy. Furthermore, for other alternatives, such as the broadcast-and-select configuration for CDC-ROADM and the AoD with switching devices at each input/output port, the blocking of demands due to the switching module constraint is equal to zero. Therefore, for benchmarking purposes the proposed strategy is compared with the basic approach, i.e., *shortest path with first fit* (SPFF), for the same fabricated network. The SPFF strategy employs the first-fit spectrum assignment technique [117] for allocating spectrum available on the very first shortest feasible path for demands.

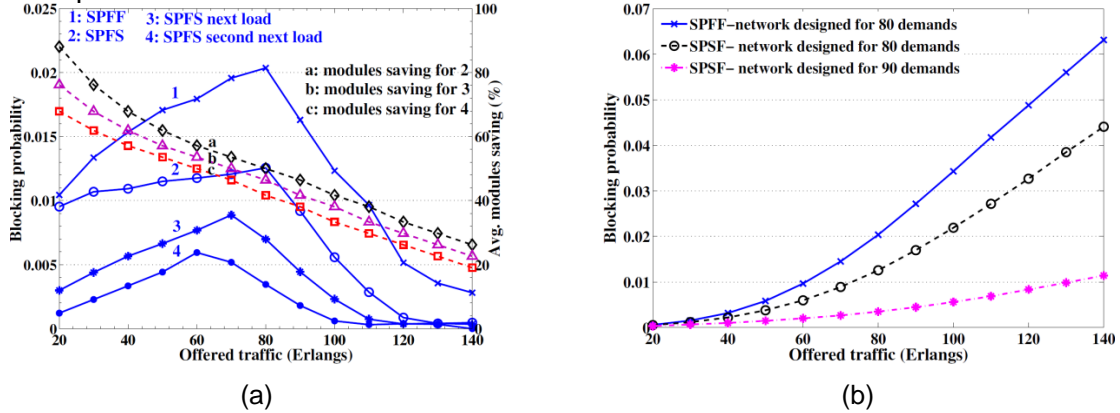


Figure 50: a) NSF topology: blocking probability (LHS) and average module savings (RHS) versus offered network load and b) NSF topology: blocking probability versus offered network load.

Figure 50a displays the network blocking probability for the NSF topology when the network designed for static load is used for dynamic traffic of the same intensity. In other words, for 20 Erlang load, the offline technique is operated for random sets of 20 connections (static traffic), and the same method is adopted for other load cases. The same figure also shows the module savings (y axis on the right) compared to AoD nodes with no fibre switch operation, i.e., nodes designed with theoretical allocation of modules. It can be seen from the figure that the proposed strategy, SPFS, exhibits better performance compared to SPFF. Furthermore, the blocking probability curve first slowly rises, and then beyond 80 Erlang it sharply declines. This anomaly is due to the fact that at low loads the number of switching modules deployed for the static case is relatively small (Figure 45a),



yielding more restricted intra-nodal connectivity and high device savings (curve a). In turn, the connectivity is not good enough to handle the routing for the stochastic arrival and departure of demands into the network. However, as the number of switching components increases with load the savings in devices decreases but the intra-nodal connectivity and consequently the blocking performance of the network improve.

To analyse the trade-off between the blocking performance and switching module savings curves 3 and 4 (Figure 50a) show the blocking probability of the network designed with supplementary switching devices. Curve 3 manifests the blocking probability when the network planned for 30 demands (static traffic) is operated for 20 Erlang traffic, and the same difference is also maintained for the other offered traffic cases. Similarly, curve 4 demonstrates the blocking performance when the network designed for 40 demands is utilized for 20 Erlang traffic, and the same inequality is also kept for the other offered traffic cases. It is obvious from the figure that the blocking performance of the network drastically improves with a slight decrease in the network module savings (curves b and c). For high loads, the blocking of demands due to the switching modules constraint approaches zero; still the savings is more than 20%.

To further investigate this trade-off, Figure 50b depicts the blocking probability of the network that is designed for a load of 80 demands (static traffic), and is operated for dynamic traffic, which varies from 20 to 140 Erlang. For this planned network, the savings in switching devices is equal to 50% at the nominal load, while the probability of rejecting demands due to the switching modules constraint (for SPFS) is very low. The results verify that by intelligently placing the switching components, and subsequently utilizing these devices effectively, the provisioning capability of the network improves significantly.

To evaluate the impact of network connectivity on the performance of the proposed dynamic algorithm, and on the trade-off between the blocking performance and switching module savings, the simulation experiments are carried out for the COST239 topology. It can be seen from Figure 51 that the blocking performance of the network in general and that of the proposed strategy in particular improves enormously for the topology with higher nodal degree. The high connectivity enhances the link sharing among the paths for different source–destination pairs. Thus, it enables the algorithm to efficiently utilize the installed switching modules, and share the fibre switches for demands with common links in their routing paths. Furthermore, the network blocking can be improved further by slightly raising the switching devices in the network nodes (i.e., curves 3 and 4). Finally, the results in the figure show that densely connected networks have the potential to attain high savings in switching modules at the expense of modest rejection of demands.

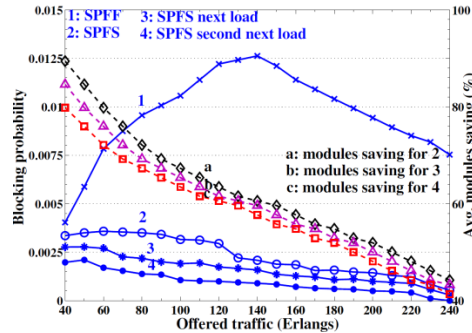


Figure 51: COST239 topology: blocking probability (LHS) and average module savings (RHS) versus offered network load.

4.4 Power Consumption

In this section, we focus on studying alternatives for power consumption reduction.

Firstly we analyse network survivability that requires the provisioning of backup resources in order to protect active traffic against any failure scenario. Backup resources, however, can remain unused most of the time while the network is not in failure, inducing high power consumption wastage if fully powered on. In D1.1 [67] we proposed IP/MPLS protection switching as failure recovery mechanisms in an IP over DWDM multi-layer optical network. MPLS protection switching is a solution interesting for current network operators to ensure fast recovery as well as fine-grained recovery treatment per LSP. In this section, we focus on IP/MPLS over DWDM multi-layer scenarios.

Next, we investigate different energy reduction strategies using power-adaptive interfaces (with on-off capabilities and/or data-rate adaptation) adjusting their power to the transmitted traffic, following predictable fluctuations. An optical grooming approach is proposed in order to use the benefits of traffic fluctuations and adaptations of the optoelectronic (OE) devices.

4.4.1 Power Consumption Reduction in Survivable Networks

In this study, we focus on IP/MPLS over DWDM multi-layer optical networks, survivable against any single link failure scenario using IP/MPLS protection switching. The interests of this solution rely on: 1) fast LSP recovery below 100 ms can be achieved [91]; 2) recovery is entirely performed at the highly reconfigurable MPLS layer and there is no need for still immature optical control plane solutions; 3) recovery can be performed with LSP granularity, allowing differentiated LSP recovery based on classes if desired. In contrast, note that optical recovery is performed with whole lightpath granularity, thus forcing all carried LSPs in a failed lightpath to be equally recovered.

We target the design of survivable multi-layer optical networks over a Single Line Rate (SLR), Mixed Line Rate (MLR) or an Elastic DWDM optical layer minimizing the total network CAPEX. This design entails significant resource overprovisioning to fit the desired survivability under the highest load period during the day (the peak traffic). Differently from many works focusing on the optical layer alone, where lightpath protection (either 1+1, 1:1 or shared protection) and restoration is adopted, in this section we investigate a network survivability scenario (resilient to 1 link failure) where both layers are jointly optimized.

When adopting SLR to implement the DWDM layer of multi-layer optical networks, all TXPs operate at the same data rate, which cannot be changed dynamically. We assume for SLR that TXPs run at 100 Gb/s with coherent Polarization Division Multiplexing Quadrature Phase Shift Keying (PDM-QPSK). Conversely, TXPs at different data rate coexist in MLR, allowing to better adjust the DWDM layer capacity to the carried traffic, while lowering the number of expensive high-capacity TXPs. We assume for MLR TXPs running either at 10 Gb/s with On-Off Keying (OOK), 40 Gb/s with partial Differential Phase Shift Keying (pDPSK) or 100 Gb/s with coherent PDM-QPSK. TXPs in MLR are also fixed and cannot change their data rate dynamically. Conversely, elastic TXPs show the capability to modify their data rate, adjusting it to the traffic they are supporting. As a short-term viable Elastic technology, we use bandwidth variable TXPs that run either at 25, 50, 75 or 100 Gb/s with coherent PDM-QPSK by adapting the symbol rate to 7, 14, 21 or 28 Gbaud, respectively [92]. This technology is compliant with the 50 GHz ITU-T grid. Table 4 details the transparent reach, TXP cost (in normalized cost units) and power consumption of the SLR, MLR and Elastic technologies. Note that given the presence of OOK signals and dispersion management in MLR networks the reach of 100Gb/s signals is lower than in Elastic scenarios [93].

Table 4: DWDM layer technology details.

	Rate [Gb/s]	Modulation format	Reach [km]	Cost [c.u.]	Power [W]
SLR	100	PDM-QPSK	1200	6	350
	10	OOK	3000	1	50
MLR	40	pDPSK	1600	3	75
	100	PDM-QPSK	800	6	350
Elastic	25	PDM-QPSK	1200	6	189
	50				207
	75				255
	100				350

To assess the advantages of the proposed method, we assume that the traffic matrix and the transparent DWDM layer topology are perfectly known. In the IP/MPLS layer, we assume a router co-located with every optical node in the DWDM layer. Moreover, we assume virtual links at the IP/MPLS layer between all router pairs that can be connected over the DWDM layer with a feasible transparent lightpath at any of the available data rates (i.e., with a physical distance shorter or equal than the maximum transparent reach of the signal). Any virtual link can be composed of multiple lightpaths. The objective of this approach is to design an IP/MPLS over SLR/MLR/Elastic DWDM optical network carrying the predicted peak traffic matrix, so that all packet flows are survivable under any single link failure through IP/MPLS protection switching with the minimum CAPEX. Figure 19 shows an example of the design. As in [94], we assume that optical nodes are equipped with a client-side Fibre Cross-Connect (C-FXC) that enables router ports to be dynamically assigned to optical transponders, thus allowing the reutilization of router ports during failure scenarios.

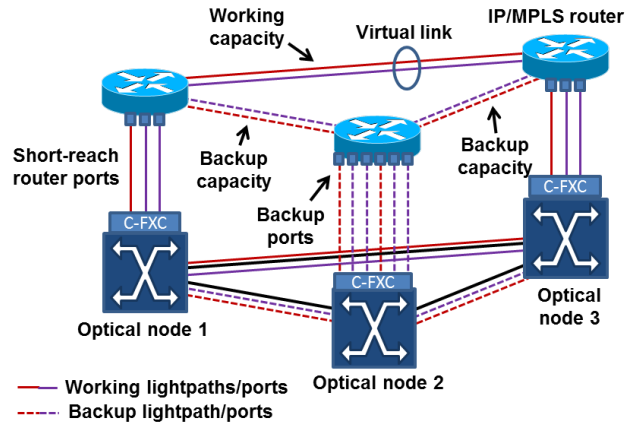


Figure 19: Survivable multi-layer optical network design example.

To solve the dimensioning of such survivable network, we used a two-step solution as described in [95], where firstly we solve the unprotected network design (primary packet flow routes and virtual link lightpaths) minimizing the total network CAPEX, and then we perform a joint optimization of the extra capacity to make the network survivable to all link failure scenarios with minimum CAPEX.

Results are obtained by dimensioning two different backbone network scenarios described in [95] (namely, a Pan-European like network of 11 nodes and 18 bidirectional links, referred as EON, and the Deutsche Telekom National Backbone network of 12 nodes and 20 bidirectional links, referred as DT, defined in D1.1 [67]. In both networks, 16 bidirectional wavelengths per fibre link have been assumed.

Regarding the offered traffic characteristics, a peak traffic matrix with 80 randomly distributed bidirectional packet flow requests at 10 Gb/s (50%), 40 Gb/s (30%) and 100 Gb/s (20%) is separately generated for each network, making up a total offered load around 6 Tb/s in both cases. To populate the set P of physical paths over G onto which virtual link lightpaths are established, only the shortest path between every pair of nodes is computed. About cost details, we consider the TXP costs previously presented in Table 4 and the router ports at 10 Gb/s with cost 1.1 (i.e., 1 c.u. + 0.1 c.u. as the average processing cost of 10 Gb/s in a router, derived from [96]).

Table 5 depicts the details of the survivable IP/MPLS over SLR/MLR/Elastic DWDM optical network design minimizing CAPEX in the evaluated EON and DT network topologies. We observe that the number of equipped devices and cost for an unprotected network design, which helps highlighting the amount of resource overprovisioning required to make the network survivable against any single link failure. The unprotected and survivable scenarios are referred as UN and SURV in the table, respectively. An additional column named Overprov. (%) is also introduced, which quantitatively depicts the percentage of overprovisioned router port and TXP capacity required in every case to make the network survivable.

Table 5: Network design: number of equipped devices and cost.

Scenario	DWDM layer	Router ports		TXPs @ 10G		TXPs @ 40G		TXPs @ 100G		Elastic TXPs		Overprov. (%)		Total cost [c.u.]	
		UN	SURV	UN	SURV	UN	SURV	UN	SURV	UN	SURV	Ports	TXPs	UN	SURV
EON (11 nodes, 18 links)	SLR	780	1052	—	—	—	—	84	140	—	—	34.9	66.7	1362	1997
	MLR	740	1020	62	54	42	80	56	94	—	—	37.8	66.3	1338	1980
	Elastic	784	1042	—	—	—	—	—	—	84	140	32.9	66.7	1366	1986
DT Network (12 nodes, 20 links)	SLR	622	902	—	—	—	—	78	142	—	—	45.0	82.1	1152	1844
	MLR	674	978	28	26	36	50	56	103	—	—	45.1	71.6	1213	1870
	Elastic	636	902	—	—	—	—	—	—	76	142	41.8	86.8	1156	1844

We remark that SLR and Elastic DWDM layer technologies lead to almost identical total network cost. Conversely, MLR DWDM layer technology is more cost-effective than SLR and Elastic. While MLR in EON yields very minor but still appreciable network design cost savings, it leads to slightly increased cost in the DT network. A result like this one can be explained by the smaller transparent reach of the TXPs at 100 Gb/s in MLR compared to SLR and Elastic (i.e., 800 km against 1200 km [93]), which prevents the establishment of high-capacity lightpaths between far-off nodes, thus demanding more router ports due to the additional hops of the packet flows at the IP/MPLS layer. This effect is not appreciated in the EON case although having longer links due to its topological characteristics, making nodes to be uniformly better connected among them. Regarding the percentage of overprovisioned router port and TXP capacity to make the network survivable, we see that less overprovisioning in terms of router port capacity than TXP capacity is required, thanks to the reutilization of router ports among in the different failure scenarios enabled by the equipped C-FXCs. In general, we identify that the DT network requires higher capacity overprovisioning to make the network survivable compared to the EON, which again can be attributed to its less favourable topological characteristics.

As a second step, we quantify the total power consumption of the EON and DT network designs in the non-failure state, that is, the normal state in which the network operates most of the time. Specifically, different offered load values have been considered (100%, 75%, 50%, and 25%) by scaling the load of the original offered demands while keeping demand routes at the IP/MPLS layer fixed. For the IP layer, it is typically assumed that routers consume between 10-20 W/Gb/s [97], including router processing and line cards. In this work, we decouple the power consumption of the processing, which is traffic dependent, from that of the router port short-reach interfaces, which is traffic independent. For the processing, 10 W/Gb/s is assumed, whereas for each short-reach interface, 56 W is considered [97]. Following the same UN and SURV nomenclature, we differentiate between unprotected and survivable network design energy consumption (i.e., outcomes of Step 1 and 2 of the proposed approach, respectively). Results are shown in Table 6 and Table 7.



Table 6: EON network power consumption vs offered load (non-failure scenario).

		Load = 25%		Load = 50%		Load = 75%		Load = 100%	
	DWDM layer	UN	SURV	UN	SURV	UN	SURV	UN	SURV
Optical layer	SLR	29.4	49	29.4	49	29.4	49	29.4	49
	MLR	25.9	41.6	25.9	41.6	25.9	41.6	25.9	41.6
	Elastic	15.9	26.5	17.3	27.9	21.2	31.8	27.9	38.5
IP layer (router ports)	SLR	43.7	58.9	43.7	58.9	43.7	58.9	43.7	58.9
	MLR	41.4	57.1	41.4	57.1	41.4	57.1	41.4	57.1
	Elastic	43.9	58.4	43.9	58.4	43.9	58.4	43.9	58.4
IP layer (processing)	SLR	33.6	33.6	67.1	67.1	100.7	100.7	134.2	134.2
	MLR	32.6	32.6	65.1	65.1	97.7	97.7	130.2	130.2
	Elastic	33.7	33.7	67.3	67.3	101	101	134.6	134.6
Total power consumption (IP + optical)	SLR	106.7	141.5	140.2	175	173.8	208.6	207.3	242.1
	MLR	99.9	131.3	132.4	163.8	165	196.4	197.5	228.9
	Elastic	93.5	118.6	128.5	153.6	166.1	191.2	206.4	231.5

Looking at Table 6, the optical layer power consumption with SLR and MLR DWDM technologies remains constant with the offered load in the EON, since TXPs are always operating at the maximum data rate, thus consuming the same. These power consumption values are substantially lower with MLR, though, which benefits from the lower power consumption of TXPs at 10 and 40 Gb/s. In contrast, Elastic adapts the data rate of the TXPs to the traffic they are carrying, saving power in this way. For instance, focusing in the SURV scenario in Table 6, Elastic can reduce the optical layer power consumption of SLR by 10.5 kW for a 100% offered load, while such a reduction increases to 22.5 kW when the offered load falls to 25%. These power consumption benefits are maximized in the SURV scenario, since the data rate of all those overprovisioned TXPs for survivability, unused in non-failure conditions, can also be lowered to the minimum power state (running at 25 Gb/s). Regarding the IP layer, we can see that the power consumption of the router ports remains constant with the offered load no matter which DWDM layer technology is employed, since short-reach interfaces are always operational. This is not the case of the power consumption due to processing, which is clearly influenced by the offered load to the network. Here, the power consumption differences perceived between mechanisms result from the length of the routes at the IP/MPLS layer in hops (i.e., MLR seems to yield slightly shorter routes at the IP/MPLS layer in the EON). Finally, looking at the total power consumption of the IP and optical layers together, we can observe that MLR is the most efficient DWDM layer technology in terms of power consumption for a load of 100%, but as the offered load starts decreasing Elastic becomes the most interesting option, thanks to its capability to adapt the data rate and thus reducing the power consumption of the TXPs.



Table 7: DT network power consumption vs offered load (non-failure scenario).

		Load = 25%		Load = 50%		Load = 75%		Load = 100%	
	DWDM layer	UN	SURV	UN	SURV	UN	SURV	UN	SURV
Optical layer	SLR	27.3	49.7	27.3	49.7	27.3	49.7	27.3	49.7
	MLR	23.7	41.1	23.7	41.1	23.7	41.1	23.7	41.1
	Elastic	14.4	26.9	15.6	28	18.9	31.3	23.8	36.3
IP layer (router ports)	SLR	34.8	50.5	34.8	50.5	34.8	50.5	34.8	50.5
	MLR	37.7	54.8	37.7	54.8	37.7	54.8	37.7	54.8
	Elastic	35.6	50.5	35.6	50.5	35.6	50.5	35.6	50.5
IP layer (processing)	SLR	29.9	29.9	59.8	59.8	89.7	89.7	119.6	119.6
	MLR	31.2	31.2	62.4	62.4	93.6	93.6	124.8	124.8
	Elastic	30.3	30.3	60.5	60.5	90.8	90.8	121	121
Total power consumption (IP + optical)	SLR	92	130.1	121.9	160	151.8	189.9	181.7	219.8
	MLR	92.6	127.1	123.8	158.3	155	189.5	186.2	220.7
	Elastic	80.3	107.7	111.7	139	145.3	172.6	180.4	207.8

Similar results but for the DT network are shown in Table 7. Regarding the optical layer power consumption, SLR is the DWDM technology providing the least efficient performance, followed by MLR and Elastic as the most efficient one even with 100% load. As for the IP layer power consumption due to the router ports, all DWDM technologies remain constant with the offered load as before (short-reach interfaces are always operational). Nonetheless, MLR consumes significantly more in this scenario, due to the large number of ports that the DT network design requires (i.e., as shown in Table 5). Moreover, since the routes at the IP/MPLS layer with MLR are longer in terms of hops, the IP layer power consumption due to processing is also higher. This eventually makes that the total power consumption of the IP and optical layers together in the DT network with MLR becomes significantly higher than Elastic for all offered loads, even for a 100% offered load.

We can also compare the power consumption of the additional resources overprovisioned for survivability in both EON and DT network scenarios, namely, the power consumption differences between UN and SURV network designs. For instance, in the EON network for a 100% load the overprovisioned capacity at the optical layer consumes 19.6 and 15.6 kW, which is reduced to 10.6 kW in Elastic by setting the data rate of the overprovisioned TXPs for survivability to the minimum power state (i.e., operating at 25 Gb/s). Similar differences are also observed in the optical layer of the DT network. Regarding the IP layer, the power consumption of the overprovisioned router ports is directly proportional to the router port capacity overprovisioning percentages disclosed in Table 5, that is, around 33-38% and 42-45% of the router port power consumption in the UN scenario with the EON and DT network, respectively. Results in Table 6 and Table 7 allow to easily quantify the daily power consumption of survivable IP/MPLS over SLR/MLR/Elastic DWDM networks under a typical daily traffic profile as the one presented in [92]. In such a traffic profile, traffic variations are discretized in 8 periods along a day, where the offered traffic fluctuates among 100% (11:30 am-19 pm, 22-23 pm), 75% (9-11:30 am, 19-22 pm, 23-24 pm), 50% (0-2 am, 7-9 am) and 25% (2-7 am) of the peak traffic. These results for the EON and DT networks are depicted in Figure 20.

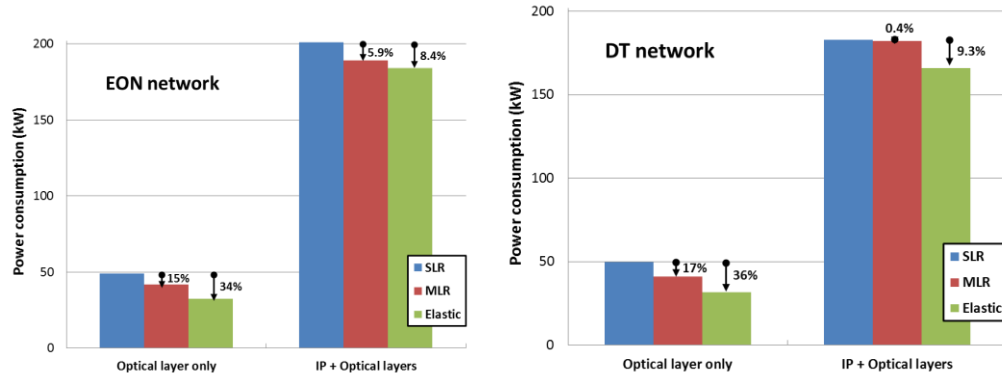


Figure 20: Daily power consumption (in kW) of the survivable IP/MPLS over SLR/MLR/Elastic DWDM layer design: EON network (left), DT network (right)

Looking at Figure 20, and focusing only on the optical layer, we observe that although MLR achieves significant power consumption reduction against SLR in both EON and DT networks, this one is clearly outperformed by Elastic, which can even double such a reduction (i.e, 34% vs. 15% in the EON network, and 36% vs. 17% in the DT network). Now, if we consider both IP and optical layers together, we can also see that the power consumption benefits that Elastic yield in the optical layer are indeed reflected in the overall network power consumption as well. For instance, by deploying Elastic DWDM technology in the EON network, an overall power consumption reduction of 8.4% can be archived, whereas this overall power consumption reduction increases to 9.3% in the DT network. Regarding MLR, it still provides appreciable benefits on the overall power consumption of the EON network (5.9%), but fails in this endeavour in the DT network, leading to only 0.4% overall power consumption reduction there, due to its additional router port necessities and longer routes at the IP/MPLS layer.

4.4.2 Elastic Optical Networks to leverage Energy Savings

In this section, we investigate different energy reduction strategies using power-adaptive interfaces adjusting their power to the transmitted traffic, following predictable fluctuations. Moreover, an optical grooming approach is proposed in order to use the benefits of traffic fluctuations and adaptations of the optoelectronic (OE) devices. The advanced solutions for power adaptability as well as periodical reconfigurations come at the expense of increased complexity within the control plane.

Two power-adaptive approaches are studied for adapting the OE energy to the carried traffic: sleep-mode and data-rate adaptation. Hereafter OE devices considered are transponders and regenerators.

Sleep-mode capability relies on the possibility of changing the power state of OE devices. They can be totally or partially powered (up and idle states, respectively) or switched off (down state). When an OE device is in idle state, it can be switched on faster in order to improve network QoS (bandwidth, latency, availability, etc.) for new traffic demands at the expense of low power consumption in idle state with respect to the up state, as in [98].

Data-rate adaptation can be implemented through modulation-format (MF) or symbol-rate (SR) adaptation [92]. MF adaptation may allow power reduction if passing from a more

complex format to a simpler one, thus being possible to bypass regenerators, as the optical reach increases with the decrease of the modulation format complexity. Moreover, sleep-mode has to be implemented as well for turning off intermediary regenerators. It has been demonstrated in [92] that with the SR adaptation, the power of an OE device scales down linearly with the reduction of its symbol-rate. However, in case of SR adaptation no regenerator is skipped as the reach is not SR dependent and thus sleep-mode is not necessary for performing SR adaptation [92]. The combination of SR and MF is also proposed, relying on the optimum choice in terms of power consumption between the two adaptations and is called ‘mixed’ in the following.

We propose also to add **grooming** capability when routing a demand. When an optical connection has to be set-up, the policy controller checks if there is an existing optical channel having the same source and destination pair. In case at least one exists and the sum of the capacity of the existing and incoming connections is lower than 100 Gb/s (channel maximum capacity) during the common lifetime of the two connections, then grooming is feasible and the incoming connection is groomed into the existing one.

If MF or SR adaptation is performed, the channel data-rate is re-configured when the connection capacity changes. In this study, the considered rate-adaptive OE devices can handle: 25, 50, 75 and 100 Gb/s, like in [98]. Table 8 provides the power consumption values of the transponders considered in this work. Regenerators consume twice the energy of the transponders working at the same rate, as they are realized by two transponders back-to-back. When OE devices are idle their consumption is 18W for a transponder and 36W for a regenerator [92]. As in this study we want to emphasize the impact of regenerators in the network power consumption, reach values are reported in Table I, but they are halved compared to values reported in [92]. In Table 8 the eight possible SR and MF configurations of an OE device for carrying the required data-rate are shown.

Table 8: Transponders power consumption as a function of symbol-rate (SR) and modulation format (MF).

Payload (Gb/s)	SR (Gbd)	MF	Reach (km)	Power (W)
100	28	PDM-QPSK	600	350
	21	PDM-QPSK	600	255
50	28	PDM-BPSK	1250	350
	14	PDM-QPSK	600	206
25	28	SP-BPSK	1500	350
	14	PDM-BPSK	1250	206
	7	PDM-QPSK	600	189

To estimate the energy savings provided when traffic over-provisioning is considered, we dimension two different network topologies, a national coverage network DT17 (an extension of the DT network with 12 nodes considered in D1.1 and a pan-European network COST37, as described in [99]. Traffic is generated in all nodes, connections are uniformly distributed and their peak capacity can be 50, 75 and 100 Gb/s. The traffic



demands are requested with an exponentially distributed mean inter-arrival time of 2 hours and hold on with an exponential distribution of a mean value of 40 hours, providing a load per node of 20 Erlang. Every link carries up to 80 wavelengths and no blocking is experienced. The traffic taken into consideration follows a weekly variation as derived from statistics in [100].

The baseline scenario sets-up all connections at 100 Gb/s, regardless of their actual requested capacity (always ≤ 100 Gb/s). When a connection is released, OE devices are turned off. To dimension the network, the peak-capacity is considered and necessary OE devices are deployed. The number of requested OE devices is independent of the chosen power-adaptation strategy. In Table III, the average number of transponders and regenerators which have to be installed in every node for enabling a zero blocking is given. The following energy savings strategies are compared:

- *On/Off* scenario (baseline scenario): only sleep-mode is used and all connections are set-up to 100 Gb/s; the minimum amount of required regenerators is configured along the path to support the 100 Gb/s connection based on a heuristic algorithm (i.e. it assigns a regenerator to every node when the distance between two consecutive regeneration hops is longer than the maximum reach of the selected MF). As traffic is dynamic, the static power management is not considered and this strategy is used for the comparison with the other energy efficient strategies.
- *MF* scenario: modulation-format adaptation is performed together with sleep-mode strategy; the less complex modulation format capable to carry the incoming traffic is chosen and only used regenerators are powered on.
- *SR* scenario: symbol-rate adaptation is performed without sleep-mode. Required regenerators are the same as for 100Gb/s connections, but their power is proportional to the used symbol-rate, which adapts to traffic variations.
- *Mixed* scenario: combines MF and SR and always chooses their combination providing the minimum power consumption for the considered path and required data-rate.
- *On/Off, MF, SR and Mixed with grooming* – all previous strategies used together with grooming algorithm.

Simulations emulate a one-year period. A number of 10 different random distributions of traffic are drawn with the purpose of minimizing, through averaging, the influence of specific random traffic patterns on the overall results.

4.4.2.1 Power consumption for different energy savings strategies

Figure 21a shows the average power consumption along the year for the DT17 network topology taking into consideration four scenarios: On/Off only, MF, SR, Mixed and Mixed together with grooming enabled. The small fluctuations for On/Off case are due to the random set-up of connections. Also, for the On/Off scenario, the increase of the energy consumption during the year is almost imperceptible because all connections are set-up at 100 Gb/s. For the rate-adaptive scenarios, the effect of the traffic growth is visible because for the MF scenario more regenerators are set-up (due to the use of more complex MF, see Figure 22a), whereas for the SR adaptation, higher and more energy hungry SR is used. For both MF and SR scenarios we observe that the energy consumption follows the traffic variations presented in [99]. Concerning MF, we observe that its average power consumption does not follow daily variations during the first half-year period and afterwards

peaks appear and correspond to peak traffic. This is due to an increase in the number of needed and powered on regenerators when peak traffic overpasses OE devices data-rate levels.

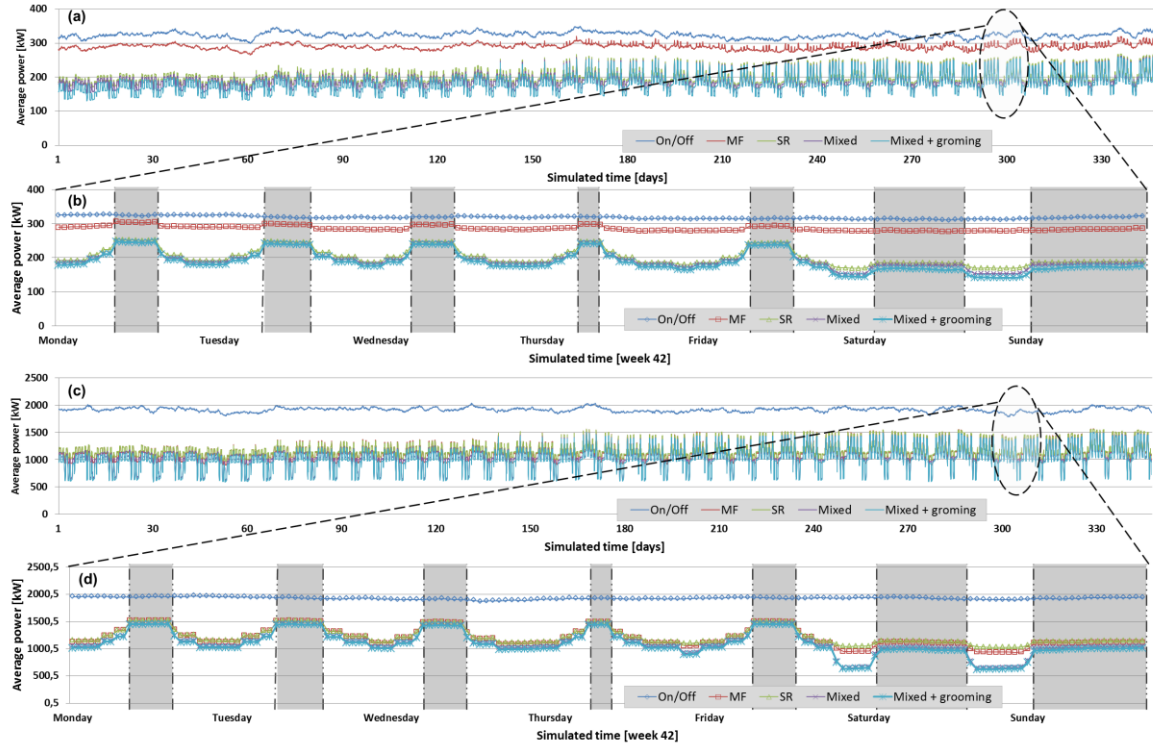


Figure 21: Average power variation per year for DT17 (a) and COST37 (c) and average power per week for DT17 (b) and COST37 (d) network scenarios.

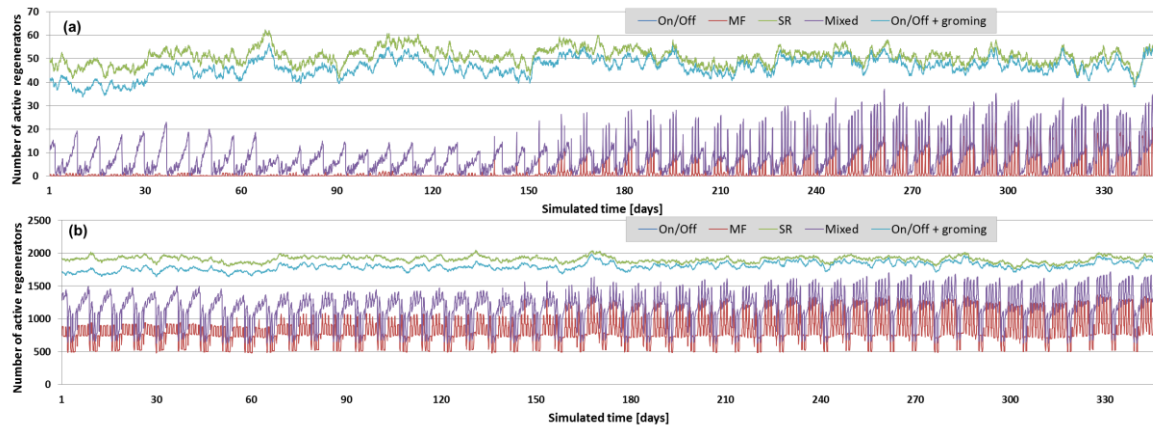


Figure 22: Number of active regenerators for DT17 and COST37 network scenarios.

Figure 21b zooms on the average power variation of week 42 for DT17 network scenario. It can be observed that power is relatively constant for On/Off strategy, which is expected since this does not involve any adaptation to fluctuations. Moreover, SR and Mixed



adaptations are able to significantly adapt the power consumption to day/night and weekend fluctuations with Mixed adaptation achieves the lowest power consumption. MF adaptation follows as well traffic fluctuations and provides reduced power consumption but significantly higher than SR and Mixed scenarios since the low number of regenerators deployed in the network does not permit significant adaptations based on the regenerators' sleep-mode capability.

Figure 21c shows the power consumption for the COST37 network scenario. Results are similar with the ones in Figure 21a as the power consumption for On/Off scenario is relatively constant but the average power is approximately six times higher than in case of DT17 network, despite the fact that COST37 network only has 2.1 times more nodes and links. The reason why power consumption grows six times is given by the average length of the links which is 3.8 times higher for COST37 compared to DT17 as well as by the number of nodes which is 2.1 times higher for COST37. Thus, the longer links demand for a higher number of regenerators while the higher number of nodes determines a higher number of transponders (since the number of connections generated in every node is constant). Moreover, Figure 21c shows that the power associated to MF scenario is not any longer relatively flat during the first half of the year as for DT17 network in Figure 21a. For COST37 network, due to the long length of the links the number of regenerators is very well adapted to the traffic fluctuations and determines a power consumption which varies with similar fluctuations as the input traffic (Figure 22b). In case of SR and Mixed adaptations results are similar, from the variation trend point of view, with the ones obtained for DT17 network.

In Figure 21d a zoom over power consumption in week 42 for COST37 network is provided. When comparing results in Figure 21b and Figure 21d it can be observed that MF, SR and Mixed adaptation (w/o grooming) achieve similar power consumption while a significant difference is given for the night periods of the weekends where Mixed adaptation achieves a significantly lower power consumption compared to MF or SR adaptations. Another consequence of the high number of OE devices (i.e. they allow for a better power adaptation which translates into higher power reductions) is emphasized by power fluctuations in Figure 21b and Figure 21d. Hence, for SR and Mixed scenarios the power fluctuations amplitude is significantly higher in the COST37 network where the power consumption varies from 700 kW to 1500 kW during week 42 (53% variation) than in DT17 network where variation is from 160 kW to 250 kW (36% variation).

Figure 22 depicts the number of active regenerators in DT17 and COST37 network topologies during the year for On/Off, SR, MF, Mixed and On/Off with grooming scenarios. It can be observed that for both networks, the number of active regenerators in On/Off and SR scenarios is the same (overlapped graphs) since both configure the connections with a number of regenerators equivalent to 100 Gb/s connections. The difference however, in terms of consumed energy is high (as seen in Figure 21) due to the possibility of tuning symbol-rate for SR scenario. Moreover, when comparing the two topologies, we observe that in DT17 network for MF scenario the number of regenerators is almost zero during the first half of the year; so mainly the transponders contribute to power adaptation for DT17 network. In case of COST37, due to the presence of high number of active regenerators, they are well adapted according to the traffic fluctuations. In case of Mixed scenario, for both networks the number of active regenerators is slightly higher compared to the MF scenario and this is due to cases when energy is reduced by reducing the symbol-rate

instead of reducing the number of active regenerators. Figure 22 emphasizes as well the case of grooming applied to the baseline (On/Off) scenario and the conclusion is that when grooming is enabled, a lower number of regenerators is required (since grooming determines fewer active connections, due to grooming). However, towards the end of the year the difference in the number of regenerators between On/Off with and without grooming becomes close to 0 since the data-rate fluctuations grow within every connection, making it less likely for two connections to be groomed without exceeding the maximum available capacity per connection of 100 Gb/s.

4.4.2.2 Power savings and energy efficiency

Figure 23 shows the relative difference between the average and the peak traffic over diverse time periods: daily, weekly, and yearly. Thus, if the input traffic model only considers daily variations, ideally 23.4% energy savings could be achieved if the power adaptation perfectly follows the traffic variation along the year. Moreover, if the weekly variations are accounted in the traffic profile, 11.9% yearly energy savings are assessed. If the yearly traffic growth is accounted as well, up to 48% savings can be reached; such values strongly depend on the considered traffic profiles.

Figure 24 shows the power savings for the diverse power-adaptive strategies deployed in the two network topologies. We notice that in all four cases (without optical grooming), the energy savings corresponding to COST37 network are higher and this is due to the considerable variation in number of regenerators for COST37 with respect to DT17. Moreover, for DT17 network SR provides higher savings compared to MF adaptation (37.5% and 10.6% respectively); about COST37 network, energy efficiency changes: 39.4% savings for MF adaptation and 37.7% for SR adaptation. When allowing a mixed rate adaptation 5% higher savings are achieved for COST37 network (40.1% for DT17 and 45.1% for COST37).

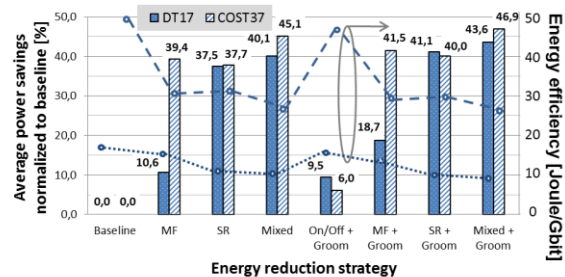
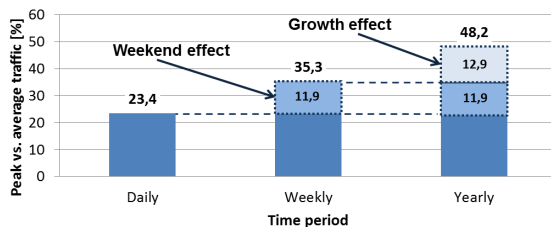


Figure 23: The normalized difference of average and peak traffic. **Figure 24: Power savings and energy efficiency for DT17 and COST37.**

The grooming approach without rate adaptation has a different efficiency for the two network topologies and it allows for 9.5% and 6% savings for the DT17 and COST37 scenarios respectively. This is explained by the fact that in the COST37 network there are more (36 compared to 16) possible destinations for a connection generated in one node, hence lower probability of having a different connection between the same source-destination nodes. Thus, while having the same load per node (same number of connections generated in one node), it is more difficult to find a connection with same source/destination pair for DT17 which allows grooming for a new incoming connection.



Thereby, when grooming is combined with the MF, SR and Mixed adaptations SR yields higher savings for DT17 compared to COST37. Generally, we observe that the proposed grooming algorithm improves the energy efficiency in networks having a limited number of nodes.

The advantage of performing only SR adaptation with grooming is the lower complexity of the control plane and the signaling process, because there is no need for regenerator set-up/release due to MF adaptation. Higher savings are reported for mixed and grooming strategies. For DT17 network they are 43.6% and 46.9% for COST37.

In terms of energy efficiency, for DT17 network the energy efficiency can be decreased from 17 J/Gbit in baseline scenario to 9.6 J/Gbit in Mixed with grooming scenario, while for COST37 network the energy efficiency decreases from 49.8 J/Gbit to 26.4 J/Gbit for the same scenarios.

5 Main algorithm results

In deliverable D1.3 [51] a number of algorithms developed during the first year of the IDEALIST project was presented and described. Table 7 (for off-line) and Table 8 (for on-line) of D1.3 include all this algorithms with their main characteristics. The progresses on the algorithms for flexgrid networks achieved in the second year of the project are described in this Section.

In particular, as far as off-line algorithms as concern, the improvements include:

- A method to improve the performance of a Column Generation method (that is, for the nature of the problem, a Path Generation method) already presented in D1.3 for solving the RSA problem, by the introduction of additional constraints in the MIP formulation, called *clique inequalities*, that strengthen the formulation and lead to the generation of more suitable paths. It is shown on some application cases that this method improve the quality of the achieved solutions if compared with pure Column Generation or Branch and Bound methods, when the budget of computation time is limited and assumed to be a practical value.
- A new algorithm for gradual network design called **GRANDE**, which propose a Path Generation method for solving the problem of the flexgrid optical network design in case of incremental design, i. e. the additional least cost upgrade in terms of equipment (new nodes, part of nodes and system on links) when an updated traffic matrix is assigned.
- A new algorithm that solve the problem of *multicast routing* (P2MP) in case of flexgrid networks. A method that apply a *sub-tree scheme* for routing the multicast flows has been found the better among the ones considered. To compare the performance of the investigated methods the P2MP-RSA problem was stated and modelled using an ILP formulation.
- A new version of the multi-layer planning problem of a flexible optical network. The examined multi-layer planning problem consists of three sub-problems at two layers: the Routing problem at the IP-layer (IP-R) and the Routing, Modulation Level (RML) and Spectrum Allocation (SA) problems at the optical layer. Compared to previous versions a more realistic model of the IP layer has been considered. In



addition a distance adaptive optical transponders (the RMLSA problem) have been introduced. Distance adaptability creates interdependencies between the routing at the optical and the IP layers and this feature, which makes the problem not decomposable and very hard to solve, is a novelty not only in the context of IDEALIST studies but, to the authors' knowledge, also in this whole research field.

Turning to online algorithms the novel findings of the project concern:

- A new *reactive algorithm* for spectrum defragmentation called **SPRING**. It applies the concept of spectrum shifting for defragmentation. Spectrum shifting means that the already established connections are shifted, not reallocated as it happens for instance in the algorithm EL-SPRESSO described in [51] to make room for a new connection). The ILP model to solve the defragmentation problem is stated and a heuristic algorithm providing much better trade-off between optimality and complexity is proposed. The performance of the SPRING algorithm was compared to that of EL-SPRESSO and the results showed a good performance and shorter computation times.
- An updated version of the algorithm that solve the **AFRO** problem (already introduced in D1.3), which is the problem to optimize the pool of established lightpaths once a link failure is repaired, with the target to improve load balancing and dynamic network performance in flexgrid networks. AFRO was defined by means of two alternative link-path IP formulations. Based on previous work on column generation for a basic RSA problem in flexgrid optical networks, specific procedures for both presented formulations were developed. Numerical results show that AFRO reduces both the amount of used capacity and the spectrum fragmentation (measured in terms of link entropy) with respect to the network state before applying AFRO. The effectiveness of AFRO in improving the performance of a dynamic network in operation is demonstrated, in particular with a study on the effect of connections holding times and failures (MTTR). Finally, it is shown that the operational overhead of AFRO can be reduced by using its mechanism only for a selected set of links without a negative impact on its performance.

Some of the developed algorithms in WP1 were selected to be implemented and tested in WP4 by means of a flexgrid network control plane prototype. They include an algorithm for spectrum defragmentation, an algorithm for the re-optimization of network resources after link failure repair (AFRO), an algorithm to perform the Inter DC Multicast connectivity and a heuristic algorithm called KSP Distance Adaptive Multifiber Spectrum Assignment (KSP-DA-MSA) for adapting the modulation format of the signal to the distance of the path.

The same algorithms are also selected to be implemented in one of the Planning and Simulation tool for flexgrid network developed within the project, which is PLATON [50]. They are all the mentioned above for implementation in WP4 control plane prototype except the last one (KSP-DA-MSA).

5.1 Network Optimization and Off-line Algorithms

5.1.1 Valid inequalities and a column-and-cut generation for RSA-based problems

The RSA optimization problem is *NP*-hard. To make large instances of the RSA-related network design problems tractable by mixed-integer programming (MIP) formulations, decomposition methods must be applied. Such methods usually involve the dynamic addition of variables (columns) and/or constraints (cutting planes, cuts) to the MIP model. In [51], a novel column generation-based algorithm (referred to as CG) for dynamic generation of lightpaths for the link-lightpath (LL)-based MIP formulations of RSA has been proposed (we refer to [51] for more details on LL). As a continuation of this work, in this section we introduce some valid inequalities (cuts), namely clique inequalities, that strengthen the LL formulation. Then, we develop and combine a clique cut generation procedure with a column generation algorithm with the aim to improve the quality of generated columns. The presented results, obtained for a set of problem instances, illustrate the effectiveness of the optimization algorithm.

Clique inequalities

The LL model assures that in a MIP solution there is at most one selected lightpath for demand d as well as there is at most one active lightpath that uses slice s on link e . It is reflected by variable x_{dl} that takes a positive value for each such lightpath (in fact, $x_{dl} = 1$ in a MIP solution) and is equal to 0 for the rest of lightpaths. However, for a solution of a linear problem (LP) relaxation, in which the integrality of variable x_{dl} is relaxed and substituted by constraint $x_{dl} \geq 0$, it is not the case. To illustrate it, in Figure 52 we present an example of two allowable lightpaths for demand 1 (denoted as 11 and 12) and one allowable lightpath for demand 2 (denoted as 21) sharing some slices in a network link. We can see that lightpaths 11 and 12 are in conflict since only one of them can be active in a MIP solution. Similarly, there is a conflict between lightpaths 11 and 21, and between 12 and 21, since they occupy common frequency slices in the link. Still, in the relaxed problem the values of variables x_{dl} can be all positive (here, equal to 0.5).

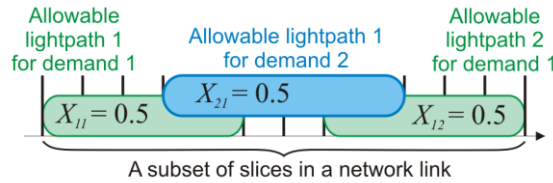


Figure 52: An exemplary clique of lightpaths.

Now, note that only one of these three lightpaths can be active in a MIP solution since they are in conflict with each other. Such a relation can be represented by inequality $x_{11} + x_{12} + x_{21} \leq 1$. Note that this inequality:

1. is valid with respect to the feasible solution space of the RSA problem,
2. when added to the MIP formulation, it can cut off a solution of the relaxed problem. Indeed, in our example this inequality is violated by the relaxed solution since we have $x_{11} + x_{12} + x_{21} = 1.5 > 1$.

In the following, we formalize the above observations. We say that a subset of allowable lightpaths $\kappa = \{l_1, l_2, \dots, l_{|\kappa|}\} \subseteq \mathcal{L}$ is in conflict if at most one of these lightpaths can appear in a MIP solution to the RSA problem. We apply a notion of clique [56] when we refer to such a set of lightpaths. For clique κ , we have the following valid clique inequality:

$$\sum_{l \in \kappa, d=d(l)} x_{dl} \leq 1,$$

where $d(l)$ denotes the demand of lightpath l .

Now, let $\bar{x} = (\bar{x}_{11}, \dots, \bar{x}_{1|\mathcal{L}(d)|}, \dots, \bar{x}_{|\mathcal{D}|1}, \dots, \bar{x}_{|\mathcal{D}||\mathcal{L}(|\mathcal{D}||)|})$ denote the current LP solution of the relaxed problem. To be useful as a cut when solving the MIP problem, a clique inequality should be able to cut off \bar{x} when it is non-integral. The separation problem for the class of clique inequalities is a maximum weighted clique (MWC) problem [57] on a conflict graph. The MWC problem is *NP*-complete [57].

Let $\mathcal{G}^c = (\mathcal{L}, \mathcal{Y})$ be the conflict graph for our MIP problem, where \mathcal{L} is the set of allowable lightpaths, and \mathcal{Y} is the set of edges connecting pairs of lightpaths being in conflict. Also, we assume that each vertex $l \in \mathcal{L}$ in graph \mathcal{G}^c has associated a weight corresponding to the value of variable $\bar{x}_{d(l)l}$ in the current LP solution. Having graph \mathcal{G}^c built, the MWC problem concerns finding in \mathcal{G}^c a complete subgraph with the maximum sum of the weights of its vertices. The vertices in this subgraph represent a clique and if the above clique inequality is violated by this clique then it can be added to LL formulation to cut off the current LP solution \bar{x} .

Optimization algorithm

The algorithm is implemented in a price-cut-and-branch (PCB) framework, which consists of two phases: first, a set of allowable lightpaths is generated using a column and cut generation (CCG) algorithm, then, the MIP problem is solved for the obtained set of lightpaths using a branch-and-bound (BB) solver. A block diagram of the PCB algorithm is presented in Figure 53.

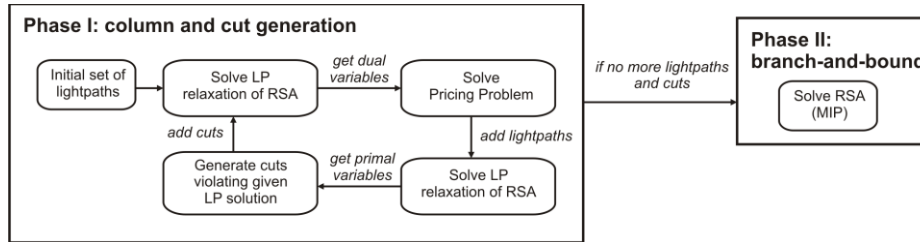


Figure 53: Optimization algorithm.

Phase I: column and cut generation

The CCG algorithm is a modified version of the CG algorithm presented in [51] and the modification concerns the application of clique cuts for strengthening the LL-based MIP formulation of RSA. In details, in CCG an LP relaxation of LL is initiated and solved with a small, feasible set of allowable lightpaths, denoted as \mathcal{L} . In our implementation, this initial set corresponds to the set of all allowable lightpaths, for each demand d , established on a shortest route. Apart from that, a set of clique cuts (i.e., violated clique inequalities),



denoted as \mathcal{K} , is generated for these lightpaths and included into the formulation. Sets \mathcal{L} and \mathcal{K} are then extended iteratively with new lightpaths and cuts. CCG terminates if neither new lightpaths nor cuts can be found and added to \mathcal{L} and \mathcal{K} , respectively.

A key element of column generation is to formulate and solve a pricing problem, which consists of finding such a new lightpath that, when included into the problem formulation, it leads to the improvement in the objective function value of the LP problem relaxation. For the LL formulation, the pricing problem reduces to the search for lightpath $l = (p, c)$ for which its reduced cost, calculated as $\lambda_{d(l)} + g(l) - \sum_{e \in p} \sum_{s \in c} \pi_{es}$, is positive. Note that for a given LP solution the values of dual variables (denoted above as λ_d and π_{es}) are also given – we use them to calculate the reduced costs. In each iteration of our CCG, for each demand, we include into set \mathcal{L} a lightpath with the largest positive reduced cost, if such lightpath exists. For more details on deriving reduced costs in MIP formulations of RSA, refer to [51].

Note that in principle the dual variables related to cut constraints (i.e., clique inequalities) need also to be respected when calculating reduced costs in the pricing problem. If they are not, the column generation procedure may re-generate “cut off” variables and end up in an infinite loop of separation and pricing [58]. However, the verification which cuts should contribute with their dual variables to the reduced cost of a new lightpath is not trivial and complicates the pricing problem. Therefore, to mitigate this problem, we consider that a (large) set of candidate lightpaths is given and a new lightpath is selected from this set under the condition that it is not yet included in $\mathcal{L}(d)$. By these means, we automatically assure that the new lightpath is not included in any cut in \mathcal{K} . A cost of neglecting the cut-related dual variables in the pricing problem is that we may lose the strength of cuts since we do not lift them with new variables [58].

To generate clique cuts, we make use of weighted conflict graph \mathcal{G}^c (as discussed in the previous subsection). Graph \mathcal{G}^c is built for an initial set of allowable lightpaths \mathcal{L} and its vertices and edges are extended for generated lightpaths at each algorithm iteration. The weights of the vertices of \mathcal{G}^c correspond to the components of a given LP solution \bar{x} , thus they change at each iteration. A key element of clique generation is the search for a maximal complete subgraph in \mathcal{G}^c for which the sum of the weights of its vertices is greater than 1. To find such subgraphs, we apply a modified version of the branch-and-bound algorithm solving a maximum clique (MC) problem [59]. The objective of the original MC algorithm is to find a clique that consists of a maximum number of vertices. To be useful for our CCG algorithm, the modification of MC consists in: a) taking into account the weights of cliques when looking for maximal cliques in \mathcal{G}^c , and b) providing a number of maximal cliques with just one algorithm run. To speed up the procedure, the MWC search is performed first on a subgraph of \mathcal{G}^c including only the vertices with positive weights, and then after finding a set of cliques on this subgraph, these cliques are lifted by considering the rest of vertices from \mathcal{G}^c (see [57] for more details).

Phase II: branch-and-bound

Once set \mathcal{L} is found using the CCG procedure, it is used as a set of allowable lightpaths in the MIP problem and this problem is now solved using branch-and-bound. Since set \mathcal{L} is much smaller than the set containing all possible lightpaths, the MIP problem can be



solved in shorter times. Note that although the optimality of the LP problem relaxation is guaranteed after column and cut generation, still the obtained set of allowable lightpaths does not necessarily lead to an optimal solution of the MIP problem.

Illustrative results

The evaluation is performed for the Telefonica reference network. We consider the flexgrid of 12.5 GHz granularity and different values of the available spectrum in network links (denoted as S). The spectral efficiency of a modulation format is equal to 2 bit/s/Hz. The demand sets are randomly generated. We have three types of demands with bit-rate requests $(h(d), H(d)) \in \{(10,40), (40,100), (100,200)\}$ Gbit/s, where $h(d)$ and $H(d)$ denote the minimum and maximum bit-rate, respectively. The percentage of each demand type in the demand set is distributed as 25%:25%:50%, respectively. The optimization objective is to minimize the number of rejected demands (primary objective) and the amount of unserved bit-rate (secondary objective).

As reference algorithms we use a column generation (CG) algorithm, which is a reduced version of CCG that does not include cut generation, and a branch-and-bound (BB) solver solving the LL formulation for a complete set of candidate lightpaths. The number of candidate routing paths $k \in \{10,30\}$ and the set of candidate lightpaths consists of all possible lightpaths established on these paths and allocating any, appropriate for given demand, segment of spectrum on the flexgrid. The algorithms are implemented in C++. We use CPLEX v.12.5.1 [18] as a LP solver in the column and cut generation phase and as a BB solver in the final algorithm phase. Numerical experiments are performed on a 2.5 GHz i5-class machine with 8 GB RAM. We set a time limit equal to either 1 or 10 hours for solving MIP problems in CPLEX.

When comparing the algorithms, our main focus is on the best integer solution found, denoted as z_Y^X , where $X \in \{BB, CG, CCG\}$ denotes the algorithm and $Y \in \{1h, 10h\}$ denotes the CPLEX run-time limit. Apart from that, we report: the number of initial allowable lightpath (L^{init}), the number of generated lightpaths (L^{gen}), the number of generated cuts (K^{gen}), the value of LP relaxation after performing the CG and CCG procedure (z^{LP}), the CPLEX gap between the best integer solution and a lower bound solution (gap), and algorithm runtime (T).

Table 16: Comparison of algorithms (best results marked in bold); $|\mathcal{D}| = 64, k = 30$.

S [GHz]	BB				CG						CCG							
	L^{init}	z_{1h}^{BB}	gap	T	L^{init}	L^{gen}	z^{LP}	z_{1h}^{CG}	gap	T	L^{init}	L^{gen}	K^{gen}	z^{LP}	z_{1h}^{CCG}	gap	T	
100	30240	20940	0%	12	400	294	20815	21175	0%	2	400	1842	1878	20921.2	20940	0%	39	
150	59040	13300	0%	351	656	633	13097.5	13350	0%	7.6	656	2366	2733	13122.5	13300	0%	299	
200	87840	7930	3.1%	3600	912	690	7680	8155	0%	1788	912	1299	1571	7680	8005	0.6%	3667	
250	116640	4275	22%	3600	1168	741	3350	3855	0%	938	1168	1301	1802	3350	3785	2.5%	3686	
300	145440	2555	20%	3600	1424	575	2040	2590	10%	3610	1424	657	830	2040	2615	9.2%	3626	
350	174240	1840	24%	3600	1680	1250	1400	1810	18%	3637	1680	1331	1953	1400	1810	21%	3713	
400	203040	4000	75%	3600	1936	638	1000	1350	0%	1933	1936	703	1308	1000	1335	1.3%	3641	
450	231840	4015	85%	3600	2192	572	600	1025	0%	637	2192	708	1168	600	865	0%	1529	
500	260640	4015	95%	3600	2448	820	200	690	8.9%	3623	2448	1016	1417	200	410	0%	1270	

In Table 16, we present numerical results obtained for a set of $|\mathcal{D}| = 64$ demands, $k = 30$, and for different values of available spectrum (S) after 1 hour of CPLEX computation. We can see that the BB algorithm is able to provide optimal (or near-optimal) solutions only for



the scenarios with low values of S . Concurrently, both CG and CCG, which require a smaller set of allowable lightpaths, are always able to find solutions close to the LP lower bound solution (z^{LP}). We can also see that CCG outperforms CG in almost all cases and it is able to find optimal results for smaller scenarios at the cost of some increment in computation time (required for column and cut generation). In smaller scenarios (with $S \in \{100, 150\}$), CCG provides better lower bounds (z^{LP}) than CG.

Table 17: Averaged results over 5 experiments for larger scenarios.

S [GHz]	$ \mathcal{D} $	k	BB		CG				CCG					CCG vs. CG	
			L^{init}	z^{LP}	L^{init}	L^{gen}	z_{1h}^{CG}	z_{10h}^{CG}	L^{init}	L^{gen}	F^{gen}	z_{1h}^{CCG}	z_{10h}^{CCG}	Δ_{1h}	Δ_{10h}
500	64	30	260640	220	2448	796.8	713	713	2448	1043.2	1476.6	575	573	24%	24%
850	112	10	269640	704	7420	2435.2	1703	1567	7420	2938.6	2474.8	1576	1407	8.1%	11.4%
1200	160	10	553200	692	15080	3578.8	1893	1626	15080	4208.8	3273.2	1829	1525	3.5%	6.6%

In Table 17, we present the results obtained for larger problem instances. In each scenario, the evaluation is performed for 5 different demand sets and the presented results are averaged. We report the best solution found (z_Y^X) after both 1 hour and 10 hours of CPLEX computation. We do not present the results for BB since either their quality is low or the algorithm is not able to find any in a given time limit, which is mainly due to the large number of allowable lightpaths (L^{init}) given in these scenarios. Both CG and CCG are still able to provide good feasible solutions and, on average, the quality of CCG solutions is better than after running the CG procedure only. The average relative difference (denoted as Δ_Y) between the results of z^{CG} and z^{CCG} is up to some tens of percents.

Concluding remarks

Valid clique inequalities for a natural MIP formulation of the RSA problem have been developed. To generate the clique cuts we have proposed a cut generation procedure. The generated cuts have been used in a column generation algorithm with the aim to strengthen the MIP problem formulation and, by these means, to improve the quality of generated columns. The presented results, obtained for a set of problem instances, show that the proposed optimization algorithm is able to provide better solutions than a basic column generation algorithm.

5.1.2 Gradual Network Design

In this section, based on [36], we study the problem of upgrading core networks to extend its capacity as the traffic to be transported increases; we call this as the GRAdual Network Design (GRANDE) problem. In contrast to consider a green-field scenario, in the GRANDE problem the already deployed equipment can be reused to reduce upgrading CAPEX cost.

In the following we first state the GRANDE problem and present its ILP formulation. As a result of its size, a method based on path generation is proposed. As an alternative, a BRKGA heuristic is also presented. Illustrative numerical results are shown for two real network examples.

GRANDE problem statement

The GRANDE problem can be formally stated as follows.



Given:

- a connected graph $G(N,E)$, where N represents the set of potential OXC locations, and E is the set of links connecting pairs of locations,
- subsets $N_{in} \subseteq N$ and $E_{in} \subseteq E$ containing already installed OXCs and links, respectively,
- the characteristics of the optical spectrum,
- a traffic matrix D ,
- the cost of installing a new OXC and a new fiber link.

Output:

- the network topology with the extra equipment needed to transport the set of demands D ,
- the route and spectrum allocation for each demand in D ,

Objective: Minimize the CAPEX cost resulting from upgrading the network to transport the given traffic matrix.

It is worth noting that to cope with new traffic requirements, manual operations causing service disruption are in general needed to implement the GRANDE solutions. In this regard, the GRANDE problem focuses on optimizing traffic routing and network dimensioning admitting service disruption.

Illustrative numerical results

For evaluation purposes we consider the 22-node 35-link British Telecom (BT) and the 30-node 56-link Telefonica (TEL) topologies depicted in Figure 54. The set of nodes is divided into two subsets, namely, the set of nodes that are source (or destination) of demands and the set of intermediate nodes (20+2 in the BT and 15+15 in the TEL topologies). Any topology solution of the GRANDE problem must contain all the source nodes, a subset of intermediate nodes, and the links needed to interconnect all the source nodes with the capacity sufficient to serve all demands. Let us assume the optical spectrum width of every link equal to 1 THz and the frequency slices of 12.5 GHz, which results in 80 slices per link. We also assume that every lightpath uses QPSK as modulation format.

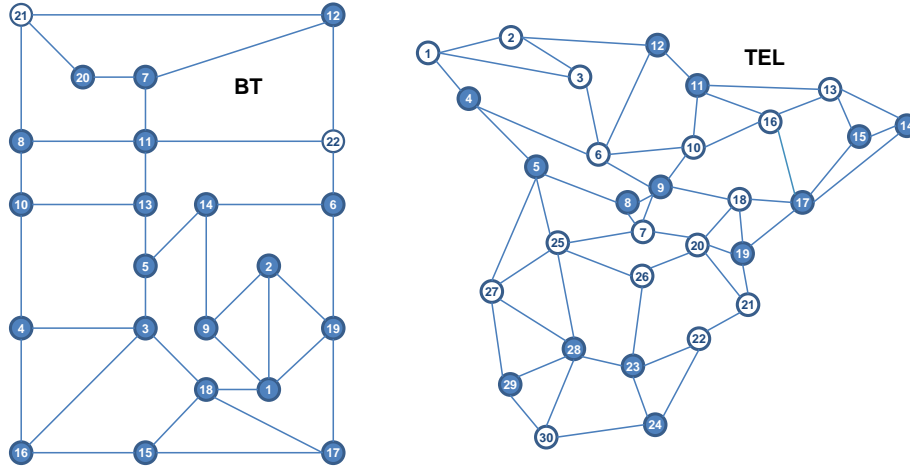


Figure 54: Network topologies under evaluation, where coloured circles represent source nodes.

A traffic matrix between source nodes with a total volume (number of demands) of 1.18 Tb/s (120) and 0.89 Tb/s (56) for BT and TEL, respectively for year 0. The 35% annual increment is assumed for the subsequent 7 years, leading to 8 traffic matrices per network where the traffic to be served in year i is obtained as $traffic_year_0 * (1.35)^i$.

Regarding the number of demands, the number for BT is almost two times larger than for TEL. This implies that, for the same number of pre-computed shortest paths or path generation iterations, the BT instances have two times more binary variables x_{dpc} than the TEL instances. We assume that the largest transponder is of 400 Gb/s, so the number of demands remains invariable annually provided that each demand remains lower than 400 Gb/s. If this threshold is exceeded, then the demand is split.

The instances of the GRANDE problem described above were solved using the solving methods described in section 8.7.1: *i*) ILP with variables generated by pre-computing k shortest paths (KSP); *ii*) ILP with variables generated by means of the proposed path generation algorithm (PG); and *iii*) the proposed BRKGA heuristic. For the sake of clarity, while using the path generation algorithm we stop when $maxIter$ iterations are completed, in order to be able to compare solutions with the same number of variables but obtained by different methods.

The ILP methods were implemented in Matlab and solved using CPLEX 12.4 [18], while the heuristic was implemented in Java. All experiments were run on a 2.4 GHz Quad-Core machine with 8 GB RAM memory under the Linux operating system. Finally, we limited the running time of each method to 10 hours.

Starting from a network topology obtained after solving the initial traffic matrix for year 0, all subsequent matrices are solved fixing as installed the nodes and links obtained in the solution for the previous year. Figure 55 plots the cumulative CAPEX of the network designed for each year using each solving method; the same cumulative CAPEX value for two consecutive years means no upgrading the network. The KSP method was solved with $k=5$ and $k=10$ shortest paths, whereas $maxIter$ was set to 5 in the case of PG.

In the light of the results in Figure 55, the PG method provides the lowest CAPEX with the smallest set of variables. To clearly validate the usefulness of the PG method, the results of KSP 5 and PG 5 can be directly compared since they were obtained using the same amount of variables. As observed, PG 5 highly outperforms KSP 5; PG finds good-quality solutions even for those loads where KSP 5 is not able to route all demands (year #7 for the BT network); to improve KSP 5 more pre-computed routes need to be added. Notwithstanding, using $k=10$ routes, KSP 10 is not able to reach the quality of PG 5 solutions. Regarding the BRKGA heuristic, one can observe that, in most of the cases (10 out of 16), it reaches the best solution while in the rest of its solutions is very close to the best (within 6% worse than the best).

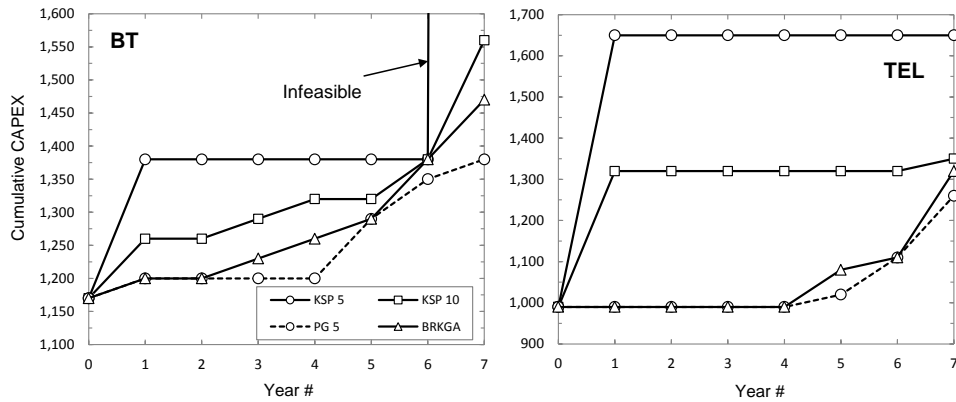


Figure 55: Cumulative CAPEX per year for BT (left) and TEL (right) networks.

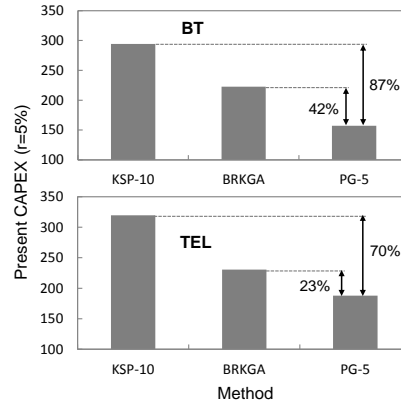


Figure 56: Present CAPEX.

Analysing Figure 55, we realize that the success of PG lies in the fact that new equipment is installed only when it is needed, not before that. Going deeper in that analysis, let us bring CAPEX cost to year 0 (present CAPEX). To this end, a discount rate r (we assume $r=5\%$) is applied to annual CAPEX costs. Figure 56 illustrates present CAPEX values for the considered topologies and solving methods. Similarly to the previous analysis, PG provides the cheapest network designs for both topologies. However, we can better appreciate the differences among methods. The difference between KSP 10 and PG 5 in terms of cumulative CAPEX leads to present CAPEX differences as high as 70% for the



TEL or even 87% for the BT network. Comparing PG 5 and BRKGA, cost savings are 23% for the TEL and more than 40% for the BT network.

It is clear in the light of these studies that the PG method provides really competitive solutions, since it installs new resources when they are needed, thus providing remarkable savings.

5.1.3 Performance evaluation of light-tree schemes

New applications require conveying huge bitrate from a source node to multiple destinations (e.g. uncompressed real time 8k transmission needs 72 Gb/s connections). As a result, point-to-multipoint (P2MP) connections need to be established on the optical layer. In this section, based on [52], the P2MP routing and spectrum allocation (RSA) problem is formally stated and modelled as an Integer Linear Program (ILP). Exhaustive numerical results show that creating several transparent sub-trees noticeably improves the performance obtained by serving the demand using one single tree.

To compare the performance of P2MP connections in flexgrid with WDM networks, the authors in [53] proposed two heuristic algorithms; they assumed broadcast-and-select based optical nodes. The authors in [54] present an ILP formulation to compute the RSA problem for a set of multicast demands. Their formulation is based on pre-computing a set of paths and ensuring the contiguity of the allocated frequency slices. As a result of its huge complexity, a heuristic is proposed.

Routing Schemes for Multicast Demands

Figure 57 illustrates the routing schemes considered in this paper to serve multicast demands in the optical layer; we assume that optical cross-connects (OXC) nodes are based on the flexgrid technology. Client routers/switches are connected to OXCs by means of SBVTs. Let us assume 400 Gb/s SBVTs that can be shared among 4 paths (SBVT flows). In the case of unidirectional paths, SBVTs can be shared by up to 4 incoming paths and up to 4 outgoing paths. In this scenario, the multicast demand $\langle A, \{B, C, D\}, 100 \text{ Gb/s} \rangle$ can be served by three lightpaths or, alternatively, by a single light-tree.

Figure 57a shows the *path scheme*, where three 100 Gb/s unidirectional lightpaths (also known as *sub-path*), are set-up. Each sub-path consumes one Tx SBVT flow from those available in the source router and one Rx SBVT flow in the destination router. Furthermore, one spectrum slot is allocated for each sub-path along the route. Although the path scheme requires many resources (SBVT flows and spectrum), every sub-path can be routed independently from the other ones serving the same demand, which provides high flexibility in the case where spectral resources are scarce or fragmented.

To reduce resource utilization, the *tree scheme* can be implemented (Figure 57b) to source a single 100 Gb/s signal, which remarkably reduces the required number of Tx SBVT flows to just one. Besides, spectral resources can be shared on some segments along the tree (e.g. from OXC-A to OXC-D), which reduces also spectral resource utilization. Notwithstanding, a single spectrum allocation is required for the whole tree, which might increase demands blocking ratio in case no continuous spectrum allocation is available along the tree, e.g. as a result of spectrum fragmentation.

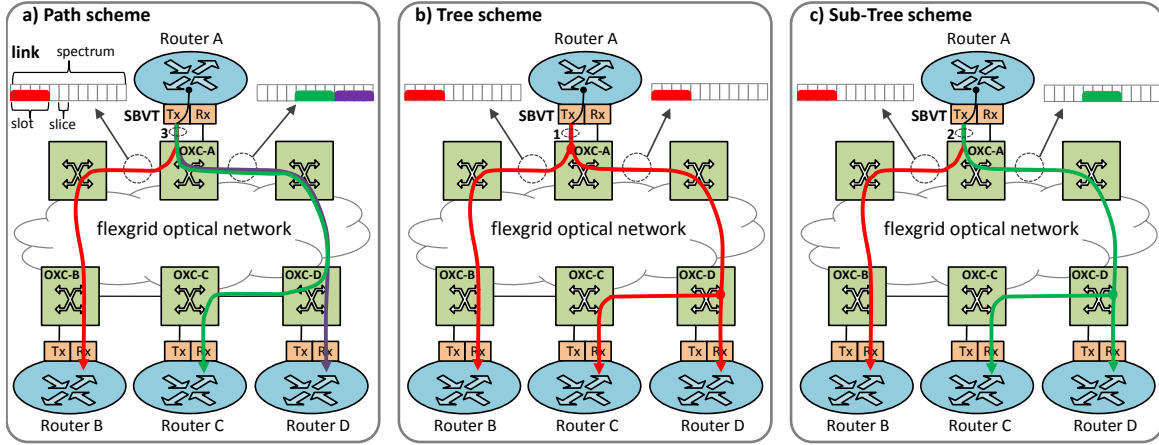


Figure 57: Routing schemes for multicast traffic.

In light of the above, we propose the *sub-tree scheme* to combine flexibility of sub-paths with resource savings of trees (Figure 57c). Under this scheme, a multicast demand can be served as a combination of light-trees, each enforcing the spectrum continuity constraint. In the example, two light-trees are created as a result of spectrum fragmentation. Thus, two Tx SBVT flows are needed at the source router, whereas spectral resources are shared in the segment OXC-A / OXC-D.

To evaluate the performance of the proposed schemes, numerical results obtained by solving the P2MP-RSA problem (detailed in section 8.7.2) are presented in the next section.

Illustrative Numerical Results

The BT, TEL and EON network topologies were considered for evaluation, where every link is configured with an optical spectrum of 2 THz divided into frequency slices of 6.25 GHz. The QPSK modulation format is considered for every sub-connection and thus the maximum reach is 2000 km.

We implemented the P2MP-RSA ILP in Matlab using CPLEX 12.5 as a solver engine. Instances consisting of a number of 100 Gb/s multicast demands with 5 destinations randomly chosen with a uniform probability were generated.

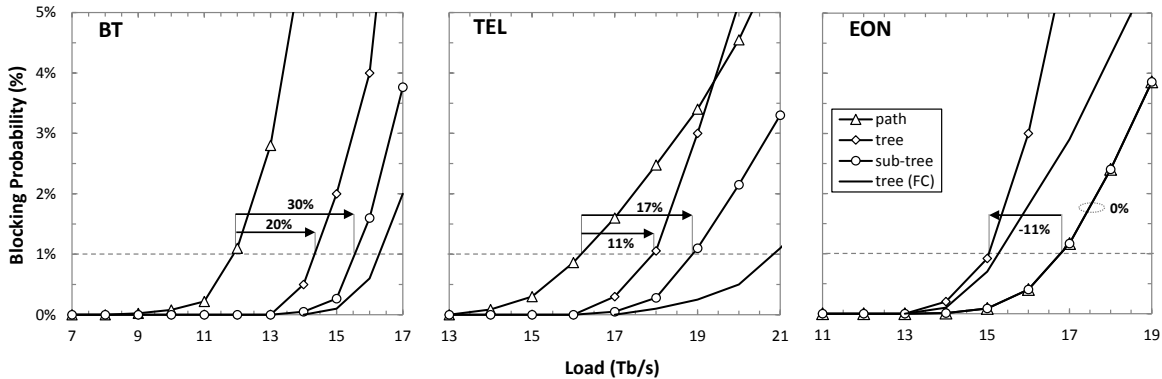


Figure 58: Blocking Probability vs Offered Load.

Figure 58 plots blocking probability (P_b) as a function of the offered load for each considered topology, where every value is the average of 10 instances; percentage of load increment at $P_b=1\%$ for tree and sub-tree schemes compared to the path scheme are also shown. For reference, the performance of the tree scheme when nodes include frequency conversion is also plotted. Remarkable differences in the performance of the tree scheme can be observed among the topologies, ranging from 20% for the BT topology to -11% for the EON topology, compared to the path scheme. This is as a result of the longer distances in the case of the EON topology, as the plot under full conversion clearly shows. In contrast, the sub-tree scheme performs the best for the three considered topologies, its load gain being as high as 30% for the BT topology, and 17% for the TEL topology. Notwithstanding the long distances of the EON topology, feasible route and spectrum allocations can be found when the constraint of a single tree is relaxed and several sub-connections are allowed, resulting in the same performance for the path and sub-tree schemes (lower bound).

Conclusions

Three different routing schemes for multicast traffic on flexgrid optical networks have been analysed: the path, the tree, and the sub-tree schemes. To compare their performance, the P2MP-RSA problem was stated and modelled using an ILP formulation.

Numerical results showed that the sub-tree scheme outperforms the path and tree schemes, increasing the capacity of national network in the range of 17 to 30%. The worst case of the sub-tree scheme was when its performance equals that of the path scheme, as a result of the long distance of the network.

5.1.4 Multi-layer planning of flexible optical networks

In this subsection we focus on the multi-layer planning problem of a flexible optical network. The examined multi-layer planning problem consists of three sub-problems at two layers: the Routing problem at the IP-layer (IP-R) and the Routing, Modulation Level (RML) and Spectrum Allocation (SA) problems at the optical layer. Given the IP end-to-end traffic matrix, the available IP/MPLS route modules, and the feasible configurations of the tuneable optical transponders we serve demands for their requested rates by selecting the routers' modules, the routes in the IP (virtual) topology and the corresponding paths and spectrum slots in the underlying optical network, together with the optical transponders transmission configuration.

As opposed to most previous works, we consider in more details the IP layer, by using a detailed modular model for the IP/MPLS routers that are deployed at the edges of the optical network, taken by IDEALIST cost model presented in Deliverable D1.1. We also take into account distance adaptive optical transponders (the RMLSA problem). Distance adaptability creates interdependencies between the routing at the optical and the IP layers, making it hard (and also inefficient) to de-couple these two problems. To the best of our knowledge this is the first time that the multi-layer planning of an IP over flexible optical network includes distance adaptability/modulation level decisions.

We used the proposed multi-layer planning algorithm to plan both flexible and fixed-grid MLR optical networks and perform a techno-economic comparison. We also distinguished



two cases: the Joint Multi-Layer Network Planning (*JML-NP*) case where IP-R and RML sub-problems are jointly solved, followed by the SA sub-problem and the Sequential Multi-Layer Network Planning (*SML-NP*) case where the IP-R, RML and SA sub-problems are sequential solved.

Using realistic cost, network, and traffic models, taken by Deliverable D1.1 of IDEALIST project [37], we found that significant savings in cost and spectrum can be obtained through joint multi-layer optimization over the IP and the optical layer (*JML-NP*), as opposed to planning the two layers sequentially (*SML-NP*). We also found that when planning the network in a jointly multilayer manner (*JML-NP* case), the flexible network outperforms the MLR network deploying fixed-grid or flex-grid optical switches and fixed optical transponders, in terms of cost for medium and high loads and in terms of maximum spectrum used in all cases. More, in the *SML-NP* case the flexible networks outperform the MLR network in terms of network cost and spectrum in all examined cases.

Network architecture and Problem Description

We are given an optical domain that consists of optical switches and fiber links. The optical switches function as ROADMS employing the flex-grid technology, and support optical connections (lightpaths) of one or a contiguous number of 12.5 GHz spectrum slots. At each optical switch, one or more IP/MPLS routers are connected, by using tunable (or non-tunable) flexible colored transceivers that are plugged to IP/MPLS routers ports. Alternatively, short reach transceivers could be plugged to the IP/MPLS routers to lead to tunable flexible transponders at the optical nodes. We will not distinguish between these two architectures, since their function and cost are almost identical. These IP/MPLS routers comprise the edges of the optical domain. We assume that the tunable transponders can control a number of transmission parameters that affect the optical reach at which they can transmit. The transponders are used to transform the electrical packets at the IP source router to the optical domain, so that the traffic which enters the optical switch, crosses over the optical network in lightpaths (all-optical connections) and reaches an optical switch. These packets are converted back to electrical signal at the optical receiver, and handled by the corresponding IP/MPLS router. This IP/MPLS router can be either (i) the final destination of the packets in the optical domain, in which case the packets will be forwarded further towards their final destination through another domain attached to the router, or (ii) an intermediate hop, in which case the packets will re-enter the optical domain to be eventually forwarded to their domain destination. Note that packet processing is performed at the electrical domain and in particular at the IP/MPLS routers, while optical switches function as transparent pipes between IP/MPLS router end-points.

We are also given the traffic matrix Λ that corresponds to the IP traffic from the domains adjacent to the IP/MPLS routers that is forwarded over the optical domain under study. Our goal is to establish lightpaths (optical connections), and route the traffic over these lightpaths and through possibly intermediate IP/MPLS routers to the end IP/MPLS router destination. For this reason the network is also referred to as IP over flexgrid network.

We assume that the IP/MPLS routers are modular and we also assume that at the optical layer the optical switches and fiber links are deployed, but not the transponders. So the goal is to decide on the modules to install at the IP/MPLS routers, and the number and placement of optical transponders, to establish the optical connections and route the IP traffic over them, so as to minimize the total cost.



As discussed above, the planning of an IP over flexgrid network consist of three subproblems which are inter-related: (i) the IP routing (IP-R), the Routing and Modulation Level (RML), and (iii) the Spectrum Allocation (SA). In the IP-R problem we decide on the modules to install at the IP/MPLS routers, how to map traffic onto the lightpaths (optical connections), and what intermediate IP/MPLS routers will be used to reach the domain destination. In the RML we decide how to route the lightpaths and also the transmission configurations of the tunable transponders. Note that optical lightpaths are bidirectional, while IP routing is not constrained to be symmetric in both directions. In the SA we allocate spectrum slots to the lightpaths, avoiding slot overlapping (assigning the same slot to more than one lightpaths). The use of tunable optical transponders, where the rate, the reach, and spectrum are not given, is the factor that makes the RML decisions to affect the two other subproblems, and significantly complicates the problem.

The optical network topology and the IP/MPLS router edges are represented by a directed graph $G=(N, L)$, where N is the set of nodes and L is the set of links. The graph consists of two types of nodes, IP (or virtual) nodes and optical (physical) nodes and two layers, the IP (or virtual) layer, where all the IP nodes are located, and the optical (physical) layer, where all the optical nodes are located. A virtual node represents an IP/MPLS router, while a physical node represents a flex-grid optical switch.

Following the IDEALIST cost model described in D1.1 [37], we assume that an IP/MPLS router is modular and is built out of (single or multi) chassis. A chassis provides a specified number of bi-directional slots with a nominal (maximum) transmission speed. Into each router slot, a line-card of the corresponding speed can be installed. Each line-card provides a specified number of ports at a specified speed and occupies one slot of the IP/MPLS router. In our performance results we consider a scalable multi-chassis router for core nodes, with up to 72 chassis, with 16 router slot capability per each chassis, but the problem definition and the algorithm is generic and can work with other models as well.

Regarding the optical network we assume the use of tunable transponders that can control a number of parameters, such as the modulation format, the utilized spectrum, the baud-rate and the rate that they transmit. The configurations of a transponder of cost c_t are indicated by transmission tuples $(d_t, r_t, b_t, g_t, c_t)$, where d_t is the reach for which a transmission of rate r_t (gpbs) using b_t spectrum slots and g_t guard-band slots of is feasible, that is, it has acceptable QoT. Note that defining a specific rate and spectrum this decision incorporates the choice of the modulation format of the transmission. Different types of transponders, with different costs have different transmission tuples, and this definition is generic so as to be able to formulate such option. A non-tunable transponder can be also expressed by a single tuple in the above form. The transponders are driven by equal rate line-cards, with each line-card supporting one or more transponders of the same type.

In the network graph we define three types of links, inter-layer (l_{ov} or l_{vo}), optical (l_o) and virtual (l_v) links. An inter-layer link connects a virtual (IP) node with an optical node and represents the use of a (tunable or non-tunable) transponder. Note that we distinguish between the two directions of the inter-layer links, with l_{ov} denoting an optical-to-virtual and l_{vo} a virtual-to-optical interlayer link. For each type (one or more) of transponders available, we create inter-layer links with different cost vectors (we will come back to this in the next section). An optical link corresponds to a fiber and connects two optical switches, while a virtual link corresponds to a lightpath that connects two IP/MPLS routers. Thus, a virtual link between two IP routes is created, when a lightpath between the two optical switches

connected to the IP/MPLS routers is established, while this lightpath can traverse one or more optical links and have zero or a number of intermediate optical nodes through which it passes transparently.

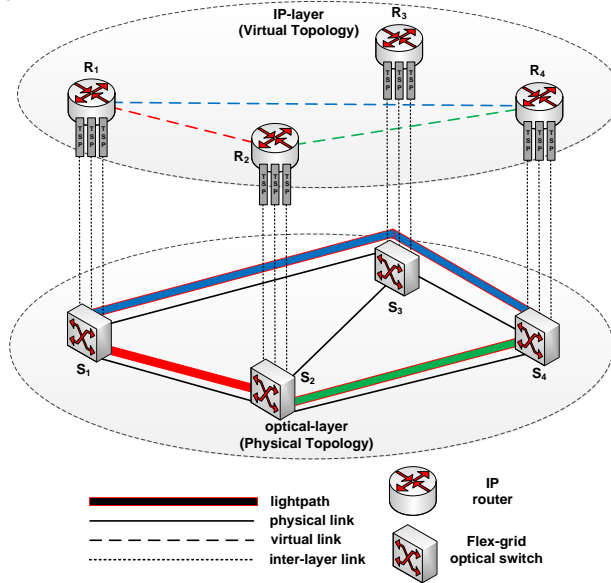


Figure 59: Auxiliary graph of IP over flex-grid optical network.

Figure 59 presents an illustrative example of the examined IP over flex-grid problem, where four IP/MPLS routers are located at the IP layer, four flex-grid optical switches are located at the optical layer and a single type of transponder is used. Also three lightpaths ($S_1 \leftrightarrow S_4$, $S_1 \leftrightarrow S_2$, and $S_2 \leftrightarrow S_4$) have been established at the optical layer, and three virtual links (one from node R_1 to node R_4 , one from node R_1 to node R_2 and one from node R_2 to node R_4) have been established at the IP layer. In the case where a new demand arrives from node R_2 to node R_4 , then this demand can be served by lightpaths already established from node S_2 to node S_4 if its remaining capacity is enough or new optical connections can be established.

The proposed algorithm is detailed in section 8.7.3.

Performance Results: Techno-economic comparison

To evaluate the performance of the proposed algorithm (see section 8.7.3) and also quantify the benefits that can be obtained by a flexible solution as opposed to a fixed-grid network, we conducted a number of simulations using Matlab.

In the experiments presented below we define the following cases of networks: (a) MLR optical network employing fixed-grid 50 GHz optical switches and fixed 40 Gbps and 100 Gbps (*fixed-grid/fixed-TSP case*), (b) MLR optical network employing 12.5 GHz flex-grid optical switches and fixed 40 Gbps, 100 Gbps and 400 Gbps transponders (*flex-grid/fixed-TSP case*) and (c) flexible (or elastic) optical network employing 12.5 GHz flex-grid optical switches and flexible transponders, also referred to as bandwidth variable transponders (*flex-grid/flex-TSP case*). The reason that in the first case we are not assuming the use of 400 Gbps transponders is that such devices are expected to require 75 GHz spectrum,

which does not fit in traditional 50 GHz fixed grid WDM systems. We also distinguished the case where a joint or sequential multi-layer network planning (*JML-NP* or *SML-NP*, respectively) is applied to each of the above networks.

The reference network used in our simulations was the DTnet for which we created traffic matrices for the period from 2014 up to 2024 assuming 35% increase per year for all commodities. The performance metrics we used for the comparisons are the maximum spectrum used (measured in GHz), the cost of transponders, the cost of routers (measured in IDEALIST Cost Units – ICU) and the total cost of network computed as the sum of transponders and routers cost (in ICU).

For the case of the flexible network, we assumed that each link has available 320 spectrum slots with 12,5 GHz width each one. Also, we assumed the use of a single type of tunable transponder that supports transmissions up to 50 GHz and modulates up to 64 QAM, so as to transmit up to 400 Gb/s. The transmission tuples (reach, rate, spectrum, guard-band, cost) of the used tunable transponders is shown in Table 18.

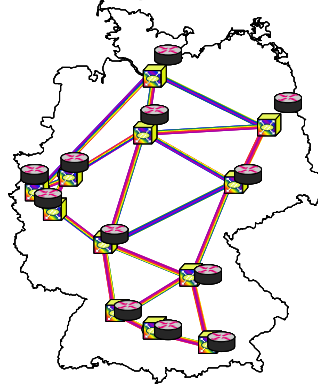


Figure 60: Generic DTnet topology. 12 nodes, 40 directed links.

Table 18: Transmission tuples of tunable transponders.

Capacity (Gb/s)	Reach (Km)	Data and guardband slots (in 12,5 GHz slots)	Cost (ICU)	Capacity (Gb/s)	Reach (Km)	Data and guardband slots (in 12,5 GHz slots)	Cost (ICU)
40	4000	5	1.76	200	2500	7	1.76
40	3000	4	1.76	200	2200	6	1.76
40	2500	3	1.76	200	1900	5	1.76
40	1900	2	1.76	200	750	4	1.76
40	600	2	1.76	200	600	3	1.76
				200	500	2	1.76
100	3500	5	1.76	400	2500	15	1.76
100	3000	4	1.76	400	2200	13	1.76
100	2500	3	1.76	400	1900	11	1.76
100	1900	2	1.76	400	750	9	1.76



100	600	2	1.76	400	600	7	1.76
				400	500	5	1.76

Table 19: Transmission tuples of fixed-grid transponders.

Capacity (Gb/s)	Reach (Km)	Required Spectrum (in GHz)	Cost (ICU)
40	2500	50	0.48
100	2000	50	1
400	500	75	1.36

For the case of the fixed-grid MLR network, we assumed that each link has available 80 wavelengths of 50 GHz width each one. Also, we assumed the use of three types of fixed-grid transponders that support 40 Gb/s, 100 Gb/s and 400 Gb/s transmissions respectively, described by the transmission tuples shown in Table 19.

The cost of linecards and core routers with 1 chassis of 16 slots and 400 Gb/s capacity per slot is shown in Table 20. Note that we assumed that flex-grid transponders are always interfaces with 400 Gb/s linecards, although they could be configured to transmit at lower loads. The cost of a multi-chassis core router is computed according to Eq. (8).

$$P = 6,02 \cdot n_{ch} + 1,76 \cdot \left\lceil \frac{n_{ch}}{9} \right\rceil + 9,11 \cdot \left\lceil \frac{n_{ch}}{3} \right\rceil, n_{ch} = \left\lceil \frac{C}{K} \right\rceil, 2 \leq n_{ch} \leq 72 \quad (8)$$

where n_{ch} is the number of chassis installed, C is the total switching capacity in Tb/s, and K is the capacity of a fully equipped shelf. The costs of transponders, linecards and routers used in our simulations are derived from the CAPEX model defined in the context of the IDEALIST.

Table 20: Cost of line cards and single chassis IP/MPLS routers

Type	Cost (ICU)
Core router with 1 chassis of 16 slots and 400 G capacity per slot	4.30
10x40G linecard for core router	2.56
4x100G linecard for core router	2.88
1x400G linecard for core router	2.74

Table 21 shows the maximum spectrum (in GHz) used for each network case examined with *JML-NP* or *SML-NP* and for years 2014 to 2024. Concerning the *JML-NP* case, we observe that the *flex-grid/flex-TSP* network outperforms both *fixed-grid/fixed-TSP* and *flex-grid/fixed-TSP* optical networks, while the *fixed-grid/fixed-TSP* network presents the lowest performance. This was expected since in the *flex-grid/flex-TSP* network the transponders used utilize exactly the amount of spectrum they require, while in the *fixed-grid/fixed-TSP* and *flex-grid/fixed-TSP* networks the transponders are fixed and utilize in the first case 50 GHz per wavelength, while in the second case they utilize either 50 or 75 GHz. Also we observe that after 2018 the connections of the *fixed-grid/fixed-TSP* network are blocked due to insufficient number of spectrum slots, since traffic increases and in this case we assumed that we cannot use 400 Gbps transponders. This shows the inefficient spectrum usage of the fixed-grid and fixed transponders compared to that of flexible optical



networks, especially at high loads. Concerning the *SML-NP* case, we observe the same findings as the *JML-NP* case in terms of spectrum. Comparing the joint network planning (*JML-NP*) case to the sequential (*SML-NP*) case, we observe that at both *fixed-grid/fixed-TSP* and *flex-grid/fixed-TSP* cases, the spectrum used is smaller during the whole period of reference (except 2014 for the *flex-grid/fixed-TSP* case), when the *SML-NP* is applied as opposed to the *JML-NP*. The *JML-NP* keeps the network cost advantage which is the optimized metric, as will be discussed at the following paragraphs. Concerning the *flex-grid/flex-TSP* network, in the case that a *JML-NP* algorithm is applied as opposed to a *SML-NP* algorithm, the spectrum used is the same at low loads (year 2014), while as load increases (after year 2014), the *JML-NP* algorithm outperforms the *SML-NP* algorithm. This shows how well the elasticity provided by flexible networks fits with the joint multi-layer network planning.

Table 21: Spectrum (in GHz) used for each case of network examined when jointly and sequentially planned and reference years from 2014 to 2020

Year	Joint multi-layer network planning			Sequential multi-layer network planning		
	fixed-grid/fixed-TSP	flex-grid/fixed-TSP	flex-grid/flex-TSP	fixed-grid/fixed-TSP	flex-grid/fixed-TSP	flex-grid/flex-TSP
2014	650	550	387,5	550	575	387,5
2016	1450	950	400	750	900	412,5
2018	2400	1125	612,5	1350	1000	650
2020	-	2075	812,5	-	1125	900
2022	-	2800	1487,5	-	1800	1537,5
2024	-	3950	2725	-	3200	3075

The transponder, router and total network cost for each network case when planned jointly (*JML-NP*) and reference years from 2014 to 2024 is shown in Table 22. The *fixed-grid/fixed-TSP* network has quite low cost for year 2014 but then it presents the highest cost as it requires a high number of low-rate fixed 40 Gbps and 100 Gbps transponders to serve the increased traffic at later years. Among the three network cases, the *flex-grid/flex-TSP* network has the smallest network cost for all the examined period, except for year 2014, where the *flex-grid/fixed-TSP* network has a slightly smaller cost. The cost difference of the *flex-grid/flex-TSP* and the *flex-grid/fixed-TSP* cases increases as the years progress. This is because at light loads, lower cost/low-rate fixed transponders are sufficient to serve the traffic, especially in the case of the *flex-grid/fixed-TSP* network, while flexible transponders used in the *flex-grid/flex-TSP* are not fully utilized, resulting in some waste of cost. Although routing at the IP layer decreases this problem, through appropriate traffic grooming, still at low load (year 2014) the *flex-grid/fixed-TSP* network turned out to be slightly better than *flex-grid/flex-TSP*. As traffic increases, the utilization and the efficiency of flexible transponders increases. Combining this with the additional flexibility of more transmission options gives the advantage to the *flex-grid/flex-TSP* which outperforms both *fixed-grid/fixed-TSP* and *flex-grid/fixed-TSP* networks, regarding the total network cost, as the load increases. Note that in the above calculations the cost of flexible transponders was 30% higher than the related cost of the equivalent rate fixed transponder. Higher savings could be obtained if this was lower.



Table 22: Transponder, router and cost (in ICU) for each case of network examined when jointly planned (JML-NP) and reference years from 2014 to 2020

Year	fixed-grid/fixed-TSP			flex-grid/fixed-TSP			flex-grid/flex-TSP		
	Transponder cost	Router cost	Network cost	Transponder cost	Router cost	Network cost	Transponder cost	Router cost	Network cost
2014	77,44	169,58	247,02	73,12	166,74	239,86	56,32	191,34	247,66
2016	129,12	248,59	377,71	106,48	247,79	354,27	77,44	275,71	353,15
2018	225,60	368,33	593,93	149,92	370,45	520,37	112,64	396,81	509,45
2020	-	-	-	226,88	654,96	881,84	183,00	649,00	832,00
2022	-	-	-	401,36	1194,45	1595,81	302,72	1191,33	1494,05
2024	-	-	-	656,70	2119,17	2775,87	545,00	2073,00	2618,00

Table 23: Transponder, router and cost (in ICU) for each case of network examined when sequentially planned (SML-NP) and reference years from 2014 to 2020

Year	fixed-grid/fixed-TSP			flex-grid/fixed-TSP			flex-grid/flex-TSP		
	Transponder cost	Router cost	Network cost	Transponder cost	Router cost	Network cost	Transponder cost	Router cost	Network cost
2014	80,2	183,6	263,8	60,19	221,38	281,57	56,32	191,34	247,66
2016	139,392	275,23	414,622	85,06	316,71	401,77	77,44	275,71	353,15
2018	227,664	392,87	620,534	128,03	468,33	596,36	116,16	402,29	518,45
2020	-	-	-	187,18	804,95	992,13	196,06	690,28	886,34
2022	-	-	-	315,52	1413,38	1728,90	332,29	1240,65	1572,94
2024	-	-	-	562,22	2451,71	3013,93	572,88	2129,73	2702,61

Table 23 shows the transponder, router and network cost for each case of network with *SML-NP* and reference years from 2014 to 2024. Again the findings are the same as those obtained when the joint planning (*JML-NP*) is applied to each case of network.

Concerning the joint (*JML-NP*) as opposed to the sequential (*SML-NP*) planning, we observe that in case of *fixed-grid/fixed-TSP* and *flex-grid/fixed-TSP* networks, the *JML-NP* algorithm outperforms the *SML-NP* algorithm for the whole examined period in terms of network cost. This is explained as follows: in the case of *SML-NP* algorithm, the connections are groomed without taking into account the reach constraints at the optical layer, and therefore the RML decisions. This results in utilizing a higher number of 400 Gbps expensive transponders and line cards (in the case of *flex-grid/fixed-TSP* network) compared to the *JML-NP* case. Concerning the *flex-grid/flex-TSP* network, we observe that at low loads (years 2014 and 2016) the *JML-NP* and *SML-NP* cases have the same network cost, while at higher loads the *JML-NP* outperforms the *SML-NP* case. Also we observe that as the load increases the difference between the network costs of *JML-NP* algorithm as opposes to *SML-NP* algorithm increases. This was expected, as in the *SML-NP* case the connections are groomed without taking into account the reach constraints at the optical layer, and therefore the RML decisions.

Conclusions



We presented a multi-layer network planning algorithm that can be applied to both flex-grid and fixed-grid optical networks. The algorithm takes as input the feasible transmission configurations of flexible transponders defined to account for physical layer limitations, a model for the modular IP/MPLS routers and the traffic matrix. It serves demands for their requested rates by jointly selecting the routes in the IP (virtual) topology and the corresponding paths, and spectrum slots in the underlying optical (physical) topology, together with the transponders transmission configuration.

Using realistic transmission specifications for flexible networks, our results show that significant savings can be obtained through joint multi-layer optimization over the IP and the optical layer, as opposed to planning the two layers sequentially. We also verified that flex-grid optical networks outperform fixed-grid in terms of maximum spectrum used, when a joint multi-layer network planning or a sequential multi-layer network planning approach is applied. In case of joint multi-layer network planning, flexible network is slightly more expensive at light load compared to fixed-MLR networks, but as load increases it becomes more efficient and cheaper. More, in the case of sequential multi-layer network planning, the flexible networks are less expensive as opposed to fixed-grid networks both at low and high loads. The above remarks magnify the need for planning the network with a joint multi-layer approach, especially in the case of elastic optical networks.

5.2 On-Line Algorithms for Real Time Operation

5.2.1 Spectrum defragmentation

Dynamic operation of flexgrid networks might cause optical spectrum to be divided into fragments, which makes it difficult finding contiguous spectrum of the required width for incoming connection requests, leading thus to an increased blocking probability. To alleviate to some extent spectrum fragmentation, in this section we propose to shift the central frequency of already established connections to create wider contiguous spectrum fragments to be allocated to incoming connections; this procedure is in general called as spectrum defragmentation.

Introduction

In dynamic scenarios, spectrum fragmentation appears in flexgrid networks subject to both the continuity and the contiguity constraints. As a consequence of the high cost of spectrum converters, they are rarely used in optical networks and, therefore spectrum fragmentation appears increasing the blocking probability of connection requests, and degrading the network Grade of Service (GoS). To improve network efficiency, defragmentation can be applied by re-configuring selected connections, thus compacting the utilized resources and facilitating that incoming connection requests can be served.

Several re-optimization strategies (including rerouting and spectrum reallocation) have been proposed so far for optical networks and can be applied to the specific case of flexgrid networks [41]-[44]; those strategies can be divided into *proactive* and *reactive* strategies considering the way they are triggered [41]. The proactive strategy focuses on minimizing fragmentation itself at a given period of time, whereas *reactive*, also known as *provisioning-triggered*, focuses on making enough room for a given connection request if it cannot be established with current resources allocation. Periodic defragmentation,

requiring long computation times as a result of the amount of data to be processed, is essentially performed during low activity periods, e.g. during nights [42]. Conversely, path-triggered defragmentation, involving only a limited set of already established connections, might provide solutions in shorter times and can be run in real time [43], [44].

To minimize traffic disruption, the standardized make-before-break rerouting technique, included in the standardized resource reservation protocol traffic engineering (RSVP-TE) [45], can be used for signalling the rerouting connection while the traffic is being transferred using the original one. After the rerouting connection is operational, traffic is switched and the original connection is finally torn down. In optical networks, make-before-break requires additional resources to support two parallel connections such as spare transponders, which entail additional costs.

Alternatively, the recently introduced *push-pull* technique [46] can be used to shift the central frequency of established optical connections without traffic disruption. Push-pull consists in re-tuning the transmitter laser from the original to the target nominal central frequency, while the receiver is automatically pulled to track the signal shifting. The limitation of the central frequency shifting performed using the push-pull technique resides on the fact that connections can be allocated to new slots as long as no other connections are established between the current and the new slot, reducing thus the set of reallocations available for each connection.

Aiming at illustrating spectrum fragmentation, Figure 61 shows an example on the small network topology depicted in Figure 61a, where each node and link is labelled. The entire spectrum width consists of 16 slices. Figure 61b represents the utilization of each frequency slice in the network, where a number of optical connections are already established. In this scenario, the connection request between nodes 4 and 7 requesting 4 slices cannot be served. Notwithstanding, each link in the shortest route of the new optical connection *newP* (through links 4-5-6) has at least 4 free slices and then, the request could be established shifting some of the established connections. In the example, connections *p1*, *p3*, *p4*, *p5*, and *p6* are using one or more of the links in the computed shortest route, and thus can be considered as candidates to be part of the defragmentation process. Finally in Figure 61c, connections *p4* and *p5* have been shifted making enough room for *newP*.

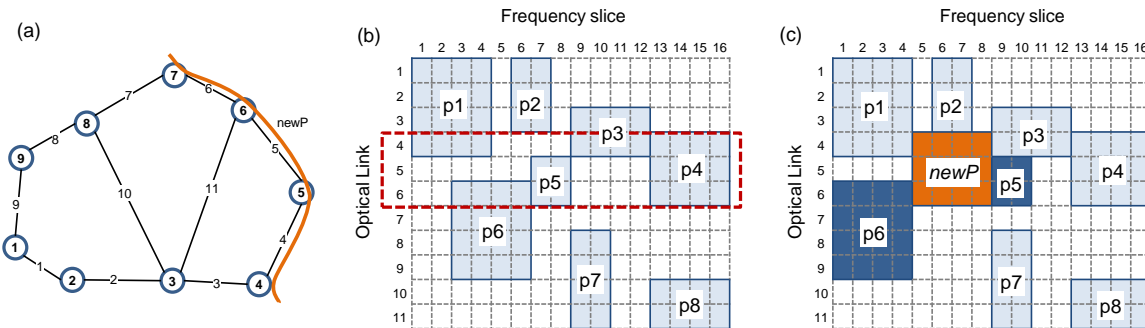


Figure 61: Example of spectrum reallocation by shifting the central frequency of some already established optical connections.



The Spectrum Shifting (SPRING) problem

As discussed above, to avoid traffic disruption we propose using spectrum shifting for defragmentation. To that end in this section we firstly formally state the problem and then present an ILP model to solve the defragmentation problem. The model is based on the work in [43], but adapted with specific constraints to ensure that connections are shifted and not reallocated, i.e. no other connections are established using slices between current and new slots. Finally, a heuristic algorithm providing much better trade-off between optimality and complexity is proposed.

The SPRING problem can be formally stated as follows:

Given:

- an optical network, represented by a graph $G(N, E)$, being N the set of nodes and E the set of optical links connecting two nodes,
- a set S of frequency slices available in each link $e \in E$,
- a set P of already established LSPs,
- a new request ($newP$) to be established in the network. A route for the LSP has been already selected but there is no feasible spectrum allocation,

Output:

- for each LSP to be shifted, its new spectrum allocation,
- the spectrum allocation for $newP$.

Objective: Minimize the amount of LSPs to be shifted to fit $newP$ in.

An ILP model for the SPRING problem is presented in section 0.

Numerical results

In this section, the performance of SPRING is compared to that of other reallocation-based defragmentation algorithms, such the SPRESSO algorithm proposed in [43].

To evaluate the performance, we consider the TEL network topology where the fiber links have the spectrum width equal to 4 THz, divided into frequency slices of 6.25 GHz. Simulations are run on an ad-hoc event-driven simulator implemented in OMNET++ [47] running on a 2.4 GHz Quad-core machine with 16 GB RAM.

Let us first analyse the results from the SPRESSO algorithm. As previously introduced, the SPRESSO algorithm allows already established optical connections to be reallocated in the spectrum, which causes traffic disruption. In some cases, however, shifting is found to be the best solution. Therefore, SPRESSO could be modified to return a solution only when hitless defragmentation could be applied and therefore, so no further algorithm (such the proposed SPRING) would be needed.

Figure 62 shows the percentage of defragmentation occurrences where reallocation was performed as a function of the network load; the rest of defragmentation were performed using shifting. As shown, a very high percentage of defragmentation was done by using reallocation. This fact, indicates that a very poor performance could be obtained in case of using SPRESSO for spectrum shifting.

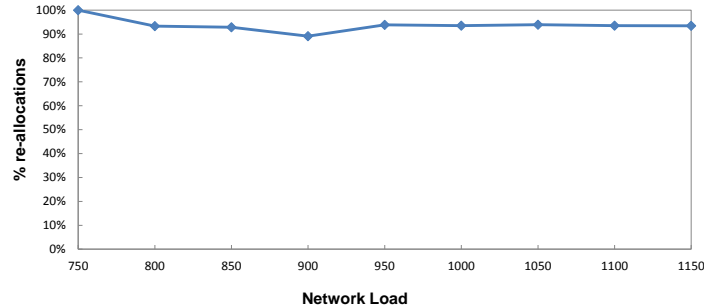


Figure 62: Percentage of reallocations using the SPRESSO algorithm.

Figure 63a shows the obtained blocking probability against the network load for the considered defragmentation algorithms. The case where no defragmentation is applied is also shown for the sake of comparison. It is clear, in light of the results, that a very nice performance can be obtained by applying SPRESSO; the gain is in the order of 15% at the load unleashing 1% blocking probability. Nonetheless, this gain is obtained at the expense of an appreciable traffic disruption. If we use the SPRESSO algorithm using only shifting, the gain drops noticeably to only 3%. The SPRING algorithm, applying spectrum shifting provides double of gain, just above 6%.

Computation time is an additional requirement for on-line algorithms where short computation times are desired. For those algorithms related to provisioning, high computation time might impact negatively on the set-up delay for those requests needing current optical connections to be shifted prior them to be served. Figure 63b compares the obtained computation times for the SPRESSO and SPRING algorithms. As shown, relaxing the capabilities of SPRESSO results in an even increased computation times with respect to the already high computation times obtained with SPRESSO. Our proposed SPRING algorithm, in contrast, provides more computation times in the order of 250ms, much shorter that those provided by SPRESSO that are above 1s.

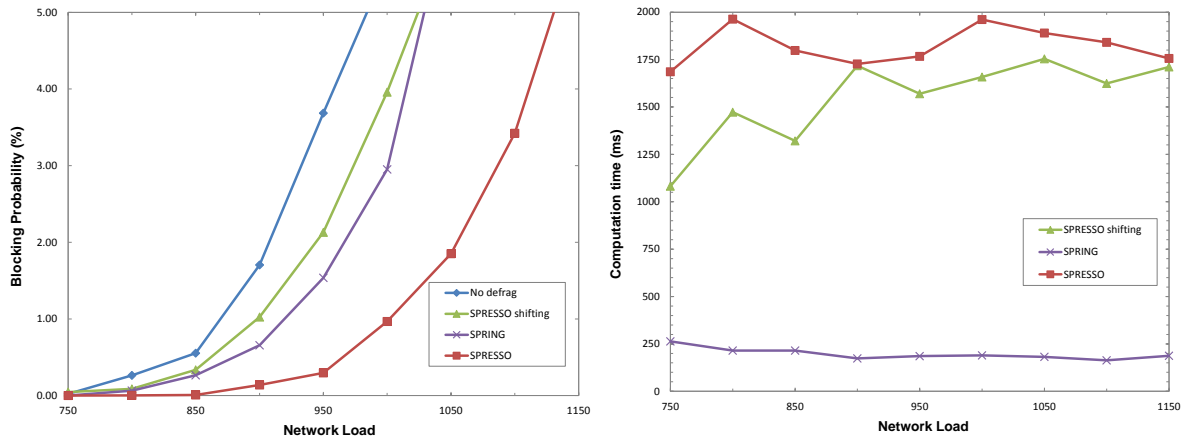


Figure 63: Performance evaluation of spectrum defragmentation algorithms against traffic load. a) Blocking probability and b) computation times.



Conclusions

In this section, the provisioning-triggered spectrum defragmentation use case was investigated. The use case, an example of in-operation planning, starts when a connection request cannot be served as a consequence of spectrum fragmentation in the links of the network. A subset of already established connections is candidate to be part of the defragmentation problem, where the central frequency of some of them will be shifted to make enough room to serve the new connection requested.

The spectrum shifting (SPRING) problem was formally stated and modelled using an ILP formulation that considered specific constraints to limit connection reallocations to just central frequency shifting that can be done in a hitless manner. An algorithm was also proposed.

The performance of the SPRING algorithm was compared to that of already existing reallocation-based defragmentation algorithms, such as SPRESSO. The results showed a good performance taking into account that only spectrum shifting was allowed. Regarding computation times, the SPRING algorithm provided much shorter times than those of the SPRESSO. This tests validate the SPRING algorithm.

5.2.2 Reoptimization of Dynamic Flexgrid Networks After Link Failure Repairs

In dynamic flexgrid optical networks, the usage of capacity may not be optimal due to the permanent process of setting up and tearing down connections, which, if not controlled, leads to spectrum fragmentation and, as a result, to increase of connection blocking. On top of this, a restoration mechanism that is launched in reaction to a link failure (cable cut) restores the affected lightpaths. Eventually, when the cable is repaired and its capacity becomes available for new connections, the unbalance between lightly and heavily loaded links increases, thus further decreasing the probability of finding optical paths with continuous and contiguous spectrum for future connection requests. In this section we study the effects of re-optimizing the lightpath connections after a link failure has been repaired (namely, the AFRO problem) as an effective way for both reducing and balancing capacity usage and, by these means, for improving network performance. To solve AFRO a column generation decomposition method, developed in [51], is applied. Illustrative numerical results show that AFRO allows to significantly decrease the request blocking probability in realistic dynamic network scenarios. Moreover, the column generation algorithm delivers quasi-optimal solutions in reasonable times. Besides, traffic disruptions resulting from lightpath rerouting are practically negligible. Finally, we show that it is sufficient to apply AFRO only for a selected set of link failures in order to achieve high network performance.

This section is based on our work in [62], [63].

AFRO problem

In order to describe AFRO, in Figure 64 we show a process a dynamic network undergoes after being subject to a certain link failure. Moreover, in Figure 65, which complements Figure 64, we show an illustrative example of the network states at different time instants. Here, we consider a dynamic network that serves a set of lightpaths under normal

operation (denoted as state t_0). After an event of a link failure, the restoration mechanism finds an alternative route for each of the lightpaths affected by the fiber cut (state t_1). Specifically, Figure 65a) shows a network with three lightpaths, namely p_1 , p_2 , and p_3 , just before the event of a fiber cut in link 2 – 7. After that, lightpaths p_1 and p_2 are rerouted by the restoration algorithm (Figure 65b)).

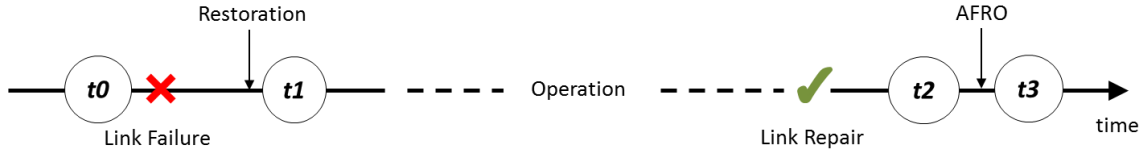


Figure 64: Dynamic evolution of network states (t) with link failure and link repair events.

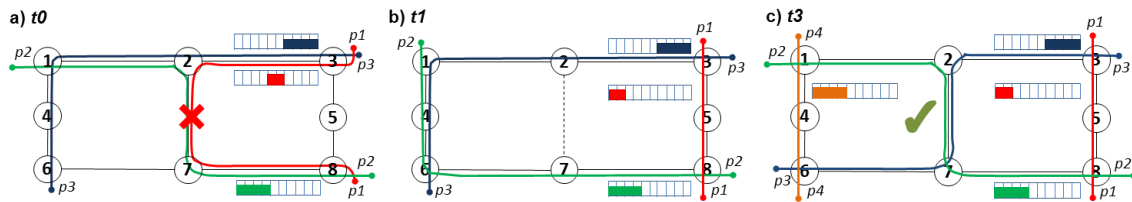


Figure 65: An example of link occupancy in different network states.

Once the solution of the restoration algorithm is implemented, the network operates without the failed link until it is restored (6 – 12 hours are usually needed to repair a fiber link), which is denoted as state t_2 . For the ease of presentation, in Figure 65 we assume that $t_1 = t_2$, i.e., neither setups nor tear-downs of lightpaths occur during the period of link restoration. At this point, we can see that after the link is repaired, the network capacity increases, however, its performance is not necessarily improved. Namely, in Figure 65b) a new connection requesting 4 slices between nodes 1 and 6 can be established neither when the link is broken nor when the capacity of the repaired link is available again.

With the aim to improve network performance and make use of restored link resources we consider that lightpath connections are reoptimized after the link is repaired (i.e., in state t_2). To this end, we solve an after-failure-repair optimization (AFRO) problem, which at the input gets the network state information, in particular, concerning routing and spectrum allocation of lightpaths established in the network, and at the output returns a list of rerouted lightpaths with their new routes and spectrum allocation. Since AFRO tries to make use of the recently restored link capacity, new routes are forced to use the repaired link. The lightpaths that cannot be routed over the restored link are kept without changes and are not considered for reoptimization. This constraint allows to reduce significantly the amount of lightpath reallocations in the network. After solving AFRO and implementing its solution (state t_3), the network is ready to operate and its performance is expected to be improved. As can be seen in Figure 65c), connection p_4 between nodes 1 and 6 can be established now.

Note that during the AFRO processing time the process of establishing new connection requests is stopped, similarly to other dynamic network reoptimization operations such as spectrum defragmentation. For this reason, the time needed for solving AFRO and



implementing network changes should be short in order to reduce the delay of that process. Indeed, to coordinate with other operations and to reduce even more the impact over new incoming traffic, AFRO might be scheduled in a time slot with less dynamic activity (e.g. night hours) and, therefore, not applied immediately after the failure repair.

The AFRO optimization problem can be formally stated as follows:

Given:

1. flexgrid optical network represented by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{D})$ where \mathcal{V} is the set of optical nodes, \mathcal{E} is the set of fiber links, and \mathcal{D} is the set of (traffic) demands,
2. optical spectrum (given at each link) which is divided into a set $\mathcal{S} = \{s_1, s_2, \dots, s_S\}$ of frequency slices of a given width,
3. set of lightpaths currently established in the network,
4. link $e^r \in \mathcal{E}$, representing the link recently repaired that was previously unavailable.

Find: a subset of demands from set \mathcal{D} which will be rerouted and, for each of these demands, a new lightpath over the network such that it uses the repaired link e^r , aiming at utilizing recently restored link capacity, with the *objective* to optimize the usage of spectrum in the network.

The mathematical model for the AFRO problem is presented in section 8.7.5.

Numerical and simulation results

In this section, we evaluate the performance of a dynamic flexgrid network operating with AFRO. To this end, first we compare different options for the AFRO optimization algorithm in order to find best performing settings. Then, considering these settings, we apply the AFRO framework in a dynamic network scenario and evaluate it by means of network simulations. Eventually, we study the impact of applying AFRO only for a selected set of network links, instead of using it in all network links, with the aim to reduce the AFRO operational complexity.

Evaluation scenario

To evaluate the performance of AFRO, we use two representative core network topologies: the 22-node British Telecom (BT) and the 28-node European (EON) networks (see Figure 66). The details of the networks are presented in Table 24. We consider the fiber links with the spectrum width equal to 2 THz, divided into frequency slices of 6.25 GHz (320 slices per link). We assume a link failure rate of $2.72 \cdot 10^{-3}$ per km per year [64]. Taking this into account and considering the total length of fiber links in BT and EON (as presented in Table 24), the average number of cuts per year is equal to 14 and 70, respectively. Accordingly, the mean time to failure (MTTF) is equal to 625 and 125 hours, respectively, for BT and EON. We assume that the mean time to repair (MTTR) is the same for both networks and equal to 12 hours. The AFRO problem is solved using the CPLEX v.12.5 optimizer [18] on a 2.4GHz Quad-core machine with 8GB RAM. Simulations are run on an ad-hoc event-driven simulator implemented in OMNET++ [47].

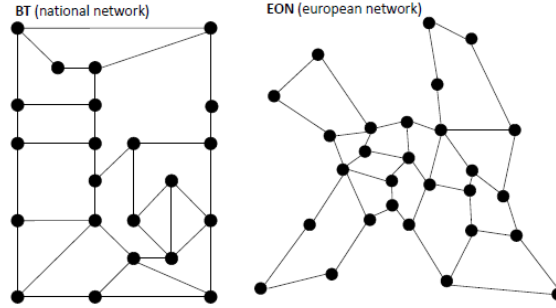


Figure 66: Network topologies.

Table 24: Network parameters.

Topology	Nodes	Links	Total Link Length [km]	Av. # of cuts per year	MTTF [h]	MTTR [h]
BT	22	35	5145	~ 14	625	12
EON	28	41	25625	~ 70	125	12

Connection requests are generated following a Poisson process and lightpaths are torn down after an exponentially distributed holding time (HT). The source and destination nodes are uniformly chosen. We consider that the requested bit-rate can be equal to 40, 100, or 400 Gbit/s, and the share of each connection type is equal to 66.7%, 26.7%, and 6.6%, respectively. Such traffic profile combines several bit-rates and takes advantage of high spectrum granularity of the flexgrid. Assuming QPSK as the applied modulation format, with spectrum efficiency equal to 2 bit/s/Hz, the number of slices requested for each connection type is 4, 8, and 32, respectively. One of the performance metrics that we use in the evaluation is the blocking probability (BP) of incoming connection requests weighted by the requested bit-rate.

Evaluation of the AFRO algorithm

Here, our main objective is to find best performing settings of the AFRO optimization algorithm. To this end, we generate network instances in the following way:

1. we run a network simulation, realizing connection setup and tear down requests, until reaching a steady state in terms of BP—after that the network is in state t_0 ;
2. we select a link, which is subject to failure, and run the bulk restoration algorithm proposed in [47] for affected lightpaths—the network is in state t_1 ;
3. the link is immediately repaired after restoration—the network is now in state t_2 ;
4. we run AFRO—after that the network is in state t_3 ;

After performing the above steps, we calculate network statistics (discussed in details later). The evaluation is performed for such traffic load, proper for each studied network, that it results in BP equal to 1% in state t_0 . Presented results are averaged over 100 randomly generated network instances.

Efficiency of IP formulations

First, we compare the results obtained for two different formulations of the AFRO problem, i.e., MIN-SUM and MIN-MAX. We remind that these formulations differ in the objective function that is subject to optimization. In this analysis, we consider that all affected lightpaths are subject to optimization and the mechanisms minimizing traffic disruptions are not used.

Table 25: Link usage and link entropy observed in different network states.

Scenario		Link usage						Link entropy					
		$t0$		$t2$		$t3$		$t0$		$t2$		$t3$	
Network	Formulation	av.	max.	av.	max.	av.	max.	av.	max.	av.	max.	av.	max.
BT	MIN-SUM	79.40	226.04	81.22	231.92	79.31	227.56	1.0755	1.7355	1.0543	1.7506	1.0519	1.7307
	MIN-MAX	79.40	226.04	81.22	231.72	81.18	231.04	1.0755	1.7355	1.0541	1.7465	1.0435	1.7386
EON	MIN-SUM	77.97	215.68	79.76	229.33	77.87	219.60	1.1232	1.7427	1.1034	1.7500	1.0930	1.7302
	MIN-MAX	77.97	215.68	79.75	229.36	79.63	228.36	1.1232	1.7427	1.1063	1.7500	1.1011	1.7450

In Table 25, we present the results of network link usage, in terms of the number of allocated slices, both averaged over all network links and obtained for the most occupied link in the network. Also, we focus on link entropy, which is a metric related to the fragmentation of the free spectrum and which is computed by means of the Shannon link entropy formula proposed in [61]. This entropy value is lower when the fragmentation of free spectrum is lower.

We can see that in all studied cases the application of AFRO (see $t3$) improves both studied metrics when compared to the network without AFRO (see $t2$). Moreover, AFRO allows to reoptimize the network so that its link statistics are similar to those observed during normal network operation (compare $t3$ with $t0$). Finally, the optimization of overall spectrum usage (MIN-SUM) is more effective than optimization of the sum of highest indices of used slices over all links (MIN-MAX). Therefore, in the remainder we consider that AFRO makes use of the MIN-SUM formulation.

Evaluation of the column generation algorithm

Next, we analyze different strategies for selecting demands to be processed, namely, *complete* and *reduced*, as well as a strategy in which all demands in the network are processed (denoted as *full*). We consider that the mechanisms minimizing traffic disruptions are not applied in the analysis.

Table 26: Performance of the column generation algorithm.

Network	Demand selection strategy	$ D' $	$ P $	$ P^r $	link entropy	gap [%]	T [s]
BT	full	255.28	723.90	13.12	1.0520	0.021	21.567
	complete	226.75	717.86	13.08	1.0516	1.010	21.686
	reduced	45.58	747.18	13.07	1.0520	0.021	22.458
EON	full	263.53	645.43	13.37	1,0928	0.027	15.946
	complete	238.14	639.62	13.12	1,0934	0.023	15.636
	reduced	43.48	625.43	13.16	1,0928	0.025	15.404



In Table 26, we present averaged results obtained for the considered strategies. We focus on the number of demands which are subject to reoptimization ($|\mathcal{D}'|$), the number of allowable lightpaths obtained in the column generation algorithm ($|\mathcal{P}|$), the number of rerouted lightpaths ($|\mathcal{P}^r|$), the link entropy, a relative optimality gap between the integer problem solution obtained for the generated set of lightpaths and the LP relaxation (gap), and the algorithm processing time (T).

We can see that the reduced strategy allows to decrease considerably the number of demands selected for optimization when compared to the other two strategies. At the same time, the number of generated and rerouted lightpaths as well as the link entropy is comparable for all strategies. Note that the optimality gap is very low (below 0.03% in almost all cases). It shows that the quality of generated lightpaths is very high and for these lightpaths near-optimal integer solutions can be found. We can also see that the processing time is very similar in all analyzed cases and is much less than 1 minute. We can conclude that the algorithm is fast enough for the scenarios with traffic dynamics that are assumed in our dynamic simulations (in the next section).

In the remainder, we assume that the column generation algorithm is implemented with the *reduced* strategy, since it provides the best trade-off between solution quality and execution time.

Analysis of traffic disruptions

Eventually, we study the effectiveness of different traffic disruption-aware mechanisms.

Table 27: Disruption analysis.

Network	Disruption-aware CG	Rerouted lightpaths	Number of disruptions	Number of bundles
BT	No	13.40	2.29	2.15
	Yes	14.19	0.47	2.49
EON	No	13.18	2.89	2.36
	Yes	13.51	0.26	2.52

In Table 27, we present averaged results of the number of rerouted lightpaths, the number of lightpath disruptions, and the number of lightpath bundles. We remind that lightpath disruption occurs if the lightpath cannot be rerouted with the make-before-break mechanism. The number of bundles corresponds to the signalling overhead of the lightpaths rerouting.

We can see that both the number of lightpaths selected to be rerouted, which is the result of the optimization algorithm, and the signalling overhead are not affected by using traffic disruption-aware mechanisms. At the same time, the number of disrupted lightpaths is significantly reduced, up to one order of magnitude for EON.

In the following experiments, the traffic disruption-aware mechanisms are applied.

Dynamic performance of AFRO

Here, we study the performance of AFRO by means of dynamic network simulations.

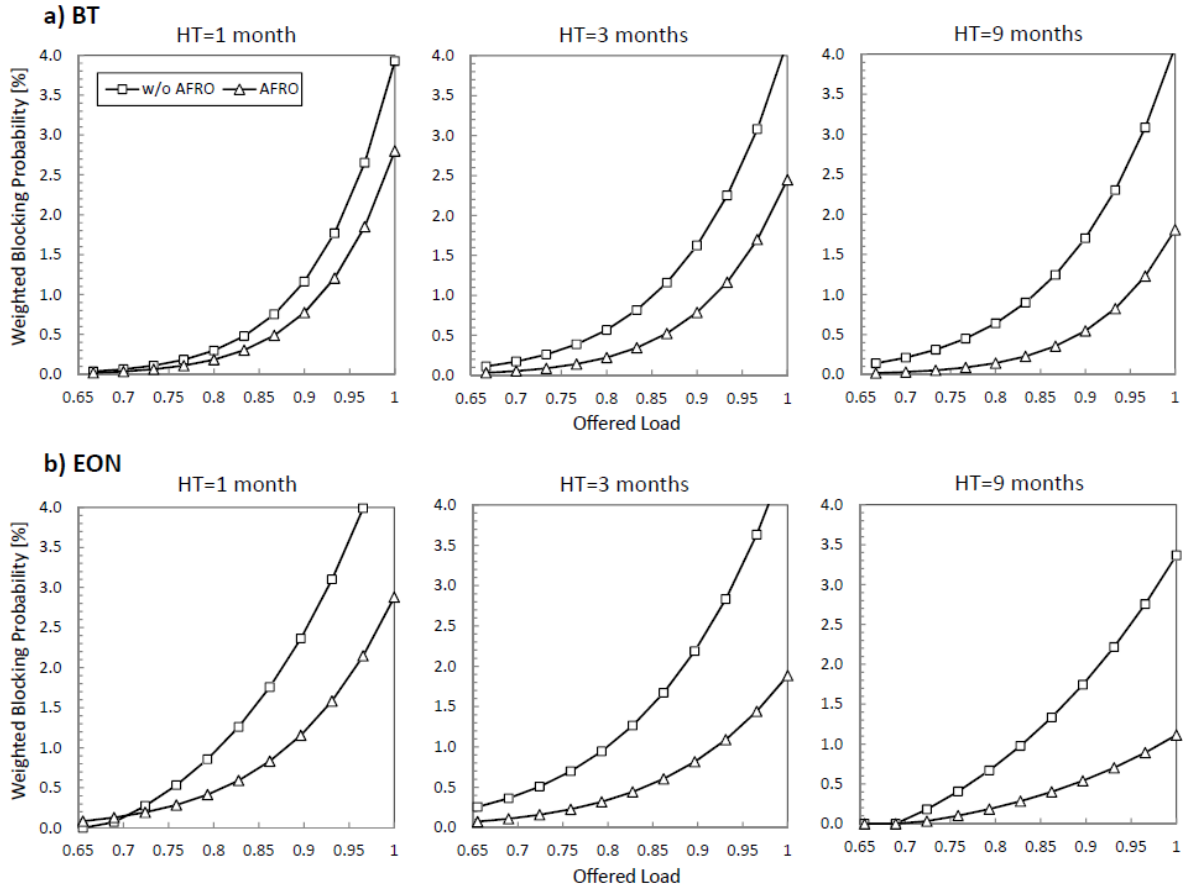


Figure 67: Weighted blocking probability vs. normalized network load.

In Figure 67 we show the results of BP in a function of the offered load, which is normalized to the load for which BP is equal to 1% in a network operating without failures (i.e., for $MTTF = \infty$). The results are obtained for different values of connection holding time (HT). We can see that the application of AFRO leads to lower values of BP in all studied scenarios when compared to a network without AFRO. Moreover, the relative gain in terms of BP after applying AFRO increases with traffic load as well as with higher values of HT. It can be concluded that if a network has to guarantee certain target level of BP (e.g., $BP=1\%$), the application of AFRO allows to increase the network throughput. Finally, we can see that the difference between the cases with AFRO and w/o AFRO is more prominent in the EON topology than in the BT topology.

In order to evaluate the gain in supported traffic load (i.e., throughput), in Figure 68 we present these results obtained for different values of HT at $BP=1\%$. Again, we can see that the higher value of HT the higher gain in network throughput after applying AFRO. Also, in order to achieve a 10% gain, the connection holding time can be shorter in EON ($HT \approx 1.5$ month) than in BT ($HT \approx 4$ months).

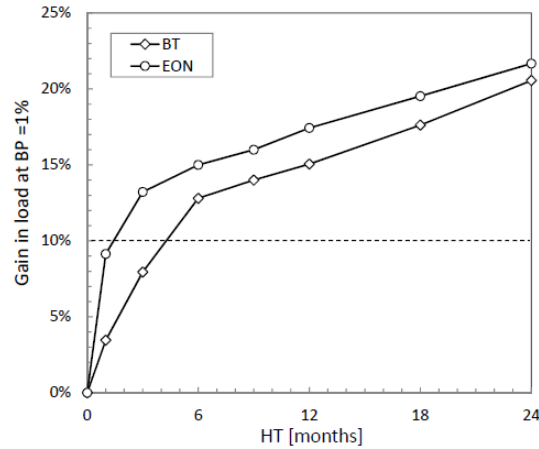


Figure 68: Gain in supported load at BP = 1% vs. connection holding time.

Table 28 Effectiveness of connection restoration in a dynamic network without and with AFRO.

Network	Scenario		% of restored connections	
	HT	HT/MTTF	w/o AFRO	AFRO
BT	4	4.6	99.58	99.65
EON	1.5	8.7	99.48	99.52

Eventually, in Table 28 we evaluate the impact of AFRO on the effectiveness of the restoration mechanism used in our network. In particular, we can see that in both BT and EON topologies the percentage of restored connections in a network operating with AFRO is not only preserved but it is even slightly better than in a network w/o AFRO.

Solving AFRO after selected link repairs

The last set of experiments concerns network scenarios in which AFRO is run only after the recovery of a selected set of links, instead of using it for all network links. In particular, in this study AFRO is applied for a set of links which are most loaded during normal network operation.

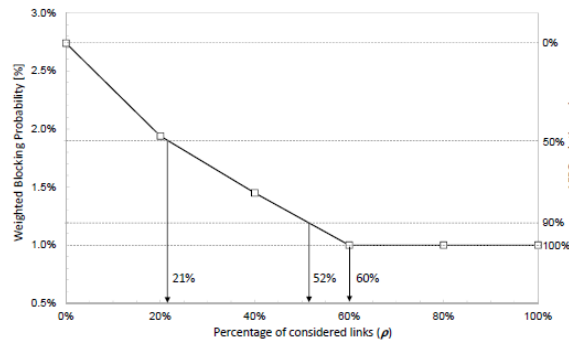


Figure 69: Weighted blocking probability and AFRO relative gain vs. percentage of links considered for reoptimization.



In Figure 69, for the BT topology and HT=9 months, we present the performance of AFRO, in terms of weighted BP, assuming different percentage of network links which are considered for reoptimization (denoted as ρ). Note that AFRO is not applied if $\rho = 0\%$, is always used if $\rho = 100\%$, and for instance for $\rho = 40\%$ it is applied for 40% of (most loaded) network links. Apart from that we calculate the AFRO relative gain, which represents the improvement in BP with respect to the scenario without using AFRO.

In Figure 69, we can see that it is not necessary to run AFRO for all network links in order to obtain its maximal efficiency. For instance, if we restrict AFRO to be applied after the recovery of links which are among 60% of most loaded links, the AFRO relative gain is already equal to 100%. Moreover, it is enough to use it only for 20% of links so that to achieve a 50% reduction of BP. These results indicate that AFRO can be applied with lower frequency, thus reducing its operational and signaling complexity, without affecting significantly its effectiveness.

Conclusion

In this section we presented AFRO—an online reoptimization problem useful for improving traffic performance of a flexgrid optical network subject to link failures and employing dynamic lightpath restoration. Specifically, when a failure is repaired, AFRO considers rerouting the traffic from some existing lightpaths to new lightpaths established through the repaired link, thus making use of the restored capacity. Through this, a better distribution of network load and, consequently, a decrease in the connection blocking probability can be achieved.

The AFRO problem was defined by means of two alternative link-path IP formulations. Due to the AFRO requirement to use the repaired link by new lightpaths and in order to avoid the solution of large problem instances, i.e., with a large set of pre-computed lightpaths, column generation was applied in a form of a time-wise acceptable procedure to obtain near-optimal solutions. Based on our previous work on column generation for a basic RSA problem in flexgrid optical networks, specific procedures for both presented formulations were developed. Moreover, several strategies for reducing the computational effort, including the reduction of the amount of lightpaths considered for reoptimization, as well as some extensions to minimize traffic disruptions when migrating from old to new lightpaths were proposed.

Numerical results obtained for a national network and a European network show that AFRO reduces both the amount of used capacity and the spectrum fragmentation (measured in terms of link entropy) with respect to the network state before applying AFRO. Additionally, the proposed column generation algorithm provides quality solutions in execution times acceptable for a real dynamic scenario. The effectiveness of AFRO in improving the performance of a dynamic network in operation is demonstrated by two main results: (i) a significant blocking probability reduction ($> 10\%$) for reasonable values of MTTF, MTTR, and HT is achieved; (ii) application of AFRO allows improving the performance of the dynamic restoration algorithm by increasing the restorability of connections affected by a failure. Finally, it is shown that the operational overhead of AFRO can be reduced by using its mechanism only for a selected set of links ($\geq 60\%$ of all links) without a negative impact on its performance.



The presented models are developed for a network operating with a single modulation format. Further extensions corresponding to adaptive transponders, i.e., making use of multiple modulation formats in accordance to lightpath characteristics, are left for further study.

5.3 Selected algorithms to be implemented

Both IDEALIST planning tool and control architectures include open interfaces so that algorithms proposed in the project could be easily developed and implemented over the IDEALIST architecture. According to it, a set of algorithms will be implemented and experimentally tested in WP4. In the following subsections we describe our proposal to implement these algorithms.

5.3.1 Spectrum Defragmentation

The main purpose of spectrum defragmentation is improving network resource utilization. From a control and management perspective, the defragmentation process maps into a set of state changes of active connections. Such state changes are reflected in the change of connections' attributes, in our case spectrum allocation by shifting the nominal central frequency of the slot allocated to a connection. Some other attributes, not covered in this work, can be updated, e.g. its allocated spectrum width (i.e., due to a change of modulation formats or bitrate).

The optimization process can be triggered either manually by a network operator through the NMS, by an automated process triggered by some threshold or in a periodical fashion. In our case, let us assume that the defragmentation procedure is triggered after the front-end PCE fails to find a suitable route for a provisioning request.

The request for a new connection is originally issued by the NMS and received by the ABNO controller through the north-bound interface. In such case, the ABNO controller is responsible for coordinating connection set-up, which composes and sends a specific request towards the front-end PCE, in charge of computing and finally coordinate connection establishment.

The workflow that represents the provisioning-triggered defragmentation use case is detailed in Figure 70. As already introduced, it starts with a network operator requesting a new connection provisioning through the NMS. Establishing a flexible optical connection includes computing and provisioning a continuous slot between two nodes in the data plane. The request is received by the ABNO controller via its north-bound interface (step 1 in Figure 70). When the ABNO controller receives the request it asks the policy agent to check about rights of the received request (2). If access is granted, the ABNO controller requests the front-end PCE to compute the route and eventually set up the optical connection (3).

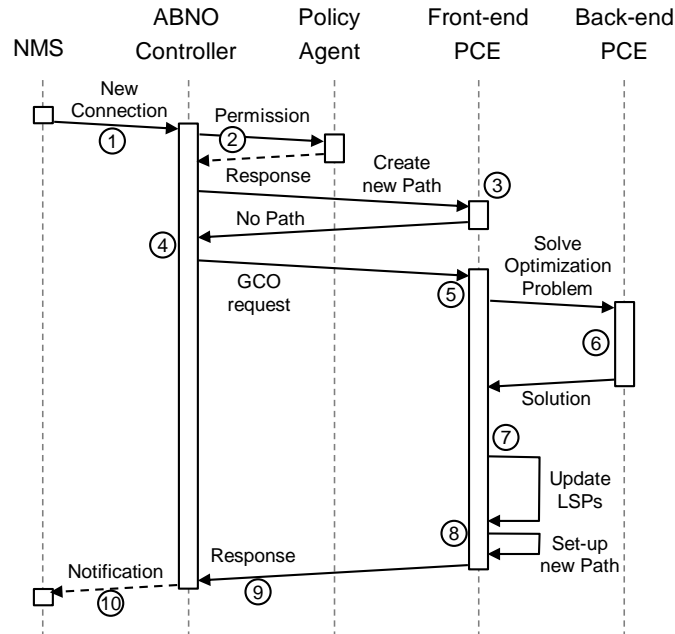


Figure 70: Defragmentation workflow.

Let us assume that as a result of spectrum fragmentation no end-to-end continuous slot is found (4). In that case, the ABNO controller may autonomously decide to perform a defragmentation process and sends a message to the front-end PCE (5). When the front-end PCE receives the request for defragmentation, it checks its feasibility and gathers information to create a GCO request that is sent towards the back-end PCE to solve the optimization problem (6). When the back-end PCE ends, it sends back the solution found. The front-end PCE proceeds then to execute the defragmentation that consists in shifting some of the candidate LSPs (7) and finally, when every LSP has been updated, the front-end PCE proceeds to establish the requested connection (8). Upon its completion, the front-end PCE notifies the ABNO controller (9), which in turn notifies the NMS (10).

Obviously, if the request in (3) finds a feasible route and spectrum allocation, the front-end PCE proceeds with step (8) to establish the connection.

All interactions between ABNO and PCEs are done by exchanging PCE protocol (PCEP) messages. In particular, Path Computation Request (PCReq) and Path Computation Reply (PCRep) messages are exchanged between front-end and back-end PCEs (step 6 in Figure 70).

5.3.2 Re-optimization After Link Failure Repairs

In this use case, the re-optimization of a set of existing LSPs is requested and therefore, an algorithm in the PCE must be in charge of computing the candidate LSPs to be re-optimized. Once computed, the list of LSP must be sent to the bPCE. This is done by a PCReq message including a Global Concurrent Optimization (GCO) request. The PCReq includes a list of LSPs in a Synchronization VECtor (SVEC) object. For each path computation request in the SVEC, there is a Request Parameter (RP) object. The RP

carries a Request-ID (32 bit integer) and the symbolic path name of the LSP within the LSPDB as a TLV. This symbolic name is always passed from fPCE to bPCE when a new LSP has been set up and identifies each LSP. Consequently, the bPCE is able to unambiguously identify the LSPs to be re-optimized using both the RP and the LSPDB

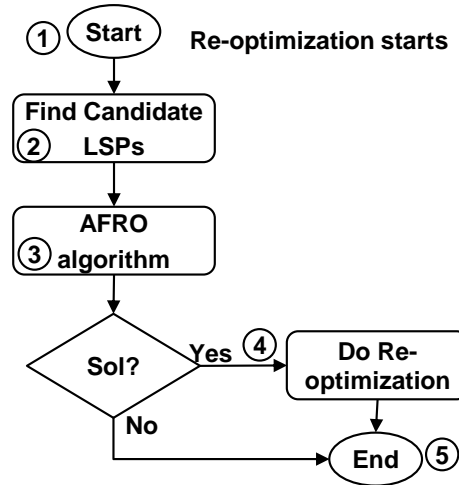


Figure 71: Workflow for the re-optimization process.

Next, the bPCE runs the re-optimization algorithm for each LSP (step 3). The AFRO algorithm re-routes the LSPs on shorter routes to enhance the spectrum efficiency. The path computation result is then sent to the fPCE via a PCEP Response (PCRep) message (step 4). In this message, for each re-optimized LSP an Explicit Route Object (ERO) is carried with the route and assigned FS.

At the fPCE, the new received ERO for each LSP is compared to the one stored in the LSPDB. If an ERO update is detected, the fPCE sends a PCEP Update (PCUpd) message to the corresponding PCC (step 5). This indicates the PCC to re-establish the LSP using the new computed ERO. Finally, the fPCE sends PCRpt messages to the bPCE to actually update the LSPDB repository with the current LSP information. The complete re-optimization is summarized in the workflow of Figure 71.

5.3.3 Inter-datacenter Multicast Connectivity

To implement multicast connectivity services in a multi-layer network, a virtual topology needs to be created connecting every source switch to every other leaf switch. Several alternative approaches can be considered to support that virtual topology: *i*) create a set of point-to-point (p2p) optical connections between each pair of source-leaf switches or *ii*) create a set of point-to-multipoint (p2mp) optical connections, one for each source router connecting all the leaves. Let us assume the latter approach.

An Application Service Orchestrator (ASO) maintains the network information from the DCs requests and interacts with the ABNO. In this use case, the ABNO architecture includes *i*) a controller, responsible for implementing workflows orchestrating operations among ABNO modules; *ii*) a Layer 0 PCE (L0 PCE), responsible for path computation on the optical topology; *iii*) a Virtual Topology Network Manager (VNTM), responsible for

maintaining a virtual topology between the DCs using resources in the optical topology; iv) a Layer 2 PCE (L2 PCE), which computes paths on the virtual topology; v) a Provisioning Manager (PM) dealing with the configuration of the network elements (switches or optical nodes) and vi) a Topology Module (TM) in charge of obtaining the TED.

Besides, the L0 PCE is split into two dedicated modules: a front-end PCE (L0 fPCE), responsible of performing impairment-aware computation and initiating the optical connections: and a back-end PCE (L0 bPCE), capable of performing computationally intensive tasks, such as solving the p2mp routing and spectrum allocation algorithm.

When multiple DC networks need to be connected, the ASO module asks the ABNO controller for unicast and multicast services between the Ethernet switches. Figure 72 shows the workflow for one single p2mp connection request. The same workflow is repeatedly executed for the rest of the requests belonging to the same multicast service request.

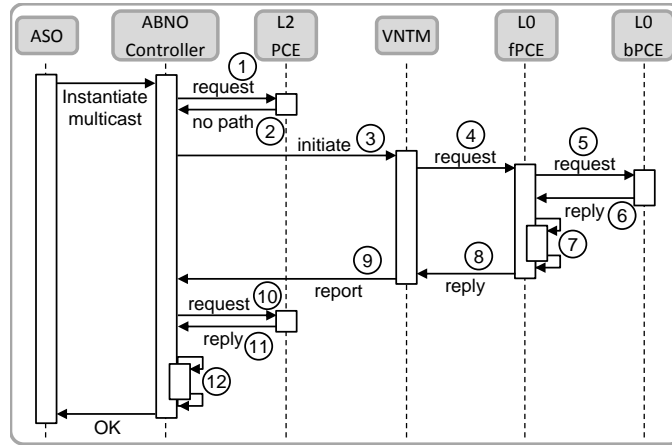


Figure 72: Workflow for multicast connectivity.

In our implementation, the L0 fPCE is active whereas L2 PCE is not. Therefore, the ABNO controller requests a L2 p2mp path computation to the L2 PCE (message 1 in Figure 72). Let us assume that no enough resources are available at this time, so the L2 PCE returns a NO-PATH message (2). The controller delegates to the VNTM module updating the virtual topology, possibly adding more resources to serve the L2 p2mp request. To that end, a request is sent containing the end points of the requested p2mp connection (3). Upon reception, the VNTM sends a request to L0 fPCE to create optical connectivity among the specified end points (4); in this particular use case, a p2mp optical connection needs to be created.

Because L0 p2mp path computation might take long time, L0 fPCE delegates it to the specialized L0 bPCE (5). When the computation ends, the L0 bPCE sends back the solution to the fPCE (6). When the L0 fPCE receives the solution, it delegates its setting-up to the PM, which sends the appropriate commands to the underlying data plane (7).

When the PM module receives the confirmation from the L0 fPCE (8), it is forwarded to the ABNO controller (9). Then the ABNO controller requests a L2 p2mp path computation to the L2 PCE (10), which now finds a solution and returns it to the ABNO controller (11).



Since L2 PCE is not active, the ABNO controller module delegates connection set-up to the PM (12).

5.3.4 Distance Adaptive RSA

EON and Bandwidth Variable Transponders (BV-T) enable transmission of different channel's capacities and different signal's format. This characteristic allows going a step further introducing the adaptation of the modulation format of the signal to the distance of the path to traverse.

In a network scenario where the routes can vary between longer and shorter reach paths, the latter can be transmitted through a more efficient modulation format such as DP-16QAM.

Algorithm's definition

To this purpose, it has been implemented a novel heuristic algorithm called: KSP Distance Adaptive Multifiber Spectrum Assignment (KSP-DA-MSA), to compare the network evolution results with a more efficient use of the spectral resources by adapting the modulation format to the path. Table 73 shows the different formats admitted by the algorithm, with the related maximum reach of each case.

<i>Modulation Format</i>	<i>Capacity</i>	<i>SE</i>	<i>FEC</i>	<i>Guard Band</i>	<i>Distance Max</i>
OOK	10	1	0.12	7	2200
DP-QPSK	40	4	0.12	7	2800
DP-16-QAM	40	8	0.12	7	800
DP-QPSK	100	4	0.12	7	2800
DP-16-QAM	100	8	0.12	7	800
OFDM-DP-QPSK	400	4	0.12	10	3560
OFDM-DP-16-QAM	400	8	0.12	10	800
OFDM-DP-QPSK	1000	4	0.12	10	3560
OFDM-DP-16-QAM	1000	8	0.12	10	800

Table 73: Signal's formats available by the distance adaptive algorithm.

The KSP-DA-MSA uses a graph with multiple fibers between each node in order to account the scalability of the network, through the annual estimation of new fibers, WSS and BV-Ts required to deploy in order to accommodate the incoming demands.

KSP-DA-MSA:

- (1): K-shortest paths are calculated, and the list of edges sending to the Spectrum Assignment Module (SAM).*
- (2): SAM takes the edge list from the first shortest path, and checks out the availability of optical channels (an optical channel here is defined as a number of contiguous frequency slots which accommodate the bandwidth demand) through the entire route).*
- (3): The order (number of contiguous available frequency slots) of the channel depends on the signal's format of the optical carrier which will transport the demand. Initially the algorithms always choose the most spectrally efficient modulation format available.*
- (4): The number of possible "spectrum assignments" solutions grows up with the number of hops. If n is the number of parallel fibers and m the number of hops, the number of possible solutions will be n^m .*
- (5): All solutions are computed. The electability's criteria among all the options, is to choose the most occupied paths in order to maximize the occupation of the network, and to delay the most as possible, the introduction of new equipment.*
- (6): When the "physical route" has been chosen, the distance and physical impairments of the route are checked out to fulfill the modulation format OSNR requirement at the reception.*
- (7): Finally the lowest starting slot (LSS) strategy is chosen to select the optical channel where allocate the request.*

Figure 74: Steps in KSP Distance Adaptive Multifiber Spectrum Assignment algorithm.

The algorithm performs the routing through a K-shortest path algorithm (with $K=3$), and uses a channel-based approach to calculate the spectrum assignment. The novelty of this algorithm is the introduction of the RSA problem into a MultiGraph environment. For this study the number of parallel fibers n is growing following the same strategy that in the previous section. When it is impossible to find a channel, for a specific demand, the availability of resources, of the links within the path, are check out. On those links with the highest occupation, a new fiber is included.

To divide the problem in two: (1) routing and, (2) spectrum assignment, gives an easier solution of the election of the fiber to route the channel by increasing the time wasted in the computation. However as this study has not the aim of working on dynamic scenarios, the timing performance of the algorithm is not a concern.

5.3.5 Algorithms to be implemented in PLATON

Finally, we have selected the set of algorithms to be implemented in PLATON. These are:

- Spectrum Defragmentation: the SPRING algorithm.
- Reoptimization of Dynamic Flexgrid Optical Networks After Link Failure Repairs: AFRO algorithm.
- Multicast Connectivity provisioning for Ethernet services.

6 Planning Tool Prototype

Several planning tools are being developed within WP1, as a result of the wide range of problems to be solved. Those problems are related to the assignment of optical spectrum across specific network routes to meet traffic demands. The RSA problem is NP-complete and therefore requires carefully crafted heuristic approaches to give solutions in realistic time scales, especially as the number of network nodes increases.

As technologies are developed to allow the network to become more agile, it may be possible to provide response to traffic changes by reconfiguring the network near real-time. In fact, some operators have deployed Generalized Multi-Protocol Label Switching (GMPLS) control planes, mainly for service set-up automation and recovery purposes. However, those control only parts of the network and do not support holistic network reconfiguration. This functionality requires an in-operation planning tool that interacts directly with the data and control planes and operator policies via Operations Support System (OSS) platforms, including the Network Management System (NMS).

Assuming the benefits of operating the network in a dynamic way are proven, the classical network life-cycle has to be augmented to include a new step focused on reconfiguring and re-optimising the network, as represented in Figure 75. We call that step in-operation planning and, in contrast to the traditional network planning, the results and recommendations can be immediately implemented on the network.

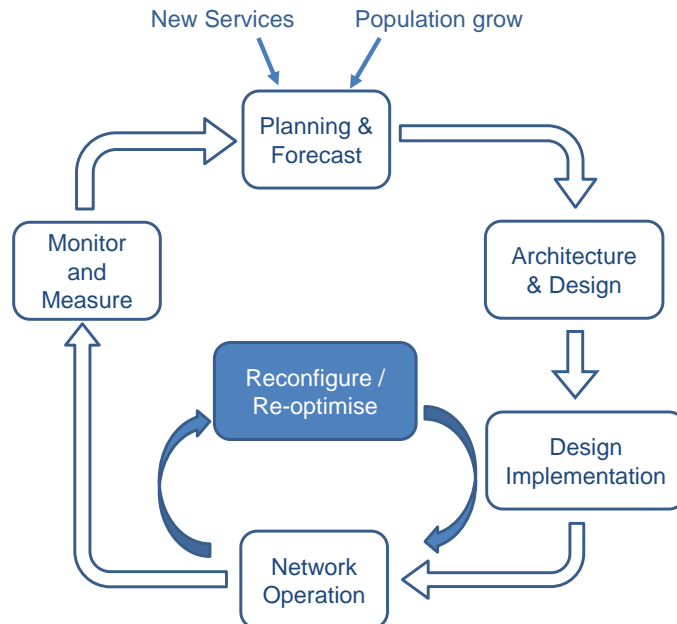


Figure 75: Augmented networks life-cycle.

Figure 76 illustrates the proposed control plane architecture to support flexgrid network re-optimisation, which also facilitates human verification and acknowledgement of network

changes. When Router A needs a new connection to Router B, it sends a request to the control plane of the optical network (1). After checking admission policies, a PCE Protocol (PCEP) message is sent to the PCE (2), which invokes its local provisioning algorithm (3). In the event of insufficient resources being available, e.g. due to spectrum fragmentation, the active PCE recommends the defragmentation of relevant nodes and connections, utilising the right algorithm to provide such re-optimisations. Let us assume that the back-end PCE providing such algorithm will perform the computation (4) upon receipt of a request. When a result is obtained (5), it is sent back to the front-end PCE (6). In case that an operator need to approve implementing the computed solution in the network, a request can be sent to the NMS/OSS (7). When the solution has been verified and acknowledged by the operator, the NMS/OSS informs the PCE (8) and existing connection reallocations are requested. Once the dependent connections have been setup, the responsible PCE will invoke the local provisioning algorithm for the original connection request between routers A and B and sends a PCEP message to the originating control plane node (9).

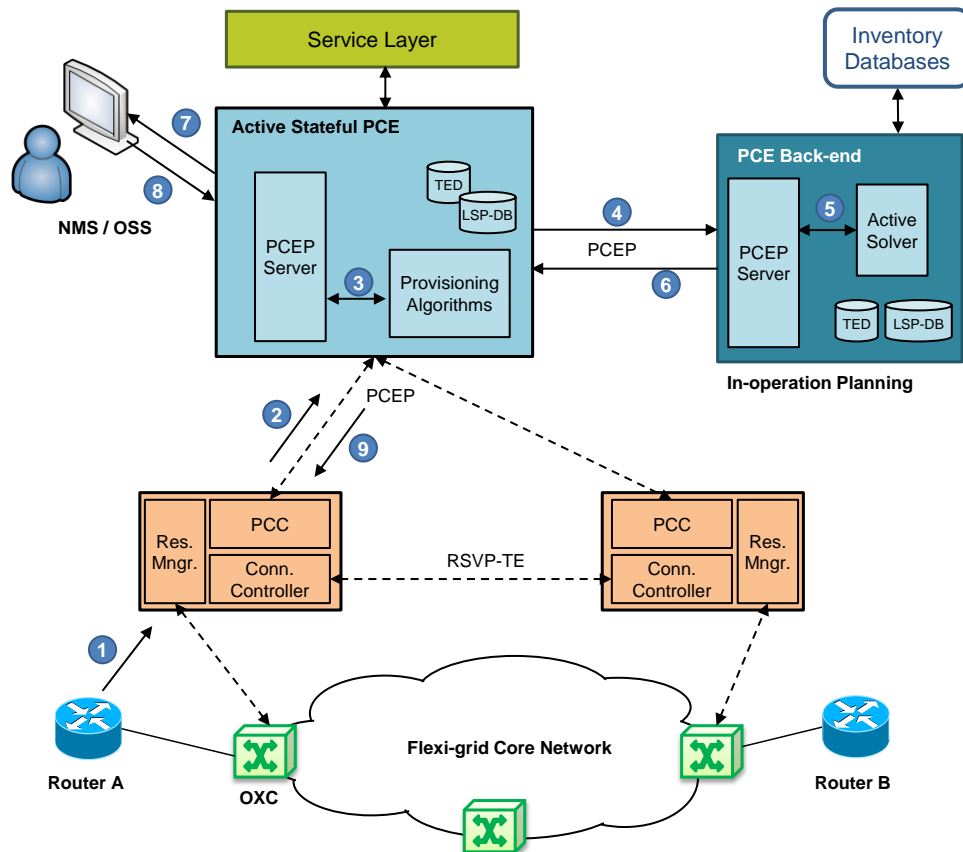


Figure 76: Proposed control plane architecture and network re-optimisation.

With the above in mind, two distinct tools are being developed to handle in-operation planning:



- **MANTIS.** Although predominantly off-line processing tool, it can be used to solve on-line problems, which are related to in-operation planning.
- **PLATON.** Addresses problems such as post-repair optimisation and spectrum defragmentation.

6.1 Mantis

Mantis is a network planning and operation tool for taking planning and operational decisions for the next generation optical networks, supporting both flexible (elastic) and mixed line rate (MLR) WDM networks. Through Mantis, the user is able to define the network topology, current and forecasted traffic matrices, CAPEX/OPEX parameters, set up basic configuration parameters, and use a library of algorithms to plan, operate, or run what-if scenarios for an optical network of interest. Mantis is designed to be deployed either as a cloud service or as a desktop application. Using the cloud infrastructures features Mantis can scale according to the user demands, executing fast and efficiently the scenarios requested. Mantis supports different cloud platforms either public such as Amazon Elastic Compute Cloud (Amazon EC2) and ~okeanos or private based on OpenStack, while its modular architecture allows other cloud infrastructures to be adopted in the future with minimum effort. The included planning and operation algorithms range from routing and wavelength or spectrum allocation, to equipment (e.g. transponders and regenerators) placement, and CAPEX/OPEX/energy analysis

In what follows we give a short description of Mantis Cloud operation and then we test and evaluate its function in solving planning and operating problems. Note that most of the planning problems, and in particularly those addressed in Mantis, are NP-hard. These problems are solved offline, and the responsiveness (the speed in finding solutions) of the related algorithms is important, but the related requirement is relaxed. Exact solutions (e.g. based on ILP) that are intractable for realistic problem instances should be avoided, especially in the case of running many what-if scenarios, but the employed algorithms can run for long times, depending on the level of optimality that we want to achieve. Mantis includes meta-heuristic planning algorithms that trade-off performance for running time and thus fit well with the above specification. On the other hand, algorithms for solving network operation problems need to be much faster, since the function of the network depends on their outcomes. It is for such algorithms that strict running time requirements, e.g. find a solution in less than one second, are posed. Mantis includes online algorithms for serving new demands and defragmenting the spectrum that also trade-off performance for running time. Thus the online algorithms can be tuned to satisfy the running time requirements at hand.

Then we examine online problems, where we focus on the running time of a single online problem and examine the responsiveness of the related algorithms and the tool.

6.2 PLAnning Tool for Optical Networks (PLATON)

Whilst a version of MANTIS existed pre-Idealist, PLATON is a new tool, being developed with in the Idealist project to specifically address problems requiring real time decisions.

The in-operation PLATON architecture, depicted in Figure 77, is divided in five functional blocks: the communications module, the manager, the optimization algorithm framework, the databases, and the set of optimization algorithms.

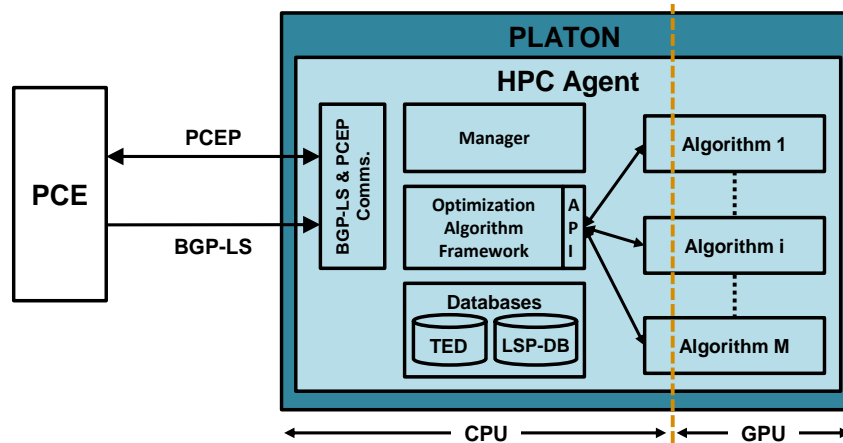


Figure 77. In-operation PLATON architecture

The communications module implements the set of standard protocol interfaces used by PLATON to synchronize its internal databases and to handle computation requests. The manager module schedules jobs being executed inside PLATON. Each request to be processed by PLATON is treated a job. The optimization algorithm framework module provides a set of methods and functions common to every optimization algorithm. The set of algorithms implements the optimization algorithms executed by PLATON when a computation is requested. The methods and functions exported by PLATON are accessible through the Algorithm's API. The databases module contains the Traffic Engineering Database (TED) and the Label Switched Path Database (LSP-DB) including appropriate methods and functions to manage them, i.e. create node, create link, connect node to link, create lightpath, allocate lightpath, etc.

Table 29 shows the intended development plan.

Table 29 PLATON development plan

Task	%Done
PLATON	100%
In-Operation Architecture	100%
REST/API	100%
PCEP	100%
BGP-LS	100%
Manager	100%
Optimization Framework	100%
Algorithms year #2	100%

	After Failure Repair Optimization	100%
	Spectrum Defragmentation	100%
Algorithms year #3		0%
	Define Algorithm Set	0%
	Implement Algorithms	0%

A set of tests were specified in [50] to test all the functionalities required to synchronize the TED and LSP-DB databases, and to receive and reply path computations. The following sections present an overview of the test-bed, the tests cases and results and the performance evaluation of PLATON.

6.2.1 Test-bed set-up

With the aim of debugging and doing realistic tests on PLATON, we have developed two additional modules to construct a distributed test-bed: a Graphical User Interface (GUI) that mimics the requests issued by a controller, and a dummy front-end PCE (fPCE) in charge of translating the GUI instructions into BGP-LS and PCEP messages for the PLATON. Figure 78 shows the distributed test-bed set-up. The GUI and the fPCE are connected through an HTTP connection, while the fPCE and the bPCE are through a BGP-LS and a PCEP ones.

The GUI takes as input XML files encoding BGP-LS and PCEP messages to be processed by PLATON. When a file is selected in the GUI, the messages contained messages are sent to the fPCE, who translates them into BGP-LS and PCEP messages and forwards them to the PLATON.

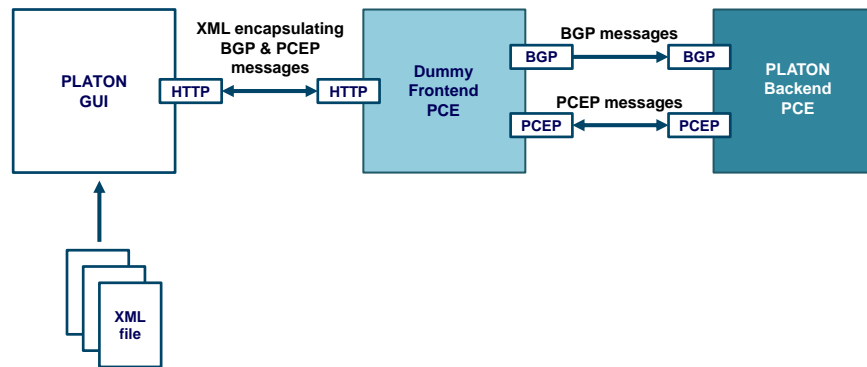


Figure 78. Test-bed set-up

When PLATON completes the path computation requests, it forwards the solution to the fPCE that encapsulates it into an XML message, and forwards it to the GUI to be shown to the user. In the case of messages corresponding to TED and LSP-DB updates, PLATON acts as a listener and does not reply any message.



6.2.2 Summary of results

The set of test cases have been done in PLATON and its results have been verified. Table 30 summarises the results of the test cases.

Table 30. Tests results summary

Id	Test Case	Short description	Result
1	Create nodes	Starting from an empty TED and LSP-DB databases populate nodes to the TED.	✓
2	Create links	Populate links to the TED	✓
3	Fail links	Starting from a populated TED a link fails.	✓
4	Repair links	The links are repaired	✓
5	Create lightpath	An LSP is created in the network and the LSP-DB is synchronized.	✓
6	Delete lightpath	An LSP is torn-down in the network and the LSP-DB is synchronized.	✓
7	Compute single path	A request to compute the RSA for a lightpath is received	✓
8	Compute bundle of paths	A request to compute the RSA for a set of lightpaths is received	✓
9	Compute elastic operation	A request to compute an elastic operation on an existing LSP is received.	✓
10	Lightpath Update	An existing LSP is updated	✓

6.2.3 Performance evaluation

To demonstrate the feasibility of the front-end back-end PCE architecture, we carried out a number of tests to evaluate its performance. In the experiments, we considered two different real backbone network topologies. The 22-node and 35-link BT network topology, and the 30-node and 56-link TEL network topology. The current status of each network is synchronized using BGP-LS. A number of bulk path computations of different sizes are generated, and can be solved in the front-end PCE or delegated to the back-end PCE.

When the front-end PCE is in charge of performing the bulk computation, one single CPU thread is used; therefore, iterations of the provisioning algorithm are performed sequentially in the front-end PCE. In contrast, when the front-end PCE delegates the computation to the back-end PCE, the former opens a new PCEP session, creates a PCReq message containing every path request and a SVEC object, and forwards the message to the back-end PCE. Upon receiving a PCReq message, the back-end PCE uses the GPU device to accelerate the computation; each iteration is executed into an



independent thread block in the GPU, and all the threads inside a block collaborate in solving the RSA problem in parallel for each demand in the bulk. The back-end PCE sends a PCRep message containing the route and spectrum allocation for each.

Table 31 details the solving times for three different bulk sizes. As observed, as soon as either the size of the bulk or the topology increases, the front-end PCE cannot perform the requested computation within 100ms. Note that bulk sizes in the order of 40-50 demands are common in restoration and re-optimization.

Table 31. Solving times

Bulk Size	BT network topology		BT network topology	
	Front-end PCE Time (ms)	Back-end PCE Time (ms)	Front-end PCE Time (ms)	Back-end PCE Time (ms)
30	21.6	11.5	248.4	38.5
40	72.4	16.9	466.3	65.4
50	514.3	74.6	889.3	142.6

The obtained solving time clearly validate the use of PLATON to compute complex optimization problems, thus liberating the PCE.

7 Conclusions

This deliverable first analyzed the impact of the adaptive network manager and an elastic-enabled control plane proposed in WP3 in terms of TCO reduction. The TCO reduction is achieved as a result of the more efficient network provisioning and resilience processes enabled by the control architecture proposed in IDEALIST.

Next, the IDEALIST architecture building blocks used for the techno-economic analysis were presented. They are based on the mid-term solution proposed by WP2 for SBVT and Flex-OXC. Models to evolve from current fixed-grid network deployments to flexgrid networks were investigated. The benefits from installing SBVTs were addressed under single and multi-layer network conditions, including point-to-multipoint traffic. In multilayer networks, the cost from SBVTs was optimized and saving as high as 15% to 30% were shown. Potential application scenarios for the introduction of the flexgrid technology in MAN (centralization of BRAS) and the viability of flexgrid in ultra-long haul networks were studied. Algorithms for the SERANO architecture proposed in D2.1 were proposed, and the design of cost-efficient architectures of networks with AoD OXCs was formalized. Power consumption reduction when the flexgrid technology is used in survivable multi-layer optical networks was quantified. Several energy reduction strategies using power-adaptive interfaces adjusting their power to the transmitted traffic were proposed.

The progresses on algorithms for flexgrid networks achieved in the second year of the project were presented afterwards. Regarding off-line algorithms, the improvements include introduction clique inequalities to strengthen mathematical formulations, a new algorithm for gradual network design, a new algorithm to solve the multicast routing



problem, and a new version of the multi-layer planning problem. As for online algorithms, the novel findings concern a new *reactive algorithm* for spectrum defragmentation and an updated version of the algorithm that solve the AFRO problem.

Some of the developed algorithms in WP1 were selected to be implemented and tested in WP4 by means of a flexgrid network control plane prototype. They include an algorithm for spectrum defragmentation, an algorithm for the AFRO problem, an algorithm to perform the Inter DC Multicast connectivity, and the KSP-DA-MSA heuristic distance adaptive multifiber spectrum assignment.

Finally, the results performed to the PLATON planning tool were presented.

8 APPENDIX: Supplemental Material

In this appendix we include detailed material related to the main part of the deliverable.

8.1 Impact of Adaptive Network Manager

8.1.1 Modeling the Field operations and Repair

The field operations and repair are a big part of the Operational Business (OBs) units expenses. Nowadays, there is a trend in outsourcing the operations and maintenance to third-party Service Providers that specialize in the maintenance of specific network technologies and supply spare parts. This way, the third-party operation and repair provider is able to leverage its field force across multiple carriers. This way, the different kinds of services that can be contracted are known, together with a categorization of the Service Level Agreements.

Whenever there is a failure in the network, it is detected by the Network Operation Center (NOC) and then, after a first diagnostic of the failure, it is decided which kind of technical support is needed and if spare parts are needed. The failures that have direct impact on the service need to be treated with extreme urgency. Also, problems that remove redundancy from the system and thus, let the network expose to business impact on a second failure, are considered urgent. As we will see, adding additional self-recovery capacities to the network by using resources of the optical layer in a dynamic way will reduce the severity of failures and thus, reduce the urgency to repair it and bringing the damaged part. In this analysis it will be shown which are the different steps in Service Levels of the hardware support and technical support, in order to determine how can reductions in severity of the failure be translated in OPEX reductions. Relaxing the SLA requirements is key to drive down the operational costs. In this subsection, the hardware repair and technical support models are explained together with a methodology to estimate the workforce needed to achieve a given Service Level guarantee. From the Multi-layer Use case analysis, the requirements in terms of Minimum Mean Time to Repair are obtained. With the knowledge of the number of failures per year, the minimum mean time to repair, it is possible to know which repair SLA is need and an estimate on the workforce needed to cope with such LSA. Using this methodology it will be possible to estimate the OPEX savings by IP-Optical integration in repair costs.



8.2 Hardware Support Model

According with each operative business requirement, two main repair models of hardware support to replace or repair damaged parts are considered: repair & return or swap. Independent of the repair model provided the Service Provider has to fulfill with a service level agreement (SLA). The prices of the repair services contracted by the Telefonica OBs depend on the SLA requested to the service provider Hardware Support: Repair and Return Model

Using the repair and return model, the service provider in charge of the hardware reparation receives the damaged spare part, repair it in their labs and return the part according the agreements and procedures established and SLAs. The service provider is responsible for the logistic (both process and costs) for the direct and reverse logistic between the operator central warehouse and the service provider repair labs. The SLAs also consider the case where the service provider does not repair in the same country where it has collected the damaged part.

Operators typically contract the repair & return services through two different methods: Flat Fee or Event Base (LPU) contract.

Flat Fee Model: There is a fixed monthly cost during the year, which includes any number of events that may exist in the month.

Event Base Model: The cost is computed by number of events multiplied by price per event. An event can be the failure of a network component (e.g. transponder, router card) or an installation.

8.3 Service Level Agreement (SLA) for the Repair & Return Model

The Service Level is accounted since the SP receives the damage part until the Prime receives a good one; this time it is called “turn around Time”

Table 32 SLA of the Repair & Return Model

Repair Time (Service Level)	Efficiency
15 working days	95 % in 15 days 100 % in 20 days
30 working days	95 % in 30 days 100 % in 35 days
45 working days	95 % in 45 days 100 % in 55 days
60 working days	98 % in 60 days 100% in 70 days
90 working days	98 % in 90 days 100% in 110 days



In the repair & return model, the damaged equipment is not repaired until the part arrives. Given the long times obtained in the repair & return model, it is not suitable for urgent repairs, but the price is much lower.

Hardware Support: Swap Model

In this case, the service provider directly exchanges the damaged equipment for new one, in a time according the agreements established and SLAs.

SLAs of the Swap Model

The Service Level is measured since the service since the notification of the failed part until the part is replaced. This time is called “swap exchange time”. Swap model is the typical one used for newly installed networks where the operator has no own stock, and needs the equipment to be quickly repaired.

Repair Time (Service Level)	Efficiency
8x5x4 To exchange to a good spare part in maximum 4 hours, within a request on business hours (8 daily) during Business Day (5 days per week)	95 % in 4 hours 100 % in 5 hours
24x7x4 To exchange to a good one in 4 hours always, so 24 hours, 7 days per week.	95 % in 4 hours 100 % in 5 hours
NBD To exchange to a good spare part in Next Business Day	98 % in NBD 100 % in 3 Business Days
7 Business Day To exchange to a good spare part in 7 Business Day	98 % in NBD 100 % in 3 Business Days

8.4 Technical Support Service Levels

Upon failure detection in the Network Operation Center, it is decided the urgency of the matter and, if needed, technical support is requested. To get a classification of the possible Service Level Agreements with the provider of the technical support, some parameters such as priority, response, recovery and resolution times and availability are defined. The operation teams define the levels of priority based on the severity of the problem. Thus, the problems with a higher severity are assigned a higher priority. There are three levels of priority defined in the Telefonica group: emergency, very urgent, urgent and ordinary, which are explained in Table 33.

Table 33 Severity of the problems

Severity Grade	Severity Definition
Priority/Severity 1 or EMERGENCY	Defects having a serious impact on the normal performance of the system, its operation,



	maintenance or administration, and/or that have a severe impact on the service provided to Operator's customers or to the Operator's ability to bill accurately.
Priority/Severity 2 or VERY URGENT	Defects that affect system performance, its operation or maintenance or administration, with intermittent or lesser degradation of the services provided to Operator's customers. Individual protocol Defects, sending interface Defects or communication Defects that cause a feature loss in the network.
Priority/Severity 3 or URGENT	A lower category function does not work properly or there has been a non-destructive error. The Defect has a reliable alternative solution and does not have an operating impact.

Severity 1 is related to emergencies. It includes major disturbances in the system functionality resulting in a 20% or more capacity decrease. Also, it considers serious reduction in the processing capacity of traffic in voice or data of more than 30 percent in some nodes. Another example of urgent issue is the complete loss of the network management system and input/output for business critical systems. A complete Exchange/ Node Failure is also considered as urgent. Finally, any Network failure resulting in the significant loss of customer service is treated as urgent.

Severity 2 is related to Very Urgent. It considers failures that cause a serious operational impact. For example, it can be related to a software Defect causing re-occurring service affecting problems for a large portion of customers. The loss of redundancy on any duplicated processing equipment is considered a very urgent issue.

Severity 3 is related to Urgent. It considers the loss of some functionality, with non-business impact. Typically, it includes failures with a minor impact on the performance of the system or parts of the system, or resulting in the intermittent failure of any function.

For each priority, the operator defines the necessary response, recovery and resolution times. Response time refers period of time from when a trouble was raised with the vendor to a technician contacting the operator. Recovery time is the period of time from when the trouble was raised with the vendor to the time that recovery is achieved. The resolution time is the period of time from when a trouble was raised with the vendor and ending on achievement of resolution.

Table 34 Technical Support Service Level Categories

Service level	Recovery Time of Severity 1	Recovery Time of Severity 2	Recovery Time of Severity 3
Iron	1 Week Day	1 Week	N/A



Bronze	14 hours	72 hours	30 days
Silver minus	8 hours	72 hours	11 week days
Silver	6 hours	48 hours	11 week days
Gold-minus	4 hours	24 hours	7 week days
Gold	1 hour	12 hours	7 week days
Platinum	15 minutes	6 hours	7 week days

Methodology to estimate Technical Support Workforce

A set of recovery and resolution times should be guaranteed by the Technical Support Vendor for several categories (from Iron to Platinum). The will charge the operator depending on the category (the tighter the requirements, the higher the price) and the estimation of failures per year that need to be solved.

In IDEALIST, we propose a model is to estimate the size of the Technical Support Workforce, in terms of number of repair teams, needed to guarantee a Minimum Mean Time To Repair (MMTTR). Each repair team has an associated monthly cost (salaries), and can give an estimate of the labour cost of repair of the communication network. The model assumes that upon each failure detected in the network, the failure should be repaired within the desired MMTTR. Note that the Minimum Mean Time To Repair will depend on the resiliency mechanism of the network. The model takes as input the number of components subject to failure, the mean time between failures (MTBF), as well as the mentioned MMTTR.

8.4.1 Elastic connectivity for Cost Optimization in Federated Datacenters

8.5 Cost Optimization in Federated Datacenters

Current inter-datacenter (DC) connections are configured as static big *fat pipes*, which entails large bitrate overprovisioning and thus high operational costs for DC operators. Besides, network operators cannot share such connections between customers, because DC traffic varies greatly over time. Those connections are mainly used to perform VM migration and DB synchronization among federated DCs, allowing thus elastic DC operations. To improve resource utilization and save costs, dynamic inter-DC connectivity is currently being targeted from a research point of view and in standardization form.

In this section, we show that dynamic connectivity is not enough to guarantee elastic inter-DC operations and might lead to poor performance provided that not enough overprovisioning of network resources is performed. To alleviate it to some extent a dynamic and elastic connectivity model taking advantage of elastic network resources allocation in flexgrid-based optical networks is proposed. The flexgrid technology enables finer spectrum granularity adaptation and the ability to dynamically increase and decrease the amount of optical resources assigned to connections.



Additionally, transfer mode requests for cloud operations are proposed. To provide an abstraction layer to the underlying network, a new stratum on top of the Adaptive Network Manager (e.g., ABNO), the carrier SDN, could be deployed, implementing a northbound interface to request transfer operations and using application-oriented semantic, liberating application developers from understanding and dealing with network specifics and complexity.

Cost savings considering both energy and communication costs are analysed in a DC federation when using the dynamic elastic connectivity and the network-driven models. This section is based on our work in [12], [13], [14], [15], and [16].

8.5.1 Dynamic connection and transfer mode requests

Evolution towards cloud-ready transport networks entails dynamically controlling network resources, considering cloud requests in the network configuration process. Hence, that evolution is based on elastic data and control planes, which can interact with multiple network technologies and cloud services. A cross-stratum orchestrator (CSO) between the cloud and the interconnection network is eventually required to coordinate resources in both strata in a coherent manner. When considering cloud-ready transport network, for the control plane to dynamically set up and tear down connections, the entry point from applications to the network is ABNO.

In contrast to static connectivity, dynamic connectivity allows DCs to manage optical connections to remote DCs, requesting connections as they are really needed to perform data transfers and releasing them when all data has been transferred. Furthermore, the fine spectral granularity and wide range of bitrates in elastic optical networks makes the actual bitrate of the optical connection closely fit connectivity needs. After requesting a connection and negotiating its capacity as a function of the current network availability, the resulting bandwidth can be used by scheduling algorithms to organize transferences.

Nonetheless, the availability of resources is not guaranteed, and the lack of network resources at request time may result in long transference times and even in transference period overlapping.

To alleviate to some extent the dependency between cloud management and network connectivity, in this section we propose a novel network-driven connectivity model. A carrier SDN controller implements a northbound interface to request transfer operations Figure 79. Those applications' operations are transformed into network connection requests. The northbound interface uses application-oriented semantic, liberating application developers from understanding and dealing with network specifics and complexity.

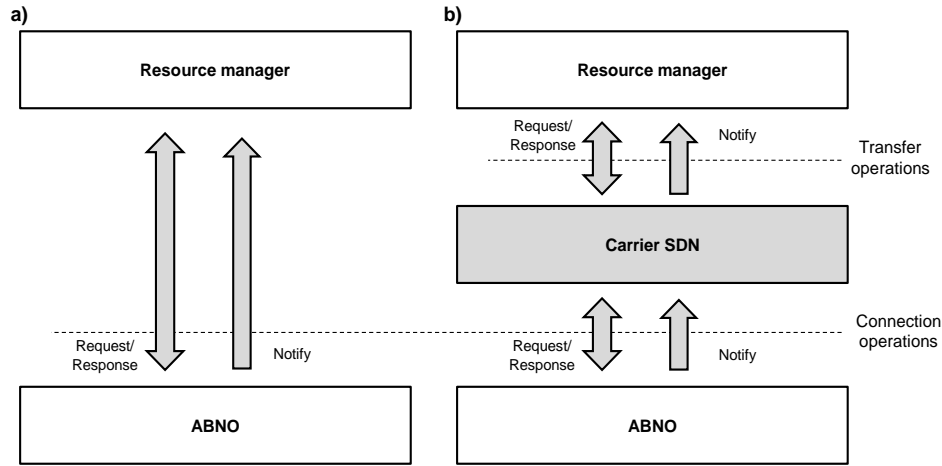


Figure 79: Control architectures supporting dynamic connections (a) and transfer mode (b) requests.

It is worth noting that the dynamic elastic model proposed before can also request connections to carrier SDN instead of the ABNO. In this case, the carrier SDN acts as a proxy between applications and network.

The SDN controller is in charge of managing inter-DC connectivity; if not enough resources are available at requesting time, notifications (similar to interruptions in computers) are sent from the ABNO to the SDN controller each time specific resources are released. Upon receiving a notification, the SDN controller takes decisions on whether to increase the bitrate associated to a transfer. Therefore, we have effectively moved from polling to a network-driven transfer mode. Figure 79 illustrates both control architectures supporting dynamic and transfer mode requests.

8.5.2 Minimizing energy expenditures

Considering a single DC, two main contributions to its power consumption can be distinguished: i) the power consumed by IT devices, P_{IT} , which comprises both the servers located in the DC as well as the switches employed to interconnect those servers; ii) the power consumption of the non-IT equipment, P_{non-IT} , such as cooling, power supplies and power distribution systems. Thus, total power consumption of a DC can be computed as $P_{DC} = P_{IT} + P_{non-IT}$. P_{IT} can be easily estimated by counting the number of servers and switches of a DC. However, it is difficult to evaluate the power consumption of non-IT devices since it depends on several details and factors which cannot be easily estimated. For instance, the power consumption of the cooling system strongly depends on the geographical location of the DC and on the building hosting that DC.

An indirect way to estimate a numerical value for P_{non-IT} is to consider the power usage effectiveness (PUE) metric [23]. PUE can be used as a measure of the energy efficiency of a DC and quantifies the amount of power consumed by non-IT equipment in that DC: $PUE = P_{DC} / P_{IT}$. Therefore, if P_{IT} and PUE can be estimated for a given DC, the total power consumed in a DC can be computed as $P_{DC} = PUE * P_{IT}$.



Regarding P_{IT} , we can distinguish between the power consumed by the servers and by network equipment. The power consumed by a server, $P_{server}(k)$, depends mainly on the CPU load (k) utilization, expressed as the ratio between the current load and the maximum capacity of the server. According to [20], the power consumption of a server can be estimated as $P_{server}(k) = P_{server-idle} + (P_{server-max} - P_{server-idle}) * k$, where $P_{server-idle}$ and $P_{server-max}$ represent the power consumed by the server when it is idle and when it operates at its maximum capacity, respectively. The power consumed by network equipment depends on the specific architecture of the DC.

Considering the fat-tree topology described before and assuming that clusters are active when one or more servers are loaded, otherwise the complete cluster is turned-off, the power consumption of cluster i , $P_{cluster}^i$, can be estimated as,

$$P_{cluster}^i = a^i \cdot \left(\frac{M}{2} \cdot (P_{agg} + P_{edge}) + \sum_{s=1}^{M^2/4} P_{server}(k_s^i) \right) \quad (9)$$

where M is the number of clusters in the DC topology, a^i indicates whether the cluster is active or not and P_{agg} and P_{edge} denote the power consumption of aggregation and edge switches. Then, the power consumption of the IT devices in the DC can eventually be computed as follows, where P_{core} denotes the power consumption of core switches,

$$P_{IT} = \frac{M^2}{4} \cdot P_{core} + \sum_{i=1}^M P_{cluster}^i \quad (10)$$

A first optimization to reduce energy expenditures is to perform consolidation, placing VMs so as to load servers as much as possible and switching off those servers that become unused. To further reduce energy consumption, consolidation can be performed by taking into account clusters structure, and switching on/off clusters as single units. Those servers in switched on clusters without assigned load remain active and ready to accommodate spikes in demand. In addition, as stated in the introduction, DC federations can perform elastic operations, migrating VMs among DCs aiming at minimizing operational costs by taking advantage from available green energy in some DCs and off-peak cheap brown energy in other while ensuring the desired QoE level, e.g latency experienced by the users of a service is used as a QoE measure. Then, federated DCs need to be orchestrated providing optimal VM placement so as to minimize operational costs. We assume that operational costs are dominated by energy and communication costs, so we focus on specifically minimizing those costs.

Let us assume that DCs are dimensioned to cover some proportion β_d of the total energy consumption for the maximum dimensioning. Then, green coverage in DC d , α_d , can be estimated as, $\alpha_d(t) = \beta_d * \delta_d(t)$, and the amount of green energy available can be estimated as $g_d(t) = \alpha_d(t) * Energy_MaxDimensioning$, where $Energy_MaxDimensioning$ represents the amount of energy consumed for the maximum dimensioning.

8.5.3 Performance evaluation

For evaluation purposes, we developed resource managers in an OpenNebula-based cloud middleware emulator. The federation orchestrator with the centralized scheduling algorithm was implemented as a stand-alone module in Java. Federated DCs are

connected to an ad-hoc event-driven simulator developed in OMNET++. The simulator implements the carrier SDN controller and the flexgrid network with an ABNO controller on the top, as illustrated in Figure 79. Finally, the algorithm described in [17] for elastic spectrum allocation was implemented.

We consider the global 11-location topology depicted in Figure 80. Each location collects user traffic towards the set of federated DCs, which consists of five DCs strategically located in Taiwan, India, Spain, and Illinois and California in the USA. A global telecom operator provides optical connectivity among DCs, which is based upon the flexgrid technology. The number of users in each location was computed considering Wikipedia's audience by regions [27] and it was scaled and distributed among the different locations in each region. Latency between location pairs was computed according to [31].

Table 35 briefly presents the value considered for some representative energy parameters. Daily PUE values were computed according to [22] using data obtained from [30]. Green energy coverage was obtained from [29], [30], and [26] and brown energy cost for each DC was estimated from their respective local electric company rates (e.g. [19] and [28]). Servers in DCs are assumed to be HP ProLiant DL580 G3 [24], equipped with four processors, 2 cores per processor, with $P_{\text{server-idle}} = 520\text{W}$ and $P_{\text{server-max}} = 833\text{W}$.

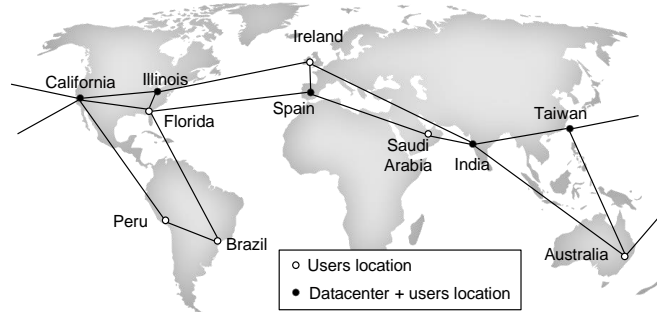


Figure 80: Scenario: federated DCs, locations and inter-DC network.

In line with [20], DCs are dimensioned assuming a fat-tree topology with a maximum of $M=48$ clusters with two levels of switches and $M2/4=576$ servers each. The number of VMs was set to 35,000, with individual image size of 5 GB; we assume that each VM runs in one single core. An integer number of clusters is always switched on, so as to support the load assigned to the DC; those servers without assigned load remain active and ready to accommodate spikes in demand. Green cover was set to ensure, at the highest green energy generation time, a proportion of energy β_d when all VMs run in DC d .

We consider a different type of switch, and thus a different power consumption value, for each layer of the intra-DC architecture. We selected the Huawei [25] CloudEngine switches series; Table 36 specifies the model, switching capacity and power consumption for each considered switch.

Table 35 Value of energy parameters

Datacenter	c_d (on/off peak) (€/kWh)	β_d	PUE (max/ avg)
Taiwan	0.0700 / 0.0490	0.5	1.671 / 1.632

India	0.0774 / 0.0542	0.9	1.694 / 1.694
Spain	0.1042 / 0.0729	0.9	1.670 / 1.457
Illinois	0.0735 / 0.0515	0.2	1.512 / 1.368
California	0.0988 / 0.0692	0.5	1.385 / 1.303

Table 36 Characteristics of Huawei CloudEngine switches

Layer	Model	Switching capacity	Power consumption
Core	12812	48 Tb/s	$P_{core} = 16,200$ W
Aggregation	6800	1.28 Tb/s	$P_{agg} = 270$ W
Edge	5800	336 Gb/s	$P_{edge} = 150$ W

Finally, we consider that each DC is connected to the flexgrid inter-DC network through a router equipped with 100 Gb/s bandwidth variable transponders. Therefore, the actual capacity of optical connections is limited to that value. To compute the real throughput, we consider headers for the different protocols, i.e. TCP, IP, and GbE. The maximum amount of bytes to transfer, was computed to guarantee that VM migration is performed in less than 40 minutes.

Finally, a dynamic network environment was simulated for the scenario under study, where background incoming connection requests arrive following a Poisson process and are sequentially served without prior knowledge of future incoming connection requests. Background traffic competes with the one generated by the federated DCs for network resources.

Figure 81 plots daily energy and communication costs as a function of the normalized background traffic intensity. We observe a clear increasing trend when the background traffic increases, as a consequence of connections' initial capacity decreases from 55 Gb/s to only 12 Gb/s on average. To try to increase that limited initial connections' capacity, elastic capacity increments need to be requested. The results obtained when each connectivity model is applied are however different.

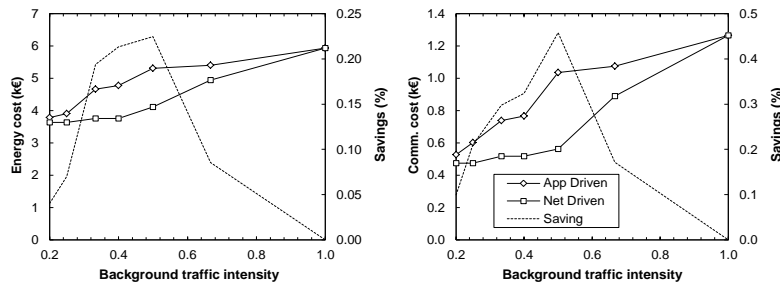


Figure 81: Daily energy cost (left) and communication cost (right).

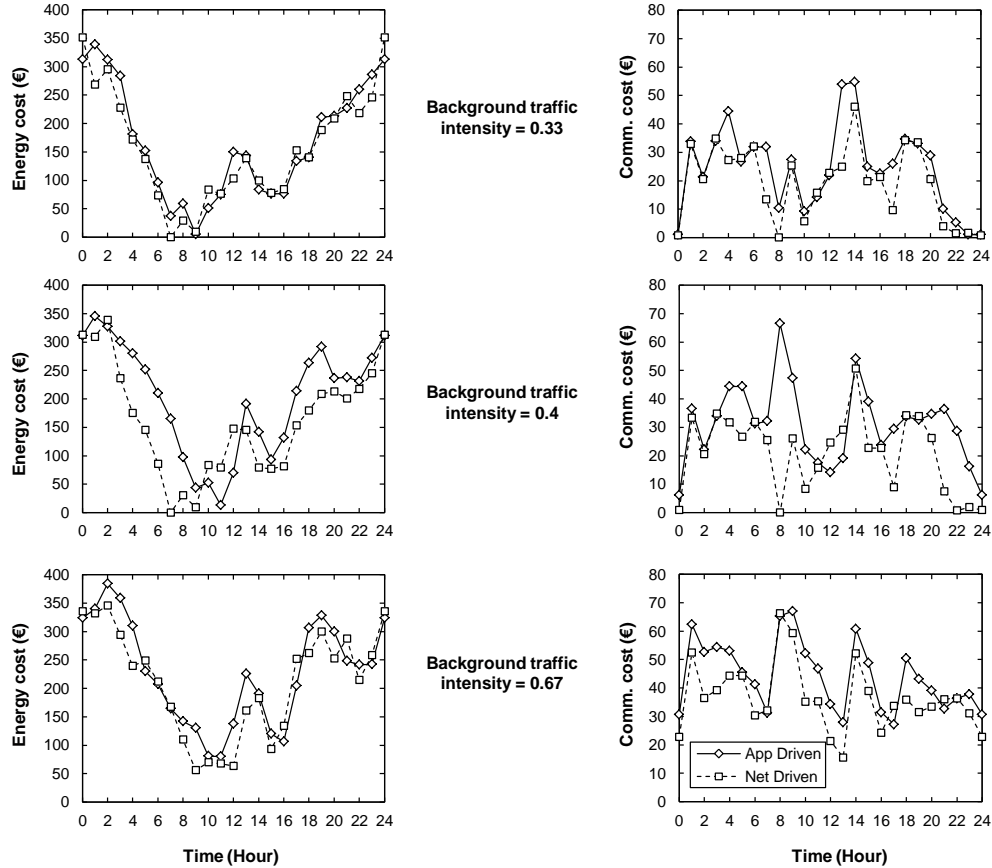


Figure 82: Hourly costs for several background traffic intensities.

Finally, Figure 82 illustrates hourly variation in the energy and communications costs when the application-driven and the network-driven models are applied, for three different background traffic loads. The behaviour of both models is basically the same and slight hourly energy cost savings can be appreciated, although they are clearly evident for the intermediate load. In contrast, there are some periods with a totally different behaviour between application-driven and the network-driven models, especially in the intermediate load. That is as a consequence of that VMs can be placed in those locations so as to minimize cost in the network-driven model so no new migrations are required, whereas massive migrations need to be performed in the application-driven model, which further increases communications needs.

8.5.4 Conclusions

A carrier SDN controller implementing a northbound interface with application-oriented semantic has been considered as a new abstraction layer between DC resource managers and the ABNO controller in the control plane of flexgrid-based interconnection networks. Each resource manager can request transfer operations specifying the destination DC, the amount of data to be transferred and the desired completion time.



That connectivity model, named network-driven, has been compared against the dynamic elastic model or application-driven, where the local resource managers are in charge of requesting connections directly to the ABNO controller. The application-driven model needs periodical retries requesting increase connection's bitrate, which do not translate into immediate bitrate increments and could have a negative impact on the performance of the network control plane.

Energy and communications costs and QoE on a DC federation were analysed. Some green energy is available in each of the locations as a function of the time, whilst the cost of brown energy shows differentiated on/off peak costs. A federation orchestrator computes periodically the global optimal placement for all the VMs in the federation so as to minimize operational costs whilst ensuring QoE.

From the results, we observed that when the network operates under low and medium traffic load costs savings as high as 20% and 40% in energy and communications, respectively can be obtained when the network-driven model is applied with respect to those of the application-driven. Besides, both connectivity models allow scheduling algorithms to provide the committed QoE.

8.6 Techno-economic and energy analysis

8.6.1 Idealist Building Blocks

8.6.1.1 Implications of variable baud rates in SBVT

An open point, that will be deepened in the last year of the Project by WP2, concerns the net and gross baud rate combinations better suited to modulate sub-carriers for traffic demands of $n \times 100 \text{ Gb/s}$ (with $1 \leq n \leq 12$), to optimize reach versus spectral efficiency. Up to date evaluations have highlighted that, considering technical feasibility in next few years, the net baud rates could range between 25 and 33.3 Gbaud/s, as shown in Table 37, where also the gross baud rates are reported, assuming a 25% FEC (other possible FEC choices could be 7% or 11%, depending on the reach we want to assure). It is worth noting that we have limited the minimum baud rate to the 25 Gaud/s value, to reduce a bit the complexity of the cards, already considered too large.

In Table 38 are reported the superchannel configurations in terms of number of carriers and net baud rates (in brackets) for the different modulation formats. For each modulation format the table shows, when available, the configuration with the minimum number of carriers and the configuration with the maximum number of carriers that match the required client bit rates (i.e. $n \times 100 \text{ Gb/s}$). A specific color code has been used for the 25 Gbaud/s (green), the 30 Gbaud/s (yellow) and the 33.333 Gbaud/s (orange) symbol rates, that all together match for almost all the required client bit rates. It can also be noticed that some client bit rates are achieved only with a sub-set of modulation formats and baud rates. For instance, the 100Gb/s can be made only with 1 carrier modulated with PM-QPSK at 25 Gbaud/s, while the 1000Gb/s can exploit all the range of modulation formats, each one in almost two ways in terms of carriers/baud rates configurations.

As already mentioned before regarding the WP2 studies on spectral efficiency w.r.t. to baud rates, a first evaluation has shown that some configurations present in Table 38 (the



ones highlighted with red color writing) could not be useful because, even if they minimize the number of carriers of the superchannel, they would not be advantageous in terms of bandwidth occupation of the flexible spectral grid as defined by ITU-T G.694.1. In particular, these configurations concern the 30,000 and the 33,333 Gbaud/s net symbol rates because, if we take into account all factors that influence the bandwidth requirements along a typical lightpath, we would often overcome the 50 GHz spectrum width for each sub-carrier and we would have a number of frequency slots at 12,5 GHz that is at least the same but often more than the number of slots needed with the 25 Gbaud/s net symbol rate. The elements that should be considered for this calculation involve the channel bandwidth (composed considering the gross baud rate that in these cases amounts to 37,500 and 41,666 Gbaud/s, increased by the roll-off factor), the cumulative filtering factor (for a given OSNR penalty it depends on the number of crossed WSS filters), the laser center frequency stability and the filter stability (that both amount to about 1 GHz). All the details of these calculations will be given in the next WP2 deliverables.

Table 37: Net and Gross baud rates required for the configurations displayed in Table 38 and assuming a 25% FEC overhead

Net Baud rate (Gbaud/s)	Gross Baud rate (Gbaud/s)
25.000	31.250
26.667	33.333
27.500	34.375
27.778	34.722
28.125	35.156
29.167	36.458
30.000	37.500
30.556	38.194
31.250	39.063
33.333	41.666



Table 38: Number of carriers and Net Baud rate (in brackets, expressed in Gbaud/s) required in a SBVT for different client bit rates multiple of 100G

Client bit rate (Gb/s)	PM QPSK		PM 8QAM		PM 16QAM		PM 32QAM	
	max#carr.	min #carr.	max#carr.	min#carr.	max#carr.	min#carr.	max#carr.	min#carr.
100	1 (25.000)	-	-	-	-	-	-	-
200	2 (25.000)	-	1 (33.333)	-	1 (25.000)	-	-	-
300	3 (25.000)	-	2 (25.000)	-	-	-	1 (30.000)	-
400	4 (25.000)	3 (33.333)	2 (33.333)	-	2 (25.000)	-	-	-
500	5 (25.000)	4 (31.250)	3 (27.778)	-	2 (31.250)	-	2 (25.000)	-
600	6 (25.000)	5 (30.000)	4 (25.000)	3 (33.333)	3 (25.000)	-	2 (30.000)	-
700	7 (25.000)	6 (29.167)	4 (29.167)	-	3 (29.167)	-	-	-
800	8 (25.000)	6 (33.333)	5 (26.667)	4 (33.333)	4 (25.000)	3 (33.333)	3 (26.667)	-
900	9 (25.000)	7 (32.143)	6 (25.000)	5 (30.000)	4 (28.125)	-	3 (30.000)	-
1000	10 (25.000)	8 (31.250)	6 (27.778)	5 (33.333)	5 (25.000)	4 (31.250)	4 (25.000)	3 (33.333)
1100	11 (25.000)	9 (30.556)	7 (26.190)	6 (30.556)	5 (27.500)	-	4 (27.500)	-
1200	12 (25.000)	9 (33.333)	8 (25.000)	6 (33.333)	6 (25.000)	5 (30.000)	4 (30.000)	-

8.6.1.2 State of the art flexgrid optical node

An example of state of the art flexgrid ROADM is depicted in Figure 83, where 1x9 WSS filters and 8x16 multicast switches, both with flex-grid support, are considered.

A colorless add/drop (A/D C-less) architecture is envisaged to be implemented by two 1xn WSS filters (one for add and one for drop direction). This solution is useful for permanent connections that don't require spatial reconfiguration flexibility. Few ports are supposed to be sufficient for these application thanks also to the already mentioned aggregation capacity inside the SBVTs.

The colorless and directionless add/drop (A/D CD-less) architecture is based on two 1xn WSS in a back-to-back configuration for both add and drop directions, followed by a set of splitter/combiners to increase the number of ports. This configuration keeps both spatial and spectral flexibility, at the prize of a greater complexity.

The colorless, directionless and contentionless add/drop (A/D CDC-less) architecture is presently implemented with a Multicast Switch (for example an 8x16 ports switch), that allows the replication of portions of the optical spectrum on all the add/drop ports and the assignment of this portions to a specific line port direction. With this solution, a limitation is imposed by the SBVT design that, as already mentioned, is supposed to have only one port for each Multiflow Optical Module (i.e. only 1 add/drop port towards/from the Multicast switch). In fact, with this choice, the contentionless capability (that is the possibility to make lambda replication) would be available only between different Multiflow Optical Modules inside the SBVT as well as between different SBVTs. However, this could not be a problem because it is expected that, in many cases, each module will support traffic from only one single client flow with a high bandwidth requirements on a specific path within the network, thus not requiring color reuse within the module. In any case, if contentionless

capability is needed for some specific demands, the A/D CD-less solution with back-to-back WSS could always be applied for other SBVT cards.

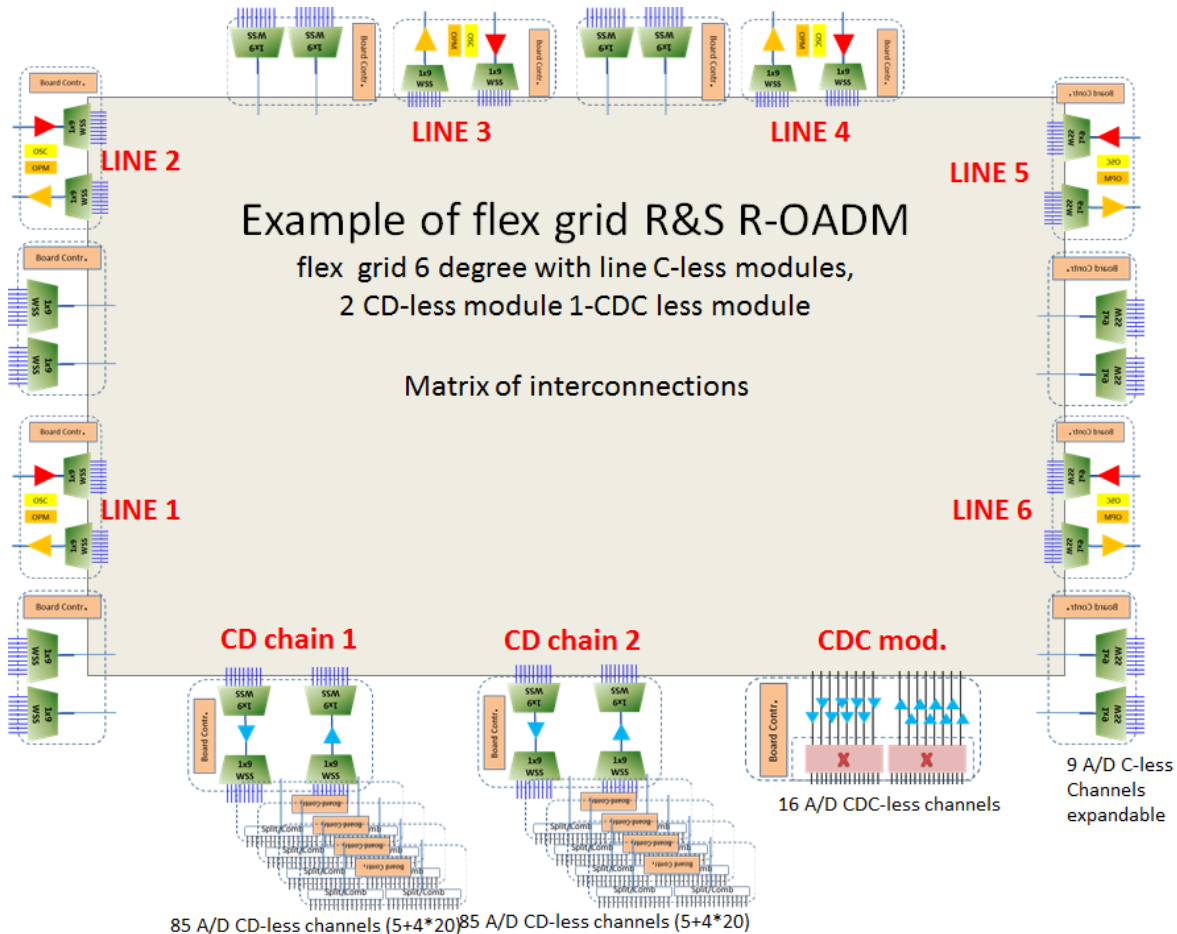


Figure 83 Scheme example of Flex-grid R&S ROADM

8.6.2 CAPEX analysis of a survivable multi-layer IP-over-Flexgrid network

In this section we introduce a mixed-integer optimization model for two-layer IP-over-Flexgrid networks, and present a cost study performed using the CPLEX package for the introduced model. The study concerns a generic German backbone network with 12 nodes and 20 links in the optical layer, and a set of six IP traffic matrices with increasing demand volumes. The study has been presented first in [101]. The optimization assumes a hot-standby mechanism in the optical layer to make the network resilient to cable cuts. The dominant cost drivers are bandwidth-variable transponders (BVTs). The numerical results reveal that due to the huge spectral capacity of a single fiber, the spectrum allocation problem simplifies significantly in practice. However, the routing problem minimizing the total cost of employed transponders becomes a challenge, but when appropriately approached and solved, can bring reductions in costs of up to 10%.



8.6.2.1 Optimization model

The Multi-Layer Restoration (MLR) optimization problem studied here is referred to as a two-layer problem (TLP). It encompasses two network layers: the lower layer that models the optical EON layer, and the upper layer modelling the IP layer. It is assumed that only the optical layer is subject to resource failures. Therefore, in this study we concentrate on failures of the optical layer links only. Moreover, each optical layer link is identified with the corresponding enlightened fibre (and, for that matter, with its inline amplifiers, connectors, etc.). Hence, in the context of MLR the lower layer link always means the fibre, and the lower layer link failure – the fibre cut.

When a failure occurs in the optical layer, the routers apply classical IP protection mechanisms for quasi-immediate recovery of high-priority IP traffic. Furthermore, after optical mechanisms have successfully recovered the lower layer (with a short time delay of up to 1-2 minutes), the remaining part of traffic, the so called best-effort traffic, is transported through the network again. Certainly the IP layer gets aware of the optical layer failure by correlation of failure patterns and exchanging status information with the optical layer.

However, for the cost study we can neglect the transition time until final optical recovery, as we focus on principles of restorability disregarding the convergence interval. Thus, we assume that the IP layer does not experience any failures. This is because the optical layer path flows that realize the IP layer links affected by the fibre cuts are immediately restored by means of a path restoration mechanism. Thus, for the modelling facilitating purposes we assume that the optical layer link failures do not propagate to the IP layer, and hence the IP links are fully protected in the optical layer.

More precisely, the capacity of each IP layer link is realized by means of a selected (optimized) primary lightpath in the optical layer. When such a primary path fails due to a failure of one (or more) of its fibres, the entire bandwidth realized on the primary path in the nominal (*i.e.*, failure-less) state is fully restored on a backup path that does not traverse the failed fibres. In general the selected backup path depends on the particular failure affecting the primary path. This protection/restoration mechanism is referred to as state-dependent backup path restoration. The name “state” in this context refers to the failure state. Each failure state is identified with a subset of fibres that are cut simultaneously. Such a subset of failing fibres is sometimes referred to as SRLG (shared-risk link group). Note that an important case of a failure scenario is the so called single link failure scenario, where we assume that each link can fail but only one at a time.

External traffic demand is offered to the IP layer of the considered two-layer network. Each demand involves two end nodes (routers). Traffic (expressed in bit-rate) to be carried between these two end nodes is realized by means of a path flow established on a selected route connecting these nodes in the IP layer graph. Once such a path flow is established for each demand, the resulting IP link loads can be calculated. Then, these link loads are treated as the IP layer link capacities, and constitute the traffic demand offered to the optical layer. For each IP layer link, its load is realized in the optical layer by means of a lightpath realized on a route (connecting the end nodes of the link in question), as already mentioned above. This makes optimization of the two considered layers dependent on each other, and makes it necessary to treat them as one joint optimization TLP problem described above.

A lightpath in the optical layer is described by two elements: the route (*i.e.*, the path in the optical layer graph) and the frequency interval (*i.e.*, a part of the spectrum assigned and dedicated to the lightpath). Thus, when a primary lightpath fails, then it is restored on a state-dependent backup lightpath that follows another route and may use another frequency interval. Consequently, to realize a lightpath and all its backup lightpaths, we need to provide two tuneable transponders at the lightpath end nodes. The cost of these transponders is taken as the cost function in the considered optimization problem.

Certainly, in the nominal state, each fibre of each optical link must be sufficient to realize all the (non-overlapping) frequency intervals assigned to the primary lightpaths that traverse the link. Also, in each failure state, all the working lightpaths and the active backup lightpaths that traverse a given non-failed optical link must in total have the frequency intervals non-overlapping.

To summarize, the considered optimization problem can be characterized as a two-layer flow allocation problem with state-dependent flow restoration in the optical layer, minimizing the cost of the transponders.

8.6.2.2 Transmission model

We assume that the optical network operates with software-defined BVTs which implement the time-frequency packing (TFP) transmission technique. Apart from achieving very high spectral efficiency (SE), TFP allows to adapt the bit-rate of transmitted data in a function of the transmission path quality. As a result, for higher quality paths, *i.e.*, having higher optical-signal-to-noise ratios (OSNRs), higher SE can be achieved than for lower-quality paths.

For estimating link OSNRs (denoted as $\sigma(e), e \in E$), we apply an approximate LOGON model that was proposed for Flexgrid EONs in [102] LOGON assumes a worst-case analysis of link OSNR by considering that the spectrum in the link is fully loaded. Such assumption simplifies significantly network design and control, since for a given network link OSNRs can be pre-computed and they are fixed and do not depend on the network state. Therefore, using linear values of link OSNRs, the OSNR of routing path p (both linear and in dB) is calculated as

$$\sigma(p) = \left(\sum_{e \in p} (\sigma(e))^{-1} \right)^{-1}, \sigma^{dB}(p) = 10 \log_{10}(\sigma(p)). \quad (11)$$

As mentioned above, the spectral efficiency of routing path p (denoted as $\varphi(p)$) is a function of path OSNR that we approximated using results of [103].

Table 39: Notation used in the mathematical formulation.

A – set of potential IP links	$b(d)$ – volume of demand d	$B(a, t)$ – bandwidth of transponder t on link a
D – set of upper layer demands	$O(a)$ – paths using link a	z_{at} – number of transponders t on link a
T – set of transponders	$\kappa(t)$ – cost of transponder t	y_p – flow on path p
$P(d)$ – paths for demand d		

8.6.2.3 Optimization procedure

As mentioned earlier, the complexity of the considered problem arises mostly from the need of simultaneous optimization of both the upper and the lower layers. Fortunately, in practical settings the tremendous available spectrum of an optical fibre reaching 4 THz allows for a simultaneous transmission of all the traffic in most national networks on a single fibre (more than 30 Tb/s can be reached on a single link on average in the tested German reference backbone network using all available spectrum). This assumption allows for an efficient decomposition of the considered problem into two phases. In the first phase protected optical paths between all pairs of nodes are calculated, while in the second phase a subset of those paths is selected to form a set of IP links. In practical setting the selected set of IP links is small enough that all of them can be realized using different parts of the available spectrum. This fact basically negates a sense of considering spectrum allocation problem in contemporary optical networks.

Phase I: Modules and Costs for the IP Links

In the first phase protected paths for each pair of nodes are computed. The paths will be used as potential links in the second phase of the optimization procedure. The paths and their protection schemes are selected in such a way that their respective OSNRs are maximized for the worst possible failure scenarios. Having an OSNR of each link $\sigma(p)$ pre-computed, we can take advantage of (11) and find a path with the lowest aggregated OSNR using any shortest path algorithm by taking $1/\sigma(p)$ as a weight of link e . In our problem single failures of all optical links can occur, thus for each link used in the shortest path we have to find a backup path that is bypassing the failed link and is the shortest in terms of the link weights. The worst, *i.e.*, with the lowest OSNR, path among all the backup paths defines the worst conditions that can be experienced on a potential IP link joining the considered pair of nodes. Those worst conditions for each potential IP link define the bandwidths for each available transponder used on each link $B(a, t)$, which are the input for the second phase of the optimization process.

Phase II: Optimizing IP Layer

The second phase takes advantage of the mathematical programming and requires a notation which is gathered in Table 39. The integer programming formulation of the considered problem can be expressed as follows:

$$\min 2 \sum_{a \in A} \sum_{t \in T} \kappa(t) z_{at} \quad (12)$$

$$\sum_{p \in P(d)} y_p = b(d), d \in D \quad (13)$$

$$\sum_{t \in T} z_{at} \leq 1, a \in A \quad (14)$$

$$\sum_{p \in O(a)} y_p \leq \sum_{t \in T} B(a, t) z_{at}, a \in A \quad (15)$$

In the above model, objective (12) is to minimize total cost of transponders such that all demands are satisfied (13), there is no more than one transponder at each link (14), and installed transponders suffice to realized the requested demands (15).

8.6.2.4 Numerical results

The above described optimization procedure has been used to evaluate how the increasing total traffic volume impacts the total bandwidth-variable transponder costs for DT's 12-node backbone network.

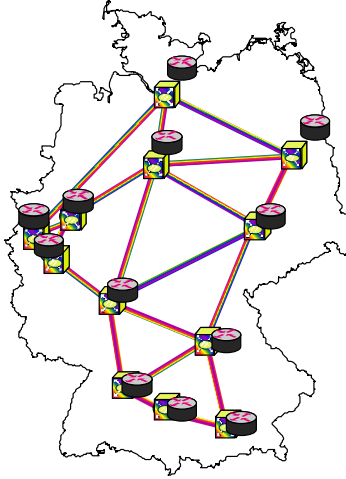


Figure 84: Generic 12-node German backbone network.

We used the 12-node generic network presented in Figure 84 with six traffic demand matrices that resemble changes in traffic demand volumes over time, ranging from 1.4 Tb/s to 6.3 Tb/s on all links in total. The set of transponder types (denoted as T) is artificially generated and consists of three elements which are characterized by different width of supported spectrum $S(t) \in \{50, 100, 150\}$ GHz and different costs $\kappa(t) \in \{10, 15, 17\}$.

Table 40: Summary of network optimization results.

Solution	Total traffic	IP-link topology	Required number of transponder pairs	Cost
1	1.4 Tb/s	center-rooted star	11 × 50 GHz	220
2	2.6 Tb/s	a tree, close to the center-rooted star, but with one additional 3-hop path	11 × 50 GHz	220
3	2.6 Tb/s	center-rooted star	10 × 50 GHz + 1 × 100 GHz	230
4	6.3 Tb/s	forced center-rooted star is infeasible	the transponder bandwidth on one of the links exceeds 150 GHz	–
5	6.3 Tb/s	center-rooted star & additional link	center-rooted star: 2 × 50 GHz + 5 × 100 GHz + 4 × 150 GHz additional link: 1 × 50 GHz	346
6	6.3 Tb/s	center-rooted star (partial) & 5 additional links	center-rooted star: 1 × 50 GHz + 3 × 100 GHz + 3 × 150 GHz additional links: 5 × 50 GHz	312



Our experiments show that the light traffic of 1.4 Tb/s could be easily accommodated in the network using a simple centre-rooted star topology (with Frankfurt as the hub node, being slightly in the south-west of gravity centre of Germany) realized using the cheapest transponders resulting in the total cost of 220 (11 pairs of the cheapest transponders of cost 50; see Solution 1 in Table 40). Taking the increased traffic (2.6 Tb/s) into account, still the theoretically cheapest solution of cost 220 is obtainable (Solution 2). However, this time it cannot simply be the centre-rooted star (Solution 3) – additional measures have to be taken to handle some high traffic volume relations. After increasing the volumes again the cheapest solutions become infeasible and the optimization method of the previous subsection has to be used to its full extent. Here we will analyse an example with the total traffic of 6.3 Tb/s. This way we will present a number of different approaches to handle the increased complexity of the problem.

The most straightforward, however definitely not the simplest one from the computational complexity point of view, is to directly apply the optimization model of the previous subsection taking into account all possible potential IP links and all possible paths (taking all possible paths into account can be easily implemented by switching from the presented link-path formulation to the node-link formulation). The approach proved to be too time consuming and after 10 minutes of computations (CPLEX 12.5 run on a 4-core 2.4 GHz CPU) returned a feasible solution of cost 330 with no perspectives for ending the optimization process within foreseeable future. Therefore, we decided to test different approaches limiting the set of available paths or the set of potential IP links.

First we limited the set of available paths to those using solely the centre-rooted star. However, with the 6.3 Tb/s traffic, such a solution is infeasible due to the huge demanded volume requested to one of the nodes (Solution 4). Another option was to limit the set of available paths less severely leaving, apart from the centre-rooted star paths, a path consisting of a direct connection for each relation (Solution 5). Such an approach resulted in an immediately obtained solution of cost 346. Unfortunately the solution was worse than the best solution obtained so far. Moreover, its quality did not improve even when more two-hop paths were taken into account for each demand. Therefore, we decided to switch to the node-link formulation and concentrate on limiting the set of potential IP links to handle the computational complexity issues. The approach proved to be very efficient allowing us to obtain, when solely 30 the most loaded potential IP links were considered, after 10 minutes of computations a solution of cost 312 (Solution 6).

8.6.2.5 Conclusion from survivable ML IP-over-Flexgrid network analysis

We have presented a two-phase approach to optimizing total bandwidth-variable transponder costs in IP-over-Flexgrid networks. We showed that the problem can be decomposed in practice into a problem of finding protected paths for each potential IP link in the optical layer in the first phase and a problem of selecting a subset of potential IP links in the second phase. Therefore the complete problem can easily be solved to optimality for light traffic cases. On the other hand, also considered medium traffic cases are difficult to solve to the proved optimality because of excessive number of binary variables. However, the obtained results are convincing and most probably optimal or are close to optimum. Finally, high traffic cases (not studied here) will require adding the spectrum allocation constraints, which will make the exact solution much more difficult to achieve.

8.6.3 BRAS Centralization network design problem

We are given a regional optical mesh network $G=(V,E)$, where V is the set of ROADM nodes (WDM or flexgrid) and E is the set of single-fiber links. We assume that we know the lengths of the links, and we denote the length of link $e \in E$ by l_e . The network has to support traffic generated by a set M of MTU switches that are lower in the hierarchy of the network and are connected to the regional optical network. Each MTU switch is connected to two different ROADM nodes (the way that the MTU switch is connected to these two nodes changes according to the different scenarios, but when we focus on the optical regional network this does not play a significant role).

To formulate the above problem, we assume the general case where an MTU switch is connected to the ROADM node $n \in V$ (this is the original location of the switch) and for backup it can be connected to one from a set BS_n of backup ROADMs sources, $|BS_n| \geq 1$ (note that the case where we are given a single backup source is a subcase, also note that we assume that all MTU switches connected to ROADM n have the same set of backup sources BS_n). In turn the two ROADM nodes, the source and the chosen backup source, are connected via two lightpaths (all-optically) to two BRAS nodes, and the chosen paths have to be node-disjoint for protection purposes. The traffic is destined to a set of ROADMs where the BRAS servers are located, we will call these ROADMs the “BRAS nodes” and denote their set by $D \subseteq V$, where $|D|$ is the number of BRAS nodes and $|D| \geq 2$ (the case where we have exactly 2 BRAS nodes is a subcase). Each connection is C Gb/s, and in particular in the examined case $C=10$ Gb/s, the spectrum occupied and the reach depends on the network architecture and the transceivers/transponders used in each scenario.

Key routing algorithmic idea: Transformation of the connection between an MTU switch and the BRAS nodes to the node-disjoint routing problem.

The traffic generated by an MTU switch enters the optical regional network via two ROADM nodes. Then traffic is forwarded via two transparent node-disjoint optical connections (lightpaths) over the regional optical mesh network to two BRAS nodes. This connection/routing design is done for protection purposes. Assume the general case where an MTU switch is connected to the ROADM source node n and for backup it can be connected to one from a set BS_n of backup ROADMs sources, and that we are given a set of BRAS nodes D . Our goal is to select the backup source from the set BS_n , select the two BRAS nodes from the set D that we are going to direct the connections and also find the paths so that they are node-disjoint. Figure 85a shows the protected BRAS connection problem for a single MTU switch.

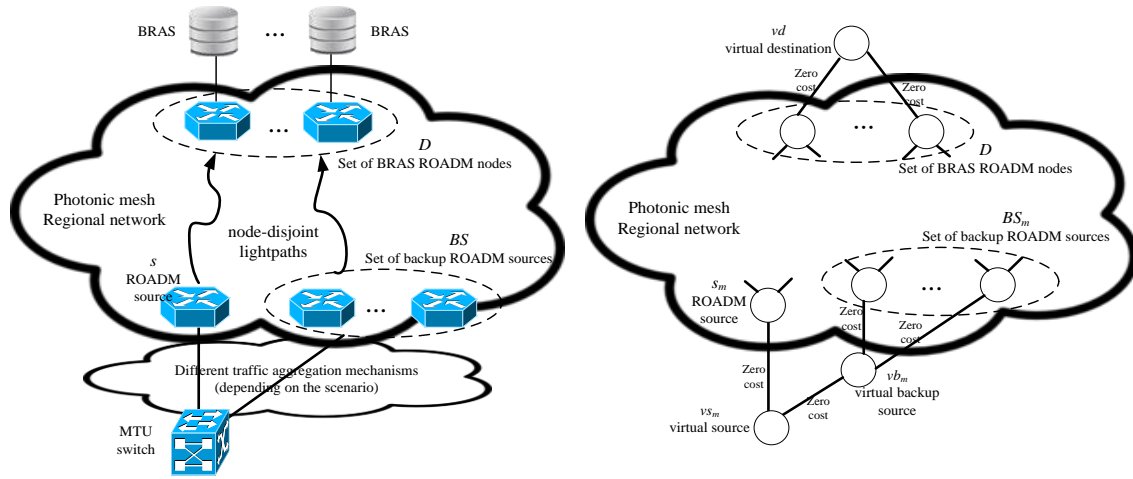


Figure 85: (a) Protected BRAS connection scenario over the regional optical mesh network. (b) Transformed problem where we need to find two node-disjoint paths from virtual source vs to virtual destination vd .

To solve this problem we transform it into a node-disjoint routing problem.

We create an expanded graph $G'=(V',E')$ that we initialize equal to the graph G that represents the regional optical mesh network. We make the following additions to G' . We create three virtual nodes: the virtual MTU switch source vs_n , the virtual backup source vb_n and the virtual destination vd . The virtual MTU source vs_n is connected to the source ROADM node n and the virtual backup source vb_n with zero cost links. The virtual backup source vb_n is connected to each ROADM node in the set BS_n of backup sources and to the virtual MTU source vs_n with zero cost links. The virtual destination vd is connected to each BRAS node in the set D with zero cost links. The cost of the links of the regional optical mesh network is taken to be their length. The problem of selecting the backup source, the two BRAS nodes, and finding node-disjoint lightpaths is now transformed to finding two disjoint paths from virtual source vs_n to virtual destination vd over the extended graph G' . This equivalent problem is shown in Figure 85b.

Combinatorial optimization: We described above how to transform and formulate the establishment of connections between a single MTU switch connected to a specific ROADM and the BRAS nodes. To formulate the combinatorial problem for all MTU switches we need for each switch that are connected to different ROADM nodes $n \in V$ to create the corresponding virtual MTU source vs_n and the virtual backup source vb_n and connect them appropriately. The virtual destination vd node is created once and is the same for all MTU switches. The combinatorial protected-BRAS routing problem is thus transformed to the combinatorial routing problem between the virtual MTU sources and the virtual destination vd .

Formulating the various types of transponders:

In the general case we can have a single or several different optical transponders that are tunable or not and can support a single or several connections (sliceable). To describe all these options we describe a feasible transmission by a (reach l , hops h , rate r , bandwidth



b , guard-band g) tuple. A transponder is then described by a (cost c , max connections mt) capability tuple and a set of feasible (l, h, r, b, g) transmission tuples that it can support. The constraint is that the transponder can support up to mt different connections (for sliceable $mt > 1$ for non-sliceable $mt = 1$), each which has to be described by a transmission tuple. This definition for the transponders is an extension of the definition used in [68], so as to take into account the sliceability of multi-flow transponders. The term feasible has to do with the physical layer feasibility and the transmission tuples are defined so as to take into account the capabilities of the transponders and physical layer impairments.

Since we are considering 10 Gb/s connections and a regional topology with diameter (longest possible path) less than 1000 km, in case of a WDM network we assume a single type of 10 Gb/s transponder, that has reach much higher than the diameter of the network. So the WDM network is described by a single transponder with ($c = cost_{10Gb/s-tx-idealists}$, $mt = 1$ connection) capability tuple and by a single $(l = \text{inf}, h = \text{inf}, r = 10 \text{ Gb/s}, b = 50 \text{ GHz}, g = 0 \text{ GHz})$ transmission tuple. The bandwidth variable transponder of [69] can be described by transmission tuples $(l = 195 \text{ km}, h = 3, r = 5 \text{ Gb/s}, b = 12.5 \text{ GHz}, g = 0 \text{ GHz})$, $(l = 100 \text{ km}, h = 2, r = 8 \text{ Gb/s}, b = 12.5 \text{ GHz}, g = 0 \text{ GHz})$, ... $(l = 60 \text{ km}, h = 2, r = 10 \text{ Gb/s}, b = 12.5 \text{ GHz}, g = 0 \text{ GHz})$. A sliceable BVT could support e.g. $mt = 4$ such connections.

Routing/Design Problem Variations

-Problem Variation 1: known BRAS nodes, minimize the max path length

Given: graph $G = (V, E)$ describing the regional optical network, the BRAS nodes placement D , the number of examined paths k (for the heuristic), the number of wavelengths (or slots) supported in the system W , the number of primary connections c_n needed to be established from each node $n \in V$ (this depends on the number of MTU switches under ROADM n and the optical regional network architecture – that is how the connections from the MTU switches are aggregated to enter the optical regional network).

We construct the expanded graph $G' = (V', E')$, where V' consists of the set V of ROADMs of the regional network and the virtual primary source vs_n , virtual backup source vb_n for each node $n \in V$, and virtual destination vd . Also E' consists of the set of edges E of the regional optical network, and zero weight edges between vs_n and vb_n , between all nodes in BS_n and vb_n , and between all nodes in D and vd . We denote by d_l the length of a link in E' .

In what follows we present an ILP and a heuristic algorithm to solve this problem variation.

ILP formulation

x_{ln} : Boolean variable equal to 1 if link l is used by the primary path from node n to the BRAS

y_{ln} : Boolean variable equal to 1 if link l is used by the backup path from node n to the BRAS

w : the worst path length of all the primary and backup paths

Min: w

- Flow conservation of primary paths



$$\text{For all } n \in V, \sum_{l \in V' \& l \in m+} x_{ln} - \sum_{l \in V' \& l \in m+} x_{ln} = \begin{cases} 1, & m = vs_n \\ -1, & m = d \\ 0, & m \neq n, d \end{cases}$$

- Flow conservation of backup paths

$$\text{For all } n \in V, \sum_{l \in V' \& l \in m+} y_{ln} - \sum_{l \in V' \& l \in m+} y_{ln} = \begin{cases} 1, & m = vs_n \\ -1, & m = d \\ 0, & m \neq n, d \end{cases}$$

- Node disjoint paths:

$$\text{For all } n \in V \text{ and all } m \in V' - \{vs_n, d\}, \sum_{l \in V' \& l \in m-} x_{ln} + \sum_{l \in V' \& l \in m-} y_{ln} \leq 1$$

- Worst path length

$$\text{For all } n \in V, w \leq \sum_{l \in V'} d_l \cdot x_{ln}, w \leq \sum_{l \in V'} d_l \cdot y_{ln}$$

- Spectrum utilization constraint

$$\text{For all } l \in E, \sum_n c_n \cdot (x_{ln} + y_{ln}) \leq W$$

Note that wavelength/slot allocation is considered easy and is not optimized. In general in the scenarios that we consider where traffic is directed to the BRAS nodes and apart from that there is no background traffic between other nodes, the links attached to the BRAS nodes become the bottleneck. There are few optimization options at these links. The wavelength/slot allocation needs to ensure that the links support the total number of wavelengths/slots that pass over them, so the number and not the actual wavelength/slot allocation matters more, in most cases.

Heuristic

For each MTU source we find two node-disjoint paths from source vs_n to destination vd . To do so we find the k -shortest paths from vs_n to vd (this are the primary paths) and then for each of these primary paths we remove the nodes of the path except for vs_n and vd and find the *new* shortest path between vs_n and vd . So, for n we calculate a set P_n that includes k primary and backup path pairs. We pass this to the RWA or RSA algorithm. The algorithm keeps an utilization vector for each link of the optical mesh network. The algorithm serves the MTU switches one by one in a particular ordering. For each switch, let n be its source ROADM. The algorithm examines the k primary-backup path pairs from the set P_n . For each of the k pairs it checks if it can allocate spectrum to both primary and backup paths. Finally, for the routable pairs (for those with available spectrum) it selects the primary-backup pair whose maximum length is the lowest.



-Problem Variation 2: given the maximum reach, find the BRAS nodes

The placement of BRAS nodes D is not given and has to be decided. What is given is the maximum transmission reach of the transponder(s), and based on that we calculate the number and best placement D of the BRAS nodes.

Given: graph G , number of examined paths k (for the heuristic), number of wavelengths (or slots) supported in the system W , maximum transmission reach L , number of connections c_n needed to be established from each node n .

We have developed an ILP algorithm and a heuristic algorithm to solve this problem. Since the ILP formulation to solve this problem variation is very similar to the one presented above for problem variation 1 (we add constraints to limit the length of the connections, constraints to count the number of BRAS nodes that are used and we change the objective function to minimize the number of utilized BRAS nodes), we skip it here and outline the heuristic algorithm.

Heuristic: we start with 2 BRAS nodes. For each combination of BRAS nodes (that is for each combination of 2 nodes in this case), we run heuristic p1 (solve the min-max length problem as described in problem variation 1). To speed up the process we do not search all source nodes, but we stop and search the next combination of BRAS nodes if we find a node whose primary or backup path has length higher than the given reach L . If we find a combination of BRAS nodes whose min-max length is lower than the given reach L we stop and select this placement. If we search all 2 BRAS nodes combination and none is successful we go to 3 nodes and repeat, and so on, until we find the first placement that satisfies the reach constraint.

8.6.4 Complete version of the study on viability of flexgrid in long distance networks

This section include the complete version of the study on long distance networks presented in section 4.4.2.

This study analyses the introduction of flexgrid technologies in the specific context of a Pan European Network (one of the reference Networks defined within IDEALIST project, which is a long haul network with a diameter of about 6000 km) and in presence of static or semi-static traffic demand. By means of a comparison between a reference network solution which employs fixgrid technology (transponders which require a fixed and assigned optical bandwidth (50 GHz) and fixgrid node) and a solution which uses SBVT and flexgrid nodes, some indications are achieved in terms of the costs that would be required for flexgrid equipment to make this solution economically viable.

The study encompasses many aspects involved in the optical network planning, design and cost evaluation, with particular reference to typical characteristics of the long haul context.

Nodes architecture are based on existing or declared as soon commercially available equipment; in particular, in addition to well-known fixgrid nodes, regarding flexgrid equipment ROADMs based on R&S architecture and flexgrid 1x20 LCoS SSS are considered, with the possibility to use in-line colorless (C) A/D blocks, colorless and



directionless (CD) A/D blocks and Colorless Directionless and Contentionless (CDC) A/D blocks based on Multi Cast Switches (MCS).

The Transponders considered for fixgrid are the fixed bandwidth transponders and the fixgrid BVTs. The last ones can switch among different modulation formats for adapting to different client bit rates; they require a fixed optical bandwidth (50 GHz) and show different reach performances depending on the operating modulation format. The transponder considered for flexgrid are the flexgrid BVT and the sliceable SBVT, which is under specification in IDEALIST WP2 . In particular the SBVT recently emerged is capable to switch an aggregate client bandwidth up to 1.2 Tbit/s and to generate up to 12 separate and independent carriers modulated with a set of formats (QPSK, 8QAM, 16QAM) that can be freely merged in separate superchannels. Additional flexibility in Baud Rates (ranging from 30 to 50 Gbaud) and in the FEC type (from Soft to Hard decision FEC, with overheads ranging from 7% to 25%) allows to make available a large combination set of superchannel types, each one characterized by its own client bandwidth (nx100 Gbit/s, OTN ODU4-Cn compliant) and OSNR tolerance (and consequently optical reach).

The cost model applied is derived from results of previous projects (STRONGEST in particular, initially adapted for IDEALIST in D1.1, and now further developed with some extensions involving the CAPEX parts, that take also into account some OPEX aspects. In particular, since the context of application is a transnational network, where the network provider does not own the greater part of the infrastructural facilities, fiber rental and equipment hosting are considered, together with energy cost, as significant OPEX components. The CAPEX and OPEX are combined together assuming an amortization period for investment, and summarizing all the cost expenditures in the synthetic annual cost.

Due to the use of coherent reception in the transponders, dispersion unmanaged transmission is assumed. To evaluate the performance of the optical channels, the well know Gaussian Noise (GN) model is used. According to this model, a set of analytical formulas can be used to estimate under certain conditions (a constant spectral power density) the equivalent OSNR at the receiver with a good degree of accuracy, using a limited set of system parameters. Equivalent OSNR estimated by GN model takes into account the impairments due to both linear (ASE noise introduced by the EDFA amplifiers) and nonlinear (four wave mixing noise due to nonlinearity) effects. For a given modulation format the feasibility of a lighthpath (i. e. the certainty that the demodulated signal has a post FEC BER under a given value, for instance 10^{-7}) is checked comparing the resulting equivalent OSNR with the sensitivity of the format, which depends on the format and on some other aspects (type of FEC and implementation penalty).

The use of RAMAN amplification in combination with EDFA is also considered. This analysis allows evaluating the trade-off between adding Raman amplification to improve the reach of some demands or increasing the number of regenerators. A real network scenario at continental scale with two average span lengths (80 and 100 km) is considered for this analysis.

The results of this study are of two types. Firstly, in the context of a Continental Network, the conditions which make viable the use of RAMAN in addition to EDFA amplification are analyzed. Secondly, the price conditions (target prices) of SBVT that make the flexgrid



solution economically viable (if compared with the reference fixgrid network solution, considering the same network context and the same network requirements) are evaluated. This part of the work is still in a preliminary version that uses a simplified approach with approximate results and will be extended in a further version.

8.6.4.1 Network Scenario

The network used for the evaluations is the Reference Pan European Network provided by Telecom Italia (called Sparkle Network) and reported in Deliverable D1.1, Sec 4.6, with some minor updates.

Topology

The graph of the topology (not in scale) is shown in **Figure 29** (Section 4.3) while the main topological and routing characteristics are reported respectively in Table 41 and Table 42.

Concerning the topology, the network has 49 nodes and 72 links. As shown in Table 41, the node degree ranges from 2 to 5 with an average value of 2.9. Links have lengths from 2 to 1251 km, with about 390 km on average. Type of fiber in the real network, where it is known, is G.655 for 22 links and G.652 for 16 links. The remaining 34 links, for which the fiber type is unknown, are assumed to be G.652 because this is the most probable kind of fiber in case of fiber rental. Concerning the span lengths, apart from the very short links, they vary from about 80 km to 120 km, with some flexibility because the amplifier sites, in many cases, can be chosen along the links in order to arrange the desired span length.

Table 41: Main Topology Data of Pan EU Network

-	Node degree	Link length [km]
Min	2	2
Average	2.9	389.5
Max	5	1251

Table 42 includes some characteristics of all paths (considering all nodes as traffic terminations) when the working and back-up paths are obtained using the shortest path algorithm, with distance as primary metric. Back up paths are link disjointed from the corresponding working path and are obtained applying the shortest path algorithm on a reduced topology that excludes the links used by the working path.

Working paths have an average length of about 1800 km and a maximum length of 5000 km. Back up paths have an average length of about 3000 km and a maximum length of almost 7000 km. In order to avoid the extensive use of signal regeneration, the transmission along such long lengths requires a careful network design and the use of transponders prepared to handle long reach systems.

Table 42: Routing characteristics on the Pan EU Network

	Working Path routes	Back-Up Path routes
--	---------------------	---------------------



	hops	length (km)	hops	length (km)
Min	1	2	1	9
Average	5.3	1794.0	8.7	2969.4
Max	13	5046	18	6667
std dev	2.5	1002.1	3.4	1301.1

Traffic demand

A set of static traffic demand has been defined. The future traffic needs at continental level are not known either as total volume or as distribution, for a network as the one presented above. But taking into account the present traffic volume and distribution, and the potentiality of optical technologies in the medium and long term perspective, a traffic demand framework has been defined.

The full mesh demand of fully meshed connections at 100 Gb/s has been assumed as a reference for the other traffic matrixes. The reference full mesh demand is a uniform pattern of 100 Gbit/s connections (1176 in total), amounting to 117.6 Tb/s of total traffic offered to the network.

Other three sets of demands are defined in addition to the uniform one. Two sets, named respectively “background 1” and “background 2”, are derived assigning different weights to the nodes (the more is the weight the more is the terminated traffic on the node), and deriving the traffic requirements with a granularity of 10 Gb/s and the constraint to have the same (117.6 Tb/s) total traffic of the uniform demand defined above (). Background 1 maintains the traffic demand full mesh configuration, but with a non-uniform distribution, while Background 2 is defined again with a method that takes into account a node weight, but reducing the number of non-null traffic relationships (more precisely, about 30% of relationships have no traffic). In addition, a traffic demand called “polarized” is defined taking into account the today traffic pattern (the present network is under development, so today it has less nodes and links than the Pan EU network considered for this case study) and projecting it to reach the total reference volume of 117.6 Tb/s..

In Table 43, the main characteristics of the four reference traffic demands are shown. In the upper part of the table the values of minimum, minimum > 0, average, maximum and standard deviation traffic per demand are reported. In background matrixes, the rate between the higher and the lower (greater than zero) values is about 90 and 30 respectively (with a standard deviation slightly bigger than the average values), while in case of polarized matrix the rate between the maximum and the minimum (> 0) is very high, being equal to 228 (with a standard deviation equal to 3 times the average value). In the lower part of the table the number of non null-node demands is reported. For background 2 and polarized matrixes, the reduction in the number of demand items with respect to the full mesh case is of the order of 30%.

Table 43: Demand characteristics on the Pan EU Network

Traffic (node to node demand in Gbit/s)				
	uniform	background 1	background 2	polarized



Min	100.0	10.0	0.0	0.0
min > 0	100.0	10.0	30.0	20.0
Average	100.0	100.0	100.0	100.0
Max	100.0	920.0	920.0	4560.0
std dev	0.0	139.1	152.4	362.4

Number of non null node-to-node demands				
	uniform	background 1	background 2	polarized
	1176	1176	832	836

In Table 44 the values of terminated traffic on a node is reported. The average terminated traffic value is of the order of 5 Tb/s, a value that is still compatible with current OXCs (f.i. an A/D capacity of 5 Tb/s could be obtained on a ROADMs with only one A/D CD chain utilized at about 60% of its full capacity, a C-band of 4 THz and 2 bit/s/Hz of spectral efficiency). In the Polarized case we have the most requiring node in terms of A/D capacity, considering that an A/D capacity of 20 Tb/s could imply a total node capacity of 120 Tb/s at most (that is the 20% A/D bandwidth case). Therefore all the values are compatible with requirements expressed for the data plane in IDEALIST D2.1 (Table 2.6 – Summary of the node specifications), where the reference value for the maximum switching capacity for the long term (2025) is 500 Tb/s (60 Tb/s for mid term/2018).

Table 44: Terminated traffic on a Node in Pan EU network

Terminated traffic on a Node [Gb/s]				
	Uniform	Background 1	Background 2	Polarized
min	4800	1120	420	210
average	4800	4800	4800	4800
max	4800	13400	14230	20540
std dev	0.0	3988.9	4422.1	6391.8

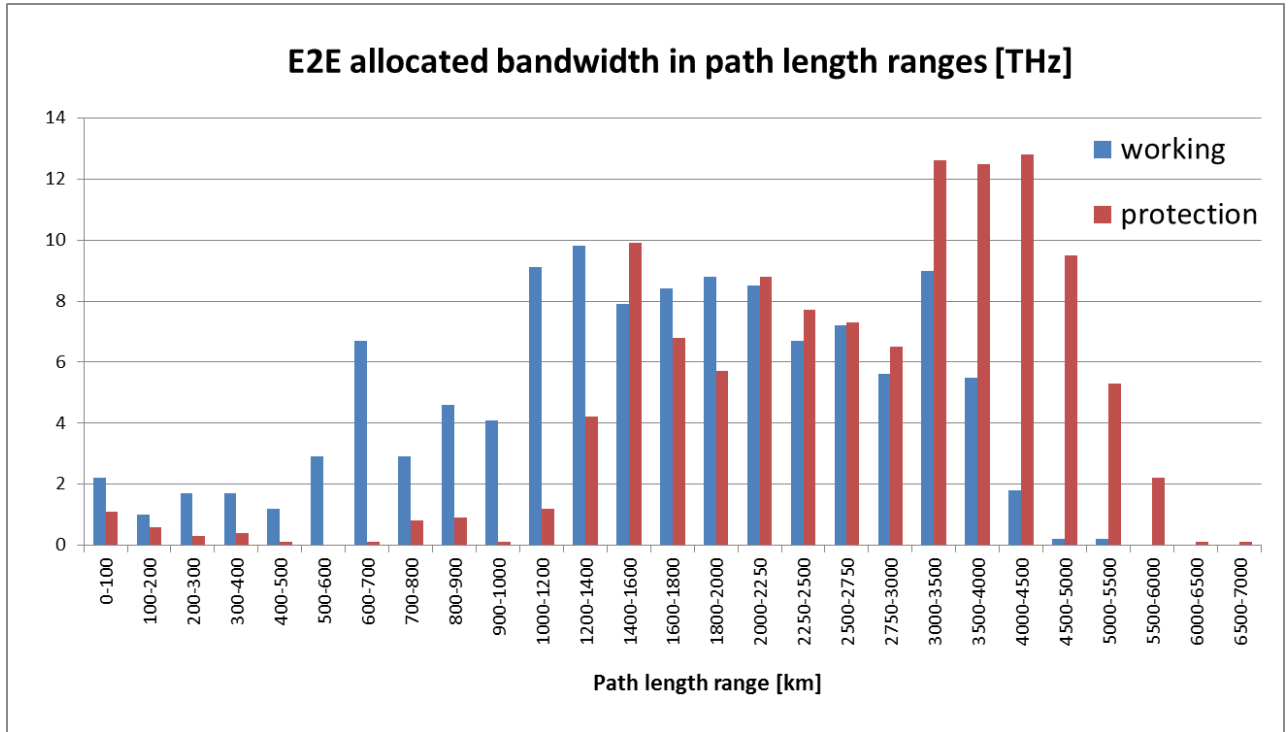


Figure 86: Allocated bandwidth (THz, on Y axis) in path length ranges for the uniform 100 Gb/s demand on the Pan EU Network.

Figure 86 shows the allocated bandwidth in terms of THz for a defined set of path length ranges, evaluated in case of the uniform 100 Gb/s traffic demand defined above. The traffic is routed using as main metric the distance, both for working and link disjoint protection paths. The spectrum allocation required for each connection is assumed to be 50 GHz, that is the typical size for a signal modulated with DP-QPSK and with the state of the art 33 Gbaud rate. This picture shows that, in case of uniform traffic, the highest spectrum allocation for the working paths is on path distances between 1000 and 3500 km while in case of protection the distances are longer, i.e. in the range between 1400 to 5000 km, with significant values between 3000 and 5000 km.

All the studies are carried on the Pan EU network, taking as references the four traffic matrixes defined above. In order to evaluate an extended range of network scenarios, the reference traffic matrixes have been scaled up or down, in order to obtain a distribution traffic pattern identical to the reference matrixes one, but with a different total volume.

8.6.4.2 Optical Nodes: architectures, cost and power consumption

The Optical cross connect architectures used for costs calculations and scenario comparisons are the fixed and flexible grid ROADMs based on WSS (fixed grid) or SSS (flexible grid). More advanced node structures, as the ones described in Section 4.1 of this Deliverable (AoD and SERANO), are not taken into consideration because, as they are still



under definition in their architectures and functionalities, it is too early for assigning them costs and power consumptions.

The general node architecture is the one presented in Figure 10 in Section 4.1 where, as an example, a R&S structure based on 1x9 WSS line modules is equipped to have 6 network degrees and 3 types of add and drop modules. In particular, dedicated Colorless A/D modules are present on each line to work as fixed A/D, while two Colorless and Directionless modules based on WSS and one Colorless, Directionless and Contentionless module based on Multicast switch are present in order to assure flexibility for demands that potentially require frequent reconfigurations.

The general description of possible fixed grid node architectures and the correspondent cost model has already been done in section 5.2 of D1.1. The model of an OXC based on ROADM structure and WSS modules has been updated taking into account the information acquired by the vendors in the last period, that lead to Many variants are considered in the model: type of signal conditioning/filtering (R&S vs. B&S); number of WSS ports in WSS modules (f. i. 1x4, 1x9 or 1x20), different types of A/D modules (C-less, CD-less or CDC-less, each one with several implementations), type of handled grid (fix grid vs. flexgrid) and others.

The general model already presented in D1.1 is confirmed, with some updates and changes. We assume that the node assembly is based on cards hosted in one or more shelves, requiring in turn different rack sizes. In general, a card includes all the basic building blocks that make possible the execution of a given functionality, including the control, monitoring and communication functionalities. Furthermore, each card requires a certain number of slots to be hosted on a shelf. The cost model assigns a cost and a power consumption to the basic building block, so that the card cost model can be derived as the sum of the cost of these building blocks. Furthermore, a cost is assigned also to the number of slots required by each card. Therefore, the overall cost and power consumption model for a node is obtained as the sum of the costs of both component parts (i.e. cards, with their building blocks) and required space occupation for each card (i.e. number of slots).

Examples of some basic elements used to build up the node cards are given in Figure 87 and Figure 88 . In particular, in Figure 87 we show the set of basic functional blocks independent from WSS technology for cards, with their cost and power parameterization, while in Figure 88 we show some examples of functional blocks implementing switching functionalities with their cost and power parameterization.

An example of cost calculation on cards is given in Figure 89 for some ROADM cards (splitters, line systems, A/D modules), showing the formulas used to compute the power consumption and the cost from the correspondent block properties.






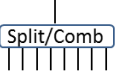




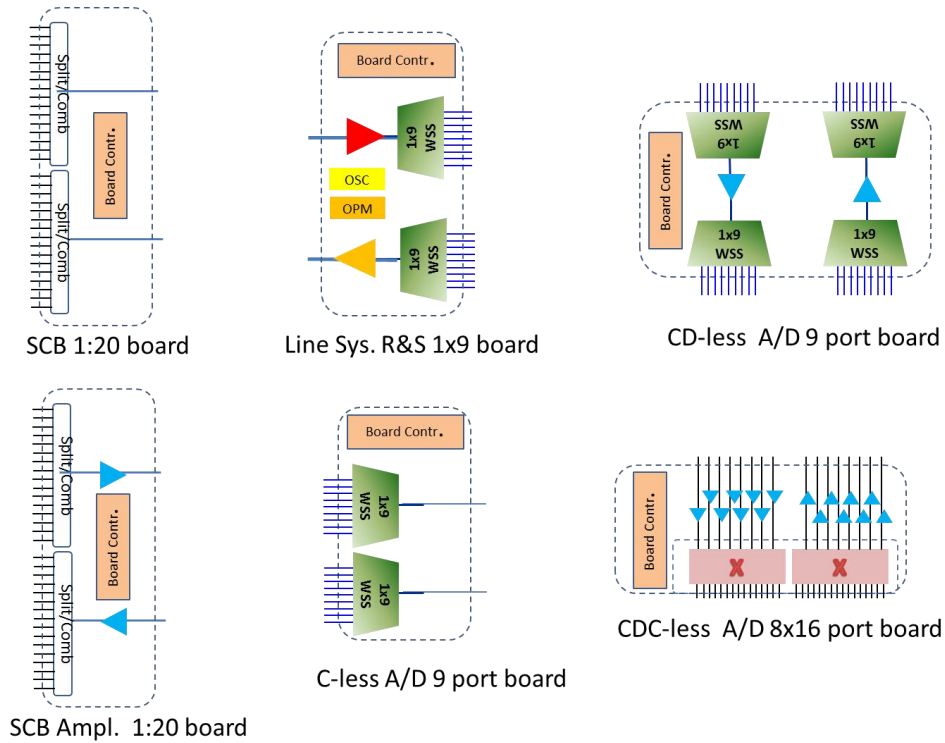
	General functionalities at board level (for both fix and flex grid)	Power	Cost
 Board Contr.	Basic control of board (only active boards)	P_BCB	C_BCB
 OPM	Optical Performance Motoring (OPM)	P_OPM	C_OPM
 OSC	Optical Supervisory Channel (OSC)	P_OSC	C_OSC
 Split/Comb	SCN 1:N (Splitter and combiner, N=4, 9, 20: cost and power are supposed not to depend on N) (one direction)	P_SCB	C_SCB
	EDFA small gain ampifier for Nodes (one model for all applications) (one direction)	P_OAN	C_OAN
	EDFA single stage fixed gain (preamplifier) (one direction)	P_OAP	C_OAP
	EDFA dual stage variable gain (booster) (one direction)	P_OAB	C_OAB
	EDFA plus RAMAN (one direction)	P_OAR	C_OAR

Figure 87: Set of basic functional blocks independent from WSS technology for cards and their cost and power parameterization

	Basic modules for flexgrid (board level)	Power	Cost
	WSS 1x4 module flexgrid (one direction)	P_WSS	C_WSS1x4_flex
	WSS 1x9 module flexgrid (one direction)	P_WSS	C_WSS1x9_flex
	WSS 1x20 module flexgrid (one direction)	P_WSS	C_WSS1x20_flex
	MSC module 4x16 (both directions)	P_MSC	C_MSC4x16
	MSC module 8x16 (both directions)	P_MSC	C_MSC8x16
	MSC module 16x16 (both directions)	P_MSC	C_MSC16x16

Figure 88: Example of functional blocks implementing Selective Switches (SS) and Multi Cast Switches (MCS) and their cost and power parameterization; SS are available in WSS (fix grid) and in SSS version (flexgrid).



Board	slots	Power	Cost
1x20 SCB (splitter combiner board)	1	P_{BCB}	C_{BCB}
1x20 SCBA (splitter combiner board with amplifiers)	1	$P_{BCB} + 2 * P_{OAN}$	$C_{BCB} + 2 * C_{OAN}$
ROADM Line System R&S WSS 1x9 with EDFA&RAMAN (2 WSS 1x9 plus 2 OArman plus OPM, OSC, BC)	3	$P_{BCB} + 2 * P_{WSS} + P_{OAP} + P_{OAR} + P_{OPM} + P_{OSC}$	$C_{BCB} + 2 * C_{WSS1x9_fix} + C_{OAR} + C_{OAB} + C_{OPM} + C_{OSC}$
C-less A/D chain board (2 WSS 1x9 plus BCB)	1	$P_{BCB} + 2 * P_{WSS}$	$C_{BCB} + 2 * C_{WSS1x9_flex}$
CD-less A/D chain board for 9 ports R&S structure (2 1x9 WSS plus 2 1x9 WSS plus 2 OA plus BCB)	2	$P_{BCB} + 4 * P_{WSS} + 2 * P_{OAN}$	$C_{BCB} + 4 * C_{WSS1x9_flex} + 2 * C_{OAN}$
CDC-less 8x16 ports A/D chain board (2 1x20 WSS plus 2 1x9 WSS plus 2 OA plus BCB)	2	$P_{BCB} + P_{MSC} + 16 * P_{OAN}$	$C_{BCB} + C_{MSC8x16} + 16 * C_{OAN}$

Figure 89: Example of ROADM cards with formulas to compute power consumption and cost from the correspondent block properties.

The complete model of the node is reported in the sheet “Equipment Model Cost & Power” of the excel file enclosed hereafter, where the basic card components and a wide set of cards are reported, with their values of cost, power consumption and slot occupancy. In



addition to line, add/drop and amplification cards, the transponders and regenerators have also to be considered, in order to have a complete configuration in terms of shelves and racks required by a node. This aspect, already mentioned before, that involves the overall power and cost model (CAPEX and part of OPEX) will be clearly explained in subsection 5, that deals with the Overall Power and Cost Model.



Cost&PowerModel_I
DEALIST_D1_5V3.xls

8.6.4.3 Transponder types

In this study we consider two types of transponders, the tunable fixed bandwidth transponders (fixed BVTs) operating in the 50 GHz fix grid and the sliceable flexgrid SBVTs operating in the flexible grid as defined by ITU-T G.694.1. The last ones are in line with the achievement got in WP2 of IDEALIST project (see Section 4.1 for an updated description of the SBVT model).

Concerning fixed transponders, the types taken into consideration are shown in the following **Table 45**.

Table 45: Fixed grid rate transponders parameters

Client rate	Typical reach on G.652	Required slots	Power consumption (W)	Cost (ICU)
40 Gbit/s	2500 km	1	100	0.48
100 Gbit/s	2000 km	2	200	1
400 Gbit/s	150 km	3	300	1.36

For SBVT, at the moment the cost and power consumption are not predictable and a “what if” analysis has to be performed in techno-economic studies. Taking as a reference the current state of development of the SBVT model reported in **Figure 10** of Section 4.1, Table 46 shows a set of characteristics and parameters of this SBVT.

Table 46: Flexible Grid SBVT characteristics and parameters.

Client rate	Reaches	Required slots on shelves	Power consumption	Cost (ICU)
Up to 12 client ports at 100 Gb/s or any	Depend on the modulation format and on network	7 slots for a fully equipped SBVT:	Unknown	Calculated as target cost in fix and flexi grid



combinations of nx100 Gb/s flows with the constraint of a total client traffic of 1.2 Tb/s	structure and parameters. Guideline values for G.652 are: 2000 km for QPSK 800 km for 8QAM 400 km for 16QAM	2 for the client interface card (any type); 3 for OTN 1.2 Tb/s flex framer card; 2 for a card hosting up to 3 multi flow modules		comparison
---	---	--	--	------------

8.6.4.4 Overall Cost and Power Consumption Model

The IDEALIST cost model has been preliminary introduced in D1.1 [37] as an extension of the STRONGEST Project model, that considered only the CAPEX for the equipment parts. In particular, the Idealist Cost Unit (ICU) is represented by the 100 Gb/s DP-DQPSK transponder with a reach of about 2000 km on SSMF fiber. All other devices and subparts of equipment are priced with reference to the ICU. On the contrary, the costs forecast for the innovative SBVTs is one of the main novelty within the IDEALIST project. The time references considered are year 2013 (i.e. the project starting period), the short term 2015 year (when flexgrid transponders at 400 Gb/s should begin to appear) and the medium term 2018 year (when it is expected that line cards and interfaces both at electrical and optical layer will reach 1 Tb/s).

In D1.1 it has already been introduced a **CAPEX** model for WDM equipment based on a rack that hosts the node subparts, i.e. the shelves where each card occupies a certain number of slots.

The model has now been developed and extended to power consumption, defining a three level model structure:

1. the functional level, including a set of sub-boards components (functional blocks), characterized by cost and power consumption;
2. the board level, with boards that include the functional blocks required to achieve a sub-system (board) functionality. Cost and power consumption of the board come from functional blocks parameters;
3. the equipment level, with the whole integrated system including all the boards and the racks (one or more). Cost and power consumption of the piece of equipment is calculated summing up the contributions of boards and racks.

Concerning the first two levels, involving blocks and cards, they are already described in the previous two subsections for node components and transponders respectively.

Concerning the equipment level, recent trends and lifelike evolution for racks, shelves and boards occupancies are taken into account to build up a general model. The model is not based on a solution from a specific vendor but it has been obtained looking at some line products from different manufacturers and considering the form factor that cards could

have in the future. The rack model can be summarized as follows. A rack or a sub-rack occupies the space of half of an ETSI tile (one ETSI tile is 60 x 60 cm). The available racks and shelves and their capacities are the following ones:

- the small rack (for OLA) is able to host 2 half shelves
- the full rack (for ROADM) hosts up to 4 shelves or 8 half shelves
- the half shelf hosts 10 slots (8 slots for lines and tributary cards and 2 slots for control/service boards)
- the shelf hosts 20 slots (18 slots for lines and tributary cards and 2 slots for control/service boards).

Figure 90 shows a view of subparts (shelves), floor needs and some example of racks totally or partially equipped.

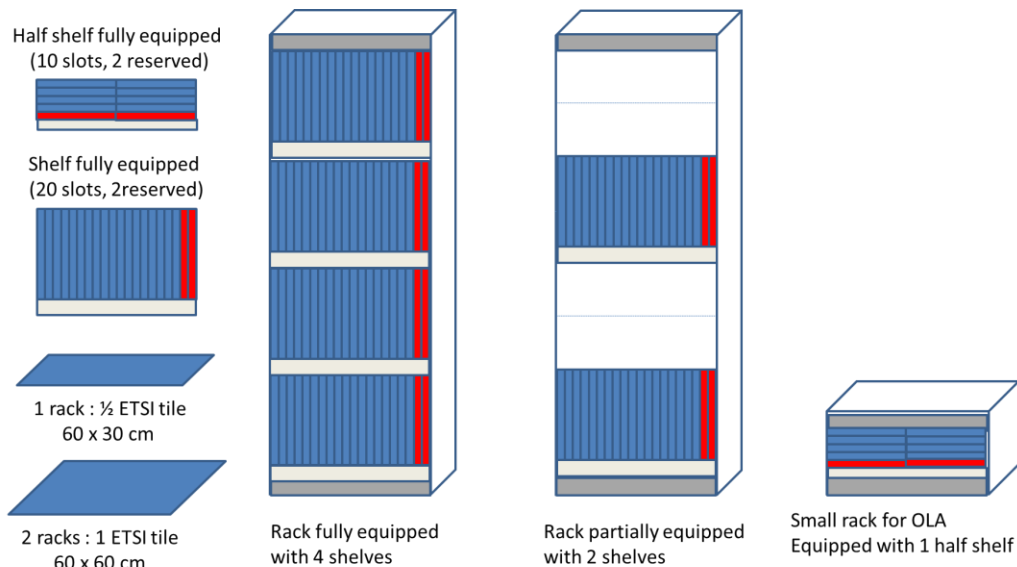


Figure 90: Example of racks equipped with shelves or half shelves.

In addition, the IDEALIST cost model for the WDM layer has been extended to include also some **OPEX** as yearly costs that at present include the following items:

- **Fiber rental cost**, i.e. the rental cost per km per year of a pair of fibers. A single value is assumed, although in reality the cost depends significantly on the area (fiber rental cost are higher in city downtown, intermediate in suburbs and lower in the countryside) and the Country (Nation).
- **Equipment hosting cost**, i.e. the cost per year for hosting a rack/subrack for OLA or ROADM. This service, for continental networks that span multiple countries, is usually provided by “telehousing” companies and normally the housing service includes floor space, conditioning, power supply and communication facilities.
- **Energy consumption cost**, specified through an average cost per KW per hour. Also in this case the price depends on some factors (the Country, the hour of the day and the day of the week in which the energy is consumed), but we apply a



single average value to perform cost calculation. Yearly power cost is the total power consumed by the network multiplied by the total hours in a year (8760 hours).

- Maintenance cost is not considered explicitly, but a simplified way to take this item into account is to increase the equipment investment cost by a given percentage (typical values are between 3% and 5%).

In order to take into account both CAPEX and OPEX at the same time, we applied the following formula that allows us to compute an annual cost of the network:

$$\begin{aligned} \text{Annual cost} = & \text{CAPEX/amortization period} + \\ & \text{fiber rental cost} + \\ & \text{equipment hosting cost} + \\ & \text{energy cost} \end{aligned} \quad (16)$$

where the amortization period is the time (in years) for which the optical layer network items can be considered paid back (normally it corresponds to a value greater than the average lifetime of the optical equipment). In particular, 5 years is assumed as CAPEX amortization period for all optical systems.

The OPEX cost parameters applied in this case study are reported in the enclosed excel file Cost&PowerModel_IDEALIST_D1_5V3.xlsx and summarized in Table 47. Sources of the data reported in Table 47 cannot be disclosed because information is confidentially collected among the involved partners. In the case of fiber cost, a public reference is the document [118].

Table 47: OPEX cost parameters (cost in ICU).

Fiber rental (couple of G.652)	minimum	0.002 per km per year
	average	0.004 per km per year
	maximum	0.008 per km per year
Equipment hosting	Small rack / 2 half shelves / 16 slots	0.12 per year
	Full rack / 1 shelf / 18 slots	0.24 per year
	Full rack / 2 shelves / 38 slots	0.48 per year
	Full rack / 3 shelves / 54 slots	0.72 per year
	Full rack / 4 shelves / 72 slots	0.96 per year
Energy consumption		3.00E-06 per kWh

8.6.4.5 Impairment Model

The main objective of this section is to provide a clear identification of the assumptions made to evaluate the performance of each channel. The exhaustive description of the assumptions, advantages and disadvantages of the performance evaluation method can

be found in section 4 of deliverable D2.1 [37]. This performance evaluation method is based on the Gaussian noise (GN) assumption for nonlinearities [106], [107], [108]. The GN model has proven to be quite accurate when transmission along dispersion uncompensated links is considered and the impact of nonlinearities is kept reduced, which is the case of interest for optical telecommunications. In summary, the GN model corresponds to a good compromise between accuracy and simulation time.

The impact of nonlinearities on the performance is taken into account using equations 4.2 and 4.3 of D2.1. The simultaneous transmission of 96 channels along each fiber is assumed when evaluating the impact of nonlinearities. An equivalent OSNR after nonlinear fiber transmission is estimated by substituting eq. 4.2 in eq. 4.1 of D2.1.

The feasibility of a data channel is assessed by evaluating the residual margin. By definition, it is assumed that a channel with negative residual margin is unfeasible. The residual margin of the transmitted channel (RM) is defined as the difference between the total equivalent OSNR at link output, $OSNR_{eq,tot}$ and the OSNR required in back-to-back (B2B) to recover the transmitted signal with sufficient quality, $OSNR_{B2B}$:

$$RM = OSNR_{eq,tot} - OSNR_{B2B} \quad (17)$$

where the total equivalent OSNR at the path output is given by:

$$OSNR_{eq,tot} = \left[\frac{1}{OSNR_{add}} + \sum_{i=1}^L \frac{1}{OSNR_{eq,i}} + \sum_{i=1}^{L-1} \frac{1}{OSNR_{psth}} + \frac{1}{OSNR_{drop}} \right]^{-1} \quad (18)$$

$OSNR_{add}$ is the contribution to OSNR of the add node, $OSNR_{eq,i}$ is the equivalent OSNR of link i , $OSNR_{psth}$ is the contribution to OSNR of the pass-through nodes (express layer of ROADMs) and $OSNR_{drop}$ is the contribution to OSNR of the drop node.

The transmission of 100 Gbit/s dual-polarization with coherent-detection quadrature phase-shift keying (DP-QPSK) channels with $OSNR_{B2B}=13$ dB is assumed. Moreover, only transmission along standard single mode fiber (SSMF) and large effective area fiber (LEAF) is considered. The modeling of these fibers is shown in Table 48:

Table 48: Optical fiber parameters.

Fiber type	Attenuation parameter α_s [dB/km]	Attenuation parameter α_p [dB/km]	Dispersion parameter D [ps/nm/km]	Nonlinear coefficient γ [1/W/km]	Raman gain coefficient g_r [1/W/km]
SSMF (G.652)	0.21	0.25	17	1.3	0.43
LEAF (G.655)	0.22	0.22	3.8	1.5	0.54

where α_s is the attenuation parameter at the data channels wavelength and α_p is the attenuation parameter for wavelengths ranging from 1400 nm to 1500 nm.

Some additional parameters of the transmission system are shown in Table 49:

Table 49: Transmission system parameters.

Average splice spacing	5 km
Average splice attenuation	0.05 dB
Additional loss at fiber input	1 dB
Additional loss at fiber output	1 dB
Additional fiber attenuation, distributed evenly along the fiber length	2 dB
ROADMS insertion loss	18 dB
Add/drop chain insertion loss	15 dB
Average optical power at input of each ROADM	1 dBm

The attenuation added in each fiber span is compensated using EDFAs. The EDFAs noise figure is modelled as dependent of gain, as occurs in real EDFAs. However, in order to keep the analysis quite simple, only a very simple modelling of the dependence of noise figure of EDFAs in gain is done. The assumptions made in this work are shown in Table 50:

Table 50: EDFAs noise figure.

EDFA gain [dB]	EDFA noise figure [dB]
≤ 15	10
[15, 20]	7
[20, 25]	5.5
> 25	4

Moreover, it is assumed that the EDFAs gain cannot be set too low. Namely, it is assumed that the minimum EDFA gain is 15 dB. In case the target EDFA gain is lower than 15 dB, the EDFA gain is set to 15 dB and an attenuator is added at the EDFA output to set the combined gain of the EDFA plus the attenuator equal to the target gain.

The launch power at each optical fiber input is always optimized so that the equivalent OSNR is maximized. If a too low power level was set at the optical fiber input, the impact of nonlinearities would be reduced. However, only a reduced OSNR would be obtained at receiver input. On the other hand, if a too high power level was set at the optical fiber input, a high OSNR would be obtained at receiver input. However, the signal distortion caused by nonlinearities would be too high.

Using the assumptions presented in this section, we find that a 100 Gbit/s DP-QPSK channel, modeled by a total baud rate of 33 Gbaud (to accommodate FEC), can be transmitted along 7 OMS, with each OMS consisting of 4 spans with 100 km of SSMF (2800 km) or 11 OMS, with each OMS consisting of 4 spans with 80 km of SSMF (3520 km), before requiring signal regeneration. These reach estimates are, however, too optimistic. Since the only transmission effects that are considered are ASE noise from EDFAs and nonlinearities, this result was somewhat expected. Indeed, real transmission systems have to deal with additional transmission impairments such as gain fluctuations along the transmission bandwidth of EDFAs, polarization dependent losses, cumulative filtering due to the use of ROADMs, power transients caused by adding or removing



channels from a network, etc. Moreover, a system margin is also usually set to guarantee that the transmission system still operates correctly even after some (long) time has passed (aging margin) and to cope with the possibility of optical fibers or any other equipment may have worse performance than the expected one. Therefore, in order to make the performance model more realistic, an additional OSNR margin of 0.1 dB is added every time an inline EDFA is crossed and an OSNR margin of 0.2 dB is added every time a ROADMs is crossed. Rewriting the RM, so that these system margins are taken into account, we have:

$$RM = OSNR_{eq,tot} - OSNR_{B2B} - 0.1 \times N_{inlineEDFAs} - 0.2 \times N_{ROADMs} \quad (19)$$

where $N_{inlineEDFAs}$ is the total number of inline EDFAs and N_{ROADMs} is the total number of ROADMs in the path of the channel under analysis.

We find that a 33 Gbaud DP-QPSK channel can be transmitted along about 1600 km of SSMF before requiring signal regeneration. We would like to highlight that this reach is somewhat conservative with respect to other results usually reported in the literature. However, by using such approach, the results presented in this work should be more aligned with the ones usually seen in real transmission systems, where channels usually operate with some safety margin.

Hybrid EDFA/Raman amplification

The performance of optical fiber transmission systems can be improved using Raman amplification, especially in presence of longer spans. The Raman amplification consists in transferring power from lower wavelengths to higher wavelengths [110]. Thus, if a high power (pump) signal is launched into an optical fiber in a smaller wavelength than the one of data channels, the data channels will experience amplification, thus mitigating the impact of fiber attenuation. However, a careful selection of the pump wavelength must be performed as the efficiency of the Raman power transfer, which translates in a fiber gain, depends on the spacing between the pump and data channels. The power transfer is maximized when the difference between the pump and the data channel wavelengths is around 13 THz [110].

The dependency of the Raman fiber gain on the spacing between the pump and data channels can be quite detrimental in optical communication systems as it leads to power fluctuations. In order to cope with this limitation, instead of transmitting just a single pump signal, several pump signals may be launched into the fiber with the objective of obtaining a flatter gain along the transmission bandwidth [111].

Figure 31 shows the two possible Raman amplification approaches: the pump signals are launched at the fiber input (co-directional propagation) or the pump signals are launched at the fiber output (counter-directional propagation). From a noise perspective, co-directional propagation leads to the best performance, since the additional noise added by Raman amplification is mostly added at fiber input and is, therefore, attenuated along the optical fiber [111]. However, co-directional propagation enhances the impact of nonlinearities because data channels are transmitted with (much) higher power levels. The counter-directional propagation of the pump signal usually leads to only negligible enhancement of nonlinearities because data channels experience higher gain amplification near the end of fiber, where the power level of data channels is usually reduced [111]. On the other hand, the noise added to data channels by the counter-directional Raman amplification is mainly

added close to the end of the optical fiber and, therefore, noise is only slightly attenuated before reaching the end of the optical fiber.

In summary, the best solution regarding Raman amplification with respect to performance consists in using several pump signals to get a flat gain in the transmission bandwidth. These pump signals should be launched in a co- and in a counter-directional propagation approach [111]. However, a careful setting of the pump powers is required to mitigate the impact of the disadvantages of each approach. This Raman amplification solution is, however, very cost ineffective since it requires the use of several pump lasers. In real transmission systems, where minimum cost is one of the main objectives, only counter-directional Raman amplification with a reduced number of Raman pumps (e.g. 2 pumps) is usually considered. In this work, only the transmission of a single pump in a counter-directional propagation approach is considered.

The Raman amplification is a distributed nonlinear effect that occurs along the entire length of the optical fiber. Additionally, the depletion of the pump signals depends on the power level of data channels. Moreover, the power level of data channels depends on the power level of pump signals. These characteristics make the accurate simulation of the impact of Raman amplification complex and time consuming [109], [110]. However, several simplifications can be done to simplify the analysis. If we admit that only a single monochromatic pump is transmitted with a weak signal (the data channel), the evolution of the pump power may be considered independent of the data signal and ASE noise. In this case, the gain at position z of optical fiber is given by [112]:

$$G(z) = \exp\left(\int_{l=0}^z g_r P_p(l) dl - \alpha_s z\right) \quad (20)$$

where $P_p(l)$ is the power level of pump laser at position l . As only counter-directional propagation is assumed in this work, the power of the pump signal at position l is given by:

$$P_p(l) = P_p^{launch} \exp(-\alpha_p(L-l)) \quad (21)$$

where P_p^{launch} is the power launched into the optical fiber by the pump laser and L is the total length of the optical fiber.

A parameter often referred when using Raman amplification is the Raman on-off gain. The Raman on-off gain is given (in the undepleted regime) by (25), but omitting the second term of the equation, which provides the accumulated fiber loss at fiber position z . As only a very simple modeling of the Raman impact is considered in this work, the Raman on-off gain is set as a constant, and imposed equal to 8 dB. The power launched into the optical fiber by the pump laser is optimized so that the required Raman on-off gain is achieved.

One of the problems of Raman amplification is adding noise to the transmission system. The noise power at the fiber output, in a single polarization state, is given by:

$$P_{ASE} = G(L) \int_{z=0}^L \frac{(1 + \eta_T) h \nu_s B_o g_r P_p(z)}{G(z)} dz \quad (22)$$

where h is the Planck's constant, ν_s is the frequency of the data channel, B_o is the bandwidth used as reference when defining the OSNR and assumed equal to 0.1 nm in this work, and η_T is the thermal equilibrium excitation given by:

$$\eta_T = (\exp(h - (\nu_p - \nu_s)/(kT)) - 1)^{-1} \quad (23)$$

where ν_p is the frequency of the pump signal, assumed equal to 206 THz in this work, k is the Boltzmann's constant and T is the temperature, assumed equal to 300 Kelvin.

In summary, it is assumed that Raman amplification is modelled by an amplification of the power level of channels of 8 dB and by an increase of the noise power level, which is given by (29).

Considering again the cases referred previously without Raman amplification: OMSs consisting of 4 spans with 100 km of SSMF and OMSs consisting of 4 spans with 80 km of SSMF, we find that Raman amplification provides an additional reach of about 25%, allowing the transmission of 100 Gbit/s DP-QPSK channels along about 2000 km before requiring signal regeneration.

In order to provide more insight on the performance model, Table 51 shows the estimated OSNR for some links of the Pan European network when only the impact of noise from EDFAs is taken into account and when the impact of nonlinear effects on the OSNR is also considered, for the cases with and without Raman amplification. The transmission fiber is SSMF.

Table 51: OSNR for some links of the Pan European network.

				EDFA only		EDFA+Raman	
Link	link length [km]	Number of uniform spans	Span length [km]	$OSNR_{eq}$ [dB]	$OSNR(EDFAs ASE noise only)$ [dB]	$OSNR_{eq}$ [dB]	$OSNR(EDFAs ASE noise only)$ [dB]
36-9	608	8	76	21.4	27.2	22.2	29.1
9-8	335	4	84	23.1	28.1	24.3	30.1
8-26	394	5	79	22.9	28.3	23.9	30.3
26-21	630	8	79	21.0	26.5	22.0	28.5

If a channel is transmitted from node 36 to node 21 (nose sequence: 36 – 9 – 8 – 26 – 21), the evolution of the equivalent OSNR along the transmission path is given in Table 52.

Taking into account the number of EDFAs (25) and ROADMs (4) crossed by a channel transmitted between nodes 36 and 21, the system margin to be considered is 3.3 dB. As the required OSNR in B2B is 13 dB, we conclude that a channel can be transmitted between nodes 36 and 21 without using signal regeneration if Raman amplification is used. On the other hand, if no Raman amplification is used, the transmission of a channel between nodes 36 and 21 requires using signal regeneration.

Table 52: Evolution of the equivalent OSNR for a channel transmitted between nodes 36 and 21.

	$OSNR_{eq,tot}$ [dB]
--	----------------------



	Add node	Link 36-9	ROADM pass- through	Link 9-8	ROADM pass- through	Link 8-26	ROADM pass- through	Link 26-21	Drop node
Without Raman	33.8	21.4	21.2	19.2	19.1	17.7	17.6	16	16
With Raman	33.8	22.2	21.9	20.2	20	18.7	18.6	17.1	17

8.6.4.6 Network Design Framework

A customized network design framework, which to a large extent mimics the dimensioning functionalities of a commercial optical network planning tool, was augmented to enable performing the case study reported in the next section. Without going into detail, the main components of the network design framework exploited within the scope of this work are:

- Routing path computation functionality: pre-compute a set of K -shortest paths or K -shortest cycles over the DWDM network layer. By configuring a smaller or larger value of K it is possible to decrease or increase the routing solution search space.
- 3R regenerator placement functionality: for a potential routing path solution the best 3R placement, that is, the 3R placement solution that results in minimum number of regenerators required for signal regeneration is assessed. This module takes into account the availability of free lightpaths at the path's links and add/drop ports at the path's nodes (along with any port usage restriction associated to the add/drop ROADM architecture being deployed. Importantly, this module requires an optical performance model to estimate the feasibility of lightpaths.
- Wavelength assignment functionality: for each lightpath of a given 3R placement solution, selects one of the available wavelength channels according to a given criteria.
- Traffic demand routing ordering functionality: for the set of traffic demands not routed in the current planning iteration, determines the most suitable traffic demand to be routed in the next iteration according to a given criteria.
- Equipment inventory functionality: based on the outcome of routing the entire set of traffic demands, the main equipment cards needed at the network nodes is calculated. Particularly, the number and type of transponders and regenerators (or generally the number and type of line interfaces) required is determined along with the number and type of the main switching cards deployed at the ROADM nodes (e.g., WSSs, splitter combiners, multicast switches) taking into account the ROADM architecture selected in advance. The number of tributary cards (cards directly connected to the transponders/regenerators) calculated is the minimum one that enables to support all transponders/regenerators deployed at every node.

Importantly, the main modification realized in the already existing (custom-made) network design framework to allow performing the case study consisted of implementing the impairment model described above and integrating it within the appropriate steps of the framework. Particularly, the impairment model was implemented and then integrated into the 3R regenerator placement module also described above. With this extension in place, it is possible to optimize the number of 3R regenerators required in a long-haul DWDM network assuming transparent connectivity.



An equipment inventory is generated after running the complete network design framework over the Sparkle Pan-European network with the assumed specific ROADM architectures and link capacities and resorting to the line interface technology whose performance is modelled via the impairment model. This equipment inventory is used as input to the cost, footprint and power consumption model. As a result, an estimation of the overall cost associated to deploying and operating a DWDM network over the Sparkle topology is provided enabling the analysis, for example, the impact of adding Raman amplification to the Sparkle network.

8.6.4.7 Results obtained on the analysed cases

The final goal of the ongoing joint work between Telecom Italia and Coriant Portugal is to assess the impact of a variety of options, currently available or available in the future (e.g., SBVTs) on cost, footprint and power consumption of the Sparkle Pan-European network. For example, the trade-offs between using more advanced transmission formats (e.g., 16QAM, superchannels) versus deploying additional fibers are of significant interest for the IDEALIST project, as they should hint at the potential benefits and limitations of these formats and associated transponder equipment in long-haul networks. However, within the timeframe of this deliverable, the support of these formats/transponders (e.g., for the impairment model and for the design framework) is still not available. Hence, we focused on validating the entire setup assuming the utilization of only state-of-the-art commercial 100G transponders. This work not only did enable to validate the network planning setup, but it also provided a benchmark regarding the capacity and cost attained with current technology, which can be later used for comparison purposes. An extension of the study has been done in order to evaluate the SBVT target cost in a flexgrid environment, with an approximate network dimensioning which does not apply any accurate impairment model but only a rough reach estimation for modulation formats different from QPSK used on 100G fixed grid networks.

Two different distributions of inline amplifiers along the links of the Sparkle topology were assumed for the case study. The first one considered a more typical amplifier spacing of around 80 km, whereas the second one assumed longer spacing values of around 100 km. Table 53 summarizes the network properties that result from both configurations. For instance, it can be seen that opting for longer spans results in a decrease of around 25% in the number of inline amplifiers, although at the expense of increasing the average span length by 20 km, which will influence the performance of the lightpaths routed over the Sparkle network, consequently impacting the number of 3R regenerators deployed.

Table 53: Properties of the two inline amplifier distributions in the Sparkle topology.

Topology	Number of inline amplifiers	Average span length [km]	Average number of spans per link
Sparkle S80	281	79.478	4.903
Sparkle S100	210	99.442	3.917

The traffic pattern utilized in the case study was the polarized pattern, whose main properties have been identified, assuming the traffic demands are unprotected. The baseline fiber topology of Sparkle was considered (no additional fiber deployment) and each link was capable of carrying 96 channels (in a 50GHz fixed grid). In all considered



cases, by appropriately setting the routing parameters of the network design framework (e.g., K parameter set to 6) it was possible to successfully set up over this network all the 384 traffic demands.

As discussed in the impairment model, the deployment of RAMAN amplification in a long-haul network such as the Sparkle one is expected to improve the transparent reach, thereby eventually providing savings in the number of expensive 3R regenerators that need to be deployed. Noteworthy, these savings will depend also on the number of traffic demands and on the routing paths that can be used across the network. Consequently, there is a trade-off between the cost of the additional RAMAN amplifiers that need to be deployed and the cost savings with 3R regenerators that can be attained. This trade-off was studied in more detail as described in the following.

For each network topology – Sparkle S80 and Sparkle S100 – the dimensioning was performed for different configurations in terms of RAMAN amplifiers placement, ranging from the case in which RAMAN is not deployed at all to the cases in which RAMAN is deployed for spans that are relatively short. The input parameter, used to decide which spans require RAMAN amplification or not, consists of the minimum span length above which RAMAN can be considered useful. A granularity of 5 km was used to obtain the results.

Figure 31 shows the number of RAMAN amplifiers and 3R regenerators required, as a function of the span length threshold above which RAMAN is deployed, where plot (a) corresponds to Sparkle S80, plot (b) corresponds to Sparkle S100, and “Inf” denotes an infinite value of the threshold which implies not to install any RAMAN amplifier.

As can be seen from the results, and more notoriously in Figure 31 (b), increasing the threshold in the region of small span length values results at the beginning in the smooth decrease in the number of RAMAN amplifiers deployed and also a minor increase in the number of 3Rs needed. However, as the average span length value is attained, there is a steep drop in the number of RAMAN amplifiers deployed at the expense of a visible increase in the number of 3R regenerators. This trend highlights the trade-offs expected between RAMAN amplification and 3R regeneration.

Based on the cost model employed, Figure 32 depicts the cost contribution (in ICU) in terms of CAPEX only, associated to RAMAN amplifiers (e.g., additional cost of deploying an EDFA/RAMAN amplifier instead of just an EDFA amplifier) and to 3R regenerators (e.g., the cost of two 100G transponders used to perform the 3R functionality). Particularly, according to the enforced cost model, adding RAMAN amplification to an existing amplifier (for the two ways of transmission) costs 30% of deploying a single 100G transponder (Raman for the two directions in a hybrid amplifier costs 0.3 ICU while a 100 G transponder costs 1 ICU).

It is important to emphasize that the conclusions drawn above about the conditions in which the RAMAN is convenient depend on the ratio between the costs of transponder and RAMAN. For instance, if adding a RAMAN amplification to an existing amplifier costs 60% more than deploying a single 100 G transponder (the double of the parameter values used to obtain results in Figure 32 the costs of RAMAN and 3R regenerators are as in Figure 91. In this case the better condition is evidently for a particular value of the RAMAN placement threshold and, above or below, the cost is always appreciably higher.

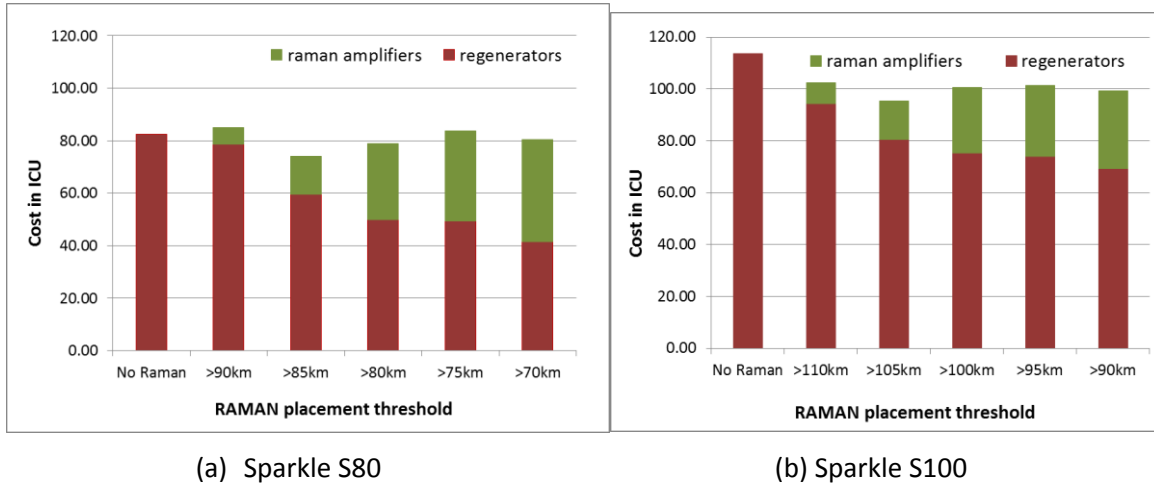


Figure 91: Cost in ICU of RAMAN amplifiers and 3R regenerators only (cost of one way RAMAN functionality on hybrid amplifiers is increased from 15% to 30% of a 100 G transponder).

Importantly, this is only a first and straightforward cost analysis made on CAPEX of Regenerators and RAMAN amplifiers only that is not considering, for example, the impact of the power consumption and footprint of RAMAN vs. 3R regeneration and the indirect implications on other network co-factors.

Figure 33 shows the total cost diagram with the contribution of several components for some of the cases analyzed above, when the concept of Yearly cost is assumed. The yearly cost includes the CAPEX components (Nodes, OLAs (EDFA), Transponders, Regenerators and RAMAN amplifiers) which are calculated with a 5 years amortization period, and the OPEX components (Fiber rental, Hosting, Energy). Diagram shows that the differences in terms of absolute yearly cost are not significant in all the analyzed cases. It also shows that the six cost components from “Fiber rental” to “Transponder” (the lower part of the bars in the diagrams) are very steady and do not depend significantly from the strategy used to place RAMAN amplifier in the network. This is due to the fact that the basic infrastructure of the network (nodes, fibers, OLAs) has to be placed regardless the traffic volume carried by the network. About the remaining two cost components, (related to the regenerators and RAMAN), their individual changes depend on the strategy applied to put RAMAN amplifiers (the threshold on span distance), but the sum of their costs is approximately constant and give in the best case an advantage of only a 6% on the total yearly cost w.r.t. the worst case.

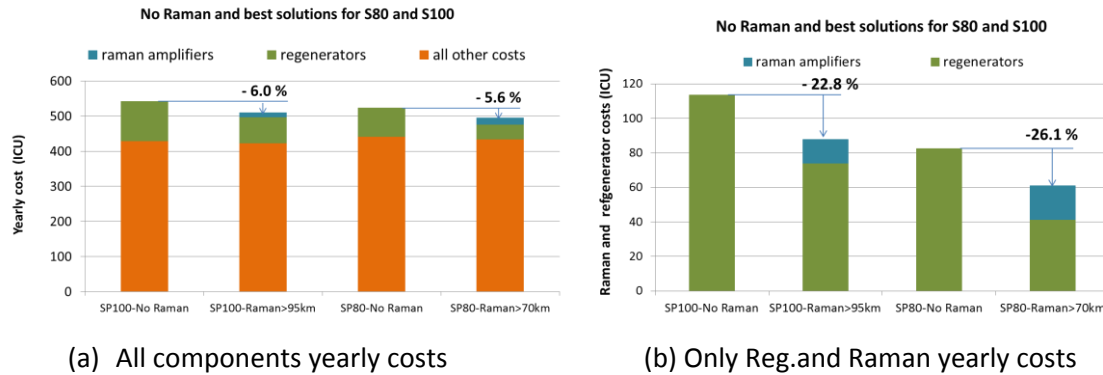


Figure 92: Yearly costs for No Raman and best solutions for Polarized reduced traffic.

Figure 92 shows the cost figures for the S80 and S100 Networks without Raman as well as with Raman at best solution (that are the solutions for spans greter than 70 km in S80 and for spans greater than 95 km in S100). On the left side of the Figure 92 there are the overall costs, partitioned into three items: Raman, regenerators and all the other ones. It is possible to see that the least overall yearly cost solution occurs for spans of 80 km (S80) with Raman used on spans greater than 70 km. This suggests that applying the IDEALIST reference cost model to the Pan European Nestwork leads to take the span short and deploy RAMAN extensively. On the right side of the Figure 92, for the same cases described above, only the Raman amplifiers and regenerators cost components are reported. The difference between the No Raman cases and the corresponding lowest cost ones is significantly greater (of the order of 25%) than the in case of total costs calculation.

8.6.4.8 Preliminary extension to flexgrid networks

An extension of the dimensioning and cost study done for fix grid network has been done in case of use of SBVT and flexible grid network. The aim of this study is to find the target cost of an SBVT, i. e. the absolute cost of this component that equals the total cost of the fixed grid network using only transponders at 100 Gb/s in 50 GHz fixed grid.

The transponder type used as tributary for the flexgrid network traffic is always the SBVT, thus also for flows at the lowest bit rates (100 Gb/s) and also for regenerations which have been supposed implemented by SBVTs connecting back to back two tributary ports. These ports have to be at the same rate of the flow to be regenerated.

The same polarized demands used in the fixgrid dimensioning are considered, but assigning the node to node bandwidth requirement to different types of client port rates. The assortment of circuits at different client port rates for the whole network are listed in Table 54. Higher port rate circuits (1 Tb/s and 400 Gb/s) concern the most demanding nodes relationships. Circuits at 100 Gb/s are highly dominant in terms of number (196 of 250), carrying 52% of the total traffic volume. The higher rate circuits (from 200 Gb/s to 1 Tb/s) share almost equally the remaining traffic demand volume (48% of the total).

Table 54: Number of circuits at different rates in the traffic demand.

Client Rate	Number of demands	Carried Bandwidth (Tb/s)	% of the total bandwidth
-------------	-------------------	--------------------------	--------------------------



1 Tb/s	5	5.0	13%
400 Gb/s	16	6.4	17%
200 Gb/s	33	6.6	18%
100 Gb/s	196	19.6	52%

Since at the moment it was not possible to perform an accurate sizing of the flexgrid network inside the planning tool, we have considered that the whole traffic carried by the network was as in the case of fixgrid and, considering that the flexgrid Spectral Efficiency is higher on average, it has been assumed that one 4.8 THz C-band system is enough also in the flexgrid case to carry the whole traffic demand on each topology network edge. As a consequence, the configuration and cost of the transmission part are the same of the corresponding fixed grid case (km of fiber, OLA CAPEX and hosting, RAMAN where required, power consumption).

The flexibility of the SBVT in terms of modulation formats allows to select, on the base of the path distance, the number of carriers to assign to a given tributary flow, if it is greater than 100 Gb/s. An analysis of the path length for the particular traffic pattern considered here shows that the higher SE formats (8QAM and 16QAM) could be applied instead of QPSK in a few cases. As a consequence, at network level, the gain in the spectrum and in the number of carriers is very little when compared to a case that implements a systematic use of QPSK (for example for the number of carriers the gain is in the range of 1%-3%). In addition regenerators, that are a considerable part of all Optical-Electro-Optical (OEO) devices resulting in about 1/3 of the whole OEO cost, involve circuits whose path lengths are always higher than the reach of the best reach performance modulation format (QPSK), and so they require always the QPSK modulation.

Table 55 reports the yearly cost of two pairs of configurations. In each pair, one configuration is based on the fixgrid technology with 100 Gb/s transponders and the other one on the flexgrid with SBVT, assuming that the other conditions (span length and length threshold applied to put RAMAN) are the same for both configurations. The S100 topology (with spans having on average 100 km length) is chosen with the two options of No Raman and Raman amplification put on spans longer than 105 km.

Concerning the costs of the fixed grid technology configurations (named 1.fix and 2.fix in Table 55), for which all the cost components are known (as the cost model is available for all the network items), two main cost categories are highlighted, the OEO devices (transponders and regenerators) cost and a part that includes all the other costs. Since the SBVT cost is unknown and it has to be evaluated as target cost, the OEO cost for flexgrid configurations is evaluated subtracting from the total cost of fixed grid configuration the cost of flexgrid configuration components other than SBVTs. An upper bound of budget expendable for SBVTs (i.e. for all items needed in the network), which assures an identical total network yearly cost between fixed grid and flexgrid options, is thus obtained and reported in Table 55 (values in red for configurations 1.flex and 2.flex).

Table 55: Yearly cost of two pairs of (fixed grid, flexgrid) configurations.

S100 configuration	Yearly cost (ICU)		
	OEO cost (transponders and regenerators or SBVT)	all other cost components	Total



1.fix	No Raman - 100G	267.20	290.41	557.61
1.flex	No Raman - SBVT	277.62	279.99	557.61
2.fix	RAMAN >105km - 100G	234.00	299.35	533.35
2.flex	RAMAN >105km - SBVT	266.58	266.76	533.35

Table 56 reports the total number of OEO elements at network level for the considered configurations. As the modularity of SBVT is 1.2 Tb/s and the traffic demand is assigned with a resolution of 100 Gb/s, the total capacity allocated for tributaries are greater in the flexgrid cases due to the scraps. The unused capacity on SBVT, that is potentially allocable to an incremental traffic demand, is of the order of 16%.

Table 56: EOE parts in considered fix and flexgrid configurations.

	100 Gb/s fixed grid solutions		1.2 Tb/s SBVT flexgrid solutions	
	Transponders	Regenerators	SBVT	Multiflow modules
No Raman 100G	768	284	132	351
RAMAN >105km 100G	768	201	119	312

Table 11 and Table 12 report the target cost in ICU for the SBVTs evaluated respectively for NO Raman and for Raman amplification on spans greater than 105 km. These costs are CAPEX absolute cost (not yearly cost as in previous tables) and are referred to a fully equipped SBVT as well to the SBVT without the Optical Multi Flow Modules (i.e. only with the system electronic part plus the client transceivers) and the single Optical Multi Flow Module (OMFM, which includes the optical line parts). In lack of predictions about the SBVT cost structure, the target cost is evaluated considering a set of very different conditions in terms of cost partitioning between SBVT electronic parts and optical line parts (OMFMs). For instance, if the OMFM cost percentage is 60% of the total cost and given that an SBVT has three OMFMs, it means that 40% of the total cost is imputed to the SBVT electronic part plus client transceivers and 20% of the total cost is imputed to each of the three OMFMs (60% for the three).

Results from Table 11 and Table 12 show that the target cost for the complete SBVT is always very close to 12, which is 12 times the cost of a 100 Gb/s DP-QPSK transponder, with a slightly better solution for the configuration which use RAMAN (it has to be noticed that the higher the SBVT target cost, the more chance there is for the cost saving at network level because any cost reduction of SBVT has an impact on total cost reduction). This means that, at this first rough evaluation, it seems that there are no particular advantages (but neither penalties) using the flexgrid technologies on transnational networks, i.e. over networks requiring long-reach lightpaths that limit the use of high efficient modulation formats.

These results have to be confirmed with a more accurate design, that takes into account all the flexgrid model peculiarities (RSA, Add Drop chain of ROADM targeted to SBVT) and assuming different traffic conditions in term of global volume, distribution and client rates assortment.



8.6.5 Traffic grooming under fixed and elastic spectrum allocation

In this section we present a novel network planning methodology efficiently exploiting traffic grooming over flexgrid and elastic rate systems. We constructed a power consumption model and use OPEX and CAPEX as performance metrics in order to show that efficient hierarchical topologies can be constructed that can exploit the bandwidth granularity of the allocated transponders reducing overall network cost and power consumption.

8.6.5.1 Planning methodology

The problem we address is the following: given a network topology and source/destination traffic demands, appropriately dimension the link capacities and node ports to optimally serve the given demand. Optimization can be achieved by reducing fibres or the number of transceivers per node minimizing the total cost. IDEALIST's objective is to employ SBVTs having the potential of providing higher-rates at lower cost and in this study we will explore this trend. Additionally, exploiting optical technologies (e.g. WSSs) also at the switching layer for transparent routing of lightpaths across transit optical nodes (selected during the routing process) introduces another level of flexibility and can contribute towards minimization of electronic ports across the network. An example and the node architecture to achieve this are shown in Figure 93 in the following sections after having described the resource allocation and optimization approaches.

In this work we wish to quantify the gains attained using EONs and to do so, we proceed with a two-stage planning: in the first stage we employ a traffic grooming algorithm to minimize the network resources needed to service the specified traffic demand. In the second stage the resulting traffic matrix after grooming is routed by means of Routing and Wavelength or Spectrum assignment algorithms (RWA or RSA). It is pointed out that the RWA/RSA alone is not sufficient to optimise the network operation since a) it postulates the availability of an arbitrary optical bandwidth channel which although this might be feasible from optical engineering point of view, this may not be supported from the discrete set of electrical (client) granularity rates; b) it does not take into account the cost evolution trade-offs between an electronic port and an SBVT. In other words, it does not take into consideration the conditions under which transmission is more favourable to (electronic) switching and c) it does not take into account other distance/capacity trade-offs e.g. to transmit shorter distances at higher-rates supported with electronic grooming or to choose longer transparent distances at lower rates. In any case, the role of grooming algorithms is indispensable in a network planning procedure since it allows to aggregate lower granularity traffic into larger flows, creating a new traffic matrix across the network which is then routed using high-capacity transparent lightpaths between two new end points. In the following sections we present a comparative study exploiting these techniques.

8.6.5.2 Traffic grooming

Traffic grooming algorithms proposed in the past are limited to operate over fixed grid wavelength allocation only and are not usually proposed for planning but only for resource allocation given an existing preplanned network. In this section we describe a methodology that overcomes the above limitations. For multigranular networks employing optical switching functionality we selected two traffic grooming approaches: 1) The heuristic multi-



hop traffic grooming algorithm, which is detailed in [73]; 2) The hierarchical traffic grooming algorithm based on node clustering which is detailed in [74]. These approaches were presented in deliverable D1.3 section 6.3.3.

In [73] grooming is potentially performed at every node (for non-transparent paths). A heuristic algorithm establishes direct shortest paths for all demands until link capacity is exhausted (e.g. in Figure 93a a potential result of [73] under the Traffic Matrix shown in Figure 38 after allocating wavelengths and paths for demands 3->6, 6->8, 5->8) and routes remaining traffic to the next node where grooming facilities have been allocated during planning over spare capacity of established lightpaths (e.g. in Figure 93a grooming of lower capacity demands 1->6, 1->7 and 1->8 over the established paths and establishment of additional lightpaths across the remaining links when required). Whenever grooming of multiple demands over a single wavelength resource is performed, then the respective traffic flows are electronically processed and multiplexed at layer 2/3 (exploiting the functionality of a multi-granular switching node as shown in Figure 93) and then forwarded over the allocated optical channel. Grooming and RWA are performed simultaneously employing shortest path routing and first fit wavelength assignment. In [73] the same number of wavelengths is assumed on all links and requests can be dropped if not enough wavelengths exist.

In [2] grooming is performed at preselected cluster hubs/gateways only. Direct paths are also possible (only for "large" demands \square wavelength). Routing is performed after grooming of requests (as an independent process). Then the so called Longest first alternate path (LFAP) RWA is employed assuming the same number of wavelengths on all links and increasing the number of wavelengths on every link until all traffic is routed. In Figure 93b a potential result of [74], under the Traffic Matrix and a potential resulting clustered network topology of Figure 38, is shown. In Figure 93b the demands of clusters A, B and C are served by establishing direct optical paths from each cluster node to its corresponding gateway i.e. 3, 5 and 7 apart from demand 1->8, which justifies a direct lightpath (i.e. achieves near 100% utilization of the allocated wavelength/spectral slot) to be established between nodes 1 and 8. Each gateway node then establishes in the same manner direct lightpaths to distant cluster gateways and recipient nodes.

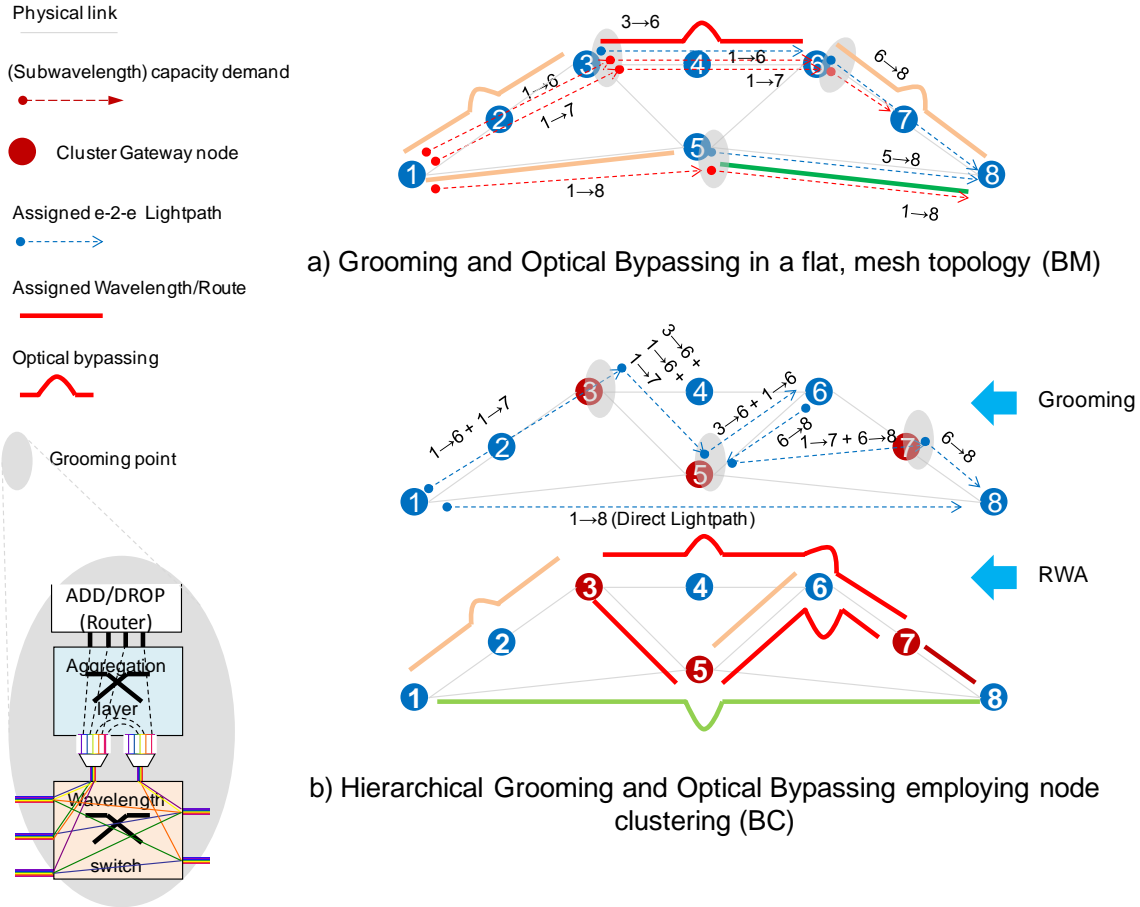


Figure 93: Example of traffic grooming based on multi-granular IP/L2/WDM employing WSSs for optical bypassing.

We extended the heuristic multi-hop traffic grooming algorithm of [1], which assumes a fixed network capacity, as a network planning using a binary search to find the minimal capacity required for the given traffic matrix. The corresponding algorithm we devised is completed in a number of successive steps:

- 1) Assuming fixed transportation granularities, the algorithm attempts to set-up a transparent lightpath between every source-destination node pair using shortest path routing, subject to constraints in the number of transceivers at the two nodes, at the designated rates, and the availability of a wavelength in a given path connecting the two nodes, in order of decreasing request capacity. In this case, the two end-points are connected by means of a single transparent lightpath.
- 2) If this is not the case, the algorithm attempts to forward the traffic of the blocked requests using the spare capacity of the existing links; that is, the unused capacity of established connections (at L2/L3).

We have modified the algorithm to operate under EON conditions using a different request order. Since the higher bit rate transceivers have limited reach, the requests that would be



served by them are searched in increasing path length order, ensuring that as many requests as possible are served by the high rate transponders.

Our traffic grooming algorithm is making use of the hierarchical implementation of [74], but it is extended to cover EONs as follows:

- a) The algorithm is decomposing the network into clusters; there is a single hub in each cluster and the hub node is used to groom the traffic from the cluster nodes that is forwarded to it. Should there is sufficient traffic demand between two end-nodes at a granularity supported by the OTN layer, then direct paths between the two end-nodes are set-up bypassing the grooming stages at source-destination hubs of the two clusters. To justify their role in hierarchical grooming, the hub nodes form a virtual cluster for grooming intra-cluster traffic. In effect, the grooming operation is a transformation of the traffic matrix from one with end-to-end requests to one with intermediate grooming points
- b) The transformed traffic matrix is then fed to the RWA algorithm that routes all requests as end to end paths over the network. Since the algorithm was developed for fixed-grid transportation it is the first time that its extension to EON is proposed. One important upgrade we performed is that the algorithm of [74] assumes that all requests are unidirectional, and the routing does not guarantee that bidirectional requests will be allocated the same path, especially where direct paths are involved. We modified the algorithm to allocate the return path of each request simultaneously with the direct path to avoid this issue.

8.6.5.3 Routing and Spectrum Assignment

To obtain consistent and realistic results, we used our custom RSA algorithm that forces a fixed number of spectral slots per fibre described in D1.3 section 6.3.3 [51]. The algorithm has two steps:

- First all requests are routed using a minimum cost network flow integer program, with first fit spectrum assignment. Subsequently, all unused fibre spans are removed from the network and not considered further.
- Then an optimization pass is used to improve the utilization of the transport network. The optimisation pass tries to reroute requests in a random order to reduce the average nodal degree (n.d.) of the network. A multiple restart stochastic hill climbing optimisation algorithm [76] is used.

8.6.5.4 Power consumption and cost modelling

The power consumption of the network is modelled using the values of [77] with a detailed model of CISCO CRS router family [78]. The model takes into account the number and kind of 400 Gb/s line cards (LC) used, the number of line card chassis (LCC) housing them and the number of fibre card chassis (FCC) necessary to interconnect more than 3 LCC. The muxponders (MXP) assumed are fixed rate for 40 Gb/s and 100 Gb/s in a 50 GHz channel and a sliceable 400 Gb/s one utilising a 100 GHz channel with a reach of 750 km. The power consumption of the three muxponders is 173.8 W, 243.4 W and 481.9 W respectively [67] and the smallest transceiver that will carry the required traffic is used in each case. Finally the model takes into account one optical amplifiers (AMP) per 80 km of



each fibre path and WSS based cross connects (WXC). The normalized cost of each component is calculated using the figures of [34] (again based on the CSR router family). The Cisco CRS router allows the connection of several line card chasses (LCC) together, using one or more fabric card chasses (FCC). Each LCC can house 16 line cards and each FCC can connect up to 9 LCCs. Up to 8 FCCs can be used, for a total of 72 LCCs and 1152 line cards. Each line card consists of two parts, the modular services card (MSC) and the physical layer interface module (PLIM). The LCC-1 uses 40 Gb/s line-cards, the CRS-3 140 Gb/s line-cards and the CRS-X 400 Gb/s line-cards. For up to 3 LCC, only 8 FCC fabric cards are required, while for more LCCs 24 FCC fabric cards are required. Finally, the CRS-3 introduced the capability to interconnect the first two LCCs back-to-back without using a FCC.

From D1.1. [67] and [79], we have an initial analysis for the CRS-1 LCC power consumption. (Note that the analysis is said to be for the CRS-3, however the power consumption for the fabric card mentioned is for the CRS-1: 206W [80]). This gives a max power consumption of 2668W. To obtain the CRS-3 LCC power consumption, we need to obtain the increased power consumption of the 8 higher data rate fabric cards, since the route processors have the same power rating and all other modules seem identical.

From [81] we have a maximum power consumption for a CRS-1 FCC with 24 fabric cards of 7036W. The same value of CRS-3 is 7660W [82]. Assuming the difference is due to the higher rate fabric cards, we get an increase of exactly 26W per card. The base value for the CRS-1 FCC fabric card is 229W [83], so the CRS-3 FCC fabric card should be rated at 255W, an 11% increase. Assuming a similar 11% increase for the CRS-3 LCC fabric card gives us 229W, and an LCC max power of 2852W.

Assuming a similar increase for the fabric cards of CRS-X and all other cards remaining the same, we get 3060W per LCC and 8356W per FCC. The results for the max power consumption are summarized in Table 57.

Table 57: Max power consumption

	CRS-1	CRS-3	CRS-X
LCC	2668W	2852 W	3060 W
FCC (without fabric cards)	1540 W		
FCC fabric card	229 W	255 W	284 W
Modular Services Cards [84]	350 W	446 W	650 W
4xOC-192 [80]	138 W		
16xOC-48 [80]	136 W		
4x10GE [85]	74 W (70W) ²		
14x10GE [85]		150 W (136 W) ¹	
1x100 Gb/s DWDM [86]		180 W	
40x10GE [87]			110 W

² Value in parenthesis is for short range XFP transceivers, which need 1 W less each



4x100GE/OTU-4 [88]			120 W
--------------------	--	--	-------

Table 58 shows the derived values for the typical power consumption. The values are taken to be 90% of the max, as in references [78] and [82].

Table 58: Typical power consumption figures (derived)

	CRS-1	CRS-3	CRS-X
LCC	2401W	2567 W	2754 W
FCC (without fabric cards)	1386W		
FCC fabric card	206 W	230 W	256 W
Modular Services Cards [84]	315 W	401 W	585 W
4xOC-192 [80]	124 W		
16xOC-48 [80]	122 W		
4x10GE [85]	67 W (63W) [†]		
14x10GE [85]		135 W (122 W) [†]	
1x100 Gb/s DWDM [86]		162 W	
40x10GE [87]			99 W
4x100GE/OTU-4 [88]			108 W

8.6.6 SERANO approach

EONs are aiming to improve spectral efficiency and allow dynamic adaptation to traffic volume changes. EONs today are employing WSSs to route spectral slots transparently in a multi-hop path. The WSS cannot provide for important on-the-fly adaptation functions like: change of carrier frequency; change of the modulation format to modify the optical bandwidth a flow may occupy and 3R regeneration, at the same time. These functions are only feasible via an electronic switch.

To overcome these limitations, without resorting to electronic switching, the SERANO architecture has been proposed (D2.1 section 3.1.2) that provides for these operations as well as a number of networking functions like: i) spectral slot conversion; ii) fine tuning of the optical reach which is necessary due to the trade-off between distance and modulation format in a multi-hop path; iii) spectral defragmentation and iv) system vendor interoperability which is important between different administrative domains. In particular, spectral slot conversion is essential to reduce blocking because of the frequent EON spectral slot expansions and contractions that may lead to similar blocking with the WDM case.

The main building block of SERANO is the bandwidth variable transponder (BVT). To efficiently introduce BVTs in EON systems, these need to be packaged, by means of PICs, into sliceable bandwidth variable transponders (SBVTs). The SBVTs are powerful subsystems, capable of handling multiple optical flows dynamically. In SERANO we take



advantage of this novel subsystem and use it as a building block offering the functionality that is currently missing in BV-ROADMs.

The architecture of the SERANO node has been previously reported in detail in D1.3 (section 5.1.1) [51]. The SERANO architecture is *flow* and *link modular*, since only the couplers need to be provisioned, and the building blocks (optical blades for new links and WDM cards for new flows) can be added progressively, as traffic increases, following a pay-as-you-grow principle. It is also worth noting that no optical or electronic switching technology is introduced since the main active building blocks are the SBVT arrays something that integrates transmission and data-forwarding allowing the industry to concentrate on optimizing a single element for both functions.

8.6.6.1 SERANO in the RMSA context

In striking contrast to typical WDM networks where traffic handling requires to consider both the grooming function as well as the RWA problem in EONs the latter function is reformulated as a Routing, Modulation and Spectrum Allocation (RMSA) problem. The RMSA algorithms/mechanisms enhance the typical RSA algorithms/mechanisms with the capability of selecting different modulation format (MF) for each network flow. This capability improves the network spectral efficiency at the cost of the need for support of different formats and a more complex problem formulation. In detail, the RMSA algorithm/mechanisms assign the appropriate modulation format among a set of alternative modulation formats based on predefined objectives (e.g., selecting the simpler modulation format that can reach the destination without regeneration and minimize the spectral occupancy as well).

In this context, the emerge of SERANO can become beneficial for EON networks in terms of improving spectral efficiency and blocking probability, cost reduction etc., because SERANO -in contrast to other solutions- can exploit the full potential of the RMSA algorithms/mechanisms. In the next part of this subsection, the capabilities of SERANO in the context of RMSA problem are presented and explained in detail.

The four major functionalities of SERANO node are:

- 3R regeneration: SERANO node can 3R regenerate (re-time, re-transmit, re-shape) a selected input signal (flow) before forwarding it to the appropriate output.
- Modulation format translation: By using SBVTs as the main building blocks, SERANO has the capability to translate the modulation format of a flow pre segment basis. That is, based on the following segment's transparent length, the SBVT may change the modulation format of a flow entering the SERANO node; either to upgrade it to a higher modulation format (e.g., from DP-QPSK to DP-16QAM) or downgrade to a simpler one (e.g., from DP-64QAM to DP-16QAM). In the first case the gain is the smaller spectral occupancy as less Frequency Slot Units (FSUs) [3] are required at the cost of limited transmission distance, while in the second case a simpler modulation format will support greater distances but with a waste in the number of FSUs (note: it is reminded that in SERANO the SBVT block consists of a BV Receiver connected back to back with a BV Transmitter so the two systems are/could be totally decoupled).
- Change of carrier central frequency: The SERANO node can modify the central carrier frequency of each flow entering the infrastructure. In this way, for example, if two or more flows compete to occupy the same frequency slot of an output fibre will not be collided, as it will happen otherwise, but they will be shifted to different non overlapping



spectral zones. This functionality is vital in order to improve network utilization and accomplish spectral defragmentation. Figure 94 schematically shows how SERANO could reconfigure two WDM cards with flows from two different input fibres towards a common output fibre, to implement spectral defragmentation. That is, the BVT transponders can flexibly shift any incoming spectral slot to any available slot in the output fibre increasing network utilization.

Moreover, an important feature of SERANO is that all three aforementioned functionalities can be concurrently applied to the same flow. For example, a congested network may resort to SERANO in order to improve the network's spectral efficiency and capacity by concurrently regenerating, modifying the modulation format and changing the central frequency of suitably chosen flows.

The aforementioned capabilities of SERANO lead to a new generation of RMSA algorithms in which the modulation and spectrum allocation part is not assigned statically from source to destination, but can potentially selectively change at one (or more) intermediate network nodes. Therefore an efficient RMSA algorithm should take into consideration the idiosyncrasy of SERANO in order to capitalise on the expected gains. The main objectives of such an RMSA algorithm/mechanism should be the following:

- *To minimize the blocking probability:* this is the main objective of network optimization operations. In the context of SERANO, the network operator would either proactively resort to the defragmentation provided from the block to reduce network-wide blocking or it will employ these features on-demand to resolve contentions in multi-hop routing schemes. For example, in a typical RMSA algorithm there is a case where a new request is not accepted because in only one link of a multi-hop path the specific spectral slot is unavailable. In this case, a SERANO module would offer a plethora of alternative operations to be used for overcoming the blocking and the waste/under-utilisation of the available resources, as we explain in more detail in the next paragraph.
- *Minimise the CAPEX/OPEX* needed to construct SERANO: the implementation of SERANO requires the deployment of BVTs that comes at an additional CAPEX/OPEX cost compared to nodes including only BW-WSSs. Therefore, there is a tradeoff between the cost due to the additional BVTs used and the network-wide gains attained from the employment of SERANO. Therefore, the RMSA algorithm would have the additional role here to fine tune of the network resources in order to maximize the gains from the adoption of employment SERANO, while retain affordable additional costs (e.g., minimize the total number of BVTs).
- *To maximize spectral efficiency:* this is also an important optimization target of the majority of RMSA algorithms. In the context of SERANO, this parameter can be further optimized by using the functionalities of SBVTs/BVTs (e.g., change of modulation format)

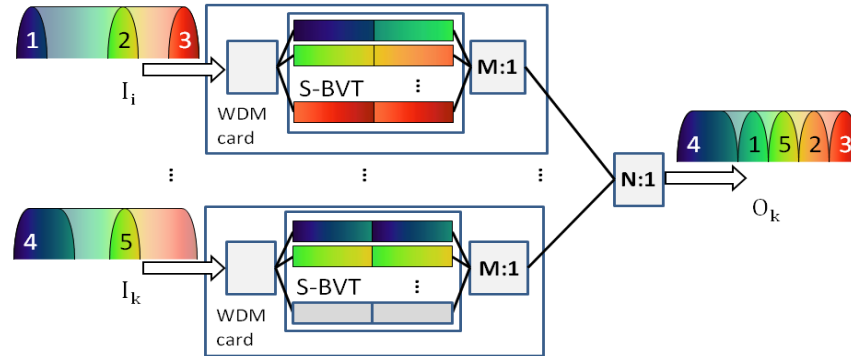


Figure 94: Exploitation of SERANO defragmentation properties.

8.6.6.2 An RMSA reinforcing SERANO's features for an agile optical layer

In this section, the most important network functions offered from SERANO will be elaborated in association to the adjacent RMSA objectives. SERANO is providing the following functions:

- Optical reach optimization
- Relaxation of frequency continuity constraint
- 3R regeneration placement

a) *The optical-reach optimisation* is implemented as follows in our RMSA algorithm:

- Firstly, the RMSA algorithm selects the most efficient modulation format for the requested transparent length (e.g., QPSK in long distances, 16QAM in short distances). This feature is already included in all typical RMSA algorithms/mechanisms and it is also incorporated in our model.
- In addition, our RMSA algorithm incorporates the features of SERANO, as they are presented allowing disintegrating the end-to-end path of a specific flow into a number of consecutive paths where different modulation formats are employed in each segment. This is schematically shown in Figure 95 featuring our RMSA algorithm that incorporates a SERANO block. The expected gains from the use of this algorithm are uncompromised (maximized) information transfer rate, optimisation of slot-slot usage and defragmentation.

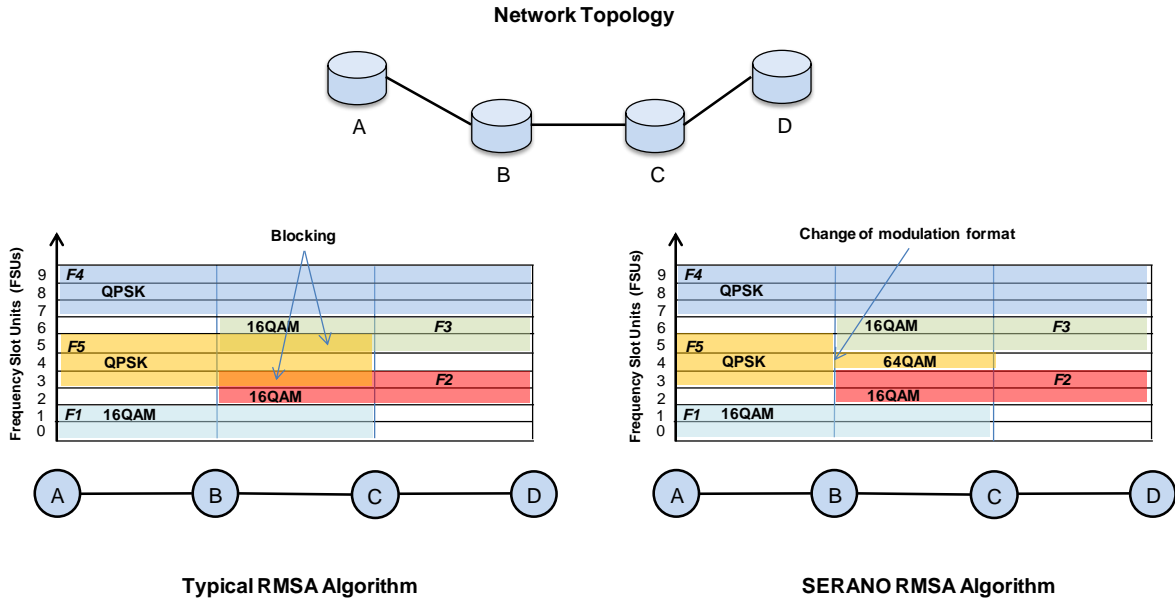


Figure 95: Exploitation of SERANO modulation format change functionality.

b) Relaxation of frequency continuity constraints: Another important feature of our RMSA algorithm is that it relaxes any frequency continuity constraints typical in other legacy RMSA counterparts. Indeed, in typical RMSAs, the central carrier frequency of a flow should be such as the flow occupies the same spectral slots throughout the path from source to destination. This is not the case in SERANO which it allows for:

- i. arbitrary selection of carrier's central frequency and
- ii. modulation format translation.

So in our RMSA, due to the former feature, some flows may occupy a group of successive spectral slot from source to an intermediate node and then, a group of different successive spectral slots from the intermediate node to the destination. In this case, the frequency continuity constrain is not requested for the entire path, but instead it is necessary only within the boundaries of the two consecutive paths. As such, new requests that will be blocked in typical RMSA mechanisms can now be granted under in SERANO (Figure 96). Due to the second feature, the use of different modulation formats per span allows to either expand or contract the number of successive spectral slots used in a path according to the RMSA objectives.

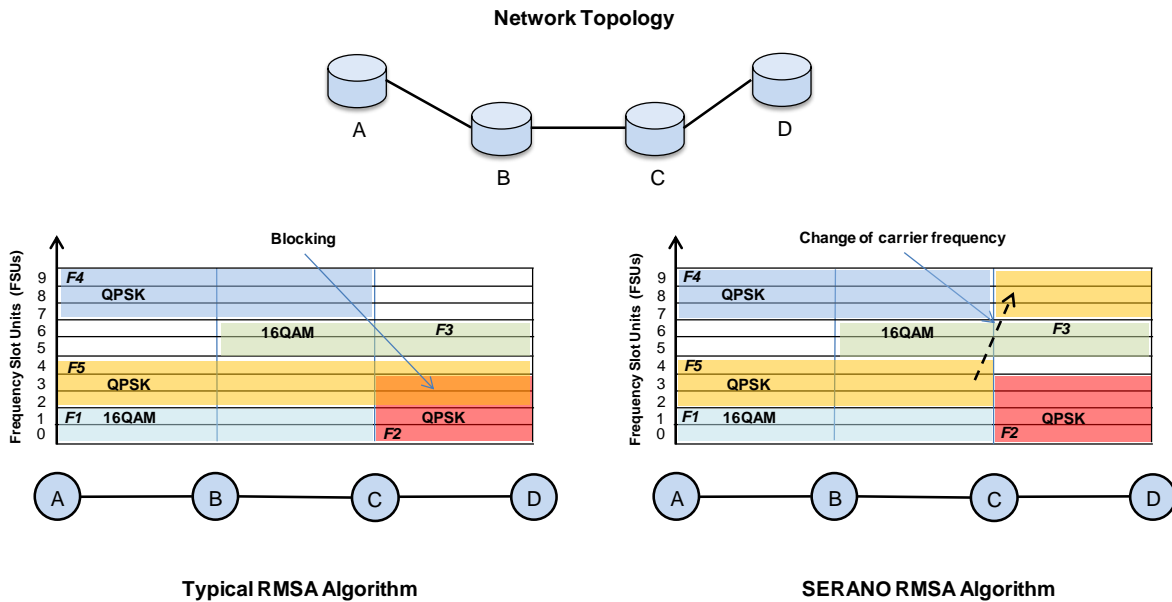


Figure 96: Exploitation of SERANO carrier frequency change functionality.

In an EON employing SERANO the change of flow's central carrier frequency is implemented pre-emptively or on-demand. In the pre-emptive mode, change of carrier's frequency is implemented on carefully selected flows during the network operation in order to retain a defragmented network. In the on-demand mode, changes are realized when new flows are emerged or existent flows need to be rerouted of the capacity of the flow is expand/contract. Then, SERANO modules are activated in intermediate nodes in order to change carrier frequencies and accept new / existing flows.

c) 3R regeneration placement: An important optimisation problem in EONs is the optimum placement of regenerators. In general, the regeneration placement problem and the RMSA problem can be studied separately or in combination. In the second case, the objectives of RMSA problem are updated with the objectives of the regeneration placement problem, typically the minimization of the number of regeneration elements in the network or/and the minimization of the number of regeneration nodes. In our RMSA algorithm, the two problems are examined as a joint optimization problem in order to take advantage of the inherent regeneration capability of SERANO (SBVTs support 3R regeneration). Therefore the RMSA decides whether a flow requires 3R regeneration, splitting the transparent length into two (or more) parts and then it decides if different modulation formats/carrier frequencies may be used in order to optimise a specific performance metric (e.g., minimum CAPEX). The first step is typical for all regeneration placement problems when the use of regenerators is inevitable (e.g., when the QPSK modulation format cannot support the source/destination distance), while the second step is possible thanks to the unique features of SERANO. As such, the combined RMSA-regeneration placement mechanism chooses to regenerate a flow at an intermediate node and also to employ a different modulation format and therefore to alter the slot occupancy of some critical network links, if, of course, the gains from this action surpass the cost of the regeneration. This is schematically illustrated in Figure 97.

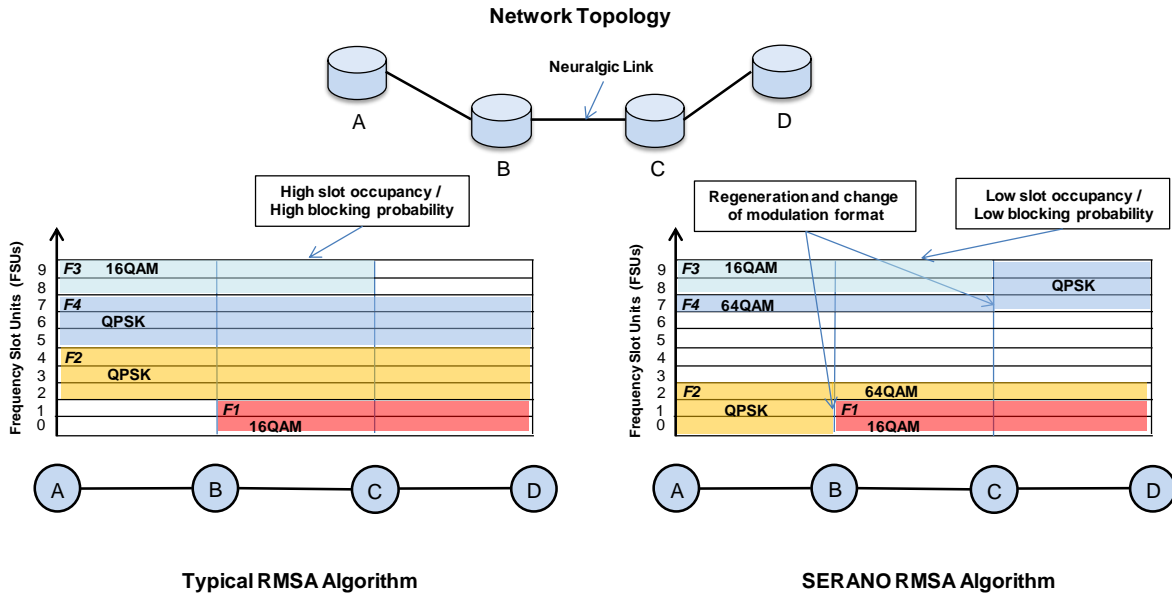


Figure 97: SERANO RMSA regeneration placement combined mechanism.

8.6.6.3 RMSA algorithms

The **off-line RMSA algorithm** starts from the assumption that all or a set of selected network nodes can support the SERANO functionality. On a first step, this global RMSA algorithm decides how many BVTs should be placed on each I/O pair of each SERANO node. This decision is based on one or more of several factors, like: a) the position of node/pair of links in the network graph (e.g., more central nodes, more BVTs), b) the nodal degree of nodes, c) the participation of node/pair of links in shortest paths or in k-shortest paths, d) take into consideration disjoint paths.

The RMSA algorithm decides on the optimal use of available BVTs in terms of 3R, change carrier frequency and change MF functionalities in order to accomplish specific objectives (e.g., maximize the spectrum occupancy). The total cost of the SERANO architecture is highly affected by the number of BVTs and not the total number of 3R, MF and frequency changes. Therefore, the RMSA algorithm dimensions the number of BVTs according to the jointly optimised requirement for 3R, MF and frequency changes. The performance indicators may include a number of parameters like blocking probability, spectral efficiency and/or cost against the number of BVTs implemented. The steps of off-line RMSA algorithm are illustrated in Figure 98.

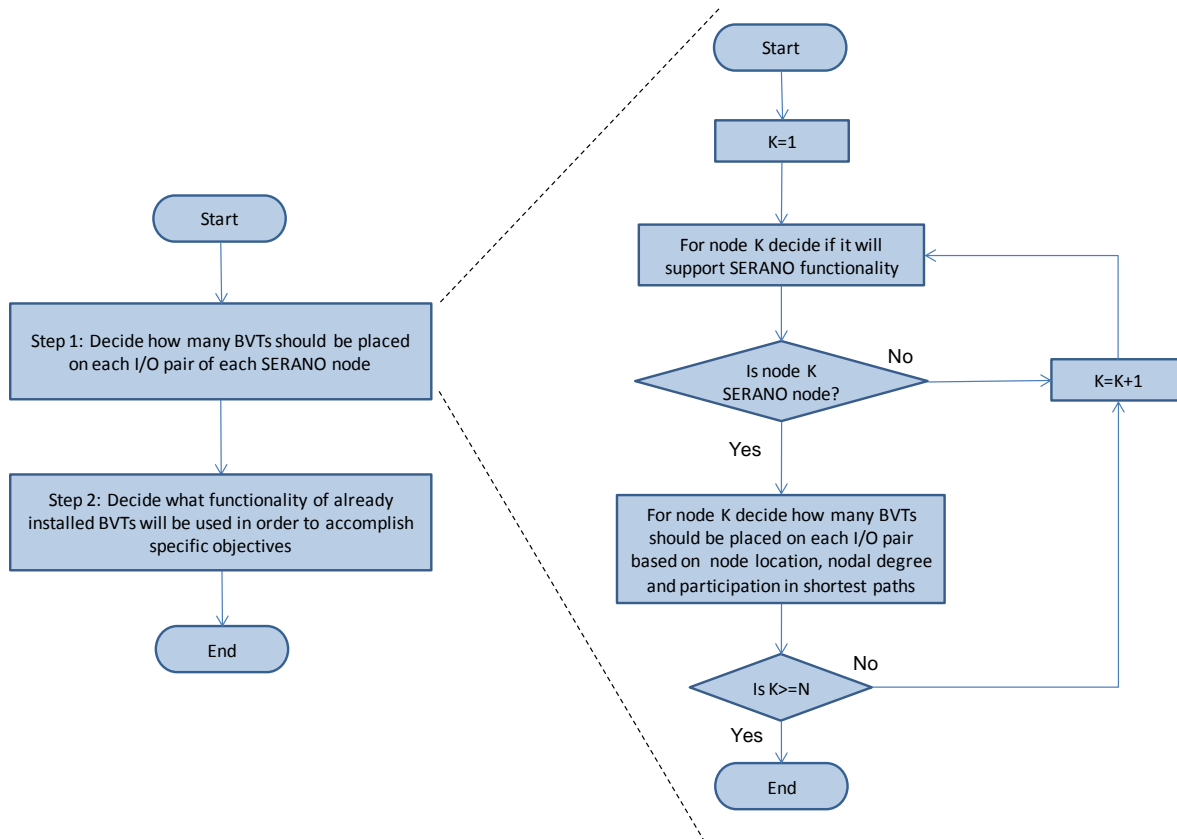


Figure 98: SERANO off-line RMSA algorithm steps.

The on-line RMSA algorithm takes as input a predefined number of SBVTs (probably from the previous design phase) and a predefined distribution of SBVTs in the network. Like the off-line RMSA algorithm, on-line algorithm decides how to use the available BVTs in order to 3R, change carrier frequency and change MF in order to accomplish specific objectives. Therefore when a new flow/request enters the network the algorithm takes actions (3R, MF change, and frequency change) in order to retain the network in an optimal state. For example in case of defragmentation functionality, the RMSA algorithm is continuously measuring the fragmentation (F) of the network (e.g., using a function to estimate the fragmentation F of a link) and takes appropriate actions in order to retain the network in a state of low fragmentation. In case of on-line RMSA algorithm we discriminate between (a) algorithm which can manipulate the previous flows and/or (b) algorithm that can only decide for the new flow.

The on-update RMSA algorithm is similar to off-line algorithm, with the addition that it receives as input the initial number and distribution of SBVTs in the network, while the estimation on additional traffic and new requirements are taken into consideration.



8.7 Main algorithm results

8.7.1 Model and solving methods for the GRANDE Problem

In this section we detail the solving methods for the GRANDE problem.

Mathematical formulation

Although the use of node-link formulation may seem more effective than using link-path formulation when topology design is involved, it is hard to solve even for moderate-size network instances as a result its large size in terms of variables and constraints. Although link-path formulations need pre-computed sets of paths for each demand, the path generation technique described in the next section can be applied to generate appropriate sets of paths that are required to achieve an optimal solution of the. For this very reason, we propose a link-path based formulation for the GRANDE problem.

The approach we follow to minimize the upgrading CAPEX consists in setting the cost of the already installed equipment to zero and minimize total network CAPEX. CAPEX is divided into different components, such as the cost of installing a new node or a new link.

We note that the constraint to ensure that the whole traffic matrix is transported could result in the problem infeasibility when not enough resources are available. This constraint makes it difficult to apply path generation, so in the proposed formulation we allow demand rejection but with a large cost penalty. Note that when too few pre-computed routes are used, the demands might be rejected.

The following sets and parameters are defined.

Topology:

- N set of OXCs, index n .
- E set of fiber links, index e .
- fn_n equal to 1 if an OXC is already installed in location n .
- fe_e equal to 1 if link e is already installed.
- cd_d penalty cost associated to demand d rejection.
- cn_n cost of installing a new OXC in location n .
- ct cost of adding a new link to an existing OXC.
- ce_e cost of installing link e .

Demands and paths:

- D set of demands, index d . For each demand d , the tuple $\{o_d, t_d, b_d\}$ is given, where o_d and t_d are the origin and target nodes, and b_d is the bitrate in Gb/s.
- P set of pre-computed paths, index p .
- $P(d)$ subset of pre-computed paths for demand d , $|P(d)|=k \forall d \in D$
- r_{pe} equal to 1 if path p uses link e .

Spectrum:

- S set of spectrum slices, index s .
- $C(d)$ set of pre-computed slots for demand d .
- q_{cs} equal to 1 if slot c uses slice s .

The decision variables are:

- w_d binary, equal to 1 if demand d cannot be served.
- x_{dpc} binary, equal to 1 if demand d is routed through path p and slot c .



y_n binary variable equal to 1 if an OXC is installed in location n .

The ILP formulation for the GRANDE problem is as follows:

$$(GRANDE) \min \sum_{d \in D} c d_d \cdot w_d + \sum_{e \in E} (1 - f_e) \cdot (c e_e + 2 \cdot c t) \cdot z_e + \sum_{n \in N} (1 - f_n) \cdot c n_n \cdot y_n \quad (24)$$

subject to:

$$\sum_{p \in P(d)} \sum_{c \in C(d)} x_{dpc} + w_d = 1, \quad \forall d \in D \quad (25)$$

$$\sum_{d \in D} \sum_{p \in P(d)} \sum_{c \in C(d)} r_{pe} \cdot q_{cs} \cdot x_{dpc} \leq z_e, \quad \forall e \in E, s \in S \quad (26)$$

$$\sum_{d \in D} \sum_{p \in P(d)} \sum_{c \in C(d)} g_{ne} \cdot r_{pe} \cdot x_{dpc} \leq |D| \cdot y_n, \quad \forall n \in N, e \in E \quad (27)$$

The objective function (24) minimizes the weighted sum of the cost of rejecting traffic (this term is assigned a coefficient c which is large with respect to the equipment costs) and the CAPEX cost of the additional equipment. It is worth highlighting that the value of the objective function is 0 when all demands can be routed using the already deployed nodes and links.

Constraint (25) ensures that either a lightpath is assigned to each demand or the demand is rejected. Constraint (26) guarantees that each slice in each link is used to convey one lightpath at most and, additionally, installs those links conveying any demand. Constraint (27) assures that an OXC is installed in those locations that are used by at least one lightpath. Note that the number of lightpaths using a node cannot exceed the number of demands.

The size of the GRANDE model is $O(|D| \cdot k \cdot |C| + |N| + |E|)$ variables and $O(|E| \cdot |S| + |E| \cdot |N| + |D|)$ constraints, in line to that of the LP-SA formulation.

As discussed above, the presence of not installed nodes and links in the original topology makes it difficult to pre-compute path sets. On the one hand, paths crossing the installed resources do not contribute to the CAPEX increment but increase probability of rejecting demands. On the other hand, paths containing non-installed resources lead to solutions with lower or zero demand rejection but increase the CAPEX cost. A way to solve this issue is presented in the next section, where a path generation algorithm designed to find the best paths to minimize the objective function by simultaneous minimization of demand rejection penalties and by reducing network CAPEX.

Path Generation Algorithm

Aiming at finding the paths leading to good-quality solutions, we propose a path generation algorithm based on the one described in [35]. We have modified that algorithm to stop when either the pricing algorithm finds no more new paths or a maximum number of iterations (*maxIter*) is reached. This allows controlling the size of the problem, making this size proportional to the computational effort needed to obtain the optimal integer solution. Thus, it could be reasonable to stop generating paths and devote more time to solve the final primal problem.

The specific pricing problem behind path generation for GRANDE is as follows. First we derive the dual of the primal (master) problem (25)-(27). We define the following dual



variables for the constraints of the relaxed GRANDE problem: λ_d unconstrained in sign for constraint (25), $\pi_{es} \geq 0$ for constraint (26), and $\mu_{ne} \geq 0$ for constraint (27). Finally, all variables in GRANDE are relaxed to be continuous $0 \leq x_{dpc}, w_d, y_n, z_e \leq 1$ in the master problem L , i.e., L contains the same set of variables and constraints as the original GRANDE formulation but with variables defined in the continuous domain.

From L , the dual problem can be easily derived from its Lagrangean function, which is obtained by moving the constraints to the objective function, multiplied by its associated dual variable. After grouping and re-ordering components, the Lagrangean function depending of primal and dual variables is as follows:

$$L(x, y, z, \lambda, \pi, \mu, \gamma, \rho) = \sum_{d \in D} \sum_{p \in P(d)} \sum_{c \in C(d)} x_{dpc} \cdot \left(\sum_{e \in E} r_{pe} \cdot \left(\sum_{s \in S} q_{cs} \cdot \pi_{es} + \sum_{n \in N} g_{ne} \cdot \mu_{ne} \right) - \lambda_d \right) +$$

$$\sum_{d \in D} w_d \cdot (cd_d - \lambda_d) + \sum_{n \in N} y_n \cdot \left((1 - fn_n) \cdot cn_n - |D| \cdot \sum_{e \in E} g_{ne} \cdot \mu_{ne} \right) +$$

$$\sum_{e \in E} z_e \cdot \left((1 - fe_e) \cdot (ce_e + 2 \cdot ct) - \sum_{s \in S} \pi_{es} \right) + \sum_{d \in D} \lambda_d \quad (28)$$

Now the dual problem (D) is defined as a maximization problem in the dual space, where the variables in brackets are the primal variables related to each constraint in the dual.

$$(D) \quad \max \quad \sum_{d \in D} \lambda_d \quad (29)$$

subject to:

$$[x_{dpc} \geq 0] \quad \lambda_d \leq \sum_{e \in E} r_{pe} \cdot \left(\sum_{s \in S} q_{cs} \cdot \pi_{es} + \sum_{n \in N} g_{ne} \cdot \mu_{ne} \right), \forall d \in D, p \in P(d), c \in C(d) \quad (30)$$

$$[y_n \geq 0] \quad |D| \cdot \sum_{e \in E} g_{ne} \cdot \mu_{ne} \leq (1 - fn_n) \cdot cn_n, \forall n \in N \quad (31)$$

$$[z_e \geq 0] \quad \sum_{s \in S} \pi_{es} \leq (1 - fe_e) \cdot (ce_e + 2 \cdot ct) \quad \forall e \in E \quad (32)$$

Since the objective of the pricing problem is to add paths not considered so far, this turns into adding new dual constraints so as to the current optimal dual solution infeasible. Looking at the formulation of the dual problem, constraint (30) is the only one defined for the path variables; therefore, the sole condition that a candidate path p^* to be added to the problem must violate this constraint. Let u_{dp^*c} be the reduced cost of demand d using new path p^* and slot c ; such p^* is a candidate path if and only if the reduced cost is strictly positive, i.e.

$$u_{dp^*c} = \lambda_d - \sum_{e \in E} r_{pe} \cdot \left(\sum_{s \in S} q_{cs} \cdot \pi_{es} + \sum_{n \in N} g_{ne} \cdot \mu_{ne} \right) > 0 \quad (33)$$

Therefore, in light of eq. (33), the pricing problem can be stated as follows: for each demand in D , find the path p^* and the slot $c \in C(d)$ that maximize the reduced cost u_{dp^*c} ,

provided that $u_{dp^*c} > 0$. Note that choosing the path with the highest positive reduced cost leads to the highest decrease rate in the objective function of the master problem.

An algorithm for solving the pricing problem is presented in Table 59. For each demand d , a shortest path is computed for each of the slots in $C(d)$. For a given slot c , link metrics (h) used to compute shortest paths are setup as follows (line 4):

$$h_e(c) = \sum_{s \in S} q_{cs} \cdot \pi_{es} + \sum_{n \in N} g_{ne} \cdot \mu_{ne} \quad (34)$$

Once the path is found (line 5), its reduced cost is computed using eq. (33) and, if it is higher than the incumbent cost, the path is stored as incumbent path. After exploring all possible slots for the demand, the incumbent path is included in the set P^* containing the generated paths for all demands (lines 6-9).

We have presented a path generation method as an alternative to pre-computing shortest paths following natural link metrics. However, when problem instances are excessively large preventing its application, alternative heuristic methods are needed. Next subsection presents the details of a BRKGA heuristic to solve the GRANDE problem.

Table 59: Pricing problem algorithm

INPUT: G, D, S , dual variables $[\lambda, \pi, \mu]$	
OUTPUT: P^*	
1:	for each d in D do
2:	$incCost \leftarrow 0$; $P^* \leftarrow \emptyset$
3:	for each slot c in $C(d)$ do
4:	Compute link metrics h_e
5:	$p^* \leftarrow$ Shortest path (o_d, t_d)
6:	Compute reduced cost u_{dp^*c}
7:	if $u_{dp^*c} > incCost$ then
8:	$incCost = u_{dp^*c}$
9:	$incPath = p^*$
10:	if $incPath \neq \emptyset$ then
11:	$P^* \leftarrow P^* \cup incPath$

BRKGA heuristic

The only problem-dependent parts of the BRKGA heuristic are the chromosome structure and the decoder algorithm. Table 60 presents the pseudo-code of the decoder algorithm. Such a decoder uses chromosomes to sort the set of demands and hence each chromosome contains one gene for each demand d in D .

Starting from the topology with all nodes and links available, the metric of each node and link, which will be afterwards used by the RSA algorithm, are properly initialized to promote the use of the installed nodes and links (lines 1-3). Next, the demands are sorted using the values of the genes in the input chromosome (line 4).

Then, the decoder finds a lightpath for each of the demands following the given order (lines 5-6). Since the RSA algorithm finds a route with the minimum cost, the installed equipment will be reused before installing new nodes and/or links. After allocating the

resources (line 9), metrics of the installed nodes and links are updated (lines 11-13). Finally, the fitness value is obtained by computing the CAPEX cost.

Table 60: Decoder algorithm

INPUT:	$G, D, \text{Chromosome } chr, costInstall$
OUTPUT:	$\text{Solution}, fitness$
1:	Initialize <i>Solution</i> with installed nodes and links
2:	Initialize metrics of nodes and links:
3:	if installed set to 0 otherwise to their cost
4:	Sort D according to genes in chr
5:	for each d in D do
6:	$d.lightpath \leftarrow \text{RSA}(G, d)$
7:	if $d.lightpath = \emptyset$ then
8:	return <i>INFEASIBLE</i>
9:	allocate(G, d)
10:	$\text{Solution}.D \leftarrow \text{Solution}.D \cup d$
11:	if <i>new nodes and/or link have been installed</i> then
12:	Set metrics of new installed equipment to 0
13:	$\text{Solution}.Equip \leftarrow \text{Solution}.Equip \cup \{\text{installed equip}\}$
14:	$fitness \leftarrow \text{ComputeCAPEX}(\text{Solution})$

8.7.2 Mathematical model for the P2MP-RSA Problem

The following ILP formulation is based on the work in [36] on RSA modelling for unicast demands. Since both, path and tree schemes are particular cases of the sub-tree scheme, the ILP formulation is based on the latter scheme with configuration parameters and data pre-processing for the solutions to enforce either path or tree schemes. Routing constraints build directed Steiner trees [55], while spectrum allocation is performed by assigning pre-computed slots to each computed Steiner tree. To obtain directed Steiner trees, one path and one slot is found for each sub-connection and those paths sharing links and slots are merged into a light-tree. To guarantee the feasibility in the use of SBVTs, the graph is augmented considering Tx and Rx SBVT modules as nodes. Links incident to Tx (E^{Tx}) and Rx (E^{Rx}) modules are configured with the same spectrum structure as links connecting OXCs, whereas those links leaving Tx and Rx modules (E^U) are un-capacitated.

The following sets and parameters are defined:

Topology and demands

V	Set of nodes, index v .
E	Set of links, index e .
E^{Tx}	Subset of links incident to Tx modules.
E^{Rx}	Subset of links incident to Rx modules.
E^U	Subset of un-capacitated links.
$l(e)$	Length of link e in km.
$\delta_{ve}^{\rightarrow}$	1 if link e leaves from node v .
δ_{ve}^{\leftarrow}	1 if link e arrives to node v .
D	Set of multicast demands, index d . Each demand d is defined by the tuple $\langle s_d,$



$T(d), b_d \succ$.

Spectrum

- S Set of spectrum slices, index s .
- C Set of spectrum slots, index c .
- $C(d)$ Set of spectrum slots for demand d .
- f_{es} Equal to 1 if slice s in link e is free.

Configuration

- B Capacity (in Gb/s) of each SBVT.
- F Maximum number of flows for each SBVT.
- R Maximum reach (in km) for any connection.
- α Cost multiplier.
- $\{p, q\}$ Binary parameters for scheme selection: path={1,0}; tree={0,1}; sub-tree={1,1}.

The decision variables are:

- x_{dec} Binary, equal to 1 if destination t of demand d is reached through link e and slot c ; 0 otherwise.
- y_{dec} Binary, equal to 1 if slot c is assigned to demand d in link e ; 0 otherwise.
- w_{dc} Binary, equal to 1 if slot c is assigned to demand d ; 0 otherwise.
- z_e Binary, equal to 1 if link e is used; 0 otherwise.
- u_d Binary, equal to 1 if demand d is blocked; 0 otherwise.

The P2MP-RSA formulation is as follows:

$$(P2MP-RSA) \quad \min \alpha \cdot \sum_{d \in D} b_d \cdot u_d + \sum_{e \in E^{Tx}} z_e \quad (35)$$

subject to:

$$\sum_{e \in E} \sum_{c \in C(d)} (\delta_{ve}^{\rightarrow} - \delta_{ve}^{\leftarrow}) \cdot x_{dec} = \begin{cases} (1 - u_d), & \forall d \in D, t \in T(d), v = s_d \\ 0, & \forall d \in D, t \in T(d), v \neq \{s_d, t\} \\ -(1 - u_d), & \forall d \in D, t \in T(d), v = t \end{cases} \quad (36)$$

$$\sum_{e \in E} (\delta_{ve}^{\rightarrow} - \delta_{ve}^{\leftarrow}) \cdot x_{dec} = 0, \quad \forall d \in D, t \in T(d), v \neq \{s_d, t\}, c \in C(d) \quad (37)$$

$$\sum_{e \in E} \sum_{c \in C(d)} l(e) \cdot x_{dec} \leq R, \quad \forall d \in D, t \in T(d) \quad (38)$$

$$\sum_{t \in T(d)} x_{dec} \leq (1 + q \cdot (|T(d)| - 1)) \cdot y_{dec}, \quad \forall d \in D, e \in E, c \in C(d) \quad (39)$$

$$\sum_{e \in E} y_{dec} \leq |E| \cdot w_{dc}, \quad \forall d \in D, c \in C(d) \quad (40)$$

$$\sum_{c \in C(d)} w_{dc} \leq 1 + p \cdot (|T(d)| - 1), \quad \forall d \in D \quad (41)$$

$$\sum_{d \in D} \sum_{c \in C(d)} q_{cs} \cdot y_{dec} \leq f_{es}, \quad \forall e \in E \cap \overline{E^u}, s \in S \quad (42)$$



$$\sum_{d \in D} \sum_{c \in C(d)} y_{dec} \leq F \cdot z_e, \quad \forall e \in E^{Tx} \cup E^{Rx} \quad (43)$$

$$\sum_{d \in D} \sum_{c \in C(d)} b_d \cdot y_{dec} \leq B \cdot z_e, \quad \forall e \in E^{Tx} \cup E^{Rx} \quad (44)$$

The objective function (35) minimizes both the amount of blocked demands and the number of used Tx modules.

Constraints (36) to (38) deal with sub-connection routing. Constraint (36) finds a directed path from source to destination using multi-commodity flow conservation equalities. Note that the whole demand is rejected if a sub-connection cannot be served. Constraint (37) ensures spectrum continuity in every intermediate node of every path. Finally, constraint (38) ensures that the maximum reach is not exceeded.

Constraint (39) is in charge of building light-trees by joining sub-paths of the same demand sharing the same slot in one or more links. By assigning parameter $q=0$, such sub-connection slot sharing is not permitted, which is one of the definitions of the path scheme. Otherwise, any amount of sub-connections serving a demand can share the same slot.

Constraints (40) and (41) control demands being split into several light-trees. Constraint (40) stores the number of distinct slots used for each demand, whilst constraint (41) makes sure that this number can be greater than one only if $p=1$ (path and sub-tree schemes). Otherwise, only one light-tree per demand is allowed (tree scheme).

Constraints (42) to (44) deal with capacity constraints. Constraint (42) makes sure that every slice in a capacitated link is assigned to at most one connection. Constraint (43) guarantees that the number of connections assigned to a SBVT module, does not exceed the maximum number of SBVT flows, whereas constraint (44) makes sure that the bitrate through each module does not exceed the SBVT's capacity.

8.7.3 Multi-Layer Network Planning Algorithm

In this section, we describe the multi-layer planning algorithm for IP over flex-grid optical networks. We decouple the problem and at the first step we jointly calculate the routing paths at the IP and optical layers (IP-R+RML), while at the second step we perform the spectrum allocation (SA). More specifically, in section A we present the IP-R+RML algorithm, and in Section B we describe the SA mechanism. In both cases the developed algorithms serve a single demand and thus in order to serve the whole traffic matrix we iteratively use them for all demands one-by-one to plan the whole network.

Section A. Step 1: Multi-layer routing planning algorithm

We now describe the multi-layer routing algorithm that solves the IP-R+RML planning problem. The planning algorithm takes as input the traffic matrix, orders the demands, and serves them one-by-one. At each step the single demand multi-cost algorithm takes as input the demand to be served (source, destination, and demand rate), the IP over flexible network described by graph G , the transponders feasible configurations (described by the related tuples), the state of the network, in terms of previously established lightpaths and IP routing decisions. The multi-cost algorithm creates for each type of link (inter-layer, optical and virtual) a cost vector, that incorporates information regarding both layers optical



and virtual (IP) layer. Then it carries out two phases: in the first phase it calculates the cost vectors of candidate paths from the source to the destination combining the cost vectors of links using an appropriate combination relationship. A domination relationship is used to prune the solution space by removing dominated paths that would not be selected by the optimization optimisation function (second phase). Then in second phase an optimization function is applied to the cost vectors of the found candidate paths, that transform the multi-cost vector into a scalar, and selects the optimum.

A.1 Computing the cost vector of a path

1) Cost vector of a link

Each link of the graph is assigned a cost vector consisting of 6 cost parameters, one of which is a vector of size $2m$, where m is the number of available transmission configurations of the tunable transponder. These parameters are the following:

- An integer variable D_l representing the length of the link. The length of a virtual (l_v) or AN inter-layer link (l_{ov} or l_{vo}) is equal to 0.
- A boolean variable T_l that shows if a tuneable transponder is used or not. T_l is equal to 1, if a transponder is used, otherwise it is 0. So, T_l is set to 1 in case of an inter-layer link (l_{ov} or l_{vo} , since we assume bidirectional transponders placed at the start and end of a lightpath), while it is set to 0, in case of optical (l_o) and virtual links (l_v).
- An integer variable C_l representing the transponder's cost. In case of an inter-layer link (both l_{ov} or l_{vo}), C_l is equal to the cost of the particular transponder used (remember that when multiple transponder types are available, we create a virtual link for each type), and 0 for other types of links (when $T_l = 0$).
- A float variable P_l representing the additive cost of the router (having as reference its cost up to this point), which again is non-zero only for an inter-layer link (both l_{ov} or l_{vo}), and zero otherwise. For the given state of the network, the IP/MPLS router at the start for l_{vo} or the end for l_{ov} of the link has a number of chassis, line-cards and free ports. A new line-card must be installed when all there are no free ports (installed line-cards are full), and a new chassis must be installed when all the slots of the chassis are used. So (i) in the case where there are free ports at the router, $P_l = 0$, (ii) when line-cards are full but the chassis is not then P_l is set to the cost of a new line-card, and (iii) when we need to add a chassis, P_l is set to the additive cost of the chassis and a line-card. Note that the cost of adding a chassis might not be linear.
- A vector $\bar{H}_l = ((r_1, d_1), (r_2, d_2), \dots, (r_m, d_m))$, defined only for a virtual to optical inter-layer link (l_{vo}), whose i -th element (r_i, d_i) records a transmission tuple, where d_i is the feasible reach for rate r_i for the specific transponder. These are taken from the transmission tuples of the corresponding transponder that the inter-layer link represents. Note that the vector \bar{D}_l is defined only for the inter-layer links from IP nodes to optical nodes, while it is zero for the other direction (optical to IP) and all other types of links. We define as r_{max} the maximum rate in \bar{D}_l .
- A boolean variable F_l which is equal to 1 if the link is a virtual link (which represents an established lightpath), and 0 otherwise.

Thus, the cost vector V_l characterizing a link l is given by:



$$V_l = (D_l, T_l, C_l, P_l, \bar{H}_l, F_l)$$

2) Cost vector of a path

A path in the graph is built link by link, so that at each step the current path is extended by adding a (virtual, physical or inter-layer) link. We assume that a path p consists of a number of sub-paths m , where each sub-path m is defined as the path between two consecutive IP nodes (a sub-path corresponds to a lightpath). A sub-path can be a virtual link (an already established lightpath), or a sequence of a virtual-to-physical inter-layer link, one or more physical links, and a physical-to-virtual inter-layer link, which corresponds to a new lightpath. For each sub-path of the latter category (new lightpath), the algorithm decides on its rate that is denoted by r_m , which is set equal to the maximum rate of the tuples available when reaching the sub-path's terminating IP node.

Similarly to a link, a path is characterized by a cost vector of 8 parameters: $V_p = (D_p, T_p, C_p, P_p, \bar{H}_p, F_p, R_p, *p)$, which are the previously described parameters of a link, plus a new parameter R_p denoting the set of the chosen rates r_m of the previous sub-paths, and a new parameter $*p$ defining the list of identifiers of the links that comprise the path p . Note that parameter D_p represents the optical length of the current sub-path, that the length of the optical links from the last IP/MPLS router. Assuming we extend path p with V_p cost vector by adding link l with V_l cost vector to obtain path p' . The cost vector of p' is calculated using the associative operator $\oplus_{l_{ov}}, \oplus_{l_{vo}}, \oplus_{l_o}, \oplus_{l_v}$, to V_p and V_l , depending on the type of the link l . To be more specific:

- if l is an optical link or a virtual link

$$V_{p'} = V_p \oplus_{l_o(l_p)} V_l = (D_p + D_l, T_p + T_l, C_p + C_l, P_p + P_l, \{(r_i, d_i) \in H_p \cap D_p + D_l \leq d_i\}, F_p + F_l, R_p, \{*p, l\})$$

- if l is a virtual-to-optical inter-layer link

$$V_{p'} = V_p \oplus_{l_{vo}} V_l = (0, T_p + T_l, C_p + C_l, P_p + P_l, \bar{H}_l, F_p + F_l, R_p, \{*p, l\})$$

- if l is an optical-to-virtual inter-layer link

$$V_{p'} = V_p \oplus_{l_{ov}} V_l = (D_p + D_l, T_p + T_l, C_p + C_l, P_p + P_l, \emptyset, F_p + F_l, \{R_p, \max_{r_i} \bar{H}_p\}, \{*p, l\})$$

Note that the values of the cost parameters of the links are also different depending on the type of link. Thus, e.g. although in both physical and virtual links the length of the new path p' is defined as the sum of the length of the path p up to that point plus the length of the additive link l ($D_{p'} = D_p + D_l$), D_l is zero for a virtual link and equal to the length of the fibre for an optical link. We used different definitions of associative operators for the links in order to perform special actions: to reset the length ($D_{p'} = 0$) and re-initialize the transmission tuple set ($\bar{H}_{p'} = \bar{H}_l$) when adding a virtual-to-optical link, to reset the transmission tuple set ($\bar{H}_{p'} = \emptyset$) and select the maximum rate ($R_{p'} = \{R_p, \max_{r_i} \bar{H}_p\}$) when adding an optical-to-virtual link.



A.2 Multi-Parametric Routing Algorithm

The proposed multi-parametric routing algorithm consists of 2 phases, the first of which computes the set of non-dominated candidate paths to serve the demand, while the second one selects the optimal path. We assume that the node where the algorithm is executed, knows the network topology, the current state of the network (lightpaths established, installed router modules) and the feasible transmission configurations of the available tunable (or non-tunable) transponders. The algorithm runs for a specific demand with source s and destination d virtual nodes of the graph G and a demanded rate r . In the case which a demand requires rate bigger than the one supported by the transponders, then it is split to sub-demands of the supported rates, and the algorithm is executed as many times. The algorithm constructs a reduced graph $G_A=(N_A, L_A)$ from G , where we keep all nodes and all links except for the virtual links (established lightpaths) that have remaining capacity lower than the demanded r .

1) Phase 1: Computing the set of non-dominated paths P_{n-d}

The goal of this phase is to find a set of candidate paths to efficiently serve the demand. The algorithm which computes the set of non-dominated paths P_{n-d} runs on the reduced graph G_A , and we denote by $s, d \in N_A$ its source and destination nodes. We also define:

- M_i as the set of cost vectors of the paths from node s to node $n_i \in N_A$.
- $M = \cup_{i \neq s} M_i$ as the set of all vectors to all nodes.
- $M_i^f \subseteq M_i$ as the set of *final* vectors to node n_i .
- $M^f = \cup_{i \neq s} M_i^f \subseteq M$ as the set of all *final* vectors to all nodes.
- $R_p^{\max} = \max_m (r_m \in R_p)$ the maximum rate among all sub-paths of p , that is among all elements of \bar{R}_p .

This algorithm also utilizes a domination relationship between paths that have the same source and destination node in order to reduce the number of paths considered. In particular, we say that a path p_1 dominates a path p_2 with the same source and destination nodes (notation $p_1 > p_2$), if the following relation holds:

$$p_1 > p_2 \text{ iff } L_{p_1} \leq L_{p_2} \text{ and } C_{p_1} + P_{p_1} \leq C_{p_2} + P_{p_2} \text{ and } F_{p_1} \leq F_{p_2} \text{ and } R_{p_1}^{\max} \geq R_{p_2}^{\max}$$

Table 61: Pseudocode of the algorithm which computes the set of non-dominated paths from the given source to the given destination.

```

 $P_{n-d} \leftarrow \text{Compute\_the\_set\_of\_Non\_Dominated\_Paths}(G_A, s, d, V)$ 
##Initialization
 $M^f = \{\}, M = \{\}$ 
for all links  $l \in L_A$  starting at  $s \in N_A$ ,  $M = M \cup V_l$ 
while  $M \neq M^f$ 
    ##Choose the optimum label
    find path  $p$  in  $M$  whose vector  $V_p$  has minimum network cost. In case of tie, select the one with
    minimum number of virtual links, etc.
     $n \leftarrow$  ending node of selected path  $p$ 
     $M^f = M^f \cup V_p$ 
    ##Obtain new paths and discard dominated paths
    for all  $n_j \in N_A$  neighbors of  $n$  (connected through  $l = (n, n_j) \in L_A$ )
        ##Obtain the cost vector of the new path  $p_j'$ 
         $V_{p_j'} = V_p \oplus V_l$ , with  $\oplus = \oplus_o$  or  $\oplus_v$  or  $\oplus_{vo}$  or  $\oplus_{ov}$  depending on the link
        ##Discard the new path  $p_j'$  if it runs out of transmission tuples
        if ( $l$  is an optical link) and  $(\max_i (r_i, d_i) \in H_{p_j'}) < r$ )
            check the next neighbor  $n_j$ 
        end if
        ##Check if the new path  $p_j'$  is dominated
        for all  $V_{p_j} \in M_i$  ( $p_j$  are paths ending at node  $n_j$ )
            if  $p_j' > p_j$  (" $>$ " is the domination relationship)
                check the next neighbor  $n_j$  (discard the new path  $p_j'$ )
            else if  $p_j' > p_j$ 
                 $M = M - \{V_{p_j}\}$  (discard the old path  $p_j$ )
            end if
        end for
         $M = M \cup \{V_{p_j'}\}$ , (add the new path  $p_j'$  to the set of paths)
    end for
end while
end algorithm (return  $P_{n-d} = M^f$ )

```

A path that dominates another path has smaller length, less additive network cost (cost of transponders and additive cost of routers), utilizes less virtual links and has higher maximum rate among its sub-paths. Since the optimization function f to be applied to the cost vectors of the paths in the second phase of the algorithm is monotonic in each of these costs, the dominated paths would have never been selected, so the solution space along with the execution time of the algorithm will be reduced, without losing optimal solutions.

The algorithm used to compute the set of non-dominated paths is described in Table 61. It is a generalization of Dijkstra's algorithm that only considers scalar link costs. It first obtains a non-dominated path between the origin and a direct neighbor node. This path is selected so as to have the smallest network cost, in case of a tie, the smallest number of

virtual links, etc. By definition, this path is non-dominated since the parameters that comprise the cost vectors are additive and non-negative and this path has at least one parameter smaller than the other paths. The algorithm marks this path as final, and extends this path through the outgoing links of its end node, so as to calculate new paths and their cost vectors using the appropriate associative operator \oplus , according to the added link. An additional check is performed when adding a physical link, where we also check if the extended path has at least one transmission tuple with higher rate than the required ($\max_i (r_i, d_i) \in H_{p_j} < r$), and discard it if this is not the case.

Then, the algorithm selects a non-final path between the origin and one of its neighbors, or between the origin and one of the neighbors of the previously considered neighbor, marks it as final, extends it using the corresponding outgoing links, calculates new paths, and so on. For each new path that is calculated, the algorithm applies the domination relationship between the new path and all the paths with the same end node that have been previously calculated. The new path is discarded if it is dominated by one of the previously calculated paths; otherwise, it is added to the set of non-dominated paths of the specific end node, and all the previously calculated paths that are dominated by it (if any) are discarded.

The basic difference with Dijkstra's algorithm is that a set of non-dominated paths between the source and each node is obtained instead of a single path; a node for which a path has already been found is not finalized (as in the simple Dijkstra case) since it may have to be considered again later. The algorithm finishes when no more paths can be extended and returns the set of non-dominated paths P_{n-d} that has been calculated between the source s and the destination d .

For example consider the network with graph G of Figure 99. We suppose a transponder configuration with 1.500 km reach for a transmission of 100 Gb/s. Also two lightpaths have been established: (a) one from node V_1 to V_3 through nodes $P_1 \rightarrow P_2 \rightarrow P_3$, with remaining capacity $R=80$ Gb/s and (b) one from node V_2 to V_7 through the nodes $P_2 \rightarrow P_5 \rightarrow P_7$, with remaining capacity $R=100$ Gb/s.

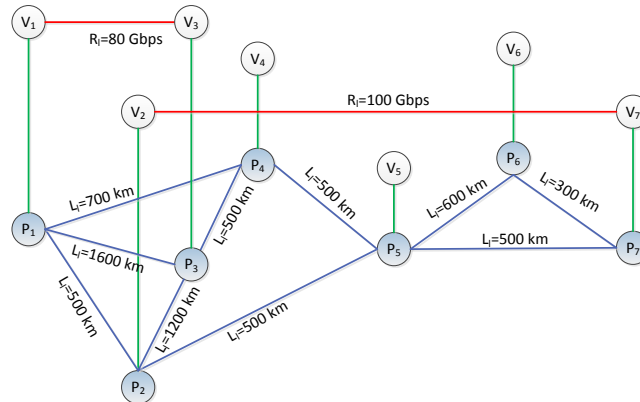


Figure 99: IP over flexible network example and related graph.

Assume that we want to serve a new demand with rate 100 Gb/s from node V_1 to node V_7 . We construct the reduced network graph G_A where we remove the virtual link $V_1 \rightarrow V_3$, as its remaining capacity is not sufficient.



The algorithm begins by initializing the set M with the link $V_1 \rightarrow P_1$. Then it examines from the set M the paths, selects $V_1 \rightarrow P_1$ (since it is the only available), and makes it final (adding it to set M^f). Then the algorithm expands $V_1 \rightarrow P_1$ to all the neighbors of node P_1 , that is, using the links $P_1 \rightarrow P_2$, $P_1 \rightarrow P_3$ and $P_1 \rightarrow P_4$. The path $V_1 \rightarrow P_1 \rightarrow P_3$ is discarded as its length is bigger than the maximum transmission reach (no transmission tuple has rate higher than the demanded). The cost vectors of the two new paths $V_1 \rightarrow P_1 \rightarrow P_2$ and $V_1 \rightarrow P_1 \rightarrow P_4$ are computed and inserted in the set M . Then the algorithm will search the set M , select $V_1 \rightarrow P_1 \rightarrow P_2$ that has the shortest length, finalize it and expand it. Links $P_2 \rightarrow V_2$, $P_2 \rightarrow P_3$ and $P_2 \rightarrow P_5$ will be examined. Path $V_1 \rightarrow P_1 \rightarrow P_2 \rightarrow P_3$ will be discarded, while the other two will be included in the set M . Note that the length of the path $V_1 \rightarrow P_1 \rightarrow P_2 \rightarrow V_2$ will be reset to zero, enabling it to reach P_3 at a later iteration. Also path $V_1 \rightarrow P_1 \rightarrow P_2 \rightarrow V_2$ when selected will also be expanded using the virtual link $V_2 \rightarrow V_3$ (previously established lightpath). The algorithm will continue its execution by selecting the best non-final path computed up to that point, and examining its neighbours. When more than one paths reach the same end node, it will apply the domination relationship to examine whether one is better than the other, and discard those that are worse. The execution of the algorithm will be completed when $M=M^f$, that is when there are no more non-final paths to expand.

2) Phase 2: Selecting the optimal path

In the second phase of the multi-parametric routing algorithm, we apply an optimization function $f(V_p)$ to the cost vector V_p of all paths calculated in the first phase. The function f yields a scalar cost per candidate path enabling us to select the optimal one. The function can be different for different demands, depending on their QoS requirements and other considerations. Note that the optimization function has to be monotonic in each of the cost components. We propose the following optimization functions:

- Optimization Function 1 (OptFun1): We select the path with the minimum length. In case of tie, select the one with the smallest number of virtual links.
- Optimization Function 2 (OptFun2): For each candidate path p we calculate the additive network cost ($C_p + P_p$) and select the one with smallest additive network cost. In case of tie, select the one with the smallest number of virtual links.
- Optimization Function 3 (OptFun3): For each candidate non-dominated path we calculate the total additive network cost. Using weights w_1 , w_2 and w_3 for the cost, the number of virtual links and the maximum supported rate, respectively, we select the path with the minimum weighted cost.

Section B. Step 2: Spectrum allocation

At this step we perform the spectrum allocation (SA) which must conform to the spectrum continuity constraint according to which a lightpath has to utilize the same spectrum segment (wavelength or spectrum slots) throughout its optical path.

The SA algorithm takes as input the selected path from the Step 1 and the spectrum utilization of the links. The selected path consists of one or more sub-paths that are defined as segments that start and end at virtual nodes. A sub-path is a new lightpath that needs to be established or an existing lightpath with sufficient remaining rate to serve the demand under study. The SA algorithm runs for each sub-path that represents a new lightpath. Note that for each such lightpath the algorithm has decided on the transmission



tuple and thus the spectrum slots that it requires. The SA algorithm calculates the slot utilization of the sub-path and finds the first placement that has enough spectrum slots to accommodate the lightpath. This SA algorithm is a variation of the heuristic proposed in [68].

Planning the whole network - iterative execution of algorithms

Steps 1 and 2 described above are designed to serve a single demand. To serve the whole traffic matrix we order the demands and serve them one-by-one. We keep track of the modules installed at the routers, the established lightpaths, the assignment of demands to them, their remaining capacity, and the spectrum utilization of the links. When a demand is served we update accordingly the above data structures and then move to serve the next demand.

Generality of the proposed algorithm

Although the algorithm described above is used for solving the joint multi-layer network planning problem at IP over flexible networks, it is quite generic and can be also used for (a) planning optical networks deploying fixed-grid or flex-grid optical switches and fixed optical transponders of single or multiple types (SLR or MLR) and (b) sequential planning of IP over flexible or fixed-grid optical networks [(what we call sequential multi-layer network planning, (*SML-NP*)). To be more specific, to plan a SLR or MLR network, the algorithm takes as input the description of the capabilities of the fixed transponders in the tuple format, discussed above. This is quite straightforward; the transponders of a SLR network are described with a single transmission tuple, while the transponders of a MLR network require as many tuples as the number of the different transponder types available. In case that we want to sequentially plan the IP and optical network (*SML-NP*), we consider that the IP routing decisions (IP-R) are taken without taking into account the distance constraints and the RML decisions. To implement this, in the case of *SML-NP* we set the reach of the flexible (or fixed) transponders equal to infinity and use an optimization function that does not account for the cost of the optical layer. When solving the RML problem we define the cost of the inter-layer links with different values than the ones defined in the *JML-NP* case: in the *SML-NP* case, inter-layer links are assumed to use transponders only at the source and the destination, while all other inter-layer links are assumed to be deployed with optical regenerators. Moreover, we use an optimization function that neglects the IP/MPLS router's costs.

8.7.4 Mathematical model for the SPRING problem

The mathematical model for the SPRING problem is presented next. The model takes as input the set of established LSPs that share common links with the shortest route for *newP*. The following sets and parameters are defined:

- E set of optical links, index e .
- P set of already established LSPs, index p .
- $E(p)$ subset of E with those links in the route of LSP p .
- $P(e)$ subset of P with those LSPs using optical link e .
- $P(p)$ subset of P with those LSPs sharing at least one link with LSP p . $P(p) = \bigcup_{e \in E(p)} P(e)$



Pm	subset of P with the candidate LSPs. $Pm = P(newP)$.
S	set of frequency slices, index s .
C	set of slots, index c . Each slot c contains a subset of contiguous slices.
$C(p)$	subset of C with those slots that can be allocated to p .
δ_{cs}	1 if slot c uses slice s , 0 otherwise.
ω_{pc}	1 if LSP p was using slot c , 0 otherwise.
η_{es}	1 if slice s in optical link e is free, 0 otherwise. Note that to compute η_{es} only LSPs in Pm are considered.
$\beta_{pp'}$	0 if originally the index of the first slice allocated to LSP p was lower than that of the first slice allocated to LSP p' , provided that p' was sharing at least one link with LSP p , i.e. $c < c'$, $c \in C(p)$, $c' \in C(p')$, $p' \in P(p)$. 1 if $c > c'$. Note that $\beta_{pp'}$ is not defined for those LSPs not sharing any link.

Additionally, the decision variables are:

x_{pc}	binary, 1 if slot c is allocated to LSP p , 0 otherwise.
y_p	binary, 1 if LSP p is shifted, 0 otherwise.

Then, the ILP for the SPRING problem is as follows:

$$(\text{SPRING}) \quad \text{minimize} \quad \sum_{p \in Pm} y_p \quad (45)$$

subject to:

$$\sum_{c \in C(p)} x_{pc} = 1 \quad \forall p \in Pm \cup \{newP\} \quad (46)$$

$$x_{pc} - \omega_{pc} \leq y_p \quad \forall p \in Pm, c \in C(p) \quad (47)$$

$$\sum_{p \in P(e)} \sum_{c \in C(p)} \delta_{cs} \cdot x_{pc} \leq \eta_{es} \quad \forall e \in E, s \in S \quad (48)$$

$$\sum_{c' \in C(p')} c' \cdot x_{p'c'} - \sum_{c \in C(p)} c \cdot x_{pc} < \beta_{pp'} \cdot |S| \quad \forall p \in Pm, p' \in P(p) \quad (49)$$

$$\sum_{c' \in C(p')} c' \cdot x_{p'c'} - \sum_{c \in C(p)} c \cdot x_{pc} > (\beta_{pp'} - 1) \cdot |S| \quad \forall p \in Pm, p' \in P(p) \quad (50)$$

The objective function (45) minimizes the number of LSPs that need to be reallocated in the spectrum so $newP$ can be served. Constraint (46) ensures that every candidate LSP and the received request for $newP$ have one slot allocated. Constraint (47) stores whether LSP p needs to be shifted comparing current and assigned slots. Constraint (48) guarantees that each frequency slice in every optical link is assigned to one path at most. Constraints (49) and (50) ensure feasible shifting by taking care of relative spectral position between LSPs pairs sharing any link, i.e. if two LSPs p and p' sharing a common link were originally allocated to slots with indexes $c < c'$ ($\beta_{pp'} = 0$), then that relative spectral position must be kept in the solution. The same must be ensured when $c > c'$ ($\beta_{pp'} = 1$). Note that when LSPs p and p' do not share any link, these constraints do not apply.

Although the size of the problem is limited –the number of variables is $O(|Pm| \cdot |C|)$ and the number of constraints is $O(|Pm|^2 + |Pm| \cdot |C| + |E| \cdot |S|)$ – it must be solved in real time, e.g. tens milliseconds, to minimize set-up delay of *newP*. For this very reason, we propose to use the heuristic algorithm described in Table 62.

Table 62: Algorithm for the SPRING problem

	INPUT: E, n_d
	OUTPUT: <i>Solution</i>
1:	$Solution \leftarrow \emptyset$
2:	$numLSPs \leftarrow INFINITE$
3:	for $s = 1.. S $ do
4:	$P^+ \leftarrow \emptyset, P^- \leftarrow \emptyset, P^s \leftarrow \emptyset$
5:	for each $e \in E(newP)$ do
6:	if $p(e, s-1) \neq p(e, s)$ then
7:	$P^- \leftarrow P^- \cup \{p(e, s^-)\}$
8:	$P^+ \leftarrow P^+ \cup \{p(e, s^+)\}$
9:	else
10:	$P^s \leftarrow P^s \cup \{p(e, s)\}$
11:	$\{shift, conn\} \leftarrow getMaxShift(P^-, P^s, P^+, s)$
12:	if $shift \geq n_d$ AND $conn < numLSPs$ then
13:	$Solution.P^- \leftarrow P^-$
14:	$Solution.P^s \leftarrow P^s$
15:	$Solution.P^+ \leftarrow P^+$
16:	$Solution.s \leftarrow s$
17:	$numLSPs \leftarrow conn$
18:	return <i>Solution</i>

The algorithm iterates on every frequency slice s to find the set of LSPs in the route of *newP* allocated using the closest slice with index lower than s (s^-), the set of LSPs allocated using the closest slice with index greater than s (s^+) and the set of LSPs allocated using s (lines 3-10 in Table 62).

Procedure *getMaxShift* find the largest continuous slot that can be generated by shifting LSPs (line 11). LSPs in set P^- are left shifted, LSPs in P^+ are right shifted, and LSPs in P^s are shifted left and right and the option generating the widest slot is chosen. If a slot with, at least, n_d contiguous slices by shifting LSPs is found and the set of LSPs involved is lower than that of the best solution found so far, the set of LSPs is stored as the best solution and the number of LSPs involved is updated (lines 12-17). The best solution found is eventually returned.

8.7.5 Mathematical models for the AFRO problem

As stated in the section 5.2.2, the AFRO problem consists in deciding, for each demand $d \in \mathcal{D}$, if it should be rerouted to a new lightpath, and if so, in determining a route over the network that uses the recently repaired link e^r . Moreover, spectrum allocation should be optimized. A goal of such reoptimization is to improve the allocation of slices on network links in order to provide the network with free spectrum resources for future connection requests.



We assume that regeneration requirements are met for all potential paths. Moreover, we assume that only one link can fail and be repaired at a time. However, the proposed method is not restricted to solely single link failure scenarios, and can be successively used to handle multiple failures, i.e., when e^r is not just a single link but a set of recently repaired links.

Notation

For the problem considered in this section, links and demands are directed. Let \mathcal{V}^2 denote the set of all two-element subsets of the set of nodes \mathcal{V} . Each link $e \in \mathcal{E}$ is represented by its end nodes $s(e)$ and $t(e)$ with $\{s(e), t(e)\} \in \mathcal{V}^2$.

The slices of a link are used by connections in the form of frequency slots. Slot c of capacity n is a set of contiguous (i.e., consecutive) slices of the form $c = \{s_i, s_{i+1}, \dots, s_{i+n-1}\}$ for some i between 1 and $S - (n - 1)$. In the sequel, \mathcal{C} will denote the set of all possible slots and $n(c), c \in \mathcal{C}$, the number of slices in slot c . Obviously, the number of all possible slots (for all $n = 1, 2, \dots, S$) is equal to $|\mathcal{C}| = \frac{S(S+1)}{2}$.

A lightpath (denoted by p) is a route in \mathcal{G} that uses the same slot c on all its links. The set of links composing the route of lightpath p is denoted by $\mathcal{E}(p) \subseteq \mathcal{E}$, and the slot assigned to this lightpath by $c(p) \in \mathcal{C}$.

Each demand $d \in \mathcal{D}$ is represented by its end nodes $s(d)$ and $t(d)$ with $\{s(d), t(d)\} \in \mathcal{V}^2$, and is characterized by demand bit-rate $h(d)$ and the currently used lightpath $p(d)$ with the route linking nodes $s(d)$ and $t(d)$. We use $n(d)$ to denote the bandwidth (i.e., a portion of the spectrum, in terms of the number of slices) able to serve bit-rate $h(d)$. Besides, each demand d is assigned a set of allowable lightpaths denoted by $\mathcal{P}(d)$. The set $\mathcal{P}(d)$ consists of the current lightpath $p(d)$ and new lightpaths. The new lightpaths have to fulfill a basic condition that each of them has to use the repaired link e^r , i.e., $e^r \in \mathcal{E}(p)$ for each $d \in \mathcal{D}, p \in \mathcal{P}(d) \setminus \{p(d)\}$.

Note that the set $\mathcal{C}(d) \subseteq \mathcal{C}$ of slots that can be used for demand $d \in \mathcal{D}$ is defined as follows: $c \in \mathcal{C}(d)$ if and only if $n(d) \leq n(c)$, which means that lightpaths allowable for demand $d \in \mathcal{D}$ can use only those slots whose capacity $b(n(c))$ is sufficient to carry the demand volume $h(d)$, i.e., $n(d) \leq n(c(p))$ for each $d \in \mathcal{D}, p \in \mathcal{P}(d)$.

IP formulations of the AFRO problem

We present two alternative IP formulations of AFRO. Although both formulations aim at improving the usage of spectrum in the network, still they implement different objective functions. The first formulation, referred to as MIN-SUM, aims at minimizing the total number of slices occupied on all network links. The second formulation, referred to as MIN-MAX, minimizes a sum of maximum identification numbers of used slices over all links. In the next section, we evaluate the effectiveness of both optimization objectives in improving the network performance.

The list of objects, parameters, and problem variables required for the problem formulation is given below.



Objects and parameters

\mathcal{V}	set of nodes, $V = \mathcal{V} $
\mathcal{E}	set of links, $E = \mathcal{E} $
$s(e)$	source node of link e
$t(e)$	termination node of link e
e^r	recently repaired link, $e^r \in \mathcal{E}$
\mathcal{S}	set of spectrum slices, $S = \mathcal{S} $
\mathcal{C}	set of frequency slots, $C = \mathcal{C} = \frac{S(S+1)}{2}$
$\mathcal{S}(c)$	set of (contiguous) slices composing slot $c \in \mathcal{C}$
$n(c)$	number of slices used by slot $c \in \mathcal{C}$
$b(n)$	bit-rate carried in a slot consisting of n slices
\mathcal{D}	set of demands, $D = \mathcal{D} $
$s(d)$	source node of demand d
$t(d)$	termination node of demand d
$h(d)$	bit-rate of demand $d \in \mathcal{D}$
$n(d)$	number of slices required to carry $h(d)$, $d \in \mathcal{D}$
$p(d)$	current lightpath used by demand $d \in \mathcal{D}$
$\mathcal{C}(d) \subseteq \mathcal{C}$	set of slots allowable for demand $d \in \mathcal{D}$, $c \in \mathcal{C}(d) \Leftrightarrow n(d) \leq n(c)$
$\mathcal{P}(d)$	set of lightpaths (new and current) allowable for demand $d \in \mathcal{D}$
$\mathcal{P} = \bigcup_{d \in \mathcal{D}} \mathcal{P}(d)$	set of all allowable lightpaths
$\mathcal{Q}(d, e, s) \subseteq \mathcal{P}(d)$	set of lightpaths for $d \in \mathcal{D}$ using slice $s \in \mathcal{S}$ on link $e \in \mathcal{E}$
$\mathcal{Q}(d, e) \subseteq \mathcal{P}(d)$	set of lightpaths for $d \in \mathcal{D}$ using link $e \in \mathcal{E}$
$\mathcal{E}(p)$	set of links traversed by lightpath $p \in \mathcal{P}$
$d(p)$	demand $d \in \mathcal{D}$ realized by lightpath $p \in \mathcal{P}$
$c(p)$	slot occupied by lightpath $p \in \mathcal{P}$, $c(p) \in \mathcal{C}(d)$
$n(p) = n(c(p))$	number of slices occupied by lightpath $p \in \mathcal{P}$
$L(p)$	cost of lightpath p , $L(p) = \mathcal{E}(p) \cdot n(p)$
$H(p)$	highest index of any slice used by lightpath p , $1 \leq H(p) \leq S$
$\mathcal{S}(p) = \mathcal{S}(c(p))$	set of slices used by lightpath $p \in \mathcal{P}$
κ_{es}	cost of using slice s in link e
\mathcal{K}	set of different sizes of demands in terms of a number of requested slices
$n(k)$	number of requested slices for demand size $k \in \mathcal{K}$

Problem variables

x_{dp} , $d \in \mathcal{D}$, $p \in \mathcal{P}(d)$	binary variables, $x_{dp} = 1$ when lightpath p carries its demand d
z_e , $e \in \mathcal{E}$	integer variables, maximum index of any used slice in link e

MIN-SUM formulation

Our first IP formulation of the problem can be written down in the following way:

$$\min \sum_{d \in \mathcal{D}} \sum_{p \in \mathcal{P}(d)} L(p) \cdot x_{dp} \quad (51)$$

$$[\alpha_d] \quad \sum_{p \in \mathcal{P}(d)} x_{dp} = 1 \quad d \in \mathcal{D} \quad (52)$$

$$[\beta_{es}] \quad \sum_{d \in \mathcal{D}} \sum_{p \in \mathcal{Q}(d, e, s)} x_{dp} \leq 1 \quad e \in \mathcal{E}, s \in \mathcal{S} \quad (53)$$

$$x_{dp} \in \{0, 1\} \quad d \in \mathcal{D}, p \in \mathcal{P}(d), \quad (54)$$

where variables in brackets, i.e. α_d and β_{es} , are dual variables.

Objective function (51) minimizes the sum of all slices used on all links. Constraint (52) assigns a lightpath (either new or the current one) to a demand. Here, we assume that



both the current paths and new possible paths for rerouting are in appropriate sets $\mathcal{P}(d)$. Constraint (53) makes sure that occupancy of the slices is not violated. Finally, constraint (54) assures that the variables are binary.

Observe that formulation (51)-(54) is based on path-flows and therefore it assumes a given set of allowable lightpaths \mathcal{P} . Clearly, the formulation is non-compact as it requires an exponential number of variables x_{dp} , $d \in \mathcal{D}$, $p \in \mathcal{P}(d)$, in order to consider all possible lightpaths for rerouting [60].

MIN-MAX formulation

Our second IP formulation of AFRO that we study is the following:

$$\min \sum_{e \in E} z_e \quad (55)$$

$$[\alpha_d] \quad \sum_{p \in \mathcal{P}(d)} x_{dp} = 1 \quad d \in D \quad (56)$$

$$[\beta_{es}] \quad \sum_{d \in D} \sum_{p \in Q(d,e,s)} x_{dp} \leq 1 \quad e \in E, s \in S \quad (57)$$

$$[\gamma_{edp}] \quad z_e \geq H(p) \cdot x_{dp} \quad e \in E, d \in D, p \in Q(d,e) \quad (58)$$

$$x_{dp} \in \{0,1\} \quad d \in D, p \in P(d), \quad (59)$$

where variables in brackets are dual variables.

The objective function (55) minimizes the sum of maximum identification numbers of used slices over all links. By minimizing the sum of z_e we assure that there are $S - z_e$ contiguous slices, with indices between $z_e + 1$ and S , available in link e and these slices can be assigned to future connection requests. Constraint (56) assigns a lightpath to a demand. Constraint (57) makes sure that occupancy of the slices is not violated. Constraint (58) finds the maximum index of any used slice in link e . Finally, constraint (59) assures that the variables are binary.

Formulation (55)-(59) describes our problem exactly. However, we formulate it also in an alternative way that allows for a simpler column generation, as it does not contain dual variables indexed by both edges and demands. The alternative formulation is as follows:

$$\min \sum_{e \in E} z_e \quad (60)$$

$$[\alpha_d] \quad \sum_{p \in \mathcal{P}(d)} x_{dp} = 1 \quad d \in D \quad (61)$$

$$[\beta_{es}] \quad \sum_{d \in D} \sum_{p \in Q(d,e,s)} x_{dp} \leq 1 \quad e \in E, s \in S \quad (62)$$

$$[\gamma_{es}] \quad z_e \geq \sum_{d \in D} \sum_{p \in Q(d,e,s)} H(p) \cdot x_{dp} \quad e \in E, s \in S \quad (63)$$

$$x_{dp} \in \{0,1\} \quad d \in D, p \in P(d), \quad (64)$$

where again, variables in brackets are dual variables, and the purpose of each constraint is as in formulation (55)-(59). Notice that now constraint equivalent to (58), i.e., (63) is indexed only by $e \in \mathcal{E}$ and $s \in \mathcal{S}$.



9 References

- [1] D. King and A. Farrel, "A PCE-based Architecture for Application-based Network Operations," IETF work-in-progress, 2014.
- [2] P. Demeester et al., "Resilience in Multilayer Networks", IEEE Communication Magazine, Vol. 37, no. 8, pp. 70-76, August 1999.
- [3] K. Lee, E. Modiano, H.W. Lee, "Cross-Layer Survivability in WDM-Based Networks", IEEE/ACM Transactions on Networking, vol.19, no.4, pp.1000-1013, Aug. 2011.
- [4] M. Ruiz, O. Pedrola, L. Velasco, D. Careglio, J. Fernández-Palacios, and G. Junyent, "Survivable IP/MPLS-Over-WSO Multilayer Network Optimization", IEEE/OSA Journal of Optical Communications and Networking, Vol. 3, pp. 629-640, 2011.
- [5] Eiji Oki, Kohei Shiimoto and Daisaku Shimazaki, "Dynamic Multilayer Routing Schemes in GMPLS-Based IP+Optical Networks," IEEE Communications Magazine, pp. 108-144, January 2005.
- [6] O. Gerstel, C. Filsfil, T. Telkamp, M. Gunkel, M. Horneffer, V. Lopez and A. Mayoral: Multi-Layer Capacity Planning for IP-Optical Networks, in IEEE Communications Magazine Feature Topic "Advances in Network Planning", January 2014, Vol. 52, Issue. 1, pp. 44-51.
- [7] E. Palkopoulou, D.A. Schupke and T. Bauschert, "Shared Backup Router Resources: Realizing Virtualized Network Resilience," in IEEE Communications Magazine, vol.49, no.5, pp.140,146, May 2011.
- [8] J.-F. Labourdette, E. Bouillet, and S. Chaudhuri, "Role of Optical Network and Spare Router Strategy in Resilient IP Backbone Architecture," 4th Int'l. Wksp. Design of Reliable Commun. Networks, Banff, Alberta, Canada, Oct. 19–22, 2003.
- [9] F. Muñoz, V. López, Ó. González de Dios and J. P. Fernández-Palacios: Multi-layer Restoration in Hierarchical IP/MPLS over WSON Networks, in Networks and Optical Communications (NOC), Jun 2012.
- [10] A. Mayoral, V.López, O.Gerstel, E. Palkopoulou, Ó. González de Dios and J. P. Fernández-Palacios: Minimizing resource protection in IP over WDM networks: Multi-layer Shared Backup Router, in Proc. Optical Fiber Conference (OFC), M3B.1, 2014.
- [11] A. Aguado, V. López J. Marhuenda, Ó. González de Dios and J. P. Fernández-Palacios: ABNO: a feasible SDN approach for multi-vendor IP and optical networks, in Proc. Optical Fiber Conference (OFC), Th3I.5, 2014.
- [12] L. Velasco, A. Asensio, J.Ll. Berral, V. López, D. Carrera, A. Castro, and J.P. Fernández-Palacios, "Cross-Stratum Orchestration and Flexgrid Optical Networks for Datacenter Federations," IEEE Network Magazine, vol. 27, pp. 23-30, 2013.
- [13] L. Velasco, A. Asensio, J.Ll. Berral, A. Castro, V. López, "Towards a Carrier SDN: An example for Elastic Inter-Datacenter Connectivity," (Invited Paper) OSA Optics Express, vol. 22, pp. 55-61, 2014.
- [14] L. Velasco, A. Asensio, J. Ll. Berral, E. Bonetto, F. Musumeci, V. López, "Elastic Operations in Federated Datacenters for Performance and Cost Optimization," Elsevier Computer Communications, vol. 50, pp. 142-151, 2014.



- [15] A. Asensio and L. Velasco, "Managing Transfer-based Datacenter Connections," IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 6, pp. 660-669, 2014.
- [16] A. Asensio, L. Velasco, M. Ruiz, and G. Junyent, "Carrier SDN to Control Flexgrid-based Inter-Datacenter Connectivity," submitted to Springer Photonic Network Communications, 2014.
- [17] A. Asensio et al., "Impact of Aggregation Level on the Performance of Dynamic Lighpath Adaptation under Time-Varying Traffic," in Proc. IEEE International Conference on Optical Network Design and Modeling (ONDM), 2013.
- [18] CPLEX, <http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>.
- [19] Europe's Energy Portal: <http://www.energy.eu/>
- [20] X. Fan, W. Weber, and L. A. Barroso, "Power Provisioning for a Warehouse-sized Computer", In Proc. of ACM International Symposium on Computer Architecture (ISCA), 2007.
- [21] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in Proc. ACM SIGCOMM, 2008.
- [22] I. Goiri, et al. "GreenHadoop: Leveraging Green Energy in Data-Processing Frameworks," in Proc EuroSys, 2012.
- [23] The Green Grid: www.thegreengrid.org/
- [24] Hewlett-Packard, <http://www.hp.com/>
- [25] Huawei, <http://www.huawei.com>
- [26] D. King, W. Boyson and J. Kratochvil, "Photovoltaic Array Performance Model," Sandia National Laboratories Report, SAND2004-3535, 2004.
- [27] Meta-Wiki.
http://meta.wikimedia.org/wiki/User:Stu/comScore_data_on_Wikimedia#Geographic_breakdown.
- [28] U.S. Department of Labor, Average Energy Prices, <http://www.bls.gov/ro5/aepchi.htm>
- [29] US Department of Energy, http://apps1.eere.energy.gov/buildings/energyplus/weatherdata_about.cfm
- [30] U.S. Department of Energy. US Energy Information Administration. 2009. Web site <http://www.eia.doe.gov/>.
- [31] Verizon, <http://www.verizonbusiness.com/about/network/latency/>
- [32] T. Tanaka et al. "Performance Evaluation of Elastic Optical Networks with Multi-Flow Optical Transponders" Tu.3.D.2. ECOC 2012
- [33] Information Resources Management Association, Networking and Telecommunications: Concepts, Methodologies, Tools and Applications.
- [34] F. Rambach, B. Konrad, L. Dembeck, U. Gebhard, M. Gunkel, M. Quagliotti, L. Serra and V. López, "A Multi-Layer Cost Model for Metro/Core Networks", IEEE/OSA Journal of Optical Communications and Networking, vol 5, pp. 210 – 225, 2013.
- [35] M. Ruiz, M. Pióro, M. Zotkiewicz, M. Klinkowski, and L. Velasco, "Column Generation Algorithm for RSA Problems in Flexgrid Optical Networks," Springer Photonic Network Communications, vol. 26, pp. 53-64, 2013.
- [36] L. Velasco, A. Castro, M. Ruiz, and G. Junyent, "Solving Routing and Spectrum Allocation Related Optimization Problems: from Off-Line to In-Operation Flexgrid Network Planning," (Invited Paper) IEEE/OSA Journal of Lightwave Technology (JLT), vol. 32, pp. 2780-2795, 2014.



- [37] Idealist Deliverable D1.1 - Elastic Optical Network Architecture: reference scenario, cost and planning, 2013.
- [38] V. Lopez et al, "Finding the Target Cost for Sliceable Bandwidth Variable Transponders," IEEE/OSA J. of Optical Comm. and Netw., vol. 6, pp. 476-485, 2014.
- [39] N. Sambo et al, "Demonstration of Data and Control Plane for Optical Multicast at 100 and 200 Gb/s with and without Frequency Conversion," IEEE/OSA J. of Optical Comm. and Netw., vol. 5, pp. 667-676, 2013.
- [40] R. Lin, M. Zukerman, G. Shen, and W. Zhong, "Design of Light-Tree Based Optical Inter-Datacenter Networks," IEEE/OSA J. of Optical Comm. and Netw., vol. 5, pp. 1443-1455, 2013.
- [41] R. Wang and B. Mukherjee, "Provisioning in Elastic Optical Networks with Non-Disruptive Defragmentation," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 31, pp. 2491-2500, 2013.
- [42] F. Agraz, et al., "Design and Implementation of a GMPLS-Controlled Grooming-capable Optical Transport Network", IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 1, pp. A258-A269, 2009.
- [43] A. Castro, L. Velasco, M. Ruiz, M. Klinkowski, J. P. Fernández-Palacios, and D. Careglio, "Dynamic routing and spectrum (re)allocation in future flexgrid optical networks", Comp. Networks, vol. 56, pp. 2869-2883, 2012.
- [44] A. Castro, F. Paolucci, F. Fresi, M. Imran, B. Bhowmik, G. Berrettini, G. Meloni, A. Giorgetti, F. Cugini, L. Velasco, L. Poti, and P. Castoldi, "Experimental Demonstration of an Active Stateful PCE Performing Elastic Operations and Hitless Defragmentation," in Proc. European Conference on Optical Communication (ECOC), 2013.
- [45] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP tunnels," IETF RFC 3209, 2001.
- [46] F. Cugini, F. Paolucci, G. Meloni, G. Berrettini, M. Secondini, F. Fresi, N. Sambo, L. Poti, and P. Castoldi, "Push-pull defragmentation without traffic disruption in flexible grid optical networks," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 31, pp. 125-133, 2013.
- [47] A. Castro, L. Velasco, J. Comellas, G. Junyent, "Dynamic Restoration in Multi-layer IP/MPLS-over-Flexgrid Networks," in Proc. IEEE Design of Reliable Communication Networks (DRCN), 2013.
- [48] Y. Sone, A. Watanabe, W. Imajuku, Y. Tsukishima, B. Kozicki, H. Takara, and M. Jinno, "Bandwidth Squeezed Restoration in Spectrum-Sliced Elastic Optical Path Networks (SLICE)," IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 3, pp. 223-233, 2011.
- [49] A. Castro, L. Velasco, J. Comellas, and G. Junyent, "On the benefits of Multi-path Recovery in Flexgrid Optical Networks," accepted in Springer Photonic Network Communications, 2014.
- [50] Idealist Deliverable D1.4 - Design and Tests of the On-Line Optimisation Framework, 2014.
- [51] Idealist Deliverable D1.3 - Preliminary Results and Improvement of Reference Scenarios, 2014.
- [52] M. Ruiz and L. Velasco, "Performance Evaluation of Light-tree Schemes in Flexgrid Optical Networks," accepted in IEEE Communications Letters, 2014.



- [53] Q. Wang and L. Chen, "Performance Analysis of Multicast Traffic over Spectrum Elastic Optical Networks," in *Proc. OFC*, 2012.
- [54] L. Gong et al., "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *IEEE/OSA J. of Optical Comm. and Netw.*, vol. 5, pp. 836-847, 2013.
- [55] M. Charikar et al., "Approximation Algorithms for Directed Steiner Problems," *Journal of Algorithms*, vol. 33, pp. 73-91, 1999.
- [56] M. Klinkowski, M. Pióro, M. Żotkiewicz, M. Ruiz, and L. Velasco, "Valid Inequalities for the Routing and Spectrum Allocation Problem in Elastic Optical Networks", in *Proceedings of the 16th International Conference on Transparent Optical Networks (ICTON)*, Graz, Austria, July 6-10, 2014.
- [57] A. Atamtürk et al., "Conflict graphs in solving integer programming problems," *Europ. J. of Operat. Res.*, vol. 121, no. 1, pp. 40–55, 2000.
- [58] J. Desrosiers and M. E. Lübbecke, "Branch-Price-and-Cut Algorithms," in *Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, 2011.
- [59] J. Konc and D. Janezic, "An improved branch and bound algorithm for the maximum clique problem," *MATCH Commun. Math. Comput. Chem.*, vol. 58, pp. 569–590, 2007.
- [60] M. Pióro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. 1em plus 0.5em minus 0.4em Morgan Kaufmann, 2004.
- [61] P. Wright, M. C. Parker, and A. Lord, "Simulation results of shannon entropy based flexgrid routing and spectrum assignment on a real network topology," in *Proc. of ECOC*, London, UK, Sep. 2013.
- [62] M. Ruiz, M. Żotkiewicz, A. Castro, M. Klinkowski, L. Velasco, and M. Pioro, "After Failure Repair Optimization in Dynamic Flexgrid Optical Networks," in *Proc. IEEE/OSA Optical Fiber Communication Conference (OFC)*, 2014.
- [63] M. Żotkiewicz, M. Ruiz, M. Klinkowski, M. Pioro, and L. Velasco, "Reoptimization of Dynamic Flexgrid Optical Networks After Link Failure Repairs," submitted to *JOCN*.
- [64] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks," in *Proc. of DRCN*, Ischia Island, Italy, Oct. 2005.
- [65] Idealist Deliverable D2.1 - Assessment of all potential flexgrid/flexrate data plane technologies including transmission & switching techniques, 2013.
- [66] Idealist Deliverable D2.2 - Modelling and Experimental Work on the Chosen Flexgrid/Flexrate Technology Options, 2014.
- [67] Idealist Deliverable D1.1 - Elastic Optical Network Architecture: reference scenario, cost and planning, 2013.
- [68] K. Christodoulopoulos, P. Soumplis, E. Varvarigos, "Planning flexible optical networks under physical layer constraints", *IEEE/OSA Journal of Optical Communications and Networking*, 2013.
- [69] F. Rambach, V.B. Konrad, L. Dembeck, U. Gebhard, M. Gunkel, M. Quagliotti, L. Serra, V. Lopez, "A Multilayer Cost Model for Metro/Core Networks", *IEEE/OSA Journal of Optical Communications and Networking*, 2013.
- [70] M. Svaluto, J. M. Fabrega, F. J. Vilchez, L. Nadal, V. López, G. Junyent, "Experimental Validation of an Elastic Low-Complex OFDM-Based BVT for Flexi-Grid



- Metro Networks”, European Conference and Exhibition on Optical Communication (ECOC), 2013.
- [71] M. Svaluto Moreolo, J. M. Fabrega, F. J. Vilchez, K. Christodouloupoulos, E. Varvarigos, V. López, J. P. Fernández-Palacios, “Assessment of Flexgrid Technologies in the MAN for Centralized BRAS Architecture Using SBVT”, European Conference and Exhibition on Optical Communication (ECOC) 2014.
 - [72] N. Sambo, A. D'errico, C. Porzi, V. Vercesi, M. Imran, F. Cugini, A. Bogoni, L. Potì, P. Castoldi, “Sliceable transponder architecture including multiwavelength source,” IEEE/OSA Journal of Optical Communications and Networking, vol.6, no.7, pp.590-600, July 2014
 - [73] K. Zhu, Bi. Mukherjee, “Traffic Grooming in an Optical WDM Mesh Network”, IEEE Journal on Selected Areas in Communications, vol. 20, no. 1, pp.122-133, 2002.
 - [74] B. Chen G. Rouskas, R. Dutta, “Clustering for Hierarchical Traffic Grooming in Large-Scale Mesh WDM Networks”, J. of Opt. Comp. Networking, vol. 2, no. 8, pp. 502, 2010.
 - [75] O. Gerstel, M. Jinno, A. Lord, S.J.B Yoo, “Elastic optical networking: a new dawn for the optical layer?” IEEE Communications Magazine, vol. 50, no. 2, pp. s12 - s20, Feb. 2012.
 - [76] A. Juels, M. Wattenberg, “Stochastic Hill climbing as a Baseline Method for Evaluating Genetic Algorithms”, University of California - Berkeley, CSD94-834 (1994)
 - [77] Ward Van Heddeghem et.al. “Power consumption modeling in optical multilayer networks”, Photon Netw Commun (2012) 24:86–102
 - [78] www.cisco.com/go/crs
 - [79] W. Van Heddeghem, F. Idzikowski “Equipment power consumption in optical multilayer networks – source data”
 - [80] “Cisco CRS-1 Carrier Routing System 16-Slot Line Card Chassis System Description”, March 2008
 - [81] “Cisco CRS-1 24-Slot Fabric-Card Chassis”, http://www.cisco.com/c/en/us/products/collateral/routers/crs-1-16-slot-single-shelf-system/product_data_sheet0900aecd80340baa.html
 - [82] “Cisco CRS-3 24-Slot Fabric-Card Chassis Data Sheet”, http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/data_sheet_c78-408226.html
 - [83] “Cisco CRS-1 Carrier Routing System Multishelf System Description”, February 2007
 - [84] “Cisco CRS Modular Services Cards Data Sheet”, <http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/datasheet-c78-730791.html>
 - [85] “Cisco CRS Carrier Routing System 16-Slot Line Card Chassis System Description”, January 2012
 - [86] “Cisco CRS 1-Port 100 Gigabit Ethernet Coherent DWDM Interface Module Data Sheet”, http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/datasheet_c78-478689_0629.html
 - [87] “Cisco CRS 40x10 Gigabit Ethernet Interface Module Data Sheet”, <http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/datasheet-c78-730792.html>



- [88] "Cisco CRS 100 Gigabit Ethernet Interface Modules Data Sheet", <http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/datasheet-c78-730788.html>
- [89] T. Orphanoudakis et. al., IEEE Communications Letters, vol. 17, pp. 785, 2013.
- [90] A. Klekamp, U. Gebhard, "Performance of elastic and mixed-line-rate scenarios for a real IP over DWDM network with more than 1000 nodes [invited]," IEEE/OSA Journal of Optical Communications and Networking, vol.5, no.10, pp.A28,A36, Oct. 2013.
- [91] A. Autenrieth, "Recovery time analysis of differentiated resilience in MPLS," Proc. DRCN 2003, October 2003.
- [92] A. Morea et al., "Datarate Adaptation for Night-Time Energy Savings in Core Networks," J. Lightw. Technol., vol. 31, num. 5, March 2013.
- [93] O. Rival, A. Morea, "Cost-efficiency of mixed 10-40-100Gb/s networks and elastic optical networks," Proc. OFC 2011, March 2011.
- [94] S.L. Woodward et al., "Intra-Node Contention in a Dynamic, Colorless, Non-Directional ROADM", Proc. OFC 2010, March 2010.
- [95] J. Perello et al, "Power consumption reduction through elastic data rate adaptation in survivable multi-layer optical networks," in Journal of Photonic Network Communications, Springer, 10.1007/s1107-014-0450-6, June 2014.
- [96] F. Rambach et al., "A Multilayer Cost Model for Metro/Core Networks," J. Opt. Commun. Netw., vol. 5, num. 3, March 2013.
- [97] G. Rizzelli et al., "Energy Efficient Traffic-Aware Design of On-Off Multi-Layer Translucent Optical Networks," Computer Networks, vol. 56, num. 10, July 2012.
- [98] A. Morea et al., "Power management of optoelectronic interfaces for dynamic optical networks," in Proc. European Conference on Optical Communications, Geneva, Switzerland, Sep. 2011, paper We.8.K.3.
- [99] I. Turus et al, "Traffic-aware Elastic Optical Networks to leverage Energy Savings," in Proceedings of IEEE LatinCom 2014.
- [100] NORDUnet traffic with Customers – Average In traffic for Week 23 2014, <http://stats.nordu.net>.
- [101] M. Pióro, A. Tomaszewski, M. Żotkiewicz, M. Klinkowski, M. Jaworski, and M. Gunkel, "Cost Analysis of an IP-over-Flexgrid Elastic Optical Network with Bandwidth-Variable Transponders", in Proc. ICTON 2014.
- [102] P. Poggiolini, G. Bosco, A. Carena, R. Cigliutti, V. Curri, F. Forghieri, R. Pastorelli, and S. Piciaccia, "The LOGON strategy for low-complexity control plane implementation in new-generation flexible networks," in Proc. OFC 2013.
- [103] G. Colavolpe and T. Foggi, "High spectral efficiency for long-haul optical links: Time-frequency packing vs. high-order constellations," in Proc. ECOC 2013, London, UK, paper P.4.1, 2013.
- [104] A. Eira, J. Pedro and J. Pires, "Optimized design of fixed/flex-rate line-cards and transceivers over multiple planning cycles," in Proc. OFC 2014.
- [105] O. Rival and A. Morea, "Cost-efficiency of mixed 10-40-100Gb/s networks and elastic optical networks," in Proc. OFC, 2011.
- [106] Poggiolini, P., Carena, A., Curri, V., Bosco, G., and Forghieri, F., "Analytical modeling of non-linear propagation in uncompensated optical transmission links", IEEE Photonics Technology Letters, vol. 23, no. 11, June 1, 2011.



- [107] Poggiolini, P, "The GN model of non-linear propagation in uncompensated coherent optical systems", *Journal of Lightwave Technology*, vol. 30, no. 24, 2012.
- [108] Poggiolini, P., Bosco, G., Carena, A., Curri, V., Jiang, Y., and Forghieri, F., "A detailed analytical derivation of the GN model of non-linear interference in coherent optical transmission systems", arXiv:1209.0394v13 [physics.optics], 2014.
- [109] Agrawal, G. P., "Fiber-optic communication systems", Wiley, 2002
- [110] Agrawal, G. P., "Lightwave technology: telecommunication systems", Wiley, 2005.
- [111] Headley, C., and Agrawal, G. P., "Raman amplification in fiber optical communication systems", Elsevier, 2002.
- [112] Kaminow, I., Li, T., "Optical Fiber Telecommunications, IV A - Components", Elsevier, 2002.
- [113] A. Stavdas, C.(T) Politi, T. Orphanoudakis, A. Drakos, "Optical Packet Routers:How they can Efficiently and Cost-Effectively Scale to Petabit/sec", *OSA Journal of Optical Networking*, vol. 7, pp. 876-894, 2008.
- [114] A. Stavdas, C. Matrakidis, C. (T) Politi, T. Orphanoudakis, J. Dunne, "A Novel Architecture for Highly Virtualised Software-Defined Optical Clouds", in *Proc. ECOC*, 2013.
- [115] A. Muhammad, G. Zervas, N. Amaya, D. Simeonidou and R. Forchheimer, "Introducing Flexible and Synthetic Optical Networking: Planning and Operation Based on Network Function Programmable ROADMs", *Journal of Optical Communications and Networking (JOCN)*, vol. 6, 2014.
- [116] J. Y. Yen, "Finding the K shortest loopless paths in a network," *Manage. Sci.*, vol. 17, no. 11, pp. 712–716, 1971.
- [117] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network," *IEEE Commun. Mag.*, vol. 48, pp. 138-145, 2010.
- [118] Porta Optica Study Project Deliverable D3.2: Economic analysis, dark fibre usage cost model and model of operations, 2007.

END OF DOCUMENT