

Chapter 3

Taking Advantage of Elastic Optical Networks

Alexandros Stavdas, Chris Matrakidis, Matthias Gunkel, Adrian Asensio, Luis Velasco, Emmanouel Varvarigos, and Kostas Christodoulopoulos

Acronyms

BV-OXC	Bandwidth Variable Optical Cross Connect
BVT	Bandwidth Variable Transponder
CO	Coherent detection
DB	Database
DC	Data centre
EON	Elastic Optical Network
FRR	Fast Reroute
FTP	File transfer protocol
MLR	Multi-layer resilience
NRZ	Non-return-to-zero
OOK	On-off-keying
RML	Routing and Modulation Level
RSA	Routing and Spectrum Assignment
RWA	Routing and Wavelength Assignment
S-BVT	Sliceable BVT
SDN	Software Defined Network
SLA	Service Level Agreement
VM	Virtual Machine

A. Stavdas (✉) • C. Matrakidis
University of Peloponnese, Tripoli, Greece
e-mail: astavdas@uop.gr

M. Gunkel
Deutsche Telekom, Bonn, Germany

A. Asensio • L. Velasco
Universitat Politècnica de Catalunya, Barcelona, Spain

E. Varvarigos • K. Christodoulopoulos
University of Patras, Rio Patras, Greece

This chapter shows how the new paradigm of Elastic Optical Networks (EON) is changing the way we perceive networking. The mode of operation of EON is described and the potential gains are outlined, followed by Use-Cases outlining how flex-grid technology can be advantageous for multi-layer resiliency, for serving metro networks and finally for data centre interconnection.

3.1 Introduction

In this section, we summarize the features of EON that are opening up new directions in the way we perceive networking. Specifically, our goal is twofold: first to demystify EON's mode of operation and then to substantiate the usefulness of this platform.

3.1.1 Flex-Grid vs. EON

Often in the literature the terms “flex-grid”, “gridless” or EON are used interchangeably to designate systems (and the networks employing such systems) that are not complying with ITU-T's fixed-grid channel spacing. EON systems differentiate from the 50 GHz/100 GHz channel spacing, which was the outcome of two important technological constraints in legacy WDM systems: (a) the use of non-return-to-zero (NRZ) on-off-keying (OOK) as the primary modulation format, and (b) the difficulty to separate fine-spectral bandwidths over the entire C-band in the third attenuation window of the silica fibre. The deployment of NRZ OOK systems at line rates of 10 Gb/s and above is accompanied with considerable channel limitations (mainly dispersion and fibre nonlinearities) whilst it necessitates a channel spacing that is at least 2–3 times the channel's optical bandwidth, to ensure sufficient optical isolation for add/drop.

As elaborated in Chap. 2, EONs may overcome these limitations. However, as it can be observed from Fig. 3.1, it is very difficult to conceive a truly “gridless” system; one can observe that in EON the channel bandwidth may become tighter occupying less bandwidth and/or central frequency of a channel might be shifted (in a process similar to what was understood as “wavelength conversion in legacy WDM systems”), but there is always a fundamental pitch, associated to the limitations of the optical demultiplexing technology used, in the way we exploit the optical bandwidth. This “spectral pitch” is usually an integer fraction of the 100 GHz standard grid spacing e.g. 6.25 or 12.5 GHz. Moreover, the term “flex-grid” is less accurate in describing the essence of those systems since, as we show in Sect. 3.1.2, it is the optical bandwidth of the channel that varies, and this optical bandwidth is more efficiently exploited in EONs because of the narrow grids used (Fig. 3.1).

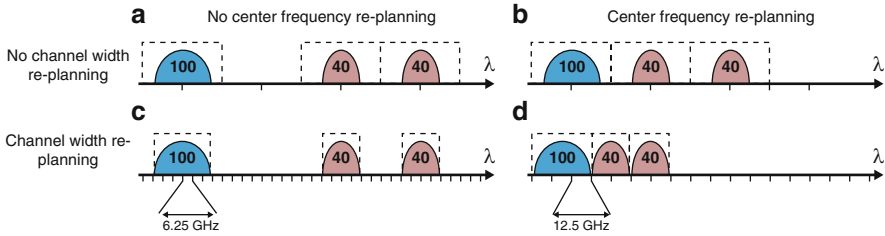


Fig. 3.1 WDM under ITU-T grid and EON modes of operation

ITU-T REC G.694.1 [1] defines the following terms that will be used throughout this book.

- *Frequency grid*: A frequency grid is a reference set of frequencies used to denote allowed nominal central frequencies that may be used for defining applications.
- *Frequency slot*: The frequency range allocated to a slot and unavailable to other slots within a flexible grid. A frequency slot is defined by its nominal central frequency and its slot width.
- *Slot width*: The full width of a frequency slot in a flexible grid.

For the flexible DWDM grid, the allowed frequency slots have a nominal central frequency on a 6.25 GHz grid around the central frequency of 193.1 THz and a slot width in multiples of 12.5 GHz as seen in Fig. 3.1.

Any combination of frequency slots is allowed as long as no two frequency slots overlap.

3.1.2 Demystifying EONs

To elaborate the properties of EONs, one needs to introduce the following quantities:

- *The baud rate*: it is the rate symbols which are generated at the source and, to a first approximation, equals to the electronic bandwidth of the transmission system. The baud rate is an important technology-dependent system performance parameter. This parameter defines the optical bandwidth of the transceiver and, following the discussion in Sect. 3.1.1, it specifies the minimum slot width required for the corresponding flow(s).
- *The modulation format*: For a given baud rate the modulation format defines the equivalent number of bits each symbol is transporting.
- *The line rate*: Effectively, it is the information-rate used for flow transportation between a source and a destination node (adjacent or remote). The dependence of the line rate on a number of sub-system parameters, mainly the nominal central frequency, the baud rate and the modulation format but also on whether or not there is polarization multiplexing and the type of the FEC used is schematically shown in Fig. 3.2a.

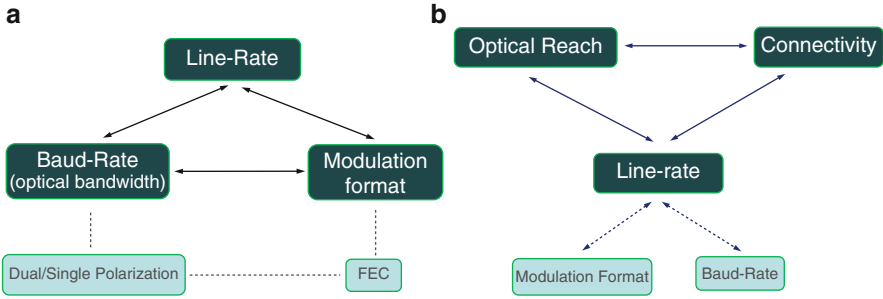


Fig. 3.2 The interrelation of parameters in EON's (a) sub-system level and (b) system level

Example: Consider a system with a 50 GSymbols/s baud rate where the QPSK modulation format is used which corresponds to 2 bits/symbol. The equivalent information rate is 100 Gb/s and the electrical bandwidth of the system is 50 GHz. The two polarizations should be simultaneously modulated, the line rate per carrier would be 200 Gb/s although the actual information rate might be somewhat less, pending on the FEC used. Given the ITU-T REC G.694.1 [1] 12.5 GHz slot width granularity, the frequency slot can be as small as 50 GHz. In practice, the frequency slot may need to be larger, with additional bandwidth allocated at each side of the signal, to ensure sufficient optical isolation to facilitate add/drop and wavelength switching in intermediate nodes. The actual guard band needed depends on the impulse response of the Bandwidth Variable Optical Cross Connect (BV-OXCs) but for LCoS-based systems, 12.5 GHz per side is sufficient.

An important difference of EONs to legacy WDM/WSON networks is the following: Assume there is a necessity to transport a 1 Tb/s flow. Considering the previous example, a five-carrier system at 200 Gb/s line rate is needed. In this case, when multiple flows are generated and transported as a single entity (this is usually called *Superchannel*) from the source to the destination node, with no add/drop in intermediate nodes, the five-carrier system need not have intermediate spectral guard-bands to separate the corresponding sub-channels.

This is part of a general trend in EONs where connectivity is traded for (aggregate) line rate: with the aid of Fig. 3.3, one can observe that there are five carriers which in the previous example would transport 200 Gb/s each. In Fig. 3.3a these carriers are used to interconnect the source node A to four other nodes: from A to D by means of two independent carriers producing a total of 400 Gb/s, A to G with one carrier supporting 200 Gb/s, A to E with one carrier supporting 200 Gb/s, and A to F with one carrier supporting 200 Gb/s. On the other hand, as shown in Fig. 3.3b, the five carriers are used to interconnect node A to node C only, but this time with a superchannel at an aggregate rate of 1 Tb/s. In this latter case, no guard band is needed to separate the five carriers, while in the former case (Fig. 3.3a), the carrier directed to nodes F and the carriers directed to node D need to have lateral guard bands due to the possibility of being optically cross connected in the transit nodes E and B, C, respectively.

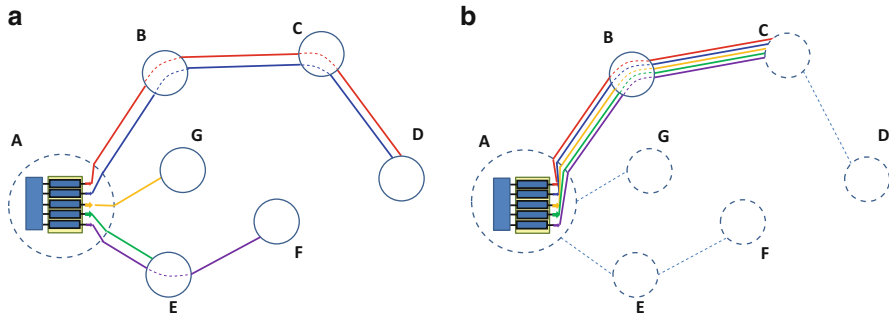


Fig. 3.3 The trade-off between node connectivity and link capacity; (a) the connectivity degree is 4 and (b) the connectivity degree is 1

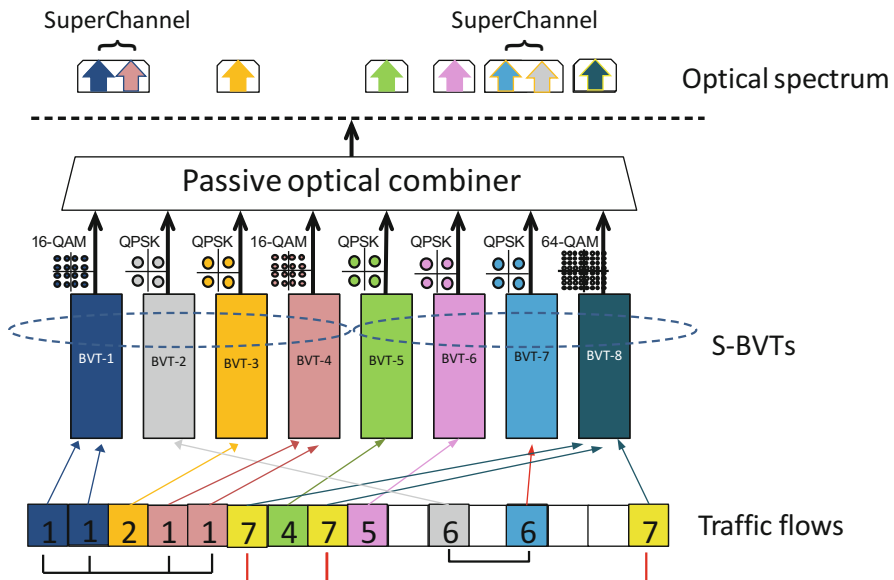


Fig. 3.4 Example of node with 8 BVTs partitioned in two S-BVTs (Thanks to Emilio Riccardi)

Finally, as schematically illustrated in Fig. 3.2b, an EON is not only trading line rate for connectivity but also for optical reach: given a baud rate, a higher constellation modulation format increases the line rate but decreases the optical reach. The outcome is the same if the modulation format is fixed but the baud rate is varying, although the physical performance degradation is due to different causes in each case. Nevertheless, a variable baud rate system—also known as variable optical bandwidth—is more technologically challenging to engineer.

To illustrate the implications of these new features introduced with EONs, consider an EON node. As shown in Fig. 3.4, the client traffic of the node is forwarded to the BVTs, typically by means of OTN/MPLS framer/switch, but the number of flows assigned to them changes dynamically.

In our example, the node is equipped with two sliceable bandwidth variable transponders (S-BVTs) incorporating four bandwidth variable transponders (BVTs) each. There are flows of different capacity so these flows are decomposed, by means of the electronic sub-systems of the S-BVT (not shown in the figure), to a number of flows all at the same nominal rate. For example, there is a four-flow group designated as “1”, a three-flow group designated as “7”, a two-flow group designated as “6” and individual flows designated as “2”, “4” and “5”. The BVTs may support different modulation formats which typically are QPSK, 16-QAM and 64-QAM, as illustrated in Fig. 3.4. To this end, the following selections are made: the flows “2”, “4” and “5” are forwarded to the BVT-3, BVT-5 and BVT-6, respectively, using QPSK. The three “7”s are re-multiplexed constructing a larger capacity flow which is, then, forwarded to the BVT-8 and for this reason this BVT is now modulated with 64-QAM. Finally, four “1”s are multiplexed in two larger flows and they are directed to BVT-1 and BVT-4, respectively, which are modulated with 16-QAM using adjacent frequency slots and in this case a Superchannel is formed. The two-“6” flows are not multiplexed; instead they are directly forwarded to BVT-2 and BVT-7 employing QPSK which are forming up, again, a superchannel. Finally, there are four more potential input traffic flows (shown in white in Fig. 3.4) that are unused since there are no further BVTs available.

In this example, one identifies the following trade-offs: we have selected the flows “7” to be grouped together to be directed to a single BVT, freeing BVTs that can be used by other flows. At network level this is equivalent to adding resources that would be used to attain higher node connectivity. However, this is made feasible by employing a 64-QAM that significantly reduces the optical reach. If on the other hand one wishes to increase the optical reach, the three “7” flows should employ 16-QAM, but this is done at the expense of an additional BVT, decreasing the “connectivity resources” of the node by one degree (the BVT). To further increase the optical reach, flows “2”, “4” and “5” are employing QPSK but that decreases the connectivity degree of the node while using wider frequency slots.

These examples illustrate the interrelations of the different parameters described in Fig. 3.2 as well as the emerging trade-offs between reach, connectivity and capacity in an EON.

Finally, to give some quantitative results, we compare two scenarios over the 22-node network shown in Fig. 3.5, with transparent end-to-end paths, with the fixed-grid and the EON case, respectively. The first scenario assumes a uniform traffic matrix, while the second one employs a traffic matrix generated using uniformly distributed random numbers.

The fixed-grid case has a channel spacing of 50 GHz and uses 100 Gb/s transceivers, while the EON case is the one described above, using only QPSK and 16-QAM modulations and polarization multiplexing (2000 km and 500 km reach, respectively) with 25 GHz guard bands between superchannels. Each link is assumed to comprise a single fibre pair where 4 THz of bandwidth can be allocated.

In this instance, the maximum capacity that can be transported over the network in each case was calculated, using shortest path routing and first fit wavelength/spectrum assignment [2]. The results are shown in Fig. 3.6. More sophisticated

Fig. 3.5 22-Node BT core network

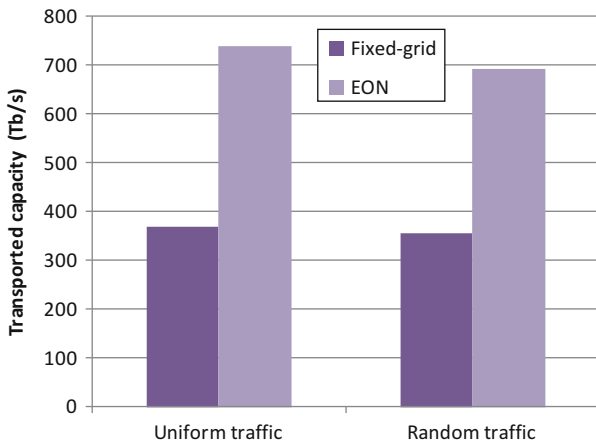
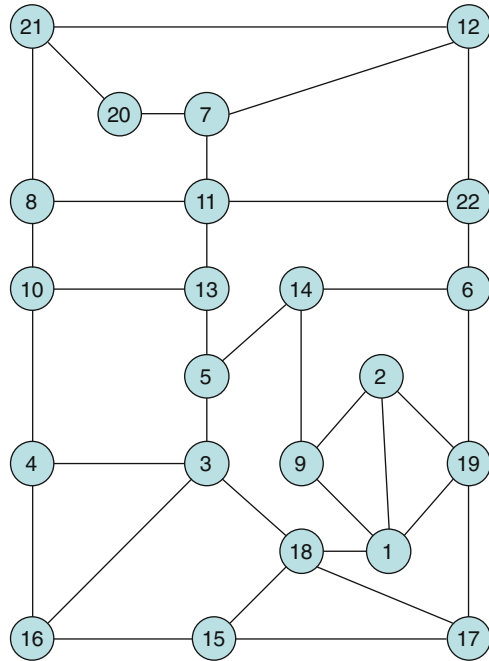


Fig. 3.6 Maximum transported capacity obtained for each scenario for the fixed-grid and EON case over the 22-node BT core network

design methodologies will be presented in the next chapter while this gives a rough estimate of the potential gains.

Figure 3.6 shows that even with this simple approach, the EON effectively doubles the transported capacity for both traffic matrix scenarios, compared to the fixed-grid network case.

3.2 Advanced Networking Features Exploiting Flex-Grid

There are multiple application scenarios possible for leveraging EONs' capabilities in modern optical communication networks. More specifically, EONs' elasticity can be brought together with the idea of multilayer cooperation in reaction to a network failure.

3.2.1 *Multi-layer Resilience in General*

Internet traffic in national backbones of European operators continues to grow with an annual rate of between 30 and 35 %. In this situation, operators are about to add new transport capacity connecting backbone routers over an agile DWDM layer. Today, the DWDM layer consists of fixed grid technology. In the future, say in 2–3 years, the technology of choice may be flex-grid based.

In traditional packet networks, it is the IP/MPLS layer which reacts to a failure, such as a fiber cut. A network failure on the optical layer may cause sudden, unavoidable and dynamic change over the transported traffic. Indeed, it is the most observed traffic dynamic in aggregation and backbone networks nowadays besides the normal constant and predictable traffic increase. At the high degree of traffic aggregation in these network hierarchies, unpredictable dynamics induced by failures happens much more frequently than dynamics induced by customers' behaviour (e.g. scheduled services).

Following the traditional approach of having reactive resilience exclusively over the packet layer might cause comparably bad utilization of both router interfaces and transponders that are equivalent to lambdas on the optical layer. The reason for this is the following: IP interfaces are loaded with traffic by adding up the packets which arrive randomly at the node. Forwarding the packets is subjected to a random process, too. Thus, multiplexing several flows together harvests a statistical efficiency gain. Though packet traffic is statistically multiplexed onto lambdas, those lambdas may be filled only up to 50 % in IP/MPLS networks relying on L3 recovery mechanisms only. The remaining 50 % capacity of a lambda is reserved for backup if a failure occurs. Without any dynamic countermeasure on the DWDM layer, optical robustness is conventionally assured by the creation of a second backup path 1 + 1 disjoint from the first primary path. This comes along with a second transponder interface leading to more wavelengths and more links in the overall network. Directly related to these intrinsic inefficiencies are high capital expenditures inhibiting overall network profitability. Therefore, we can find that resilience mechanisms relying on the IP/MPLS layer exclusively are simple and widely adopted, but they are not cost-efficient.

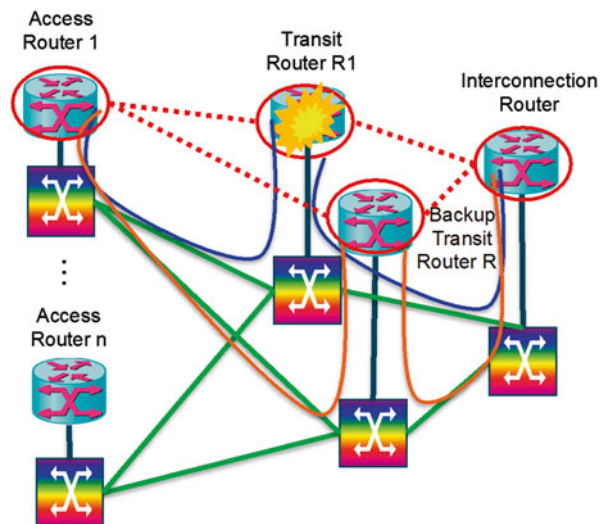
The recently proposed converged transport network architecture (in the sense of an integration of packet layer devices, e.g. routers, and optical transport equipment, e.g. OXC or ROADMs), operators aim at building a network that has

automated multi-layer resilience enabling considerably higher interface utilization. A key enabler for this new approach is the split of traffic into different service classes. For example, television, VoIP services, or financial transactions are critical applications that need to be delivered on time and with low packet loss ratio. Together with further services, they are pooled and assigned to high-priority traffic class. In contrast, a typical representative of best-effort traffic class is email or file transfer protocol (FTP) traffic which may experience some delay or packet loss.

A key idea behind multi-layer resilience (MLR) is to accept longer periods of packet losses for best-effort traffic than today. However, high-priority traffic maintains priority and is served in the same way as today. By treating different service classes unequally, capacity for best-effort class traffic is optically recovered after a considerable delay. Realistically, this recovery time is in the range between some seconds and up to a few minutes. This is huge when compared to IP resilience mechanisms like Fast Reroute (FRR) on one hand. On the other hand, it is very short when compared to the persistence of an optical cable cut or physical interface failure.

At the transmission layer, the MLR concept relies on agile photonic technology realized by multi-degree ROADMs with colourless and directionless ports (Fig. 3.7). That is where protection switching is related with the least cost when compared to packet layer protection switching. Several studies proved the feasibility and techno-economic superiority of this MLR concept under practical network scenarios (e.g. [3]). Currently, several operators are following this approach and roll-out MLR-based core networks. The general concept works fine already with conventional fixed-grid technology but EON can bring further improvements.

Fig. 3.7 Multi-layer restoration using a remote backup transit router



Case 1: Multi-layer Resilience Based on Elastic Flex-Rate Transceivers

As reaction to the failure, the conventional optical network tries to recover by finding a backup path that detours the non-functioning link. When routing the backup path through an optically rigid network, this new restoration path might exceed the reach specification of the previous light path. With a fixed and given modulation format the new light path might turn out to be infeasible. The consequence of that is the loss of the entire wavelength capacity (e.g. 100 Gbit/s). This is a quite big handicap for an operator and might induce further bottlenecks elsewhere in the network.

One promising application case of future software-configurable transceivers is to benefit from the reach-capacity trade-off (Fig. 3.2), to adjust the modulation format dependent on the reach requirement. Practically, this means a reduction of the original interface capacity.¹ Assuming the original capacity to be 400 Gbit/s realized by 2×200 Gbit/s with 16-QAM each, then the switch-over down to QPSK allows for a capacity of 2×100 Gbit/s instead of losing the entire interface capacity.

Thus, automatically lowering the existing line rate provides at least some part of the original capacity to be further exploited by IP routers accordingly. This type of photonic elasticity may save the operator a considerable amount of CapEx when compared to the case of a fixed-reach transceiver with its occasional necessity for intermediate signal regeneration. That is why flex-rate interfaces are considered as being of high importance to lower the network deployment cost.

Depending on the traffic volume and service classes, the operator may sometimes need the full capacity on the restoration path again. This can be accomplished by using an additional flex-rate BVT with again 2×100 Gb/s QPSK. But this additional interface will not be always needed, especially when the packet loss of best effort traffic is already beneath the acceptable threshold in accordance to the Service-Level Agreement (SLA). Therefore, after the failure repair, the additional restoring interface might be switched off or be used for other purposes.

Recently, a new approach was published [4] which leverages on the symbol-wise modulation adaptation in the time domain. With this approach, the considerable reach discrepancy between pure QPSK and pure 16-QAM (e.g. 2000 km and 500 km reach, respectively) can be closed by intermediate reach values. This is especially important for incumbent operators in smaller and medium countries where typical core distances fall just into the aforementioned reach gap. Techno-economic studies to quantify the benefit of this kind of flex-rate transceivers are on the way.

Case 2: Transceiver Sliceability for Multi-layer Resilience

Due to the unpredictability of traffic arising from, for example, cloud services on one hand and the general burstiness of packet traffic on the other hand, there is always a significant amount of stranded capacity in today's networks which provide fixed-capacity

¹ Without further limitations (e.g. spectral scarcity), it is supposed here that in real network deployments optical interfaces run at their maximum capacity.

circuits between nodes. This leads to an under-utilization of interfaces on the optical layer. Especially in the early days of an optical transport network, the utilization of interfaces is inherently low.

In this situation, a high-capacity transponder (e.g. at 400 Gbit/s or 1 Tbit/s) being able to be logically and physically sliced into several virtual transponders targeting at different destinations with electronically adaptive bit rates may be highly beneficial for improving network economics. At day one, even a single S-BVT may provide enough total capacity for serving all destinations at a comparably low bit rate, say 100 Gbit/s [5]. Later when the traffic increases, it may serve only a few destinations each at a higher bit rate. Finally, it may support only a single huge flow to a single destination. All this is expected to become electronically controlled and adjustable.

The spectral slicing process is to be monitored and driven by the associated IP routers. They control the different service classes and assign them to spectral bandwidth slices enabling significant provisioning flexibility. In case of a failure in the optical layer, an adapted treatment of service classes is expected to be beneficial. An integrated control plane optimizes the assignment of high-priority and best-effort traffic unequally to optical subcarriers. In addition to parameters like path cost, latency, and shared risk groups, the optimization may also be accomplished dependent on various EON-specific parameters like availability of spectral slices or fibre fragmentation status.

Consequently, the next evolution step might be the investigation of the appropriateness of S-BVTs within a realistic MLR concept based on a service class differentiation.

For both application cases introduced above, the network architecture requires three main ingredients at least:

1. Firstly, an *agile DWDM layer* with colourless, directionless and flex-grid ROADMs providing faster service provisioning. Those photonic devices are also used for resilience switching at the physical layer (LO).
2. Secondly, an *integrated packet-optical control plane* offering a software-based flexibility including routing and spectrum assignment (RSA) instead of routing and wavelength assignment (RWA), signalling functionalities as well as elastic path selection in case of a network failure. An efficient integrated control plane solution also comprises just the right amount of information exchange between the packet and elastic optical layer.
3. Thirdly, a standard interface allowing configuring all IP routers using the same protocol. The advent of multi-layer control plane may ease the configuration of the MPLS and the GMPLS equipment through an UNI. However, IP layer services require configurations that go beyond control plane functionalities. After a failure, adding new routes or changing the metrics in the IP layer can help to optimize the IP topology. There are standardization efforts made, such as in the Internet Engineering Task Force (IETF), Network Configuration Protocol (NETCONF) and Interface to the Routing System (I2RS) working groups, but this is not yet supported by all router vendors.

3.3 Flex-Grid in Metro-Regional Networks: Serving Traffic to BRAS Servers

Metro architectures are currently composed of two main levels of aggregation: (1) The first level (named multitenant user or MTU level) collects traffic from the end users (optical line terminals, OLTs), while (2) the second level (named Access level) aggregates the traffic from the MTUs mainly through direct fiber connections (i.e. dark fiber). The IP functionality, namely traffic classification, routing, authentication, etc., is implemented in the Broadband Remote Access Servers (BRASs) that are usually located after the second level of aggregation.

In the last years, the main European network operators have been expanding their optical infrastructure to regional networks. As a consequence, it has been proposed to create a pool of BRASs located at, for example, two transit sites per region (the second site is for a redundancy purpose) as opposed to multiple regional BRAS locations. In this way, the needs for IP equipment and the associated CAPEX and OPEX costs would be reduced. These centralized servers can be *physical* or *virtual*, following the Network Function Virtualization (NFV) concept and instantiated at private or public data centres in the region. In this centralized scenario (Fig. 3.8a), the regional photonic mesh provides the transport capacity from the second aggregation level towards the remote BRASs. A representative scenario is Region-A of the Telefonica Spanish Network [6], which has 200 MTU switches, access level with 62 aggregation switches, and an optical transport network comprising 30 ROADMs that are connected to 2 BRAS sites over that network.

As flex-grid technology is considered an alternative to WDM [6, 7] for the core/backbone network, BRAS regional centralization appears to be an appropriate and interesting use case for examining the applicability of flex-grid technology in regional and metropolitan area networks (MAN). The examined scenario assumes the use of flex-grid transponders at the MTU switches and Sliceable Bandwidth Variable Transponders (S-BVT) at the BRAS servers. The aggregation level is thus removed, exploiting the finer spectrum granularity of flex-grid technology. This scenario is presented in Fig. 3.8b.

Since the metro network is quite different from the backbone network, the coherent (CO) transceiver technology that is used in backbone network applications might not be appropriate for the metro. It is worth examining alternatives to the CO technology in the metro, such as direct detection (DD) transponders.

We now investigate the detail of the BRAS centralization scenario. We assume a regional optical network consisting of ROADM nodes (WDM or flex-grid) and fiber links. The network supports the traffic generated by a set of MTU switches that are lower in the hierarchy and are connected to the regional optical network. Each MTU switch is connected to two different ROADMs for fault tolerance. The way an MTU switch is connected to the ROADMs changes according to the network scenarios (WDM or flex-grid), but since we focus now on the optical regional network, this does not play a role. In turn, the two ROADMs, the source and the chosen backup

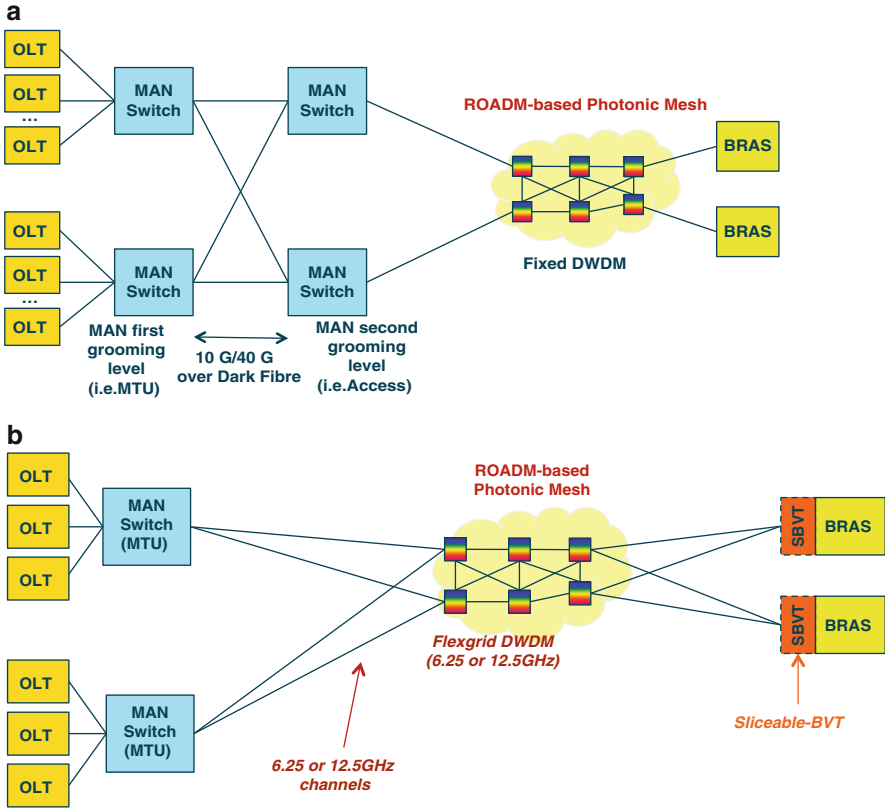


Fig. 3.8 (a) WDM approach and (b) flex-grid approach to enable BRAS regional centralization

source, are connected via two light paths to two central BRAS nodes, and the chosen two paths are node-disjoint for protection purposes. There are two problem variations: (1) the location of the BRAS nodes is given and the goal is to minimize the maximum path length of all primary and backup paths; and (2) the maximum path reach is given (constrained by the transponders used) and the goal is to minimize the number of BRAS locations (which can be higher than 2). To solve the related problems, we developed optimal (ILP) and heuristic algorithms and we applied them to find the transmission reach requirements for the Region-A of Telefonica network [6].

We used our results as transmission requirements to develop an appropriate DD transponder model. The first attempt was an OFDM-based S-BVT presented in [8]. An improved and robust physical impairment approach with multiband (MB) OFDM was then proposed in [9], where the performance was numerically assessed according to the transmission requirements and experimentally validated within the 4-node ADRENALINE photonic mesh network. The proposed S-BVT is able to serve $N \times M$

MTUs: the array of modulators generates N slices that can be directed towards different destinations and each BVT building block (each of the N modulators) serves M MTUs with a single optoelectronic front-end using only one laser source combined with simple and cost-effective DD. The number M of MTUs served per MB-OFDM slice (corresponding to a single optical carrier) is limited by the bandwidth of DAC and optoelectronic components (10 GHz in [9]), and $N=8$ and $M=5$ are considered feasible with current technology (assuming components with, e.g. 20 GHz).

To compare the two network scenarios presented in Fig. 3.8, we performed a techno-economic analysis. In particular, the scenarios studied are: (1) based on WDM technology, and (2) based on flex-grid, with two variations: (2a) with coherent (CO), and (2b) with direct detection (DD) transponders. We calculate the cost of the first scenario (WDM) and use that as a reference to calculate the target costs of components used in the flex-grid scenarios to achieve 30 % cost savings. The cost calculations are based on the cost model presented in [8] (an extension of [6]).

Scenario i—WDM solution: The optical network consists of WDM ROADMs and the traffic from the MTU switches is aggregated at a second level of aggregation switches before it enters the WDM network. Taking the reference TID Region-A domain, we have 200 MTU switches and 62 aggregation switches. Each MTU switch forwards $C=10$ Gbps and is connected to two different aggregation switches for protection. Assuming equal distribution of load to the aggregation switches, each switch needs in total 14×10 Gbps ports (for down- and up-link). We also assume that the communication between the MTUs and aggregation switches is done with 10 Gbps grey short-reach transceivers and that the communication between the aggregation switches and the BRAS servers is done with 10 Gbps coloured transceivers (routed over the WDM regional network). In the cost calculation, we add the costs of: (1) the MTU switches' transceivers (grey short-reach); (2) the aggregation switches (140 Gbps capacity, grey transceivers facing the MTUs and colored transceivers facing the WDM network); and (3) transceivers (coloured) at the BRASs. According to our cost model, the total cost of the WDM case [8] is 219.82 ICU (IDEALIST Cost Unit [6, 10]).

Scenario ii—Flex-grid solution: We now assume that we replace the WDM network with a flex-grid network, and thus we include in the cost calculation the cost of the flex-grid wavelength selective switch (WSS). We examine two variations for the transceivers: (a) CO and (b) DD.

For case (a), we assume the use of 10 Gbps grey short-reach transceivers at the MTUs and coherent 100 Gbps flex-grid BVTs (10×10 Gbps muxponders) at the add/drop ports of the flex-grid regional network, and coherent 400 Gbps S-BVT at the BRAS servers (functioning as 4×100 Gbps). We assume that each ROADM is equally loaded, so it serves 14 MTU switches. Thus, each ROADM utilizes two flex-grid BVTs. On the BRAS nodes, one 400 Gbps S-BVT serves two different ROADMs, which means that in total we need 15 S-BVTs. To obtain 30 % savings over the WDM scenario, the cost of 400G-SBVTs must be ≤ 0.01 ICU, which is impossible to achieve.

For case (b), we assume direct detected 10 Gbps MB-OFDM transceivers at the MTUs and DD MB-OFDM S-BVT at the BRAS servers (supporting $N=8$ flows, each flow serving $M=5$ MTUs at 10 Gbps). As above, we will assume that each ROADM is equally loaded, so it serves 14 MTUs. Thus, each ROADM needs three flows of the DD S-BVT, and in total we need 10 DD S-BVTs. To achieve cost savings of 30 % over the WDM network, the cost of 400G-S-BVT-DD must be ≤ 2.9 ICU, which is quite promising, since a same rate coherent S-BVT is expected to cost around 3 ICU in 2015 [6], and the cost of the DD S-BVT would be much lower.

Concluding, our comparison between a WDM and a flex-grid MAN network, showed that CO flex-grid transponders envisioned for core networks are expensive, while direct DD transponders seem a viable solution for the application of flex-grid solution in the MAN.

3.4 Multi-layer Network Planning for Cost and Energy Minimization

The increased flexibility of EONs can be exploited not only to save on spectrum but also on the cost and the energy consumption of the network. Currently, the optical network and the IP routers at its edges are not typically dimensioned with consideration of both technology domains, but this practice seems very inefficient for EONs. EONs envision the use of tunable (bandwidth variable) transponders (BVTs) that have a number of configuration options, resulting in different transmission parameters such as the rate, the spectrum, and the feasible reach of the light paths. Thus, decisions taken at the optical network often affect the dimensioning of the IP routers at its edges. This comes in contrast to the traditional fixed-grid WDM systems where the interdependence between the two layers was weaker and less dynamic. In EONs, we could still dimension both layers independently, for example sequentially, but there is vast space for optimization and gains if both layers are planned jointly. These gains come in addition to those obtained in multi-layer network operation such as restoration.

To give a closer look to the multi-layer planning problem, we model the network as follows. We assume that the optical network consists of ROADMs employing the flexgrid technology, connected with single or multiple fibres. At each optical switch, zero, one or more IP/MPLS routers are connected, and these routers comprise the edges of the optical domain. Short-reach transceivers are plugged to the IP/MPLS routers leading to flexible (tunable) transponders at the ROADMs. A transponder is used to transform the electrical packets transmitted from the IP router to the optical domain. We assume that a number of transmission parameters of the flexible transponders can be controlled, affecting their rate, spectrum and optical reach at which they can transmit. At the destination of a light path, the packets are converted back to electrical signal and are forwarded and handled by the corresponding IP/MPLS router. This can be: (1) the final destination, in which case traffic is further forwarded to lower hierarchy networks towards its final destination; or (2) an intermediate IP

hop, in which case traffic will re-enter the optical network to be eventually forwarded to its domain destination. Note that connections are bidirectional and thus in the above description opposite directed light paths are also installed, and the transponders used act simultaneously as transmitters and receivers.

Following the cost model in [6] and related datasheets (Cisco CRS-3 [11]), we model an IP/MPLS router as a modular device, built-in (single or multi) chassis. A chassis provides a specified number of slots (e.g. 16) with a nominal transmission speed (e.g. 400 Gbps). Into each slot, a line card of the corresponding speed can be installed, and each line card provides a specified number of ports at a specified speed (e.g. 10 ports of 40 Gbps). A reference router model presented in [6] has 16 slots per line card chassis, and several types of chassis (fabric card, fabric card chassis) that can be used to interconnect up to 72 line card chassis in total.

Following the above model, the multi-layer network planning problem consists of problems in two layers: (1) the IP routing (IPR), (2) the Routing and Modulation Level (RML), and (3) the Spectrum Allocation (SA). In the IPR problem, we decide on the modules to install at the IP/MPLS routers, how to map traffic onto the light paths (optical connections), and which intermediate IP/MPLS routers will be used to reach the domain destination. In the RML problem, we decide how to route the light paths and also select the transmission configurations of the flexible transponders to be used. In the SA, we allocate spectrum slots to the light paths, avoiding slot overlapping (assigning the same slot to more than one light path) and ensuring that each light path utilizes the same spectrum segment (spectrum slots) throughout its path (spectrum continuity constraint). A variation of the IPR problem is also referred to as traffic grooming, while the RMLSA problem is also referred to as Distance-adaptive RSA. As discussed previously, distance adaptivity creates interdependencies between the routing at the optical (RML) and the IP layers (IPR) and the spectrum allocation (SA), making it hard to decouple these sub-problems, unless we can afford to sacrifice the efficiency and pay a higher CAPEX and OPEX.

Several algorithms have been developed to solve variations of the multi-layer planning problem, described above, that include both optimal and heuristic algorithms [12–14]. The objective is to minimize the CAPEX cost, while some algorithms also consider the spectrum used and/or the energy consumption [15].

Findings in [12] show that we can obtain saving of about 10 % for reference year 2020 for a national (Deutsche Telecom, DT) and a pan-European (GEANT) network, when comparing an EON to a WDM network. Note that these savings were calculated for conservative scenarios that assume a 30 % higher cost for the flexible (BVT) transponders over the cost of the fixed transponders of equal maximum transmission rate (400 Gbps). Planning the multi-layer network as a whole, as opposed to planning the IP and the optical network sequentially, can yield cost savings of 15 % for topologies that exhibit wide variation in path lengths, such as in GEANT, while the savings are less for national networks, where the majority of connections are established with similar transponders configuration parameters and the increased transmission options of BVTs are not utilized. The percentage of energy savings that was observed in [15] was 25 %, higher than the related savings in cost, since in this case we assumed similar reference energy consumption levels for BVTs and the fixed transponders.

3.5 Interconnecting Data Centres

3.5.1 Motivation

As introduced above, transport networks are currently configured with big static fat pipes based on capacity overprovisioning. The rationality behind that is guaranteeing traffic demand and the committed quality of service. Considering a cloud-based scenario with federated data centres (DC) [16], the capacity of each inter-DC optical connection is dimensioned in advance according to some volume of foreseen data to transfer. Once in operation, scheduling algorithms inside the DC's local resource manager run periodically trying to optimize some cost function and organize data transfers among DC (e.g. Virtual Machine (VM) migration and database (DB) synchronization), as a function of the bit rate available. Since DC traffic varies greatly over time, static connectivity configuration adds high costs as a result of large connectivity capacity that remains unused in periods where the amount of data to transfer is low.

Figure 3.9 illustrates the normalized required bit rate along time of the day for VM migration and DB synchronization between DCs. Note that DB synchronization is mainly related to users' activities, whereas VM migration is related to particular policies implemented in scheduling algorithms running inside local resource managers.

Figure 3.10 represents the bit rate utilization along day time for DB synchronization, when the bit rate of the static optical connection between two DC is set to 200 Gb/s. As observed, during idle periods such connection is clearly under-utilized,

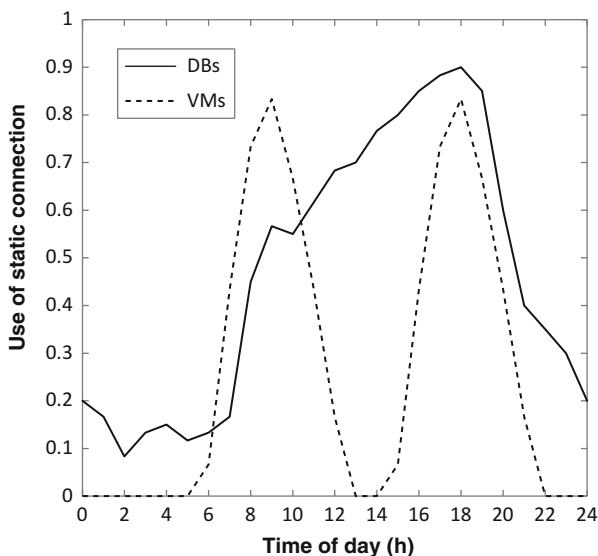


Fig. 3.9 Use of static connections for DB synchronization and VM migration during time of day

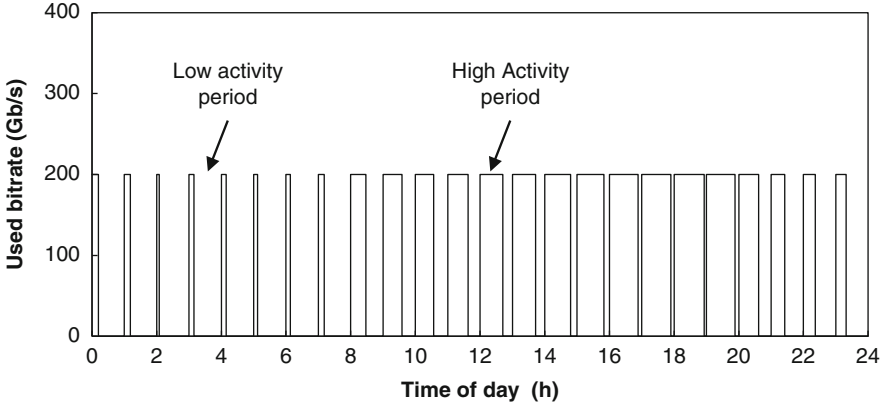


Fig. 3.10 Used bit rate over a static 200 Gb/s optical connection during time of day

whereas during high activity periods the connection is continuously used. Note that network operators cannot re-allocate non-utilized resources to other customers during idle periods, whereas more resources would be used during high activity periods to reduce time-to-transfer.

In light of the above, it is clear that connectivity among DC requires new mechanisms to provide reconfiguration and adaptability of the transport network to reduce the amount of overprovisioned bandwidth. To improve resource utilization and save costs, dynamic inter-DC connectivity is needed.

3.5.2 *Dynamic Connection Requests*

To handle dynamic cloud and network interaction, allowing on demand connectivity provisioning, the cloud-ready transport network was introduced in [17]. A cloud-ready transport network to interconnect DC placed in geographically dispersed locations can take advantage of EON providing bit rate on demand; optical connections can be created using the required spectral bandwidth based on users' requirements. Furthermore, by deploying EON in the core, network providers can improve spectrum utilization, thus achieving a cost-effective solution to support their services.

Figure 3.11 illustrates the reference network and cloud control architectures to support cloud-ready transport networks. Local resource managers in the DCs request connection operations to the EON control plane specifying the required bit rate.

In contrast to current static connectivity, dynamic connectivity allows each DC resource manager to request optical connections to remote DCs as data transfers need to be performed, and also to request releasing the connections when all data transfers completed. Furthermore, the fine spectral granularity and wide range of bit rates in EON make the actual bit rate of the optical connection match with the connectivity

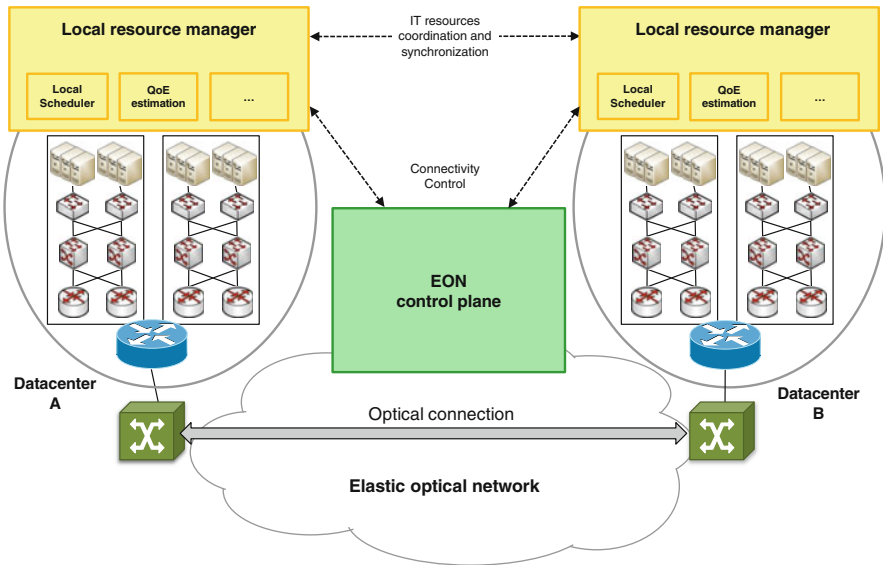


Fig. 3.11 Architecture to support inter-DC connections

needs. After requesting a connection and negotiating its capacity as a function of the current network resources availability, the resulting bit rate can be used by scheduling algorithms to organize transfers.

Using the architecture shown in Fig. 3.11, local resource managers are able to request optical connections dynamically to the control plane controlling EON and negotiating its bit rate. Figure 3.12a illustrates the message sequence between the local resource manager and the EON control plane to set up and tear down an optical connection. Once the corresponding module in the local resource manager finished computing a data transfer to be performed, a request for an optical connection to a remote DC is sent. In the example, the local resource manager sends a request for an 80 Gb/s connection. Internal operations in the EON control plane are performed to find a route and spectrum resources for the request. Assuming that not enough resources are available for the bit rate requested, an algorithm in the EON control plane finds the maximum bit rate and, at t_1 , it sends a response to the originating resource manager with that information. Upon the reception of the maximum available bit rate, 40 Gb/s in this example, the resource manager recomputes the transfer and requests a new connection with the corresponding available bit rate. When the transfer completes, the local resource manager sends a message to the EON control plane to tear down the connection, and the used resources are released (t_2) so that they can be assigned to any other connection.

Nonetheless, the availability of resources at the time of request is not guaranteed in the dynamic connectivity approach, and the lack of network resources may result in long transference time since transfers cannot be completed within the desired time. Note that connection's bit rate cannot be renegotiated and remains

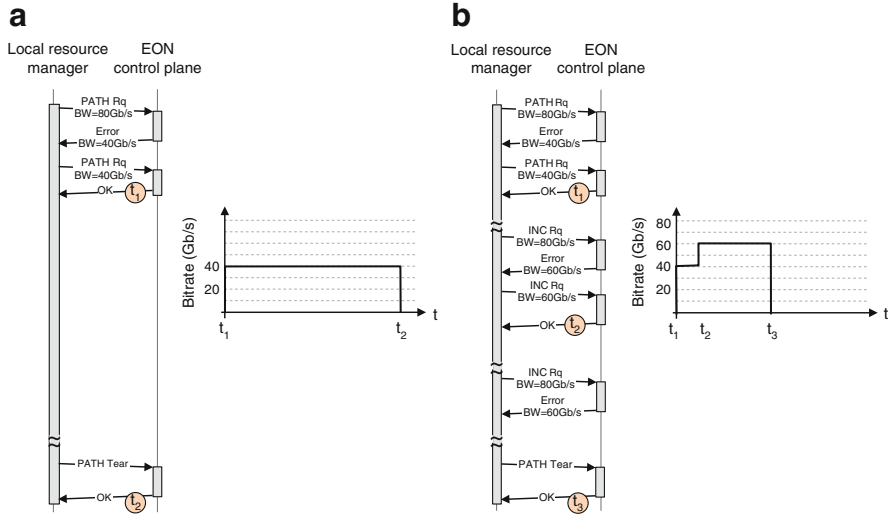


Fig. 3.12 (a) Dynamic and (b) elastic connectivity

constant during the connection’s holding time. To reduce the impact of the unavailability of required connectivity resources, authors in [18] proposed to use dynamic elastic connectivity.

In the dynamic elastic connectivity model, each local resource manager manages connectivity to remote DC so as to perform data transfers in the shortest total time. The resource manager requests not only connection set up and tear down operations but also elastic operations; the local resource manager is able to request increments in the bit rate of already established connections. Figure 3.12b illustrates messages exchanged between the local resource manager and the EON control plane to set up and tear down an optical connection as well as to increase its bit rate. In the example in Fig. 3.12b, after the connection has been established at t_1 with less of the initially requested bit rate (i.e. 40 Gb/s instead of 80 Gb/s initially requested), the local resource manager sends periodical retrials to increment connection’s bit rate. In this example, some resources are released after the connection has been established. A request for increasing connection’s bit rate is received and additional resources can be assigned to the connection (t_2), which reduces total transfer time; in this example, connection is increased to 60 Gb/s. Note that this is beneficial for both DC operators, since better performance could be achieved, and the network operator, since unused resources are immediately utilized.

It is worth noting that, although applications have full control over the connectivity process, physical network resources are shared among a number of clients; therefore, connection set up and elastic operations could be blocked as a result of lack of resources in the network. Hence, applications need to implement some sort of periodical retries to increase the allocated bit rate. These retries could impact negatively on the performance of the EON control plane and do not ensure achieving higher bit rate.

3.5.3 Transfer Mode Requests

With the aim to improve both network and cloud resource management, authors in [19] proposed to use Software-Defined Network (SDN), adding a new element, named as Application Service Orchestrator (ASO), in between local resource managers and the EON control plane. Resource managers can request transfer operations using their native semantic (e.g. volume of data and maximum time to transfer) liberating application developers from understanding and dealing with network specifics and complexity. The ASO stratum deployed on top of the EON control plane provides an abstraction layer to the underlying network and implements a north-bound interface to request those transfer operations, which are transformed into network connection requests. It is worth noting that in the case where connection requests are sent to the ASO, the ASO would act as a proxy between the local resource manager and the network. Figure 3.13 illustrates both control architectures supporting dynamic and transfer mode requests.

The ASO translates transfer requests into connection requests and forwards them to the EON control plane which is in charge of the network. If not enough resources are available at the time of request, notifications (similar to interruptions in computers) are sent from the EON control plane to the ASO each time specific resources are released. Upon receiving a notification, the SDN takes decisions on whether to increase the bit rate associated to a transfer or not and notifies changes in the bit rate to the corresponding local resource manager. This connectivity model for inter-DC effectively moves from polling in the dynamic elastic model to a network-driven transfer model.

Figure 3.14 shows an example of the messages requesting transfer operations exchanged between a local resource manager, the ASO and the EON control plane.

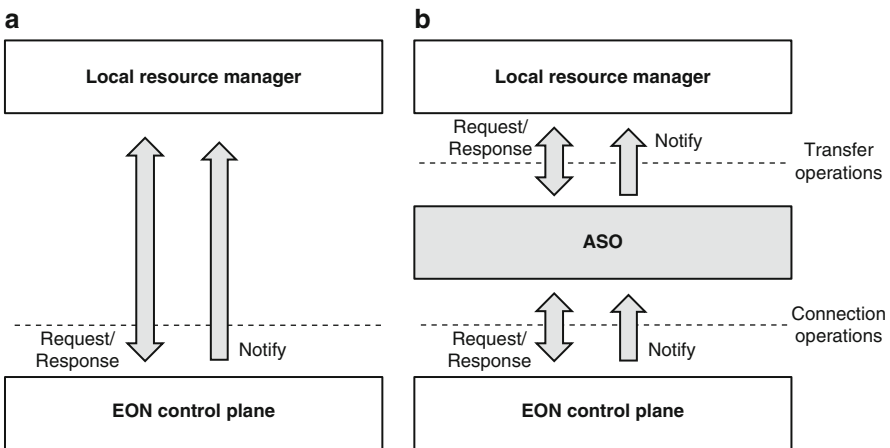


Fig. 3.13 Control architectures supporting: (a) dynamic connections, and (b) transfer mode requests

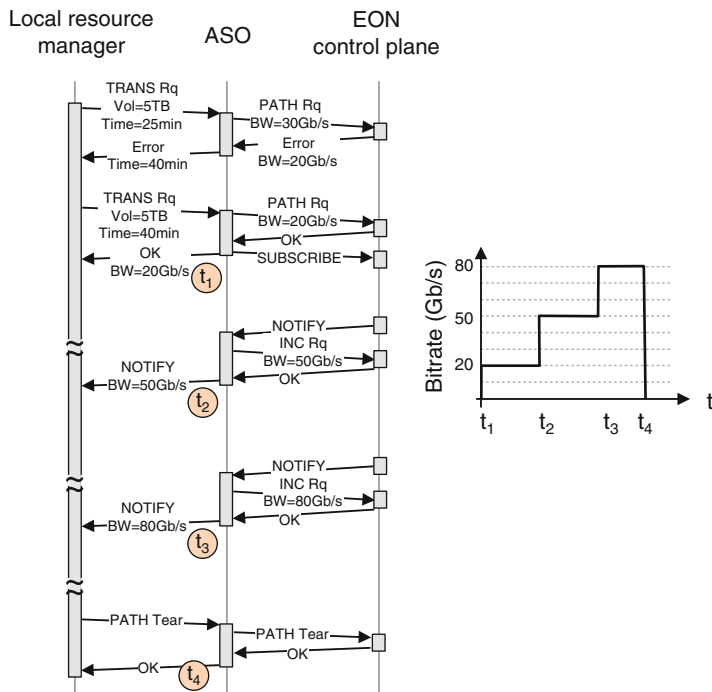


Fig. 3.14 Transfer mode request

A local resource manager sends a transfer mode request to the ASO for transferring 5 TB of data in 25 min. Upon receiving the request, the) ASO translates it to network connection’s semantic, sends a request to the EON control plane to find the biggest spectrum width available, taking into account local policies and current SLA, then sends a response back to the resource manager with the best completion time for transferring the required amount of data, which is 40 min in this example. The resource manager organizes data transfers and sends a new transfer request with the suggested completion time. A new connection is established and its capacity is sent in the response message at t_1 ; in addition, the SDN requests the EON control plane to keep informed upon more resources left available in the route of that connection. Algorithms deployed in the control plane monitor spectrum availability in the corresponding physical links. When resource availability allows increasing the allocated bit rate of some connection, the SDN controller autonomously performs elastic spectrum operations so as to ensure committed transfer completion times. Each time the SDN controller modifies bit rate of any connection by performing elastic spectrum operations, a notification is sent to the source resource manager containing the new throughput information. In this example, some network resources are released and the control plane notifies the SDN. Provided that those resources can be assigned to the connection, its bit rate is increased up to 50 Gb/s and the EON notifies the resource manager of the increased bit rate at t_2 . The DC resource manager can then optimize transferences as a function of the

actual throughput while delegating and ensuring completion transfer time to the SDN. Similarly, at t_3 more network resources are released and become available; the connection capacity is then increased to 80 Gb/s. Finally, at t_4 , the connection is torn down and the corresponding network resources are released.

Finally, the proposed network-driven model opens the opportunity to network operators to implement policies so as to dynamically manage connections' bit rate of a set of customers and fulfil simultaneously their SLAs [20].

To conclude, EON can provide the infrastructure for interconnect DCs. The requirements for that relate to the capacity to support dynamic and elastic connectivity, both in the control and the data plane, since inter-DC traffic highly varies along the day.

3.6 Concluding Remarks

This chapter summarized the features of EON that are opening up new directions in the way we perceive networking. EON's mode of operation was demystified and the potential gains were outlined. The usefulness of the EON paradigm was substantiated by means of Use-Cases, outlining how flex-grid technology can be advantageous for multilayer resiliency, for serving metro networks and finally for data centre interconnection.

References

1. ITU-T G.694.1, Spectral grids for WDM applications: DWDM frequency grid, February 2012
2. R. Ramaswami, K.N. Sivarajan, Routing and wavelength assignment in all-optical networks. *IEEE/ACM Trans. Netw.* **3**(5), 489–500 (1995)
3. M. Gunkel, A. Autenrieth, M. Neugirg, J. Elbers, Advanced multilayer resilience scheme with optical restoration for IP-over-DWDM core networks—how multilayer survivability might improve network economics in the future, in *International Workshop on Reliable Networks Design and Modeling (RNDM)*, 03.10.2012, St. Petersburg, Russia, 2012
4. Curri et al., Time-division hybrid modulation formats: Tx operation strategies and countermeasures to nonlinear propagation, in *Proceedings OFC, 2014*
5. A. Autenrieth, J. Elbers, M. Eiselt, K. Grobe, B. Teipen, H. Grießer, Evaluation of technology options for software-defined transceivers in fixed WDM grid versus flexible WDM grid optical transport networks, in *Proceedings ITG Photonic Networks*, 2013
6. F. Rambach, B. Konrad, L. Dembeck, U. Gebhard, M. Gunkel, M. Quagliotti, L. Serra, V. López, A multilayer cost model for metro/core networks. *J. Opt. Commun. Netw.* **5**, 210–225 (2013)
7. O. Gerstel, M. Jinno, A. Lord, S.J. Ben Yoo, Elastic optical networking: a new dawn for the optical layer? *IEEE Commun. Mag.* **50**(2), S12–S20 (2012)
8. M. Svaluto et al., Experimental validation of an elastic low-complex OFDM-based BVT for flexi-grid metro networks, in *ECOC*, 2013
9. M. Svaluto et al., Assessment of flexgrid technologies in the MAN for centralized BRAS architecture using S-BVT, in *ECOC*, 2014
10. Idealist deliverable D1.1, Elastic optical network architecture: reference scenario, cost and planning

11. Cisco CRS-3, http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/data_sheet_c78-408226.html
12. V. Gkamas, K. Christodouloupoulos, E. Varvarigos, A joint multi-layer planning algorithm for IP over flexible optical networks. *IEEE/OSA J. Lightwave Technol.* **33**(14), 2965–2977 (2015)
13. O. Gerstel, C. Filsfil, T. Telkamp, M. Gunkel, M. Horneffer, V. Lopez, A. Mayoral, Multi-layer capacity planning for IP-optical networks. *IEEE Commun. Mag.* **52**(1), 44–51 (2014)
14. C. Matrakidis, T. Orphanoudakis, A. Stavdas, A. Lord, Network optimization exploiting grooming techniques under fixed and elastic spectrum allocation, in *ECOC*, 2014
15. V. Gkamas, K. Christodouloupoulos, E. Varvarigos, Energy-minimized design of IP over flexible optical networks. *Wiley Int. J. Commun. Syst.* (2015). doi:[10.1002/dac.3032](https://doi.org/10.1002/dac.3032)
16. I. Goiri, J. Guitart, J. Torres, Characterizing cloud federation for enhancing providers' profit, in *Proceedings IEEE International Conference on Cloud Computing*, 2010, pp. 123–130
17. L. Contreras, V. López, O. González De Dios, A. Tovar, F. Muñoz, A. Azañón, J.P. Fernández-Palacios, J. Folgueira, Toward cloud-ready transport networks. *IEEE Commun. Mag.* **50**, 48–55 (2012)
18. L. Velasco, A. Asensio, J.L. Berral, V. López, D. Carrera, A. Castro, J.P. Fernández-Palacios, Cross-stratum orchestration and flexgrid optical networks for datacenter federations. *IEEE Netw. Mag.* **27**, 23–30 (2013)
19. L. Velasco, A. Asensio, J.L. Berral, A. Castro, V. López, Towards a carrier SDN: an example for elastic inter-datacenter connectivity. *OSA Opt. Express* **22**, 55–61 (2014)
20. A. Asensio, L. Velasco, Managing transfer-based datacenter connections. *IEEE/OSA J. Opt. Commun. Netw.* **6**, 660–669 (2014)