We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

6,500  Open access books available
176,000  International authors and editors
190M  Downloads

154  Countries delivered to
TOP 1%  Our authors are among the most cited scientists
12.2%  Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 2

Network Virtualization Over Elastic Optical Networks: A Survey of Allocation Algorithms

Danilo Bórquez-Paredes, Alejandra BegHELLi and Ariel Leiva

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.68340

Abstract

Network virtualization has emerged as a paradigm for cloud computing services by providing key functionalities such as abstraction of network resources kept hidden to the cloud service user, isolation of different cloud computing applications, flexibility in terms of resources granularity, and on-demand setup/teardown of service. In parallel, flex-grid (also known as elastic) optical networks have become an alternative to deal with the constant traffic growth. These advances have triggered research on network virtualization over flex-grid optical networks. Effort has been focused on the design of flexible and virtualized devices, on the definition of network architectures and on virtual network allocation algorithms. In this chapter, a survey on the virtual network allocation algorithms over flexible-grid networks is presented. Proposals are classified according to a taxonomy made of three main categories: performance metrics, operation conditions and the type of service offered to users. Based on such classification, this work also identifies open research areas as multi-objective optimization approaches, distributed architectures, meta-heuristics, reconfiguration and protection mechanisms for virtual networks over elastic optical networks.

Keywords: optical fibre networks, flexible-grid/elastic networks, network virtualization, resource allocation algorithms

1. Introduction

Cloud computing has emerged as a new network paradigm [1]. Built on the success of grid computing applications, cloud computing implements the idea of ‘computing as a utility’ in a more commercially-oriented vision. Thus, the customer pays per use of computing facilities
under the conditions stated in a service level agreement (SLA), having dynamic scaling of resources and transparent access to network services, unaware of the location and hardware/software characteristics of the required resources [2]. Apart from high bandwidth, cloud computing applications require the following functionalities from the underlying physical network [1]:

- **Abstraction**: The technology/implementation specific details of the physical network resources are hidden to the users, due to the “computing as a utility” philosophy.
- **Isolation**: Different cloud computing applications should not interfere with each other in the access to common physical resources.
- **Flexible resource granularity**: The amount of resources (storage, processing power and bandwidth) required by different cloud computing applications might vary significantly.
- **On-demand setup/tear down**: For efficiency, network resources should be set-up/torn down with a highly dynamic, rapidly reconfigurable and programmable network environment, something not possible with the current status of Internet [3].
- **Resiliency**: Grid/cloud computing applications should continue running in spite of failures affecting the optical network.

Network virtualization, which extends the well-known concepts of server and storage virtualization to networks, is envisaged as a key enabling technology for cloud computing services. As such, the benefits of running cloud applications on top of virtual networks (as opposed to on top of virtual servers alone, as usually done [4, 5]), was evidenced by several preliminary studies on network virtualization for cloud computing. In Ref. [6], resource allocation of cloud-based data centres services was proposed by abstracting the service requests as virtual network requests. In Ref. [7], a network virtualization platform that acts as a mediator between the cloud user requirements and the physical resources was proposed. In Ref. [8], a new network architecture based on network virtualization was proposed for cloud computing applications where the geographic location of servers is relevant. In Ref. [9], a network operator perspective was given about the convenience of network virtualization as an enabler for cloud computing. Nowadays, the benefits of network virtualization for cloud services are well identified in terms of cost, agility, resilience and multi-tenancy [10–12].

The underlying network over which network virtualization takes place is of fundamental importance to guarantee a good service. Arguably, the two most important requirements regarding the underlying network are the bandwidth capacity and the variety in the bandwidth granularity of connections, to allow for a high number of cloud computing applications with different bandwidth requirements. Both requirements would be naturally provided by flexible-grid optical networks [13, 14]. By overcoming the rigid spectrum allocation of current fixed-grid wavelength-division multiplexing (WDM) networks, elastic optical networks would make better use of the band C by allocating each connection the bandwidth just required. Depending on the bit rate and the modulation format, a gain in bandwidth usage between 33 and 100% could be achieved by using flexible-grid networks instead of a fixed one operating with a spectral width of 50 GHz [15]. Finally, flex grid would allow a wide bandwidth granularity of connections: bit rates from 10 Gbps to 1 Tbps are envisaged [13].
Given the impact that network virtualization is expected to have on the ever-increasing cloud computing area and the potential for significant bandwidth increase and bandwidth granularity offered by flexible-grid optical networks, in this survey, we review the efforts on network virtualization over optical flexible-grid networks.

The remaining chapter is as follows: Section 2 reviews the fundamental concepts of network virtualization and flexible-grid optical networks; Section 3 discusses the main challenges in the area of network virtualization over flexible-grid optical networks; Section 4 presents a taxonomy of the proposals found in the literature to allocate virtual networks over a flexible-grid underlying transport network; and Section 5 concludes the chapter highlighting the open research lines in the area.

2. Fundamental concepts

2.1. Network virtualization

Network virtualization refers to the creation of different isolated virtual networks on top of a common physical substrate. The isolation feature means that the information transmitted through a particular virtual network cannot be retrieved or affected by other existing virtual networks and the operation of the different virtual networks cannot affect the operation of the physical substrate [16].

Among the main features of network virtualization environments, we found several of the requirements imposed by cloud computing applications, namely, coexistence of different virtual networks, isolation between coexisting virtual networks, programmability, dynamicity, flexibility and heterogeneity [17].

By implementing cloud applications on virtual networks (i.e. one virtual network for each different cloud computing application), several benefits can be identified:

- Resource allocation based on maximum load could be avoided, leading to a more cost-effective operation, as the virtual network associated to the cloud application would request just the resources needed for proper operation. Some virtual network environments have even considered the possibility of reconfiguring the virtual network during operation (e.g. exploiting the feature of on-line virtual server migration) to adapt to time-variant requirements from the applications [2, 18].

- Isolation between different cloud applications for access to common physical resources

- Resiliency against node/server failures, due to the server-migration feature of virtualization environments

- Implementation of proprietary non-standard protocols for specific cloud applications requirements

Network virtualization has been envisaged as a very useful tool in network research and industry. In research, the test of new routing algorithms, network protocols or network controllers can be done by establishing a virtual network, without interrupting the normal
operation of a physical network or deploying a physical network for tests. Thus, the production network may become the testbed [19]. An early example of this type of use was PlanetLab [20–22], established in 2002 for distributed systems and network research. Other efforts have been GENI in USA [23], FEDERICA and OneLab2 in Europe [24, 25], Akari in Japan [26] and FIBRE in a joint effort between Brazil and Europe [27]. For a review of several precursor experimental initiatives, see Ref. [17]. In an industry, network virtualization can offer separate networks for different units in a company, differentiation of services based on bandwidth usage (e.g. voice and video) or a rapid and flexible creation of sub-networks for different projects [28, 29]. For example, in a data centre each client can have its own topology and control its traffic flows. Finally, different service providers can share the same network infrastructure being unaware of the others.

As a way of illustration, Figure 1 shows a schematic of a network virtualization system. The lower part shows the physical substrate, made of five nodes and six bidirectional links. The available capacity of physical links, measured in capacity units (c.u.), is shown next to each link. The upper part shows two of the virtual networks (three-node rings with three bidirectional links each) that have been established on the physical network. The capacity unit required by the virtual links are shown near to each link. Dotted lines represent the mapping between virtual and physical links. Both virtual networks can have virtual links established over the same physical link and a virtual link can require more than one physical link to be established. For the sake of clarity, the mapping of the nodes is not shown but it can be deduced by identifying the physical nodes at the extreme of the physical links associated to the virtual links. The decision about whether establishing a new virtual network is possible or not and what virtual link/node is established in what physical link/node is made by a virtual network allocation algorithm.

![Figure 1. Two virtual networks established on the same physical substrate.](image-url)
2.2. Mathematical modelling for network virtualization

The physical network is modelled by a directed graph \( \mathcal{P} = (N_p, L_p, R^p, C_p) \), where \( N_p \) and \( L_p \) are the sets of physical nodes and links, respectively; \( R^p \) is the set of resources of type \( t \) in the physical nodes (for example, storage and processing resources; \( t \in \mathbb{N} \)) and \( C_p \) the set of resources at the physical links (optical bandwidth).

Analogously, the \( i \)-th virtual network can be modelled by a directed graph \( \mathcal{V}_i = (N^v_i, L^v_i, R^{v_i}, C^{v_i}) \), where \( N^v_i \) is the set of virtual nodes and \( L^v_i \) the set of virtual links; \( R^{v_i} \) is the set of resources of type \( t \) required by each virtual node of the virtual network \( \mathcal{V}_i \) (e.g. storage and processing resources) and \( C^{v_i} \) is the set of resources required by the virtual links (optical bandwidth).

The information required to execute the resource allocation algorithm is as follows:

1. \( \mathcal{A} = \{ \mathcal{V}_i \} \): Set containing the identification of all virtual networks \( \mathcal{V}_i \) already established over the physical network \( \mathcal{P} \).
2. \( \mathcal{S}_{n}^v_k \): Set of all virtual nodes already established on the physical node \( k \in N_p \).
3. \( \mathcal{S}_{l}^v_m \): Set of virtual links already established on physical link \( m \in L_p \).

Every time the resource allocation algorithm must process a new virtual network request, at least the following two constraints must be met to be able to accept such request:

\[
\begin{align*}
    r^n_k t & \geq \sum_{v \in \mathcal{S}_{n}^v_k} r^v_{nt} ; \quad \forall t \tag{1} \\
    c_m & \geq \sum_{l \in \mathcal{S}_{l}^v_m} c_{lv} \tag{2}
\end{align*}
\]

where \( r^n_k \) is the total number of resources of type \( t \) in physical node \( k \), \( r^v_{nt} \) is the number of resources of type \( t \) allocated to virtual node \( v \), \( c_m \) is the total number of resources in physical link \( m \) and \( c_{lv} \) is the number of resources allocated to virtual link \( l \).

Eqs. (1) and (2) forbid that the number of resources allocated to the virtual nodes/links established in a particular physical node/link exceed the capacity of that node/link.

Additionally, depending on the type of physical network, extra constraints might appear on the allocation of resources to the virtual links. In the case of an optical network, fixed and flexible-grid networks impose different constraints. We review these two types of optical networks and their associated constraints in the following.

2.3. Fixed-grid optical network

In a circuit-switched optical network, each circuit is carried by an optical channel/carrier, based on the wavelength division multiplexing (WDM) technique. Currently, such optical channels operate in the range 1530–1565 nm, known as band C.

In a fixed-grid optical network, the optical carriers are determined by their central frequency and use a fixed amount of spectrum. According to the specification ITU-T G.694.1 [30], the selectable...
spectrum widths are 12.5 GHz, 25 GHz, 50 GHz and 100 GHz. Once a spectrum width is selected, all optical channels in a link are established with such spectral width. Depending on the selected spectral width, the central frequency used by the \( n \)-th optical channel is given by the following equation:

\[
193.1 + n \times W \text{THz}
\]  

(3)

where \( W \in \{0.0125; 0.025; 0.05; 0.1\} \) denotes the spectral width selected and \( n \) is an integer number whose range depends on the spectral width as follows: \( n \in [-123, 227] \) for \( W = 0.0125 \); \( n \in [-61, 113] \) for \( W = 0.025 \); \( n \in [-30, 56] \) for \( W = 0.05 \); \( n \in [-15, 28] \) for \( W = 0.1 \).

Figure 2 shows an example of the spectral usage of a fixed-grid link where six optical channels have been established: two optical channels at 10 Gbps using the on-off keying (OOK) modulation format, three channels at 40 Gbps modulated with dual polarization-quadrature phase shift keying (DP-QPSK) and one channel at 100 Gbps, also modulated with DP-QPSK. The spectral width of each channel is equal to 50 GHz and the central frequencies are determined by Eq. (3). It is common practice to identify the channels by their equivalent wavelength as well. Thus, in Figure 2, the corresponding wavelength of each channel has been written between brackets under the central frequency.
In fixed-grid optical networks (in the absence of wavelength converters), the wavelength continuity constraint must be met. That is, the optical channel used by the virtual link must use the same central frequency and spectral width in all the physical links used. In networks operating with multiple transmission rates (as shown in Figure 3), additional constraints to deal with the signal degradation of higher bit rates channels mainly due to cross-phase modulation [31–33] may be required: for example, some channels should be left unused as guard bands or an optical reach (the maximum distance an optical signal can travel without exceeding a threshold on the bit error rate) be established.

The main drawback of fixed-grid optical networks is the inefficient spectrum usage [34], as observed in Figure 2, where channels are allocated more spectrum than effectively required: both a 10 Gbps OOK-modulated channel and a 40 Gbps channel modulated with DP-QPSK require a bandwidth equal to 25 GHz [34, 35], whereas a 100 Gbps channel modulated with DP-QPSK requires just 37.5 GHz [34]. To increase the spectrum usage, the flexible allocation of it has been proposed [14, 34]. This type of networks is known as flexible-grid or elastic optical networks.

### 2.4. Flexible-grid optical networks

In a flexible-grid optical network, the spectral width of a channel can be varied depending on the data transmission requirements [36]. Thus, the spectrum is divided in small units, typically of 12.5 GHz, known as frequency slot units (FSU) [34]. By using a different number of contiguous FSUs, different spectral widths can be achieved [37, 38] depending on the transmission requirements of the signal, such as the modulation format and the bit rate.

As a way of illustration, Figure 3 shows the same six channels of Figure 2, now operating in a flexible-grid system. The numbers of 12.5 GHz FSUs required are 2, 2 and 3 for the 10, 40 and 100 Gbps channels, respectively. Thus, the flexible-grid allocation uses just 54.2% of the spectrum originally required (162.5 GHz instead of 300 GHz).

Figure 3. Frequency allocation to different transmission rate optical channels in a flexible-grid link.
Single-carrier and multi-carrier (super-channel) can be used to create an optical connection. In the latter, the overall bit rate is achieved through lower-rate sub-carriers. Examples of these systems are Co-WDM, Nyquist-WDM and time frequency packing [34, 39, 40]. In general, multi-carrier systems require a lower number of FSUs and exhibit a longer optical reach than single-carrier systems with the same total bit rate and modulation format [41, 42].

Regarding the modulation formats, there are bi-level and multi-level types. In a bi-level modulation format, as OOK and binary phase shift keying (BPSK) [42], the symbol rate equals the bit rate. In a multi-level modulation format, as QPSK and x-quadrature amplitude modulation (x-QAM) [41, 42], the symbol rate is lower than the bit rate of the bi-level type, leading to a lower requirement of FSUs. However, the optical reach of multi-level modulation formats is lower than that of bi-level [34, 36], highlighting a trade-off between number of FSUs and optical reach [34, 43, 44].

Once the number of FSUs required by a virtual link has been determined, the establishment of such link must meet at least two additional constraints: FSU continuity and FSU contiguity constraints. The FSU continuity constraint is analogous to the wavelength continuity constraint (exactly the same FSUs must be used in every physical link selected to establish a virtual link). The FSU contiguity constraint imposes that, if more than one FSU is required to establish a virtual link, then these FSU must be contiguous in the spectrum [45].

The sequence of physical links used to establish a virtual link meeting the FSU continuity and contiguity constraints is known as a spectrum path.

3. Research challenges in virtual network allocation over flexible-grid optical networks

In the following, the main challenges in the research area of network virtualization over flexible-grid optical networks are discussed.

3.1. Performance metrics

A performance metric allows defining the quality of an algorithm to carry out its task. Thus, usually the (single) objective of an algorithm is the maximization or minimization of a performance metric. However, for a complex algorithm such as a virtual network allocation algorithm, there are several performance metrics that could be optimized.

Most published results have focused on minimizing the virtual network request rejection rate [46–56, 64]. The main advantage of using the performance metric is that it allows evaluating the ability of the algorithm to accommodate new virtual networks on the physical substrate. However, given that the blocking depends on many parameters (the physical and virtual network topologies, the capacity availability in physical nodes and links, the capacity requirements of virtual nodes and links [50–52]), to identify the best algorithm is necessary knowing exactly the network configuration where the algorithm will operate (something difficult to achieve in dynamic scenarios) or running extensive simulation experiments with different network configurations (a time-consuming task).
Instead of registering the blocking ratio, a computationally simpler metric consists on registering the number of virtual network establishment requests received when the first blocking (rejection) occurs [48, 49]. A good algorithm would aim at registering such event at the latest possible instant. If used in conjunction with the blocking ratio, the first blocking metric can give information about the instant when the network starts saturating (when the first blocking occurs) and the dynamic of the system once such saturation state is reached.

Maximizing the traffic carried by the physical network due to the established virtual networks has also been the objective of some algorithms [54, 57–59]. As with the blocking ratio, the value of this performance metric depends on the topologies of the physical and virtual networks as well as the capacity of physical nodes/links and capacity requirements of virtual nodes/links, which makes difficult drawing general conclusions about the quality of different algorithms. Additionally, the lack of information about the number of virtual networks rejected does not allow measuring the quality of the service offered to the users. Thus, it should be used in conjunction with the blocking ratio.

Guaranteeing a given level of availability (e.g. 0.99999) to a virtual network has not been addressed by the proposed virtual network allocation algorithms to date, although availability (the fraction of time that a service is in operative state) is one of the most important quality of service metrics in a service level agreement (SLA). However, some efforts have been carried out in guaranteeing operation under specific failure conditions [49, 53, 59, 61, 62].

All previous performance metrics somehow aim to evaluate the capacity of the algorithm to offer a good quality of service. However, the main challenge in evaluating the performance of complex algorithms is selecting a performance metric that can capture the quality of the service offered to the user as well as the cost in achieving such quality.

To offer physical resources to a virtual network, the service provider incurs expenditure and operational costs due to the acquisition and maintenance of transponders, regenerators, optical cables, optical amplifiers and ROADMs (reconfigurable optical add drop multiplexer) [60]. Thus, algorithms aiming at minimizing the cost have also been studied. This metric has been mostly used in static scenarios [46, 56, 61, 62], and it is useful for the network planning stage. In dynamic scenarios, it can be used to determine the cost per virtual network, the total cost of providing the network virtualization service during a period of time or the cost incurred to achieve a given performance in terms of blocking ratio or traffic carried.

To date, quality-of-service-related metrics and cost have been studied separately. The algorithm is designed to minimize/maximize one of them whilst the other one is just measured. Thus, a multi-objective optimization approach that evaluates quality (as blocking or availability) and the cost incurred to achieve the required quality would deliver more realistic information about the best algorithm alternative from a network operator perspective.

### 3.2. Network virtualization dynamics characterization

To date there are no commercial network virtualization systems over flexible-grid optical networks. In Ref. [63], an experimental system is reported, but traffic is artificially generated. Therefore there are no empirical statistics that help to model the structure (virtual topologies and their capacity requirements) and dynamic of such system. In terms of structure, it would
be useful knowing how to model the virtual topologies and their capacity requirements. Such knowledge would facilitate the evaluation of allocation algorithms in terms of simulation, the only technique used so far to evaluate performance of dynamic systems.

In terms of structure, different works make different assumptions regarding the topologies of the virtual networks and their capacity requirements. Table 1 summarises the main models used to characterize the virtual topologies. In it, the name of each physical and virtual topology is given along with its number of nodes ($|N_p|$, $|N_v|$) and links ($|L_p|$, $|L_v|$). When a number lower than one is provided for $|L_v|$, it means that the probability interconnection between a node pair is given. The column ‘Node/Link requirement’ corresponds to the percentage of usage of the physical node and link by any virtual node and link, respectively. The symbol ‘‐’ implies that such information is not found in the chapter.

As most works (15 of 17) use a medium-sized physical network (NSFNet or DTNet) for evaluation, future works should consider at least one of these topologies as the physical substrate to facilitate comparison among different proposals. No pattern can be observed in terms of the virtual topologies, with most works using mesh topologies with different degrees of connectivity. Regarding resource requirements, all proposals require no more than 10% of the physical node/link resources. The rest uses percentages of a few units.

Regarding dynamism, the most used distribution to model the virtual network request inter-arrival time is the exponential [46, 50–56, 64]. The holding time is usually modelled by an exponential distribution [52, 55], a deterministic value [64] or infinite (to model incremental traffic) [48, 49].

### 3.3. Physical impairments

It is expected that flexible-grid optical networks can accommodate channels (used to implement virtual links) at rates from 10 Gbps to 1 Tbps. Such channels, in the same way as fixed-grid channels, will be affected by several physical impairments that degrade the quality of the signal transmission. Additionally to typical physical impairments, as attenuation, chromatic dispersion, four-wave mixing (FWM) and amplified spontaneous emission (ASE) noise [65], in elastic optical networks the non-linear effect of cross phase modulation (XPM) takes relevance because of the existence of channels with different modulation formats in the same link. Due to the XPM effect, channels using intensity-based modulation formats (e.g. OOK typically used in 10 Gbps channels) interfere negatively in the quality of the signal of phase-modulated channels (e.g. BPSK and QPSK, used for higher bit rate channels) [66].

Most previous works have not considered this situation, with some of them assuming an ideal physical substrate [50, 54] whereas others have resorted to simplified models. For instance, in Refs. [48, 49, 51, 53, 57, 61, 62], the degradation is summarized in the figure of the maximum optical reach of signals, in Refs. [46, 56, 58, 59], the use of guard bands to all channels is used to simulate an ideal substrate, whereas in Refs. [47, 52, 55, 64], guard bands (to all channels or selectively added to channels most affected by the XPM degradation) are added to the limitation of the optical reach.
### 3.4. Resource allocation to virtual networks

The selection of the physical nodes and links to be allocated to a virtual network is a \( \text{NP}- \)Hard problem [67]. Thus, most proposals solving this problem over flexible-grid optical networks have resorted to heuristics [46–59, 61, 62, 64] and a few of them have proposed integer linear models [51, 58, 59, 61, 64], but mostly in the context of a static scenario where the random nature of the virtual network requests is not a problem.

| Physical topology \((|N_p|, |L_p|)\) | Virtual topology \((|N_v|, |L_v|)\) | Node/link requirement | Work |
|----------------------------------------|---------------------------------|------------------------|------|
| NSFNet (14, 21)                        | Mesh (2–6, 0.5)                | 0.35–3.5%              | [46] |
|                                        | Mesh (2–5, 0.5)                | 0.32–2.5%              | [48] |
|                                        | Mesh (5–7, 7)                  | 0.62–0.93%             | [52] |
|                                        | Ring (5–7, 5–7)                | 0.62–0.93%             | [53] |
|                                        | Mesh (2–5, 0.5)                | 0.29–0.86%             | [54] |
|                                        | Ring (3–7, 3–7)                | 1%0.63–15.94%          | [55] |
|                                        | Mesh (3–5, 3–10)               | 1%0.625–3.125%         | [56] |
|                                        | Mesh (4–7, 4–14)               | 1%0.25–5%              | [58] |
|                                        | Ring (2–4, 2–4)                | -/2–16%                | [59] |
|                                        | Mesh (5, 7)                    | 2–4%/2–6%              | [62] |
| 6-node (6, 8–10)                       | Mesh (2–3, 0.5)                | 2–6%/1%                | [51] |
|                                        | Ring (3, 3)                    | 10%/1.25%              | [58] |
|                                        | Mesh (3–4, 3–3.8)              | 1%0.63/1%              | [59] |
|                                        | Mesh (3–4, 0.5)                | 1%0.625–3.125%         | [56] |
|                                        | Mesh (2–5, 1–15)               | 0.85–3.12%             | [61] |
|                                        | Mesh (3–4, 0.5)                | 0.5–5%/0.52–5%         | [64] |
| DTNet (14, 23)                         | Mesh (3–4, 0.5)                | 0.5–5%/0.52–5%         | [51] |
|                                        | Mesh (3–10, 0.5)               | 1%/1%                 | [58] |
|                                        | Ring (2–4, 2–4)                | /2%                  | [59] |
|                                        | Mesh (2–5, 1–15)               | 0.85–3.12%             | [61] |
|                                        | Mesh (3–4, 0.5)                | 0.5–5%/0.52–5%         | [64] |
| US network (24, -)                     | Mesh (2–5, -)                  | 0.06–2%/0.31–2.5%      | [50] |
| Random (50, 141)                       | Mesh (2–10, 0.5)               | 0.5–10%/0.5–10%        | [51, 52] |
| CORONET (75, 99)                       | Mesh (3–7, 0.5)                | 2–1%/2–2%              | [54] |
| ARPANET (20, 32)                       | Mesh (3–7, 0.5)                | 0.2–1%/0.5–5%          | [58] |

Table 1. Characteristics of virtual network requests used in the literature.
Much work is still needed in identifying the features of good performing heuristics to allocate virtual networks as well as evaluating the performance of meta-heuristics.

3.5. Spectrum fragmentation

Under dynamic operation, as a result of the resource release from virtual network that depart from the network, voids in the spectrum are generated. A void is a set of contiguous available FSUs between portions of allocated FSUs (or between a portion of allocated FSUs and the beginning/end of the band), as shown in Figure 4.

Due to the FSU contiguity constraint, the existence of these voids is problematic, as they fragment the spectrum. As a result, a virtual link could not be implemented due to the lack of enough contiguous FSUs, leading to a higher blocking ratio. For example, in the situation depicted in Figure 4, although three FSUs are available, a virtual link requiring three FSUs could not be established because of the contiguity constraint.

![Figure 4. Spectrum fragmentation.](image-url)
To decrease the spectrum fragmentation, the re-allocation of FSUs to the different channels in a link has been proposed in the area of flexible-grid networks by Refs. [68–72]. In Ref. [73], the impact of avoiding fragmentation on the blocking ratio can be seen.

In Ref. [54], a technique of spectrum defragmentation in the area of virtual networks over flexible-grid optical networks is reported, showing that the blocking ratio decreases with respect to an algorithm without defragmentation. However, defragmentation is costly as computation time and additional resources must be used to apply it. This highlights a trade-off between the blocking ratio decrease and the frequency of defragmentation. Further research on the interplay of allocation algorithms and defragmentation techniques is required.

4. Taxonomy

Figure 5 shows a comprehensive classification of the resource allocation algorithms in the area of network virtualization over flexible-grid optical networks. The taxonomy includes current proposals, but it is generic enough as to include algorithms not studied yet.

In the taxonomy, each possible algorithm is defined by three main dimensions: its performance metric, its operation conditions and the type of service offered to the user. In the following each of these dimensions are described as well as the different choices available in each one of them.

4.1. Performance metric

The most commonly used performance metric in the literature is the blocking ratio [46–56, 64]. Although the use of the same metric would facilitate comparison, due to the different assumptions made on the physical and virtual topologies, a direct comparison is not always possible.

The variant of blocking, first blocking, has been used in Refs. [48, 49]. Remaining metrics used in reported works are the traffic carried by the physical networks [54, 57–59] and cost-related metrics [46, 56, 61, 62].

Although published works do not explicitly mention the performance metric of availability, few works make assumptions on the operation conditions of the network that allow guaranteeing 100% availability. In Refs. [53, 59, 61, 62], only single link failures are assumed. Thus, the allocation of two link-disjoint spectrum paths to implement each virtual link is enough to ensure the operation of every virtual network. In Ref. [49], single link/node failures are assumed and, then by allocating two node/link-disjoint spectrum paths to each virtual link, a 100% availability is provided. Note that if the system violates the assumptions on the type of failure that can occur (e.g. a double link failure occurs in a system designed to tolerate single link failure), 100% availability cannot be guaranteed anymore.
Guaranteeing availability under any type of failure has not been researched in the area of network virtualization over flexible-grid networks neither the combination of quality and cost performance metrics.

4.2. Operation

4.2.1. Information management

To date, all proposals implicitly assume centralized information management [46–59, 61, 62, 64]. That is, a central entity has global knowledge of the network status and the resource allocation algorithm is executed every time a new virtual network request is generated. In fact, the first proposals for architecture with a virtual network controller are based on a centralized scheme,
as the one proposed in Ref. [74]. Centralized systems are suitable when the time between successive requests is long enough for the central controller to execute the resource allocation algorithm.

In Ref. [75], different distributed virtual network allocation approaches are discussed in the context of packet networks. Results show that a distributed operation reduces the delay in mapping a virtual network and the number of messages required to be exchanged to coordinate the allocation. In Ref. [76], the impact of a distributed virtual network reconfiguration approach on the interruption time of the service is studied in the context of fixed-grid networks. Although the distributed operation has advantages in terms of resilience against failures, lower computation times and network congestion due to message exchange, it has increased complexity in terms of control plane network (more controllers), synchronization of messages and a potential decreased performance due to the obsolescence of information. These aspects are yet to be studied in network virtualization systems over flexible-grid networks.

4.2.2. Resource allocation strategy

The virtual network allocation strategy must consider two aspects: the method used to solve the problem of embedding the virtual network on the physical network and the model used to characterize the constraints of the physical substrate.

There are three general methods to solve the problem of the virtual network embedding:

a. Exact methods: These are the techniques that find the global optimal solution to a problem. However, they are computationally complex and thus, they are usually applied only to small instances of the problems with slow dynamics. In real dynamic systems, where a solution must be found in short-time scales, this type of method is not feasible. However, in simulation environments, an integer linear programming (ILP) model can be solved for each virtual network request to be used as a benchmark, as done in Ref. [51]. In the area of virtual network allocation over flexible-grid networks, most ILP models have been used to solve the problem in a static scenario (virtual networks permanently established, not allocated on-demand). Works in Refs. [51, 58, 59, 61, 64] apply ILP to allocate a set of predefined virtual networks on a small physical network (six nodes) with the objective of minimizing cost.

b. Meta-heuristics: These are the generic algorithms capable to adapt to different problems by adjusting their parameters and configurations. Usually, they find very good quality solutions, but cannot guarantee the optimum solution as exact methods do. Work in Refs. [77, 78] proposed the use of genetic algorithms and ant colony to solve the problem of virtual network embedding in conventional networks, respectively. No works have been reported on flexible-grid networks as the physical substrate, neither in static nor in dynamic scenarios.

c. Heuristics: These are ad-hoc algorithms that do not guarantee a global optimum solution, designed for a specific problem. However, they are computationally simpler than the previous techniques. Most works in the area of network virtualization over flexible-grid networks resort to heuristics [46–59, 61, 62, 64], mainly focused on dynamic scenarios.
Normally, heuristics designed to solve complex problems, divide the original problem into sub-problems easier to solve separately. This approach is applied in this area as well. The original problem of mapping a virtual network is divided into node mapping (allocation of a physical node to a virtual node) and link mapping (allocation of a spectrum path to a virtual link). Most proposals map nodes first to then establish the virtual links connecting them [46–48, 50, 51, 53, 54, 56, 58, 62, 64].

To map the nodes and links, the heuristic must define the order in which the virtual and physical nodes/links are processed. To do so, a ranking is elaborated for each set of physical/virtual nodes/links and the first element in the ranking of virtual nodes/links is attempted to be mapped in the first element of the ranking of the physical nodes/links. The most common criterion to build the physical node ranking is the amount of available resources [48, 49, 58]. A function of the computing capacity and the nodal degree [50], a function of the number of sub-carriers of each transponder in the physical node and the slice capability of the physical node [46] and the node index [64] have also been used. Criteria to rank the virtual nodes are the amount of resources required [48, 58], the nodal degree [50] or the node index. The case where the virtual nodes must be established in specific physical nodes (defined in the virtual network establishment request), as in Ref. [47], is a particular case of a node/link mapping, as all virtual nodes are established in the specified physical nodes (if enough resources are available) before establishing the virtual links.

Physical links can be ranked in terms of their distance [48, 50, 58], cost [64] or number of available FSUs. Finally, virtual links are ranked in terms of their FSU requirements [47, 53, 58, 64].

Given that the solution found by solving the node/link mapping sub-problems sequentially is expected to be of lower quality than solving the original problem, an attempt to solve both problems jointly was proposed in Refs. [53, 55, 57, 59]. In these works, a sub-set of all possible mapping patterns for the nodes of a virtual network are evaluated and the one using the lowest slot layer (slot layer of a mapping pattern is the highest FSU used) [57], lowest cost [53] or best Hamming-inspired distance [55] is selected.

Finally, the approach of alternating the allocation of virtual nodes and links (mixed) has also been studied in Refs. [48, 49, 52, 61, 62, 64]. For example, in Ref. [61], the virtual nodes at the ends of each virtual link are mapped to then map the virtual link, showing results close to the ILP approach in a static scenario.

Apart from the FSU continuity and contiguity constraints, the solution methods can use one of several models to characterize additional constraints of the physical substrate. To date, the following models have been used:

a. Ideal, where no signal degradation is assumed [50, 54].

b. Optical-reach-based, this is the simplest model where the maximum distance covered by a spectrum path is determined solely by the modulation format and the bit rate, as in Refs. [48, 49, 51, 53, 57, 61, 62].

c. Guard-band-based, where a given number of FSUs might be left unused between channels of different bit rate, as in Refs. [46, 56, 58, 59].
d. Optical reach and guard band, where the optical reach is determined by the modulation format and the bit rate. Since the optical reach can decrease due to effect of neighbouring signals, by adding (selectively or not) guard bands between channels [47, 52, 55, 64] such detrimental effect can be mitigated.

4.2.3. Traffic management

In the context of packet networks, the split of traffic of a virtual link into several paths in the physical substrate has been proposed as a way of increasing the probability of accepting a virtual network establishment request [79]. In a flexible-grid optical network where a virtual link requiring $M$ contiguous FSUs must be established but no path has more than $x < M$ contiguous FSUs, such situation could be solved by establishing the virtual link along several spectrum paths in such a way that the total number of FSUs used along all the paths equal $M$. Such mechanism could be enabled by recently introduced sliceable or multi-flow transponders [80, 81]. This approach has not been explored in the area of network virtualization over flexible-grid networks.

4.3. Type of service

4.3.1. Service nature

The service provider can offer a static or dynamic service. In the former case, the virtual network demands are known a priori and they are established permanently, whether they are used to transmit information or not [58, 59, 61, 62, 64]. In the latter case, virtual networks are established and released on demand.

In a dynamic service, spectrum experiences fragmentation. As a result, even when there is an enough number of FSUs to accommodate a new virtual network, these FSUs might not meet the contiguity constraint, leading to the rejection of requests. To decrease spectrum fragmentation, some dynamic systems reconfigure the established connections. Several works have evaluated the impact of reconfiguration on point-to-point connections on flexible-grid optical networks [68–72]. As expected, reconfiguration decreases blocking [54] at the expense of higher complexity of the control plane.

There are two types of reconfiguration techniques: proactive or reactive [82]. The former re-allocate resources before a blocking condition occurs, either in a synchronous or asynchronous way. In Refs. [69–71], pro-active reconfiguration algorithms are presented for point-to-point connections over flexible-grid optical networks. Reconfiguration may take place every time a given number of virtual networks request has been received. No proactive systems have been reported in network virtualization over flexible-grid networks. Reactive reconfiguration techniques re-allocate resources only when a new request cannot be accepted. In Ref. [54], a reactive reconfiguration method to re-allocate virtual networks over fixed-grid networks is presented, getting lower rejection rates than not reconfiguring at low-medium loads.

Reconfiguration can be applied at two different levels: re-allocation of complete virtual networks or re-allocation of a sub-set of virtual links/nodes, as in Ref. [54] in flexible-grid
networks or [83] in fixed-grid networks. None of these cases has been studied in network virtualization systems over flexible-grid optical networks.

4.3.2. Fault tolerance

A network virtualization service can offer different levels of fault tolerance: zero, specific or guaranteed. Most works reported to date have studied systems without fault tolerance at all [46–48, 50–52, 54–58, 64]. In that case, the occurrence of any type of failure interrupts the operation of the virtual networks operating over the physical component affected by the failure. A specific survivability system is capable of continuing operation in spite of the occurrence of specific types of failures. Normally, these systems are designed to survive the most common failure events (e.g. a cable cut) and remain unprepared for unlikely events (as a node failure). In the area of network virtualization over flexible-grid networks, the algorithm proposed in Refs. [53, 59, 61, 62] can survive only to single link failures, whereas Ref. [49] can survive single link or node failure. Finally, a guaranteed survivability system ensures that limits on downtime are not exceed, no matter what the type of failure, as done in Refs. [84, 85] in a context different from network virtualization. If such condition is violated, the service provider is enforced to pay an economic compensation to the user. Such approach has not been explored in the area of network virtualization over flexible-grid networks.

Fault tolerance mechanisms can also be classified as proactive (protected systems) or reactive (restored systems). Protected systems allocate backup resources when the primary resources for the virtual network are allocated [49, 53, 59, 61, 62]. Therefore, upon failure occurrence, the time to recover from failure is shorter than reactive systems. Protected systems can allocate a complete backup virtual network (total protection) [49] or backup to some components (partial protection, e.g. only virtual links have backup resources) [53, 59, 61, 62]. Protected systems can also be classified as dedicated or shared. In the former, backup resources are dedicated to the corresponding primary resource. In the latter, a backup resource is shared among several primary resources. No research has been reported on the area of shared protection for virtual networks over flexible-grid networks. Restored systems allocate resources to the virtual networks affected by a failure only once the failure has occurred; as a result, the recovery time is longer, but a lower amount of backup resources are required. Restoration can be carried out for complete virtual networks or only for the part of them affected by the failure. Restoration on virtual networks over flexible-grid networks has not been researched yet.

4.3.3. Revisitation

Revisitation allows the establishment of two virtual nodes from the same virtual network in the same physical node [16]. Revisitation has been proposed in the context of overlay networks [86] as a way of emulating larger networks on small testbeds. In virtual network systems over flexible-grid networks, revisitation has been used in Ref. [64] and the impact of it on blocking was studied in Ref. [52] showing a decrease of blocking ratio of two orders of magnitude with respect to the same algorithm without revisitation.
Revisitation has been little researched in the literature, probably because a real application for it has not been found yet. For example, for research on new Internet protocols, delay and bandwidth utilization are two key metrics that could not be measured if two virtual nodes are hosted in the same physical node. For cloud replication services would not be useful either, as the replicas must be allocated to geographically different sites. However, it is mentioned as one of the four key architectural principles of network virtualization in Ref. [16], where it would be useful to help the service providers to manage highly complex tasks and facilitate virtual networks management.

In **Table 2**, a summary of the virtual network resource allocation proposed to date is presented. For each algorithm, all the dimensions presented in the taxonomy of **Figure 5** are specified.

<table>
<thead>
<tr>
<th>Performance metric</th>
<th>Operation</th>
<th>Resource allocation strategy</th>
<th>Traffic management</th>
<th>Type of service</th>
<th>Service nature</th>
<th>Failure tolerance</th>
<th>Revisitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[46] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Guard band</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[48] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Mixed Optical reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[49] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Optical reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Specific, total protection</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[50] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Ideal</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[51] Blocking ratio</td>
<td>Centralized</td>
<td>Exact Nodes-links Optical reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[52] Blocking ratio</td>
<td>Centralized</td>
<td>Mixed Optical reach + Selective guard band</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Specific, partial protection</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>[53] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Joint Optical reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration in virtual links</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td>[54] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Ideal</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
<td></td>
</tr>
</tbody>
</table>
5. Conclusions

Network virtualization has emerged as an enabling technology for cloud computing services. Such services would push even further the limits on bandwidth utilization, where flexible-grid optical networks will be the key to increase the network capacity of actually deployed optical networks. As a result, a new area of research focused on network virtualization over flexible-grid networks has emerged.

On such area, the research efforts focus on three main lines: design of flexible and virtualized devices, definition of network architectures and virtual network allocation algorithms.

In this chapter, a survey on the virtual network allocation algorithms over flexible-grid networks has been presented along with a classification of all possible proposals of algorithms by

<table>
<thead>
<tr>
<th>Performance metric</th>
<th>Information management</th>
<th>Resource allocation strategy</th>
<th>Traffic management</th>
<th>Service nature</th>
<th>Failure tolerance</th>
<th>Revisitaiton</th>
</tr>
</thead>
<tbody>
<tr>
<td>[56] Blocking ratio</td>
<td>Centralized</td>
<td>Nodes-links Guard band</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
</tr>
<tr>
<td>[57] Traffic carried</td>
<td>Centralized</td>
<td>Joint Optical Reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Without</td>
</tr>
<tr>
<td>[58] Traffic carried</td>
<td>Centralized</td>
<td>Exact Nodes-links Guard band</td>
<td>Without split</td>
<td>static</td>
<td>Zero</td>
<td>Without</td>
</tr>
<tr>
<td>[59] Traffic carried</td>
<td>Centralized</td>
<td>Exact Joint Guard band</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Specific</td>
<td>Without</td>
</tr>
<tr>
<td>[61] Cost/revenue</td>
<td>Centralized</td>
<td>Exact Mixed Optical reach</td>
<td>Without split</td>
<td>Static</td>
<td>Specific, partial protection</td>
<td>Without</td>
</tr>
<tr>
<td>[62] Cost/revenue</td>
<td>Centralized</td>
<td>Nodes-links Mixed Optical reach</td>
<td>Without split</td>
<td>Dynamic without reconfiguration</td>
<td>Specific</td>
<td>Without</td>
</tr>
<tr>
<td>[64] Blocking ratio</td>
<td>Centralized</td>
<td>Exact Nodes-links Mixed Optical reach Guard band</td>
<td>Without split</td>
<td>Static Dynamic without reconfiguration</td>
<td>Zero</td>
<td>Total</td>
</tr>
</tbody>
</table>

Table 2. Summary of the characteristics of the algorithms reviewed.
means of taxonomy. Such classification allowed the identification of several aspects that must be further investigated in the area:

- Multi-objective optimization approaches that allow to select resource allocation algorithms with a good compromise between quality and cost.
- The design and evaluation of distributed virtual network allocation algorithms.
- The application of meta-heuristics (as genetic algorithms, ant colony, etc.) to solve the virtual network allocation problem over flexible-grid networks.
- The study of the impact of traffic split on the performance of virtual network allocation algorithms.
- The effect and complexity of reconfiguration on the performance of network virtualization systems.
- The design and evaluation of shared protection mechanisms.
- The design and evaluation of shared protection and restored fault tolerance mechanisms.

Acknowledgements

Financial support from CONICYT through the Fondecyt Project 1130388 and 11140386 and the PhD Scholarship, CONICYT-PCHA/Doctorado Nacional/2015-21150525 is gratefully acknowledged.

Author details

Danilo Bórquez-Paredes*, Alejandra Beghelli† and Ariel Leiva‡
*Address all correspondence to: danilo.borquez@uai.cl
†Science and Engineering Faculty, Universidad Adolfo Ibañez, Chile
‡Electrical Engineering School, Pontificia Universidad Católica de Valparaíso, Chile

References


[52] Barra E, Salinas R, Mora F, Tarifeno M, Beghelli A, Sambo N, Castoldi P. Virtual network provisioning over multi-line rate networks with fixed or flexible grid. In: 16th International Conference on Transparent Optical Networks; 6-10 July 2014; Graz, Austria. IEEE; 2014. pp. 1–4. DOI: 10.1109/ICTON.2014.6876709


