Routing and Spectrum Allocation in Elastic Optical Networks: A Tutorial

Bijoy Chand Chatterjee IEEE, Member, Nityananda Sarma IEEE, Member, and Eiji Oki, IEEE, Fellow

Abstract—Flexgrid technology is now considered to be a promising solution for future high-speed network design. In this context, we need a tutorial that covers the key aspects of elastic optical networks. This tutorial paper starts with a brief introduction of the elastic optical network and its unique characteristics. The paper then moves to the architecture of the elastic optical network and its operation principle. To complete the discussion of network architecture, this paper focuses on the different node architectures, and compares their performance in terms of scalability and flexibility. Thereafter, this paper reviews and classifies routing and spectrum allocation (RSA) approaches including their pros and cons. Furthermore, various aspects, namely — fragmentation, modulation, quality-of-transmission, traffic grooming, survivability, energy saving, and networking cost related to RSA, are presented. Finally, the paper explores the experimental demonstrations that have tested the functionality of the elastic optical network, and follows that with the research challenges and open issues posed by flexible networks.

Index Terms—Elastic optical networks, node architecture, spectrum management, routing and spectrum allocation, and sliceable bandwidth-variable transponder.

NOMENCLATURE

AG Auxiliary Graph
AoD Architecture on Demand
AR Adaptive Routing
AST Average Setup Time
BP Blocking Probability
BPSK Binary Phase-Shift Keying
BV-SSS Bandwidth-Variable Spectrum Selective Switch
BVT Bandwidth-Variable Transponder
BV-WXC Bandwidth-Variable Cross-Connect
CF Central Frequency
DWDM Dense Wavelength Division Multiplexing
EDFA Erbium-Doped Fiber Amplifier
EF Exact Fit
FAR Fixed Alternate Routing
FF First Fit

FLF Fast-Last-Fit
FR Fixed Routing
Gb/s Gigabit Per Second
GMPLS Generalized Multi-Protocol Label Switching
HOPS Hitless Optical Path Shift
IEEE Institute of Electrical and Electronics Engineers
ILP Integer Linear Programming
IP Internet Protocol
ITU International Telecommunication Union
LCoS Liquid Crystal on Silicon
LCR Least Congested Routing
LSP Label-Switched Path
LU Least Used
MC-RSA Multicast-Capable Routing and Spectrum Assignment
MEMS Micro-Electro Mechanical System
MILP Mixed-Integer Linear Programming
MUX Multiplexer
NSFNET National Science Foundation Network
NP Non-Deterministic Polynomial Time
OFDM Orthogonal Frequency-Division Multiplexing
OSPF Open Shortest Path First
PCE Path Computation Element
PIC Photonic Integrated Circuit
PLI Physical Layer Impairment
PSK Phase-Shift Keying
QAM Quadrature Amplitude Modulation
QoS Quality-of-Transmission
QPSK Quadrature Phase-Shift Keying
R Random
RML Routing and Modulation Level
RODAM Reconfigurable Optical Add-Drop Multiplexer
RSA Routing and Spectrum Allocation
RSVP-TE Resource Reservation Protocol - Traffic Engineering
RWA Routing and Wavelength Assignment
RX Receiver
SA Spectrum Allocation
SBVT Sliceable Bandwidth-Variable Transponder
SDM Space Division Multiplexing
SDN Software-Defined Networking
SF Smallest Fit
SSS Spectrum Selective Switch
TAP-BR Time-Aware Provisioning with Bandwidth Reservation

Manuscript received xxxx; revised xxxx; accepted xxxx. Date of publication xxxx; date of current version xxxx. This work was supported in part by Strategic Information and Communication R&D Promotion Program of the Ministry of Internal Affairs in Japan and National Institute of Information and Communications Technology, Japan.

B. C. Chatterjee and E. Oki are with the Department of Information and Communication Engineering, The University of Electro-Communications, Tokyo - 182-8585, Japan. e-mail: {bijoycc, eiji.oki}@uec.ac.jp.

N. Sarma is with the Department of Computer Science and Engineering, Tezpur University, Assam-784028, India. e-mail: nitya@tezu.ernet.in.
HE rapid growth in world-wide communications and the rapid adoption of the Internet has significantly modified our way of life. This revolution has led to a vast growth in communication bandwidth in every year. An optical network has the potential to support the continued demands for communication bandwidth. The high capacity of dense wavelength division multiplexing (DWDM)-based optical networks [1], [2] is assisted by the use of upper layers to aggregate low-rate traffic flows into lightpaths in mechanism of traffic grooming [3]–[6]. DWDM-based systems with up to 40 Gb/s capacity per channel have been deployed in backbone networks, while 100 Gb/s interfaces are now commercially available, and 100 Gb/s deployment is expected soon. TeleGeography [7] expects that international bandwidth demands will be approximately 606.6 Tb/s in 2018 and 1,103.3 Tb/s in 2020. Therefore, optical networks will be required to support Tb/s class transmission in the near future [8], [9]. Unfortunately, conventional optical transmission technology has inadequate scaling performance to meet the growing traffic demands as it suffers from the electrical bandwidth bottleneck limitation, and the physical impairments become more serious as the transmission speed increases [10]. Moreover, the traffic behavior is changing rapidly and the increasing mobility of traffic sources makes grooming more complex. Therefore, researchers are now focusing on new technologies for high-speed optical networks.

To meet the needs of the future Internet, optical transmission and networking technologies are moving toward the goals of greater efficiency, flexibility, and scalability. Recently, elastic optical networks [11]–[15] have been shown to be a promising candidate for future high-speed optical communication. An elastic optical network has the potential to allocate spectrum to lightpaths according to the bandwidth requirements of clients. The spectrum is divided into narrow slots and optical connections are allocated a different number of slots. As the result, network utilization efficiency is greatly improved compared to DWDM-based optical networks. In elastic optical networks, a certain number of transmission parameters, such as — optical data rate, modulation format, and wavelength-spacing between channels, which are fixed in currently deployed networks, will be made tunable. Given that future demands indicate that high-speed optical connections are needed to optimize data transport, elastic optical networks are a suitable replacement for DWDM-based optical networks.

As the elastic optical network has been widely accepted as the next generation high-speed network, researchers have focused on its architecture, spectrum allocation mechanism, and future scope. Accordingly, a comprehensive survey on orthogonal frequency-division multiplexing (OFDM)-based optical high-speed transmission and networking technologies, with a specific focus on OFDM technology, was provided by G. Zhang et al. [15]. S. Talebi et al. [16] presented a review of spectrum management techniques for elastic optical networks. Recently, a survey of the current ongoing research efforts into defining elastic optical network control plane architectures was presented in [17].

I. INTRODUCTION

A. Contribution

This paper provides an integrated tutorial that covers different aspects of elastic optical networks. In this tutorial paper, we explore routing and spectrum allocation in elastic optical networks. We begin with the basic concept of the elastic optical network and its unique characteristics, and then the architecture of the elastic optical network and its operation principle are presented. We elucidate the functionalities of the bandwidth-variable transponder (BVT) and the sliceable bandwidth-variable transponder (SBVT). SBVT architecture and its advantages in the future optical network are detailed. Immediately after discussing network architecture, our discussion focuses on the different node architectures, namely — broadcast and select, spectrum routing, switch and select with dynamic functionality, and architecture on demand, along with their functionalities. We compare different node architectures in order to clarify their performances. Next we look into the basic concept of the routing and spectrum allocation (RSA) approach in elastic optical networks. We discuss the differences between RSA and routing and wavelength assignment (RWA) in optical networks. Then, we move to different routing approaches along with their pros and cons. We compare different routing approaches in order to analyze their performances. Thereafter, we discuss different spectrum allocation policies. In addition, we separate the spectrum allocation policies into two categories based on spectrum range allocation for connection groups and spectrum slot allocation for the individual connection request.

Various aspects, namely — fragmentation, modulation, quality-of-communication, traffic grooming, survivability, energy saving, and networking cost related to RSA, all of which have been reported in the literature, are presented in this tutorial paper. We classify the fragmentation-aware approaches, or de-fragmentation approaches, and explain how these approaches deal with fragmentation problems. Distance adaptive RSA that considers the modulation technique is one of the features of the elastic optical network, and is also presented in this tutorial paper. We discuss and compare traffic grooming in wavelength-division multiplexing (WDM)-based optical networks, elastic optical networks with BVTs, and elastic optical networks with SBVTs. Different survivability techniques, namely — protection and restoration, are addressed in this paper. We discuss the networking cost reduction made possible by the use of SBVT in elastic optical networks. In addition, our discussion on various aspects related to RSA is summarized.

Finally, we explore the experimental demonstrations that have been published to confirm the functionality of the elastic optical network. We address the research challenges and open issues posed by flexible networks.
B. Organization

The rest of the paper is organized as follows. Section II provides the basic concept of the elastic optical network. Section III presents the architecture of the elastic optical network. Section IV explains the different node architectures. Section V discusses the basic RSA approach in the elastic optical network. Section VI discusses and compares the major routing problems of the elastic optical network. Section VII focuses on different spectrum allocation policies. Various issues related to RSA are presented in Section VIII. Section IX presents the experimental demonstrations that have tested the functionality of the elastic optical network. Research issues and challenges faced by optical network researchers are highlighted in Section X. We close this tutorial in Section XI, where we draw our conclusions.

II. Concept of Elastic Optical Networks

The traditional WDM-based optical network divides the spectrum into separate channels. The spacing between adjoining channels is either 50 GHz or 100 GHz, which is specified by international telecommunication union (ITU)-T standards as shown in Fig. 1. The frequency spacing between two adjacent channels is relatively large. If the channels carry only low bandwidth, and no traffic can be transmitted in the large unused frequency gap, a large portion of the spectrum will be wasted, which is reflected in Fig. 2.

![Fig. 1. ITU-T grid.](image)

To overcome the limitations of traditional optical networks, M. Jinno et al. [11] presented a spectrum efficient elastic optical network based on OFDM technology [15], [18]. Optical OFDM allocates the data to several low data rate subcarrier channels. As the spectrum of adjacent subcarrier channels are orthogonally modulated, they can overlap each other as shown in Fig. 3, which increases transmission spectral efficiency. Furthermore, optical OFDM can provide fine-granularity capacity to connections by the elastic allocation of low rate subcarriers. A bandwidth-variable OFDM transponder generates an optical signal using just enough spectral resources, in terms of subcarriers with appropriate modulation level, to satisfy the clients requirements. OFDM signals are usually generated in the radio-frequency domain, so many transmission properties can be freely set, i.e., different subcarriers can be assigned different numbers of modulated bits per symbol. To establish a connection, each bandwidth variable cross-connect on the route allocates a cross-connection with sufficient spectrum in order to create an appropriately sized end-to-end optical path. This end-to-end path is expanded and contracted according to the traffic volume and user requests, as necessary. The main characteristics of an elastic optical network are bandwidth segmentation, bandwidth aggregation, efficient accommodation of multiple data rates, elastic variation of allocated resources, reach-adaptable line rate, etc. These are discussed in more detail below.

- Bandwidth segmentation: Traditional optical networks require full allocation of wavelength capacity to an optical path between an end-node pair. However, elastic optical networks provide a spectrum efficient bandwidth segmentation (sometimes called sub wavelength) mechanism that provides fractional bandwidth connectivity service. If only partial bandwidth is required, elastic optical network can allocate just enough optical bandwidth to accommodate the client traffic, as shown in Fig. 4, where a 40 Gb/s optical bandwidth is segmented into three sub wavelengths, such as — 5 Gb/s, 15 Gb/s, and 20 Gb/s. At the same time, every node on the route of the optical path allocates a cross-connection with the appropriate spectrum bandwidth to create an appropriate-sized end-to-end optical path. The efficient use of network resources will allow the cost-effective provisioning of fractional bandwidth service.

- Bandwidth aggregation: Link aggregation is a packet networking technology standardized in IEEE 802.3. It combines multiple physical ports/links in a switch/router into a single logical port/link to enable incremental growth of link speed as the traffic demand increases beyond the limits of any one single port/link. Similarly, the elastic optical network enables the bandwidth aggregation feature and so can create a super-wavelength optical path contiguously combined in the optical domain, thus ensuring high utilization of spectral resources. This unique feature is depicted in Fig. 4, where three 40 Gb/s optical bandwidths are multiplexed with optical OFDM, to provide a super-channel of 120 Gb/s.

- Efficient accommodation of multiple data rates: As shown in Fig. 4, the elastic optical network has the ability to provide the spectrally-efficient direct accommodation of mixed data bit rates in the optical domain due to its flex-
Fig. 4. Unique characteristics, namely — bandwidth segmentation, bandwidth aggregation, accommodation of multiple data rates, and elastic variation of allocated resources, of elastic optical networks.

Inelastic spectrum assignment. Traditional optical networks with fixed grid leads to wastage of the optical bandwidth due to the excessive frequency spacing for low bit rate signals.

- Reach-adaptable line rate: The elastic optical network has the ability to support reach-adaptable line rate, as well as dynamic bandwidth expansion and contraction, by altering the number of subcarriers and modulation formats.
- Energy saving: It supports energy-efficient operations in order to save power consumption by turning off some of the OFDM subcarriers while traffic is slack.
- Network virtualization: It allows optical network visualization with virtual links supported by OFDM subcarriers.

In elastic optical networks, spectrum resources are allocated in a contiguous manner. Accordingly, allocation of resources to short duration connections can increase the number of non-aligned available slots, unless the allocation is well organized. These non-aligned available slots can be reduced by partitioning the total set of subcarrier slots as shown in Fig. 5. This will increase the number of acceptable connections and hence decrease the blocking probability.

Sometime, partitioning also negatively impacts the blocking probability due to the lack of statistical multiplexing-gain [19]. We explain this with a simple example wherein partitioning the total set of subcarrier slots decreases the number of channels in each partition, which increases the networks blocking probability. First, we calculate the blocking probability using the Erlang B loss formula [20] under a simple traffic model (a Poisson arrival process and exponential distribution of the connection holding times). If the number of channels is 100 and the offered traffic is 100 Erlang, the blocking probability is 0.0757. Dividing the same channel resources among four partitions and splitting the traffic among the partitions (i.e., 25 channels with offered traffic volume of 25 Erlang), the blocking probability for each partition becomes 0.1438, which is higher than that of the non-partitioned case.

Therefore, based on the above discussion, partitioning has an advantage in reducing the bandwidth blocking probability in elastic optical networks, provided the number of partitions is minimized. Taking this direction R. Wang et al. [21] and W. Fadini et al. [22] presented spectrum partitioning schemes in order to utilize the spectrum resources efficiently and manage the fragmentation issue; both are discussed later in the fragmentation subsection (Section VIII-A).

III. Elastic Network Architecture

To fulfill our ever-increasing bandwidth demands, the elastic optical network is indispensable. This is due to its many desirable properties including flexible data rate and spectrum allocation, low signal attenuation, low signal distortion, low power requirement, low material usage, small space requirement, and low cost. This section discusses the architecture of the elastic optical network. Figure 6 shows the typical architecture of the elastic optical network, which mainly consists of BVTs and BV-WXCs. These basic components and their working principle are explained in the following subsections.
A. Bandwidth-variable Transponder

BVTs [15] are used to tune the bandwidth by adjusting the transmission bit rate or modulation format. BVTs support high-speed transmission using spectrally efficient modulation formats, e.g., 16-quadrature amplitude modulation (QAM), with 64-QAM used for shorter distance lightpaths. Longer distance lightpaths are supported by using more robust but less efficient modulation formats, e.g., quadrature phase-shift keying (QPSK) or binary phase-shift keying (BPSK). Therefore, BVTs are able to trade spectral efficiency off against transmission reach.

However, when a high-speed BVT is operated at lower than its maximum rate due to required reach or impairments in the optical path, part of the BVT capacity is wasted. In order to address this issue, an SBVT [23]-[27] has been presented that offers improved flexibility; it is seen as a promising transponder technology. An SBVT has the capability to allocate its capacity into one or several optical flows that are transmitted to one or several destinations. Therefore, when an SBVT is used to generate a low bit rate channel, its idle capacity can be exploited for transmitting other independent data flows. An SBVT generates multiple optical flows that can be flexibly associated with the traffic coming from the upper layers according to traffic requirements. Therefore, optical flows can be aggregated or can be sliced based on the traffic needs. Figure 7 distinguishes BVT and SBVT functionalities.

Fig. 6. Architecture of elastic optical network.

The SBVT architecture [23], [24] was introduced in order to support sliceability, multiple bit rates, multiple modulation formats, and adaptive code rates. Figure 8 shows the architecture of an SBVT; it mainly consists of a source of $N$ equally spaced subcarriers, a module for electronic processing, an electronic switch, a set of $N$ photonic integrated circuits (PICs), and an optical multiplexer. In this architecture, the $N$ subcarriers are generated by a single multi wavelength source. However, such a source may be replaced by $N$ lasers, one per subcarrier. Each client is processed in the electronic domain (e.g., for filtering) and then is routed by the switching matrix to a specific PIC. The generated carriers are equally spaced according to the spectral requirements and transmission technique adopted. Generated subcarriers are selected at the multi wavelength source, and they are routed the appropriate PICs. Each PIC is utilized as a single-carrier transponder that generates different modulated signals, such as 16-QAM and QPSK, in order to support multiple modulation formats. Finally, subcarriers are aggregated by the optical multiplexer in order to form a super channel. Sometime, subcarriers may be sliced and directed to specific output ports according to the traffic needs. A detailed description of PIC generation of different modulated signals is given in [24].

Fig. 7. (a) Functionalities of (a) BVT, and (b) SBVT.

Fig. 8. Architecture of SBVT.
B. Bandwidth-variable Cross-connect

The BV-WXC [11] is used to allocate an appropriate-sized cross-connection with the corresponding spectrum bandwidth to support an elastic optical lightpath. Therefore, a BV-WXC needs to configure its switching window in a flexible manner according to the spectral width of the incoming optical signal.

![Diagram of BV-WXC](image)

**Fig. 9.** Architecture of BV-WXC.

Figure 9 shows an implementation example of a BV-WXC, where bandwidth-variable spectrum selective switches (BV-SSSs) in the broadcast-and-select configuration are used to provide add-drop functionality for local signals as well as groomed signal, and routing functionality for transit signals. Typically, a BV-SSS performs wavelength demultiplexing/multiplexing and optical switching functions using integrated spatial optics. The light from an input fiber is divided into its constituent spectral components using a dispersive element. The spatially-separated constituent spectra are focused on a one-dimensional mirror array and redirected to the desired output fiber. Liquid crystal on Silicon (LCoS) or Micro-Electro Mechanical System (MEMS)-based BV-SSSs can be employed as switching elements to realize an optical cross-connect with flexible bandwidth and center frequency. As the LCoS is deployed according to phased array beam steering, which utilizes a large number of pixels, LCoS-based BV-SSSs can easily provide variable optical bandwidth functionality. A detailed description of a BV-WSS employing LCoS technology can be found in [18], [28]. Similarly, details of an MEMS-based BV-SSS can be found in [18], [29].

IV. NODE ARCHITECTURES

This section discusses various node architectures [30], [31], which are the building blocks of spectrum efficient elastic optical networks.

A. Broadcast-and-Select

The broadcast-and-select architecture has been used to determine the elastic optical node architecture that uses spectrum selective switches [31]. Figure 10 shows the node architecture of broadcast-and-select, which is implemented using splitters at the input ports. Splitters are used to generate copies of the incoming signals that are subsequently filtered by spectrum selective switches in order to select the required signals at the receiver side. The add/drop network may implement colorless, direction-less, and contention-less elastic add/drop functionality, thus allowing the addition of one or more wavelength channels to an existing multi-wavelength signal automatically. It can also drop (remove) one or more channels from the passing signals to another network path dynamically. The main drawbacks of the broadcast-and-select node architecture are as follows - (i) it requires synchronization and rapid tuning, (ii) it cannot support wavelength reuse and hence a large number of wavelength channels is required, (iii) the signal power is split among various nodes, so this type of node cannot be used for long distance communication. The broadcast and select architecture is mostly being used in high-speed local area networks and metropolitan area networks. It must noted that the broadcast-and-select architecture struggles to support additional functionality to cope with dynamic requirements, e.g., spectrum defragmentation [32]–[34].

![Diagram of Broadcast-and-Select](image)

**Fig. 10.** Node architecture of broadcast-and-select.

B. Spectrum Routing

The spectrum routing node architecture is being designed to overcome the problems with the broadcast-and-select node architecture. It is basically implemented with arrayed waveband gratings [18] and optical switches as shown in Fig. 11. In spectrum routing, both switching and filtering functionalities are controlled by the spectrum selective switches. The basic
advantage of this architecture, compared to the broadcast-and-select architecture, is that the through loss is not dependent on the number of degrees. However, it requires additional spectrum selective switches at the input fibers, which makes it more expensive to realize. Furthermore, the additional functionality needed to cope with dynamic requirements, e.g., spectrum defragmentation [32]–[34], is still difficult to implement in this architecture.

\[ \text{SSS}_1 \]

Fig. 11. Node architecture of spectrum routing.

C. Switch and Select with Dynamic Functionality

We have already observed that the broadcast-and-select architecture and spectrum routing architecture are unable to support dynamic requirements, such as, spectrum defragmentation, time multiplexing, regeneration, etc. To overcome these limitations, the switch and select architecture with dynamic functionality has been introduced. In this architecture, an optical switch is used to direct copies of the input to a specific spectrum selective switch or to a module $(f)$ that provides additional functionalities, such as — defragmentation, time multiplexing, and regeneration. The outputs of the modules connect to spectrum selective switches, where the required signals are filtered for delivery to the corresponding output fiber. Figure 12 shows the node architecture of the elastic optical network with the dynamic functionalities that support dynamic requirements, namely — spectrum selective switch, fast switch, erbium-doped fiber amplifier (EDFA), spectrum defragmenter, splitter, etc. The inputs and outputs of the node are connected via the optical backplane as shown in Fig. 13. The different arrangements of inputs, modules, and outputs are realized by setting appropriate cross connections in the optical backplane. Therefore, it provides greater flexibility than the architectures explained above. This is mainly due to the non-mandatory nature of the components (such as — spectrum selective switch, power splitters and other functional modules) unlike static architectures, but they can be interconnected together in an arbitrary manner. The number of spectrum selective switches and other processing devices is not fixed but can be determined based on the specific demand for that functionality. Thus, savings in the number of devices can balance the additional cost of the optical backplane, and hence this type of architecture provides a cost-efficient solution. Furthermore, AoD provides considerable gains in terms of scalability and resiliency compared to conventional static architectures.

\[ \text{SSS}_1 \]

Fig. 12. Node architecture of switch and select with dynamic functionality.

D. Architecture on Demand

The architecture on demand (AoD) [35] consists of an optical backplane that is implemented with a large port-count optical switch connected to several processing modules, namely — spectrum selective switch, fast switch, erbium-doped fiber amplifier (EDFA), spectrum defragmenter, splitter, etc. The inputs and outputs of the node are connected via the optical backplane as shown in Fig. 13. The different arrangements of inputs, modules, and outputs are realized by setting appropriate cross connections in the optical backplane. Therefore, it provides greater flexibility than the architectures explained above. This is mainly due to the non-mandatory nature of the components (such as — spectrum selective switch, power splitters and other functional modules) unlike static architectures, but they can be interconnected together in an arbitrary manner. The number of spectrum selective switches and other processing devices is not fixed but can be determined based on the specific demand for that functionality. Thus, savings in the number of devices can balance the additional cost of the optical backplane, and hence this type of architecture provides a cost-efficient solution. Furthermore, AoD provides considerable gains in terms of scalability and resiliency compared to conventional static architectures.

\[ \text{SSS}_1 \]

Fig. 13. Node architecture on demand with $N$ input/outputs, and signal processing modules.
E. Comparing Node Architectures

Table I summarizes the above discussed node architectures in terms of total power loss, port count of switch/backplane, routing flexibility, port count of spectrum selective switches, defragmentation capability, time multiplexing, and regeneration capability. The calculation of total power loss [31] is determined by the type of node architecture implemented. In case of AOD, total power loss depends on the architecture implemented and the number of cross connections used in the optical backplane. The total power loss in the switch and select with dynamic functionality architecture depends on the spectrum selective switches, backplane, and modules used. However, the total power losses of the broadcast-and-select architecture and spectrum routing architecture mainly depend on the spectrum selective switches.

Port count of switch/backplane [31] varies the networking cost. The switch and select node and AOD node architectures need optical switches. However, the number of SSSs and other processing devices may be tailored to suit the specific demand. Therefore, as savings in the number of devices can offset the additional cost of the optical backplane, it provides an overall cost-effective solution. The number of SSS ports required by an AoD node is not strictly related to the node degree. This is because several small-port-count SSSs may be connected together in order to increase the number of available ports.

Routing flexibility is the capability of the system to carry signals from source to destination along different routes. This type of flexibility is required when strengthening system resilience to failures along working paths; signals may be directed to their backup paths. Time multiplexing is used to transmit and receive independent signals over a common signal path by synchronized switches at each end of the transmission line. As a result, each signal appears on the line only a fraction of the time in an alternating pattern. On the other hand, all-optical 3R (Re-amplification, Re-shaping, and Re-timing) signal regeneration is needed to avoid the accumulation of noise, crosstalk and non-linear distortion, and to ensure good signal quality for transmission over any path in an optical network. Spectral defragmentation is a technique to reconfigure the network so that the spectral fragments can be consolidated into contiguous blocks.

V. BASIC CONCEPT OF RSA

This section presents the basic concept of RSA in the elastic optical network. RSA is considered one of the key functionalities due to its information transparency and spectrum reuse characteristics. RSA is used to (i) find the appropriate route for a source and destination pair, and (ii) allocate suitable spectrum slots to the requested lightpath.

The RSA problem in elastic optical networks is equivalent to the RWA problem in WDM-based optical networks. The problem of establishing lightpaths for each connection request by selecting an appropriate route and assigning the required wavelength is known as the routing and wavelength assignment (RWA) problem [1], [2]. In WDM-based optical networks without wavelength converters, the same wavelength must be used on all hops in the end-to-end path of a connection. This property is known as the wavelength continuity constraint.

The difference of RSA and RWA is due to the capability of the elastic optical network architecture to offer flexible spectrum allocation to meet the requested data rates. In RSA, a set of contiguous spectrum slots is allocated to a connection instead of the wavelength set by RWA in fixed-grid WDM-based networks. These allocated spectrum slots must be placed near to each other to satisfy the spectrum contiguity constraint. If enough contiguous slots are not available along the desired path, the connection can be broken up into small multiple demands. Each one of these smaller demands would then require a lower number of contiguous subcarrier slots. Furthermore, the continuity of these spectrum slots should be guaranteed in a similar manner as demanded by the wavelength continuity constraint. If a demand requires \( t \) units of spectrum, then \( t \) contiguous subcarrier slots must be allocated to it (due to the spectrum contiguity constraint), and the same \( t \) contiguous slots must be allocated on each link along the route of the demand (due to the spectrum continuity constraint).

The concept of the contiguity and continuity constraints of the spectrum allocation is explained with an example. For this purpose, we consider the network segment shown in Fig. 14. We assume a connection request that requires a bit-rate equivalent to two slots for RSA from source node 1 to destination node 4. The connection request cannot be established through the shortest route 1-2-4 because the links from 1-2 and 2-4 have two contiguous slots that are not continuous, so the continuity and contiguity constraints are not satisfied. However, the continuity and contiguity constraints are satisfied if the connection uses the route 1-2-3-4, and spectrum slots 5 and 6.

RSA in WDM-based optical networks is an NP-hard prob-
These algorithms are discussed below.

1) Fixed Routing: In fixed routing (FR) [1], [38], a single fixed route is precomputed for each source-destination pair using some shortest path algorithm, such as Dijkstra’s algorithm [43]. When a connection request arrives in the network, this algorithm attempts to establish a lightpath along the predetermined fixed route. It checks whether the required slot is available on each link of the predetermined route or not. If even one link does not have the slot desired, the connection request is blocked. In the situation when more than one required slot is available, a spectrum allocation policy is used to select the best slot.

2) Fixed Alternate Routing: Fixed alternate routing (FAR) [1], [38] is an updated version of the FR algorithm. In FAR, each node in the network maintains a routing table (that contains an ordered list of a number of fixed routes) for all other nodes. These routes are computed off-line. When a connection request with a given source-destination pair arrives, the source node attempts to establish a lightpath through each of the routes from the routing table taken in sequence, until a route with the required slot is found. If no available route with required slot is found among the list of alternate routes, the connection request is blocked. In the situation when more than one required slot is available on the selected route, a spectrum allocation policy is used to choose the best slot. Although the computation complexity of this algorithm is higher than that of FR, it provides comparatively lower call blocking probability than the FR algorithm. However, this algorithm may not be able to find all possible routes between a given source-destination pair. Therefore, the performance of FAR algorithm in terms of call blocking probability is not optimum.

3) Least Congested Routing: Least congested routing (LCR) [1], [38] predetermines a sequence of routes for each source-destination pair similar to FAR. Depending on the arrival time of connection requests, the least-congested routes are selected from among the predetermined routes. The congestion on a link is measured by the number of slots available on the link. If a link has fewer available slots, it is considered to be more congested. The disadvantage of LCR is its higher computation complexity; its call blocking probability is almost the same as that of FAR.

4) Adaptive Routing: In adaptive routing (AR) [38], [39], routes between source-destination pairs are chosen dynamically, depending on link-state information of the network. The network link-state information is determined by the set of all connections that are currently active. The most acceptable form of AR is adaptive shortest path routing, which is well suited for use in optical networks. Under this approach, each unused spectrum in the network has a cost of 1 unit, whereas the cost of each used spectrum in the network is taken to be \( \alpha \). When a connection arrives, the shortest path between a source-destination pair is determined. If there are multiple paths with the same distance, one of them is chosen at random. In AR, a connection is considered blocked mainly when there is no route with a required slot between the source-destination pair. Since AR considers all possible routes between source-destination pair, it provides lower call blocking probability, but its setup time is comparatively higher than other routing algorithms. AR requires extensive support from control and
management protocols to continuously update the routing tables at the nodes. AR suits centralized implementation rather more than the distributed alternative.

Fig. 15. Fixed (solid line), alternate (dotted line) and adaptive (dashed line) routes are shown from source city CA to destination city L.

The functionality of the above mentioned routing algorithms is explained with the help of the sample example network, see Fig. 15. It consists of 14 nodes (representing cities) and 21 bi-directional optical links. The fixed shortest route, alternate route, and adaptive route from source city CA to destination city L are shown by the solid, dotted, and dashed lines, respectively. Furthermore, the congested links are denoted as α. If a connection request for a connection from source city CA to destination city L arrives, only AR is able to find a route between CA and L.

5) Comparisons of Routing Algorithms: A significant amount of work on the different issues of routing has been reported. Table II summarizes the major routing algorithms, and compares their performance in terms of blocking probability, average setup time, and time complexity [127]. The blocking probability [44], [45] in the network is defined as the ratio of the number of blocked connection requests to the number of connection requests in the network. The average setup time [6] in the network is defined as total execution time required to establish all the connections in the network to the number of successful connections. From Table II, we observe that FR has the lowest average setup time and time complexity of all routing algorithms. However, its blocking probability is the highest. AR provides the best performance in terms of blocking probability, but its time complexity is the highest. FAR offers a trade-off between time complexity and blocking probability.

B. Routing with Elastic Characteristics

An elastic optical network has the capability to slice the spectrum into slots with finer granularity than WDM-based networks. M. Jinno et al. [12] presented, for first time, the spectrum into slots with finer granularity than WDM-based protocols to continuously update the routing tables at the nodes. AR suits centralized implementation rather more than the distributed alternative.

Fig. 16. Frequency slot approach for elastic optical networks.

In elastic optical networks, single path routing via the RSA approach can create the spectrum fragmentation problem in and thus inefficiency. The spectrum fragmentation issue is explained in detail in Section VIII-A. To overcome this problem, multi-path routing [79], [80], [126] has been considered for the elastic optical network. An example of this routing is shown in Fig. 17. Let us consider connection request \( C(S, D, F) \), where \( S \) and \( D \) are source, destination, and the number of required contiguous sub-carrier slots, respectively. We assume connection requests arrive in time order, and each guard-band is assumed to be two sub-carrier slots wide. The group of consecutive frequency slots that are available after spectrum allocation for \( R_1 \), \( R_2 \), \( \cdots \), \( R_6 \) is referred to as a spectrum fragment. In this context, if connection request \( R_7 \) arrives at node 1 for destination node 3 with demand of four consecutive sub-carrier slots, it is rejected as they are unavailable. However, two spectrum paths, i.e., \( P_1 \) and \( P_2 \), can be established in parallel for \( R_7 \) by using multipath routing. Multi-path routing can be used to handle the spectrum fragments that are very common in dynamic traffic scenarios.

VII. Spectrum Allocation

With the aim of better fitting the bandwidth requirements at each moment, the lightpaths established in a network may dynamically change their allocated spectrum. This capability
is defined as elastic spectrum allocation [46], [47] and its implementation in future flexgrid networks is expected to provide better network performance. Spectrum allocation may be performed either after finding a route for a lightpath or in parallel during the route selection process. This section discusses the different spectrum allocation policies. We categorize the spectrum allocation based on spectrum range for connection groups and spectrum slot for individual connection request, as presented below.

### A. Spectrum Range Allocation for Connection Groups

Policies used to allocate spectrum range for connection groups can be categorized into three types, namely — (i) fixed spectrum allocation, (ii) semi-elastic spectrum allocation and (iii) elastic spectrum allocation, based on the changes allowed to the resources allocated to lightpaths in terms of central frequency (CF) and spectrum width.

1) **Fixed Spectrum Allocation**: In the fixed spectrum allocation (fixed SA) policy [46], [47], both CF and assigned spectrum width remain static for ever. At each time period, demands may utilize either whole or only a fraction of the allocated spectrum to convey the bit rate requested for that period. Therefore, this policy does not provide any elasticity. Under this policy, the spectrum allocation of lightpaths is independent of variations of bandwidth requirements. When the bandwidth demand of a connection is lower than the capacity of the assigned spectrum, the connection request is established. In this case, the utilized spectrum for carrying traffic is lower than that of allocated spectrum. This can lead to a sub-optimal use of network capacity. When the bandwidth demand is higher than the capacity of assigned spectrum, some bandwidth is not served.

Figure 18 shows the concept of the fixed-SA policy. The bandwidth demand for the lightpath at time $T$ is equal to the capacity of the assigned spectrum. In an underused spectrum condition, the bandwidth demand is lower than the capacity of the assigned spectrum in time $T'$ as shown in Fig. 18(a). Similarly, the bandwidth demand is higher than the capacity of the assigned spectrum in time $T'$ under insufficient spectrum condition as shown in Fig. 18(b).

2) **Semi-Elastic Spectrum Allocation**: In the semi-elastic spectrum allocation (semi-elastic SA) policy [46], [47], the CF remains fixed, but the allocated spectrum width can vary in each time interval. The frequency slices are allocated to a lightpath so as to suit the required bandwidth at any time. As a result, the unutilized slots can be used for subsequent connection requests. Therefore, this spectrum allocation policy provides higher flexibility than the fixed SA policy. To explain semi-elastic SA, two scenarios are considered below.

(i) If the required bandwidth demand is reduced, the capacity of the allocated spectrum can also be reduced. The unnecessary spectrum slices at each end of the allocated spectrum can be released and may be used for subsequent connection requests. Figure 19(a) represents a spectrum slot reduction condition, where both utilized and allocated spectra of the channel occupy the same number of slices at time $T'$.

(ii) If the required bandwidth demand is increased, new contiguous spectrum slices can be allocated at both ends of the CF. The capacity of the allocated spectrum can be increased in order to serve the maximum required bandwidth. Figure 19(b) represents a spectrum slot expansion condition, where a lightpath increases its required bandwidth from six to eight slots.

3) **Elastic Spectrum Allocation**: In the elastic spectrum allocation (elastic SA) policy [46], [47], both assigned CF and the spectrum width can be changed in each time interval. This spectrum allocation policy adds a new degree of freedom to the previous policy. It not only allows the number of slots per lightpath to be varied at any time, but also CF can be changed. Furthermore, we distinguish two conditions that differ in the grade of flexibility of CF movement as follows.

(i) **CF movement within a range**: As CF movement is limited to a certain range, the spectrum reallocation is
to the lightpath that varies its requirement at time $T$ and $T'$.
In this figure, both the assigned spectrum width and the CF are varied, but the CF is varied within the range specified.

(ii) Elastic spectrum reallocation: This condition reallocates the spectrum completely, and there is no CF movement limitation. The elastic spectrum allocation policy offers the best spectrum utilization performance among all spectrum allocation policies. Figure 20(b) represents a typical elastic spectrum allocation policy, where two lightpaths are reallocated and their allocated spectrum widths are varied.

A comparison of the CF movement approach within a range and the reallocation approach is given below. The first one limits CF movement and so has lower hardware requirements. However, the limitation yields only limited flexibility, while the elastic SA policy with reallocation approach is completely flexible.

B. Comparison of Spectrum Range Allocation policies for Connection Groups

Table III summarizes the different policies available to allocate spectrum range for connection groups in terms of hardware requirement, control plane, complexity of signal procedures, computation complexity, spectrum utilization, and nature of spectrum allocation. We observe that both CF and assigned spectrum width remain static in fixed SA policy. However, the control plane can be configured in order to allocate a fixed channel consisting of a fixed number of slices. The main drawback of the fixed spectrum allocation policy is the sub-optimal use of capacity, which makes this policy less convenient.

Compared to the fixed SA policy, the semi-elastic SA policy requires additional hardware. Moreover, the control plane can be configured in such way that it can allow established lightpaths to be modified. Since the amount of frequency slices are assigned to each lightpath dynamically, extension of the RSVP-TE protocol [48] should be designed so as to notify all BV-WXCs along the path to adjust their filter bandwidth, and modify the number of slices allocated to a path. Furthermore, some hardware is required to increase or decrease the utilized spectrum as needed. Therefore, bandwidth variable transponders and bandwidth variable switches should work with frequency steps in accordance with the frequency slice width. The semi-elastic SA policy has better performance than of the fixed SA policy in terms of spectrum efficiency at the cost of some extra hardware resources.

In the elastic SA policy, extra hardware and a control plane are required to vary both CF and spectrum width dynamically. As both CF and spectrum width vary dynamically, the elastic SA policy provides best performance in terms of spectrum utilization. However, the computation complexity and extra hardware requirements are high compared to other spectrum allocation policies.

C. Spectrum Slot Allocation for Individual Connection Request

The spectrum slot allocation of an individual connection request can be performed using one of the following allocation policies.

1) First fit: In the first fit spectrum allocation policy [49], [50], the spectrum slots are indexed and a list of indexes of available and used slots is maintained. This policy always attempts to choose the lowest indexed slot from the list of available slots and allocates it to the lightpath to serve the connection request. When the call is completed, the slot is returned to the list of available slots. By selecting spectrum in this manner, existing connections will be packed into a smaller number of spectrum slots, leaving a larger number of spectrum slots available for future use. Implementing this policy, does not require global information of the network. The first fit spectrum allocation policy is considered to be one of the best spectrum allocation policies due to its lower call blocking probability and computation complexity.

2) Random fit: In the random fit policy [1], [49], a list of free or available spectrum slots is maintained. When a connection request arrives in the network, this policy randomly selects a slot from the list of available slots and allocates it to the lightpath used to serve the connection request. After assigning a slot to a lightpath, the list of available slots is updated by deleting the used slot from the free list. When a call is completed, its slot is added to the list of free or available slots. By selecting spectrum in a random manner, it can reduce the possibility of multiple connections choosing the same spectrum which is possible if spectrum allocation is performed in a distributed manner.
TABLE III
SUMMARIES OF DIFFERENT POLICIES TO ALLOCATE SPECTRUM SLOTS TO A SINGLE CONNECTION REQUEST

<table>
<thead>
<tr>
<th>Spectrum Allocation Policy</th>
<th>Reference</th>
<th>Time Complexity</th>
<th>Applicable network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least used</td>
<td>[1], [2]</td>
<td>$O(L_1EZW)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>Most used</td>
<td>[1], [2]</td>
<td>$O(L_1EZW)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>Random fit</td>
<td>[1], [49]</td>
<td>$O(L_1WZ)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>First fit</td>
<td>[49], [50]</td>
<td>$O(L_1WZ)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>Last fit</td>
<td>[1], [22]</td>
<td>$O(L_1WZ)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>Exact fit</td>
<td>[49]</td>
<td>$O(L_1WZ)$</td>
<td>Single/multi-fiber networks</td>
</tr>
<tr>
<td>First-last fit</td>
<td>[50]</td>
<td>$O(L_1WZ)$</td>
<td>Single/multi-fiber networks</td>
</tr>
</tbody>
</table>

$N$: Number of fibers in the network, $M$: Number of links in the network, $W$: Number of connection requests in the network

TABLE V
NUMERICAL RESULTS OF DIFFERENT SPECTRUM ALLOCATION POLICIES IN RSA IN TERMS OF BLOCKING PROBABILITY [129]

<table>
<thead>
<tr>
<th>Spectrum Allocation Policy</th>
<th>Blocking probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-last fit</td>
<td>0.005 0.007 0.010 0.014 0.017</td>
</tr>
<tr>
<td>First-exact fit</td>
<td>0.006 0.007 0.012 0.015 0.017</td>
</tr>
<tr>
<td>First fit</td>
<td>0.005 0.008 0.014 0.014 0.017</td>
</tr>
<tr>
<td>Last fit</td>
<td>0.007 0.008 0.015 0.020 0.020</td>
</tr>
<tr>
<td>Random fit</td>
<td>0.017 0.022 0.031 0.043 0.055</td>
</tr>
</tbody>
</table>

3) Last fit: This policy [1], [22] always attempts to choose the highest indexed slot from the list of available slots and allocates it to the lightpath to serve the connection request. When the call is completed, the slot is returned back to the list of available slots.

4) First-last fit: In the first-last fit spectrum allocation policy [50], all spectrum slots of each link can be divided into a number of partitions. The first-last fit spectrum allocation policy always attempts to choose the lowest indexed slots in the odd number partition from the list of available slots. For the even number partitions, it attempts to choose the highest indexed slots from the list of available spectrum slots. Use of first fit and random fit spectral allocation approaches always attempt to choose the lowest indexed slots for each partition and randomly selects slots, respectively, from the list of available spectrum slots. This may lead to a situation where spectrum slots may be available, but connection requests cannot be established due to unavailability of contiguous aligned slots. The first-last fit allocation policy is expected to give more contiguous aligned available slots than the random fit and first fit allocation policies.

5) Least used: The least used spectrum allocation policy [1], [2] allocates a spectrum to a lightpath from a list of available spectrum slots that have been used by the fewest fiber links in the network. If several available spectrum slots share the same minimum usage, the first fit spectrum allocation policy is used to select the best spectrum slot. Selecting spectrum in this manner is an attempt to spread the load evenly across all spectrum slots.

6) Most used: The most used spectrum allocation policy [1], [2] assigns spectrum to a lightpath from a list of available spectrum slots, which have been used by the most fiber links in the network. Similar to the least used spectrum allocation policy, if several available spectrum slots share the same maximum usage, first fit spectrum allocation policy is used to break the tie. Selecting spectrum in this way is an attempt to realize maximum spectrum reuse in the network.

7) Exact fit: Starting from the beginning of the frequency channel, the exact fit allocation policy [49] searches for the exact available block in terms of the number of slots requested.
for the connection. If there is a block that matches the exact size of requested resources, this policy allocates that spectrum. Otherwise, the spectrum is allocated according to the first fit spectrum allocation policy. By selecting spectrum slots in this way, we can reduce the fragmentation problem in optical networks.

Fig. 21. Spectrum slot usage pattern for a network segment.

To illustrate the functionality of the above mentioned spectrum allocation policies, we use the example shown in Fig. 21. If two connection requests arrive that use link 2 and link 3 with one slot demand for establishing lightpaths, the strategies proceed as follows. First fit spectrum allocation policy selects spectrum slot 2 for the first connection request, and slot 3 for the second connection request. First-last fit spectrum allocation policy selects spectrum slot 2 for the first connection request, and slot 12 for the second connection request. Slot 6 and slot 4 have been used three times and two times, respectively, in this example. Therefore, slot 6 and slot 4 are used by most used spectrum allocation policies. As slot 2 and slot 9 have not been used so far, the least used spectrum allocation policy selects these two slots for the two connections requests. Random fit allocation policy selects any two of slot 2, slot 3, slot 4 and slot 12 with equal probability. Exact fit spectrum allocation policy selects spectrum slot 6 for the first connection request, and slot 2 for the second connection request.

D. Comparisons of Spectrum Allocation policies for Individual Connection Request

A significant amount of published research has addressed different policies to allocate spectrum slots to individual connection requests. Table IV summarizes some major spectrum allocation policies. It is observed from Table IV that the least-used and most-used allocation policies have higher time complexity than the other allocation policies. These two spectrum allocation policies also require global information of the network. As random fit, first fit, last fit, exact fit, first-last fit spectrum allocation policies have lower time complexity, we present the performance of these spectrum allocation policies in terms of blocking probability under different traffic loads (in Erlang), see Table V. These numerical results [129] were obtained from a simulation study performed on NSFNET. The simulation parameters followed those of [22]. We observe from the numerical results that first-last fit is lower blocking probability than the other spectrum allocation schemes, as it provides less fragmentation than the other spectrum allocation policies. The blocking probability of first-exact fit is higher than that of first-last fit, but its blocking probability is lower than those of other spectrum allocation policies. First fit and last fit spectrum allocation policies provide almost similar performance. Finally, the blocking probability of random fit is highest among all spectrum allocation policies.

E. Joint RSA

We have already discussed routing and spectrum allocation in elastic optical networks separately in Section VI and Section VII, respectively. However, many researchers have presented joint RSA [51]–[53] by considering routing and spectrum allocation at the same time. They usually employ a matrix to describe link or path spectral status by considering spectrum continuity and contiguity constraints, and choose the best performance from among all available matrix candidates.

In this direction, X Liu et al. [51] presented a layer-based approach to design integrated multicast-capable routing and spectrum assignment (MC-RSA) algorithms for serving multicast requests efficiently and minimizing the bandwidth blocking probability in the elastic optical network. For each multicast request, the presented algorithms decompose the physical topology into several layered auxiliary graphs according to the network spectrum utilization. Then, based on the bandwidth requirement, a proper layer is selected, and a multicast light-tree is calculated for the layer. These procedures realize routing and spectrum assignment for each multicast request in an integrated manner. Similarly, two joint routing and spectrum allocation algorithms [52], namely — (i) fragmentation-aware RSA and (ii) fragmentation-aware RSA with congestion avoidance, have been presented by Y. Yin et al, to alleviate spectral fragmentation in the lightpath provisioning process.

There are some works [14], [54], [55] that focus on solving joint RSA with modulation selection. This type of problem is referred to as the routing, modulation, and spectrum allocation (RMSA) problem. Taking this direction, X. Zhou et al. [54] introduced a RMLSA problem for elastic optical networks. In their work, the authors introduced an integer linear programming (ILP) for RMLSA algorithm that minimizes the spectrum used to serve the traffic matrix, and presented a decomposition method that breaks RMLSA into its two substituent subproblems, namely (i) routing and modulation level and (ii) spectrum allocation, and solved them sequentially. Finally, they presented a heuristic algorithm, which serves connections one-by-one in order to solve the planning problem. In [55], the authors investigated the principle of how dynamic service provisioning fragments the spectral resources on links along a path, and presented corresponding RMSA algorithms to alleviate spectrum fragmentation in dynamic network environments.

VIII. VARIOUS ASPECTS RELATED TO RSA AND SBVT

The performance of the elastic optical network depends on not only its physical resources, like — transponders, physical links, usable spectral width, optical switches, etc., but also how the network is controlled. The objective of an RSA algorithm is to achieve the best performance within the limits set by the physical constraints. Recently, an increasing number of studies
have investigated solutions to the RSA problem in the elastic optical network. The RSA problem can be cast in numerous forms. This section discusses various issues related to RSA and SBVT as follows.

A. Fragmentation

Elastic optical networks allocate spectrum on contiguous subcarrier slots. As the size of contiguous subcarrier slots is elastic, it can be a few GHz or even narrower. Therefore, dynamically setting up and tearing down connections can generate the bandwidth fragmentation [14], [56] problem. It is the condition where available slots become isolated from each other by being misaligned along the routing path or discontinuous in the spectrum domain. Thus, it is difficult to utilize them for upcoming connection requests. If no available slot can fulfill the required bandwidth demand of a connection request, the connection request is considered to be rejected/rejected. This is called bandwidth blocking and drives network operators to periodically reconfigure the optical paths and spectrum slots. This is referred to as network defragmentation. Figure 22 shows the fragmentation problem.

To overcome the bandwidth fragmentation problem, many RSA approaches [12], [14], [34], [36], [57] have been published. In this direction, Kadohata et al. [32] and Zhang et al. [33] developed bandwidth defragmentation schemes by considering the green field scenario, where connections are totally rerouted. A. N. Patel, et al. [58] formulated the defragmentation problem in an elastic optical network as an ILP formulation to provide optimal defragmentation with consideration of the spectrum continuity and continuity constraints. They presented two heuristic algorithms, namely — (i) greedy-defragmentation algorithm and (ii) shortest-path-defragmentation for large-scale networks in order to maximize the spectrum utilization. Fragmentation-aware RSA algorithms or defragmentation approaches can be classified into two categories, namely — (i) proactive fragmentation-aware RSA and (ii) reactive fragmentation-aware RSA, which are discussed in Section VIII-A1 and Section VIII-A2, respectively.

1) Proactive Fragmentation-aware RSA: When a new request is admitted to the network, the proactive fragmentation-aware RSA technique attempts to prevent or minimize spectrum fragmentation in the network. In this direction, R. Wang et al. [50] have presented four spectrum management techniques for allocating spectrum resources to connections of different data rates. In their approaches, all connections share the whole spectrum using the first-fit spectrum allocation policy. A similar concept of spectrum reservation has been presented by K. Christodoulopoulos et al. [14]. In their approach, a block of contiguous subcarriers is reserved for each source-destination pair. In addition, subcarriers that are not reserved may be shared among all connections on demand.

2) Reactive Fragmentation-aware RSA: In a dynamic environment, the fragmentation problem cannot be completely eliminated. Therefore, reactive fragmentation-aware RSA algorithms attempt to restore the network’s ability to accommodate high-rate and long-path connections. The main objective of defragmentation is to reconfigure the spectrum allocation of existing connections in order to consolidate available slots into large contiguous and continuous blocks that may be used for upcoming connection requests. In this direction, R. Wang et al. [50] have presented a set of reactive defragmentation strategies that exploit hitless optical path shift (HOPS). HOPS technology shifts a connection to a new block of spectrum as long as the route of the connection does not change and the movement of the new spectrum does not affect other established connections. They presented a scheme that consolidates the spectrum slots freed by a terminated connection with other blocks of spectrum available along the links of its path.

Most of the approaches [12], [14], [34], [36], [50], [56], [57] in the literature perform bandwidth defragmentation after bandwidth fragmentation occurs in subcarrier-slices. This means that the traffic is disrupted by connection rerouting, as the bandwidth defragmentation is performed. Bandwidth defragmentation that uses connection rerouting increases the traffic delay and system complexity.

To overcome this serious issue, R. Wang and B. Mukherjee [21] have presented a scheme that prevents bandwidth fragmentation without performing any connection rerouting. Typically, when connection requests with lower-bandwidth and higher-bandwidth are not separated during spectrum allocation, it is more likely that the higher-bandwidth connection requests are blocked. In order to circumvent this drawback, they explore an admission control mechanism that captures the unique challenges posed by heterogeneous bandwidths. They adopt a preventive admission control scheme based on spectrum partitioning to achieve higher provisioning efficiency. As a result, it prevents the blocking of connections due to the unfairness of bandwidth issue.

Similarly, W. Fadini et al. [22] have presented a subcarrier-slot partition scheme for spectrum allocation in elastic optical networks; it reduces the number of non-aligned available slots without connection rerouting. Thus the bandwidth blocking probability in the network is reduced. In their approach, they define a connection group as a set of connections whose routes are exactly the same. When the spectrum resources of two connections from different connection groups sharing a common link are allocated to adjacent slots, some available slots might be non-aligned with each other. When another connection request arrives in the network and its route needs these available slots, the connection request is rejected if these available slots are non-aligned. The partitioning scheme divides the subcarrier slots into several partitions. Disjoint connections whose routes do not share any link are allocated to
the same partition, while non-disjoint connections are allocated to different partitions. In this way, their approach increases the number of aligned available slots in the network and hence the bandwidth blocking probability is reduced.

B. Modulation

The traditional WDM-based optical network assigns spectrum resources to optical paths without considering the appropriate modulation technique, which leads to an inefficient utilization of the spectrum. However, the OFDM-based elastic optical network allocates optical paths with consideration of adaptive modulation and bit rate to further improve the spectrum efficiency. In the modulation-based spectrum allocation scheme of [12], [59], the necessary minimum spectral resource is adaptively allocated to an optical path. The adaptation considers the physical conditions while ensuring a constant data rate. The modulation-based spectrum allocation scheme improves the spectrum efficiency, as the allocated spectral bandwidth can be reduced for shorter paths by increasing the number of modulated bits per symbol.

In this direction, M. Jinno et al. [12] have presented a distance-adaptive spectrum allocation scheme that adopts a high-level modulation format for long distance paths, and a low-level modulation format to shorter paths. As the optical signal-to-noise ratio (OSNR) tolerance of 64-QAM is lower than that of QPSK, it suits shorter distance lightpaths as shown in Fig. 23.

![Modulation level versus transmission distance](image)

Fig. 23. Modulation level versus transmission distance.

The modulation based spectrum allocation schemes can be classified into two categories, namely — (i) offline modulation based spectrum allocation, and (ii) online modulation based spectrum allocation, and are discussed below.

1) Offline Modulation based Spectrum Allocation: K. Christodouloupolos et al. [14] have presented an offline modulation based spectrum allocation scheme, where a mapping function is provided as input to the problem. In their scheme, each demand is mapped to a modulation level according to the requested data rate and the distance of the end-to-end path. Initially, they presented a path-based ILP formulation for their scheme, and then decomposed the problem into two sub-problems, namely (i) routing and modulation level (RML), and (ii) spectrum allocation. They solved the sub-problems sequentially using ILPs. Finally, a sequential algorithm was presented to serve connections one-by-one, and to solve the planning problem sequentially.

2) Online Modulation based Spectrum Allocation: Most of the studies on online modulation based spectrum allocation [60]–[63] have introduced heuristic algorithms, which deal with randomly arriving connection requests. Initially, these algorithms compute a number of fixed-alternate paths for each source-destination pair, and arrange them in decreasing order of their end-to-end path length. In the second step, a spectrum allocation policy is used to allocate a lightpath to each connection request by considering alternate path routing and modulation.

Recent studies [15], [60], [61] on modulation based spectrum allocation claim that this type of spectrum allocation scheme increases the spectral utilization by approximately 9%–60% compared to fixed-modulation based spectrum allocation in the elastic optical network. Fixed-modulation based spectrum allocation schemes do not consider the most appropriate modulation technique for different connection requests according to their lightpath distance. Typically, they select, conservatively, one modulation technique for all connection requests regardless their lightpath distance. As an example, a fixed-modulation based spectrum allocation scheme adopts BPSK modulation format for all connection requests regardless their lightpath distance. As a result, this type of spectrum allocation schemes does not utilize the spectrum efficiently. On the other hand, modulation-based spectrum allocation schemes determine the modulation technique that best suits each lightpath distance. As an example, a modulation-based spectrum allocation scheme adopts BPSK for long distance lightpaths, and 16-QAM for shorter distance lightpaths. This minimizes the numbers of spectrum slots that must be assigned, which yields better utilization of spectrum resources compared to fixed-modulation based spectrum allocation schemes.

C. Quality of Transmission

The elastic optical network architecture offers the ability to choose the modulation format and channel bandwidth to suit the transmission distance and quality of transmission desired. One version of the online modulation based spectrum allocation scheme is referred to as quality-of-transmission aware RSA. In this direction, H. Beyranvand et al. [63] have presented a quality of transmission (QoT) aware online RSA scheme for the elastic optical network. Their approach employs three steps, namely — (i) path calculation, (ii) paths election, and (iii) spectrum assignment to construct the complete framework. The Dijkstra and k-shortest path algorithms have been adapted for computing paths, while fiber impairments and non-linear effects on the physical layer are modeled to estimate the QoT along the given route. S. Yang et al. have presented [62] a QoT-aware RSA scheme in order to select a feasible path for each requested connection and allocate subcarrier slots by using modulation formats appropriate for the transmission reach and requested data rate.

D. Traffic Grooming

In WDM-based wavelength-routed optical networks, traffic grooming [1], [5], [64]–[70] is used to multiplex a number of low-speed connection requests into a high-capacity wavelength
channel for enhancing channel utilization. Traffic grooming improves the resource utilization by aggregating multiple electrical channels (packet or circuit flows) onto one optical channel.

The elastic optical network allocates spectral resources with just enough bandwidth to satisfy the traffic demands. However, traffic grooming [26], [67], [71]–[73] is still essential for the following reasons. (i) BVT is normally designed so as to maximize the traffic rate in the network, and it does not support slicing at a very early stage [74]. Electrical traffic grooming is applied in order to use transponder capacity efficiently. (ii) Generally speaking, a filter guard band between two adjacent channels should be assigned to resolve optical filter issues. Traffic grooming can minimize filter guard band usage by aggregating traffic electrically. The electrical switching fabric is still needed for traffic grooming in the elastic optical network, similar to WDM networks. The main difference is that the transponder used in elastic optical networks does not strictly follow the ITU-T central frequency. As a result, it can provide flex-optical channels.

To further improve the flexibility and eliminate the electrical processing, researchers have designed SBVT for the elastic optical network. In SBVT-based elastic optical networks, the traffic grooming [26], [75] function can be partly offloaded from the electrical layer to the optical layer. Multiple electrical channels are groomed onto one sub-transponder channel, in which each sub-transponder channel is associated with a flex-optical channel. Multiple sub-transponder channels (flex-optical channels) are groomed optically onto one transponder by using an optical switching fabric (e.g., BV-OXC), which is called optical traffic grooming. Figure 24 distinguishes between traffic grooming in WDM-based optical networks, traffic grooming with BVT in elastic optical networks, and traffic grooming with SBVT in elastic optical networks. Spectrum efficiency and transponder usage are improved in WDM-based optical networks, whereas traffic grooming with BVT in the elastic optical network improves transponder usage and reduces guard band usage. Finally, traffic grooming with SBVT in the elastic optical network eliminates electrical processing by offloading parts of the grooming function to the optical layer.

Y. Zhang et al. [71] have incorporated, for the first time, a grooming approach for the RSA problem in the elastic optical network with BVT. In their approach, multiple low-speed connection requests are groomed into elastic optical paths by using electrical layer multiplexing. They presented a mixed integer linear program (MILP) formulation to reduce the average spectrum utilization in the traffic-grooming scenario. S. Zhang et al. [67] have presented a multi-layer auxiliary graph to implement various traffic-grooming policies by properly adjusting the edge weights in the auxiliary graph. With their approach, they have shown that there is a tradeoff among different traffic-grooming policies, and that the spectrum reservation scheme can be incorporated into various traffic-grooming scenarios. Recently, J. Zhang et al. [26] have presented dynamic traffic grooming in SBVT-enabled elastic optical networks. In their approach, a three-layered auxiliary graph (AG) model has been presented to address mixed-electrical-optical grooming under the dynamic traffic scenario. By adjusting the edge weights of AG, various traffic-grooming policies can be achieved for different purposes. Furthermore, two spectrum reservation schemes have been introduced in order to efficiently utilize transponder capacity. Finally, they compared different traffic-grooming policies under two spectrum reservation schemes, and the tradeoff among the policies was shown.

E. Survivability

The elastic optical network has the capability to support individual data rates from 400–1000 Gb/s [76]. It also aggregates the throughput per fiber link to approximately 10–100 Tb/s. Therefore, failure of a network component, such as optical fiber or network node, can disrupt communications for millions of users, which can lead to a great loss of data and revenue. As an example, in 2004, the Gartner Research Group had lost approximately 500 million dollars due to failure of a optical network [77]. Thus, survivability against failure has become an essential requirement of the elastic optical network. Failure recovery [78] is defined here as “the process of re-establishing traffic continuity in the event of a failure condition affecting that traffic, by re-routing the signals on diverse facilities after the failure”. A network is defined as survivable [78], if its recovery can be secured rapidly. Similar to WDM-based optical networks, the survivability mechanisms [79]–[81] for the elastic optical network can be classified into two broad categories, namely — (i) protection and (ii) restoration, which are discussed briefly in the following subsections.

1) Protection: The protection techniques of [81]–[85] use backup paths to carry optical signals after fault occurrence. The backup paths are computed prior to fault occurrence, but they are reconfigured after fault occurrence. In this direction, M. Klinkowski et al. [85] have presented an RSA approach with dedicated protection for static traffic demands. Although dedicated protection can provide more reliability, it is unable to utilize the spectrum slots properly as some of the spectrum slots are reassigned prior to fault occurrence. To overcome this problem, M. Liu et al. [81] presented a shared protection scheme to enhance the spectrum utilization by sharing backup spectrum slots between two adjacent paths on a link, if the corresponding working paths are link-disjoint. They explored the opportunity of sharing enabled by tunable transponders. The elasticity of the transponder enables the expansion and contraction of paths. As a result, the backup spectrum is used by only one of the adjacent paths at a time. Similarly, G. Shen et al. [82] addressed a shared protection technique for elastic optical networks to minimize both required spare capacity and maximum number of used link spectrum slots.

2) Restoration: In restoration [80], [86]–[91], backup paths are computed dynamically on the basis of link-state information after fault occurrence, and hence can provide more efficiency in terms of resource utilization compared to protection. In this direction, F. Ji et al. [92] presented three algorithms for dynamic preconfigured-cycle (p-cycle) configuration in order to provide the elastic optical network with 100% restoration against single-link failure. The first algorithm configures the
working path and p-cycles of a request together according to the protection efficiencies of the p-cycles. In order to reduce the blocking probability, they presented a spectrum planning technique that regulates the working spectrum and protection resources, and finally, two algorithms based on the protected working capacity envelope cycles and Hamiltonian cycles.

F. Paolucci et al. [91] have presented a restoration technique enabling multipath recovery and bit rate squeezing in the elastic optical network. They exploited the advanced flexibility provided by sliceable bandwidth-variable transponders that support the adaptation of connection parameters in terms of the number of sub-carriers, bit rate, transmission parameters, and reserved spectrum resources. They formulated their problems as an ILP model and finally, presented an heuristic algorithm that efficiently recovers network failures by exploiting limited portions of the spectrum resources along multiple routes. As restoration finds the backup paths after fault occurrence, it offers slower recovery than protection. Depending on the type of rerouting used, restoration can be considered as consisting of three categories, namely — (i) link restoration, (ii) path restoration and (iii) segment-based restoration. Link restoration discovers a backup path for the failed connection only around the failed link. In path restoration, the failed connection independently discovers a backup path on an end-to-end basis. Segment-based restoration discovers a segment backup path of the failed connection.

ESF: Electrical switch fabric
OSF: Optical switch fabric

Fig. 24. Comparison of traffic grooming in (a) WDM-based optical networks, (b) elastic optical networks with BVTs, and (c) elastic optical networks with SBVTs.
F. Energy Saving

The energy consumption of telecom networks is drastically increasing with the increase in traffic. The IP router consumes the maximum amount of energy in IP-over-WDM-based optical networks [94]. When transmission rate increases, the optical transponder associated with the IP router is a huge energy consumer in optical networks [95]. Therefore to minimize energy consumption, it is essential to reduce the number of IP router ports and optical transponders. By using the advantages of SBVT, the elastic optical network can offer some new features for traffic grooming (shown in Fig. 24) and optical layer bypass, which can help to reduce the energy consumption. In this direction, J. Zhang et al. [96] studied the power consumption of IP-over-elastic optical networks with different elastic optical transponders. The results of their studies show that significant energy savings are possible if SBVT is used rather than the fixed BVT.

Recently, researchers [93], [97], [98] have focused on energy efficient RSA schemes for the elastic optical network. In this direction, A. Fallahpour et al. [93] have presented a dynamic energy efficient RSA algorithm that considers regenerator placement to suppress the total network energy consumption. In their work, a virtual graph is designed based on the given network topology, where the cost functions of the virtual graph are computed according to the energy consumption of the corresponding links and intermediate routers. Furthermore, a newly arrived connection request is served by finding the most energy-efficient path among the possible candidate paths. Similarly, S. Zhang et al. [98] have presented energy-efficient dynamic provisioning in order to significantly reduce energy consumption and make efficient use of spectrum resources. In their research work, they adopt an auxiliary graph, and from it created a dynamic provisioning policy called time-aware provisioning with bandwidth reservation (TAP-BR). The TAP-BR policy incorporates the two important factors of time awareness and bandwidth reservation, in order to facilitate energy-efficient provisioning.

Several studies on energy saving in the elastic optical network are anticipated, and more research work is needed to develop truly effective energy-saving RSA schemes.

G. Networking cost of SBVT

This subsection discusses the networking cost reduction made possible by the use of SBVTs in the elastic optical network. We have observed from the literature that SBVTs allow the reuse of hardware and optical spectrum by transmitting data to multiple destinations. SBVTs enable point to multiple point transmission where the traffic rate to each destination and the number of destinations can be freely set to satisfy the request. On the other hand, the non-sliceable transponder requires at least one interface for each destination, which increases networking cost.

In this direction, V. López et al. [25], [27] have presented two node models, namely — (i) non-sliceable transponder model, and (ii) SBVT model in order to compare the networking cost, please see Fig. 25. The main difference between these two models is that the non-sliceable transponder model requires at least one interface for each destination, while the SBVT transponder reuses hardware and optical spectrum to transmit to multiple destinations. The model without sliceable transponders considers coherent modulation formats, such as — 40 Gb/s, 100 Gb/s, 400 Gb/s and 1 Tb/s, whereas only 400 Gb/s or 1 Tb/s SBVTs are considered by the SBVT model.

Finally, the result of the research work [25] has claimed that using 400 Gb/s and 1 Tb/s SBVTs reduces transponder costs in the network by at least 50% in a core network scenario. This reduction was calculated relative to BVTs of 400 Gb/s and 1 Tb/s in the non-sliceable scenario.

A significant number of works have addressed various aspects of the RSA problem in the elastic optical network. Table VI summarizes these different aspects, namely — fragmentation, modulation, quality-of-transmission, traffic grooming, survivability, energy saving, and networking cost reductions by SBVT.

<table>
<thead>
<tr>
<th>Various issues</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragmentation</td>
<td>Proactive [14],[21],[22]</td>
</tr>
<tr>
<td></td>
<td>Reactive [12],[14],[34],[36],[30],[36],[57]</td>
</tr>
<tr>
<td>Modulation</td>
<td>Online [60]–[63]</td>
</tr>
<tr>
<td></td>
<td>Offline [14]</td>
</tr>
<tr>
<td>Quality of transmission</td>
<td>[62],[63]</td>
</tr>
<tr>
<td>Traffic grooming</td>
<td>[67],[71],[72],[91],[26],[73]</td>
</tr>
<tr>
<td>Survivability</td>
<td>Protection [81]–[84]</td>
</tr>
<tr>
<td></td>
<td>Restoration [80],[86]–[90],[92]</td>
</tr>
<tr>
<td>Energy saving</td>
<td>[93],[97],[98]</td>
</tr>
<tr>
<td>Networking cost caused by SBVT</td>
<td>[25],[27]</td>
</tr>
</tbody>
</table>
IX. EXPERIMENTAL DEMONSTRATIONS

This section presents the experimental demonstrations that have tested the functionality of the elastic optical network. The first experimental demonstration was conducted by M. Jinno et al. [99]. They demonstrated a spectrum-efficient elastic optical path network for 100 Gb/s services and beyond; it used flexible rate transceivers and BV-WXCs. Subsequently, many researchers moved from the conventional WDM-based optical network to the elastic optical network architecture.

In this direction, B. Kozicki et al. [100] presented an experimental demonstration on the architecture of an elastic optical network. They used the optical OFDM modulation format and BV-WXCs to generate, transmit and receive optical paths with bandwidths of up to 1 Tb/s. They experimentally demonstrated elastic optical path setup and spectrally-efficient transmission of multiple channels with bit rates ranging from 40 to 140 Gb/s between six nodes in a mesh network. Furthermore, they demonstrated multi-hop transmission in a single-mode fiber over 400 km with 1 Tb/s data rate. Finally, they investigated the filtering properties and the required guard bandwidth for the spectrally-efficient allocation of optical paths in the elastic optical network. Similarly, their other research work [101] experimentally demonstrated path aggregation in a spectrum-sliced elastic optical network. Multiple OFDM 100-Gb/s optical paths are aggregated in the optical domain to form a spectrally continuous 1-Tb/s super-wavelength optical path and transmitted over a network of BV-WXCs. Finally, they evaluated the potential implementation issues and concluded that OFDM paths can be optically aggregated with an optical signal-to-noise ratio penalty of less than 1 dB.

A flexible-bandwidth network testbed with a real-time adaptive control plane has been demonstrated by D. J. Geisler et al. [102]. In their experiment, the modulation format and spectrum-positioning were adjusted to maintain QoS and high spectral efficiency. They have presented a performance monitoring method that dynamically reconfigures the modulation format in such a way that the optical signals maintain the required QoS and bit error rate even for signals that experience time-varying impairments. J. Zhang et al. [103] demonstrated enhanced software-defined networking (SDN) over elastic grid optical networks for data center service migration.

J. Zhang et al. [104] presented an experimental demonstration of the openflow-based control plane for elastic lightpath provisioning in flex-grid optical networks. Dynamic lightpath establishment and adjustment were implemented in their control plane testbed. Additionally, they reported the overall latency including signaling and hardware for lightpath setup and adjustment.

A path computation element (PCE) architecture for flexible optical networks has been demonstrated in [105] that maximizes the spectral efficiency. Confirmation was provided through two different experiments; they successfully showed the PCE ability to trigger dynamic rerouting with bit-rate or modulation format adaptation. In particular, the experiments demonstrated, in a testbed, the dynamic allocation of spectrum slots and adoption of modulation formats from 16-QAM to QPSK at 100 Gb/s, and bit-rate adaptation with 16-QAM from 200 Gb/s to 100 Gb/s.

Finally, S. Ma et al. [106] have presented and experimentally demonstrated a control-plane framework to realize online spectrum defragmentation in software-defined elastic optical networks in order to reduce the call blocking probability in the network.

In conclusion, a significant number of experimental demonstrations have been reported in the literature. Table VI summarizes different experimental demonstrations related to the elastic optical network. However, the practical implementation of the elastic optical network remains in development, and the full commercial deployment of the elastic optical network will be seen in the next few years.

<table>
<thead>
<tr>
<th>Experimental demonstrations</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>[99], [100], [107]</td>
</tr>
<tr>
<td>Control/management plane</td>
<td>[101], [103], [105], [108]–[110]</td>
</tr>
<tr>
<td>QoS</td>
<td>[102]</td>
</tr>
<tr>
<td>Survivability</td>
<td>[111]</td>
</tr>
<tr>
<td>Modulation</td>
<td>[102], [110], [112]</td>
</tr>
<tr>
<td>Defragmentation</td>
<td>[106], [113], [114]</td>
</tr>
</tbody>
</table>

X. RESEARCH ISSUES AND CHALLENGES

Flex-grid technology or elastic optical network is a promising concept but its implementation remains some way off. There are several issues and challenges, which need further research to resolve [115], [116]. Taking this direction, I. Tomkos et al. [115] reviewed the recent developments on the research topic of flexible/elastic networking and highlighted the future research challenges. In [116], the authors provided a comprehensive view of the different pieces composing the “flexible networking puzzle” with special attention given to capturing the occurring interactions between different research fields. In their work, physical layer technological aspects, network optimization for flexible networks, and control plane aspects were examined. Furthermore, future research directions and open issues were presented. Figure 26 summarizes the different areas demanding further work. In the following, we identify some interesting research opportunities for the elastic optical network.

![Fig. 26. Different research areas in elastic optical networks](image-url)
A. Hardware Development

Innovative and sophisticated devices and components must be developed in order to achieve high capacity spectrum efficient elastic optical networks. Novel optical switching and filtering elements need to be developed in order to provide efficient client protocol data unit mapping procedures that extract the incoming client signal via client-specific physical coding sublayers and media access controller layers, high resolution and steep filtering performance, optimum modulation format for bandwidth variability and higher nonlinear impairment tolerance, etc.

One of the important challenges faced by the research community is to develop a new sliceable bandwidth-variable transponder, which supports sliceability, multiple bit rates, multiple modulation formats, and code rate adaptability. However, 100 Gb/s OFDM transponders can adapt to lower bit rates. Therefore, fractional-bandwidth services may be provided with the use of the same device. This kind of technology and the use of optical integrated circuits may offer compact and cost-effective implementation. Similarly, the need for multiple temperature-stabilized, frequency controlled lasers can be improved by phase-locked carrier generation from a single laser source.

Recently, space division multiplexing (SDM) technology has been incorporated into elastic optical networks in order to develop high-capacity, next-generation, and few-mode/multicore fiber infrastructures. The realization of this type of infrastructure should be enabled by the development of novel multi-dimensional spectral switching nodes, which can be fabricated by extending the designs of existing flexible SSS nodes, through the addition of advanced mode/core adapting techniques.

To achieve long transmission reach, optical signals must be amplified at periodic regeneration points along the fiber span to compensate the power loss experienced in multi-core fiber. For regeneration, one technique is to demultiplex the SDM signals into multiple single-core fibers and then amplify the signals in each fiber using conventional single-core EDFAs. The amplified signals are then recombined and injected back into the multi-core fiber span, which increases the system delay. Therefore, to develop a single-core power transient-suppressed EDFA (TS-EDFAs) is one of the challenging research goals that must be accomplished; a key issue is to reduce the time taken to adjust the operation point of the amplifier since a newly added signal may suddenly change the total power at the EDFA input.

B. Network Control and Management

Traditional optical networks use a set of protocols (such as — generalized multi-protocol label switching (GMPLS), open shortest path first (OSPF), and resource reservation protocol - traffic engineering (RSVP)) for network control and management. Although these control protocols are well designed and standardized for traditional optical networks, evaluations that encompass the elastic technology are still at an early stage. The control plane of the elastic optical network must support many unique properties, such as — (i) optical channels are allowed to be flexible in size or width, (ii) optical channels can support various modulation formats, (iii) sub and super wavelength concept, (iv) support of multi-path routing of the composing waveband members of a split spectrum super-channel, (v) fast restoration upon failure, with or without resource reservation, etc. Therefore, new control protocols need to be developed or the existing protocols should be extended in order to support these unique properties of the elastic optical network.

During the last decade, the GMPLS protocol has been broadly standardized in different aspects including packet switching, label-switched paths (LSPs), and time-division multiplexing. However, its applicability to the elastic optical network has not been completely described yet. For this, GMPLS needs to address frequency slots instead of wavelengths. The control plane protocols are required to maintain coherent global information representing up to 320 or 640 slots, for 12.5 GHz or 6.25 GHz slot granularity, respectively [119]. G. Shen et al. [120] have indicated in their research work that spectrum granularity may become even finer reaching 3 GHz, resulting in 1280 possible slots. Some research works [119]– [121] have been addressed management of the control plane. Extensive research is need to adequately manage the control plane of the elastic optical network.

From the operation and control viewpoint, extending software-defined networking (SDN) toward transport networks, while retaining flexibility, is a challenging issue that needs to be addressed properly. In this direction, the Open Networking Foundation group [122] are addressing SDN and OpenFlow-based control capabilities for optical transport networks. In their research work, many activities such as — (i) identifying use cases, and (ii) defining reference architecture in order to control optical transport networks by incorporating the OpenFlow standard, have been performed to develop OpenFlow protocol extensions. Based on their model, the extended OpenFlow protocol is responsible for interfacing with network elements. The control virtual network interface is developed in order to provide the required bridge between the data center controllers. Future extensions and additional standardization activities are needed in order to realize the SDN controllers that can manage TDM circuit and wavelength-based architectures (such as — generic packet/TDM/fiber switching, bandwidth aggregation and segmentation). Finally, an efficient technology needs to be developed in order to support a combination of centralized and distributed control of a multi-layer network.

C. Energy Consumption

The increase in the traffic in carried by optical networks will increase the energy consumption. The elastic optical network has the ability to significantly reduce energy consumption. In combination with SBVTs, the elastic optical network presents some new features as regards optical traffic grooming and optical layer bypass [26], which can help to reduce the energy consumption. However, we still are unable to groom traffic optically in a very early stage, and this omission must be tackled.
The elastic optical network offers lower blocking probability compared to traditional optical networks and so can accept higher volumes of traffic. This clearly is a significant advantage in terms of energy efficiency, as the deployment of additional network elements would not only increase cost, but also increase the overall energy consumption. Several energy saving schemes are anticipated, setting some of the network elements into sleep mode when the traffic is below a certain threshold. Another interesting topic for future researchers is to analyze the energy efficiency of new protection and restoration schemes for the elastic optical network.

D. Physical Layer Impairments

As optical connections may span over many long links, physical layer impairments (PLIs), such as — dispersion, interference, noise, and nonlinear effects accumulate and degrade the signal quality, which affects the quality-of-transmission (QoT). Accounting for PLIs is a challenging issue for network designers, especially if we consider exact models and the interdependencies. Many studies [45], [123]–[125] on PLIs have been carried out for WDM-based optical networks. PLIs have distinct impact on both WDM-based optical networks and elastic optical networks. With the introduction of coherent detection and digital signal processing, impairments that are related to dispersion can be substantially reduced or fully compensated. However, high levels of flexibility make the minimization of these effects more complicated from an algorithmic perspective, which needs further research.

E. Spectrum Management

One of the most important problems in elastic optical networks, both for planning and operation, is allocating network resources in a dynamic environment. The problem of establishing connections in fixed-grid WDM-based networks and the allocation of network resources is well addressed in the literature [1], [2]. However, connection establishment in the elastic optical network is more complicated for several reasons. First, in contrast to WDM networks where each connection is assigned a single wavelength, in elastic optical networks, spectrum slots can be allocated in a flexible manner. Apart from the difference in spectrum resource allocation, the choice of the transmission parameters of the tunable transponders present in flexible networks directly or indirectly impacts the resource allocation decision and makes the problem even more complicated.

Some proposals [22] found in the literature partition the entire spectrum in an advance in order to handle spectrum resources efficiently. Most of these schemes assume the traffic demand in an advance. However, considering the fact that the traffic profile will change over time, connections with larger size slots demand will have to be accommodated over the same partitions. In this case, it is obvious that there is no other way than dropping these connections. As a result, the blocking performance will decrease significantly. Even worse, larger slot connections are the ones that will be dropped most often. Therefore, it is very important to take account of possible changes in the traffic profile. Therefore, partitioning the spectrum in a dynamic traffic environment is a challenging issue, which needs further research.

F. Disaster Management

Recovery of the network resources after large scale disasters such as — tsunami, hurricanes, floods, etc, is becoming increasingly important. The approach to disaster recovery is different from the approach to network failures such as fiber cuts or node failures. Network failures can be planned for by providing sufficient resources to deal with all network failures. However, there is no way to anticipate the impact of a disaster, and therefore no cost effective way to plan for full recovery from it. The recent development of SBVTs represents a new tool for handling disaster recovery. As SBVTs can play an important role in providing sufficient flexibility, elastic optical networks with SBVTs are an interesting research topic for disaster management.

XI. Concluding Remarks

The elastic network paradigm is an emerging research area, and a promising technology for future high-speed transmission due to its flexible properties. This tutorial has introduced and discussed different aspects of the elastic optical network.

We started with the basic concept of the elastic optical network and its unique properties, and then turned to its architecture and operation principle. The architecture of SBVT and its advantages in the future optical networking were detailed. Immediately after discussing network architecture, our discussion focused on the different node architectures, namely — broadcast and select, spectrum routing, switch and select with dynamic functionality, and architecture on demand, along with their functionalities. Next we looked into the basic concept of the RSA approach in elastic optical networks. Then, we moved to different routing approaches along with their pros and cons. Thereafter, the different spectrum allocation policies were addressed. In addition, we distinguished the spectrum allocation policies into two categories based on spectrum range allocation for connection groups and spectrum slot allocation for the individual connection request.

Various aspects, namely — fragmentation, modulation, quality-of-transmission, traffic grooming, survivability, energy saving, and networking cost related to RSA, which have been reported in the literature, were addressed in this tutorial paper. We classified them as either fragmentation-aware approaches or defragmentation approaches and explained how these approaches deal with fragmentation. Distance adaptive RSA that considers the modulation scheme is one of the new possibilities of the elastic optical network, and was also been presented in this paper. We have discussed and compared traffic grooming in WDM-based optical networks, elastic optical networks with BVTs, and elastic optical networks with SBVTs. Different survivability techniques, namely — protection and restoration, were also covered. We discussed the networking cost reduction made possible by the use of SBVTs in elastic optical networks.

Finally, we explored the experimental demonstrations that have tested the functionality of the elastic optical network.
Research challenges and open issues that flexible networking poses were presented to guide future research.

REFERENCES


Bijoy Chand Chatterjee received the PhD degree from the Department of Computer Science & Engineering, Tezpur University, Assam in 2014. Presently, he is working as a Postdoctoral Researcher in the Department of Communication Engineering and Informatics, the University of Electro-Communications, Tokyo, Japan. During his PhD, he was awarded an IETE Research Fellowship by The Institute of Electronics and Telecommunication Engineers (IETE), India. His research interests include QoS-aware protocols, cross-layer design, optical networks and elastic optical networks. He is a member of IEEE, and a life member of IETE.

Dr. Nityananda Sarma received B E (Computer Science & Engineering), M Tech (Computer Science & Engineering) and Ph D (Computer Science & Engineering) degrees from Dibrugarh University, IIT Kharagpur and IIT Guwahati respectively. In 1992, he served as a faculty member in the Department of Computer Science & Engineering, North Eastern Regional Institute of Science & Technology, Itanagar, India. In 1999 he has joined Tezpur University, India, where he is currently working as Professor in the Department of Computer Science & Engineering. Before joining Tezpur University, he worked as Assistant Professor in the Department of Computer Sc. & Engineering at Jorhat Engineering College, Jorhat, India. His current research interests include - Quality of Service and Cross-layer design in Wireless Ad Hoc Networks, Cognitive Radio Networks design, Optical networks and Network Security. He has published more than 70 journal/conference papers and edited 2 books. He is a professional member of ACM, IEEE and a Fellow of IETE. He was awarded with Best paper award in ADOC 2007, IETE K S Krishnan Memorial Award in 2009 and Amiya K Pujari Best Paper Award in ICTIT 2014.