FlexWAN: Software Hardware Co-design for Cost-Effective and Resilient Optical Backbones

FlexWAN: Software Hardware Co-design for Cost-Effective and Resilient Optical Backbones

Congcong Miao, Zhizhen Zhong [*], Ying Zhang [+], Kunling He, Fangchao Li, Minggang Chen, Yiren Zhao, Xiang Li, Zekun He, Xianneng Zou, Jilong Wang [◇○]

Tencent [*] Massachusetts Institute of Technology [+] Meta [◇] Tsinghua University [○] Quancheng Laboratory

ABSTRACT

The rising demand for WAN capacity driven by the rapid growth of inter-data center traffic poses new challenges for costly optical networks. Today, cloud providers rely on fixed optical backbones, where all hardware devices operate on a rigid spectrum grid, leading to the waste of expensive optical resources and subpar performance in handling failures. In this paper, we introduce FlexWAN, a novel *flexible* WAN infrastructure designed to provision cost-effective WAN capacity while ensuring resilience to optical failures. FlexWAN achieves this by incorporating spacing-variable hardware at the optical layer, enabling the generated wavelength to optimize the utilization of limited spectrum resources for the WAN capacity. The configuration of spacing-variable hardware in a multi-vendor optical backbone presents challenges related to spectrum management. To address this, FlexWAN leverages a centralized controller to achieve coordinated control of network-wide optical devices in a vendor-agnostic manner. Moreover, the flexibility at the optical layer introduces new algorithmic problems. FlexWAN formulates the problem of provisioning WAN capacity with the goal of minimizing hardware costs. We evaluate the system performance in production and share insights from years of production experience. Compared to existing optical backbones, FlexWAN can save at least 57% of transponders and reduce 36% of spectrum usage while continuing to meet up to 8× the present-day demands using existing hardware and fiber deployments. FlexWAN further incorporates failure resilience that revives 15% more bandwidth capacity in the overloaded optical backbone.

CCS CONCEPTS

• **Networks** → **Wide area networks**; *Physical topologies*; *Network design and planning algorithms*; **Network measurement**; **Network management**; • **Computer systems organization** → *Fault-tolerant network topologies.*

KEYWORDS

Optical backbone networks; Network modeling; Network planning; Optical restoration; Network optimization

ACM SIGCOMM '23, September 10, 2023, New York, NY, USA

ACM ISBN 979-8-4007-0236-5/23/09…$15.00
https://doi.org/10.1145/3603269.3604824

ACM Reference Format:
Congcong Miao, Zhizhen Zhong, Ying Zhang, Kunling He, Fangchao Li, Minggang Chen, Yiren Zhao, Xiang Li, Zekun He, Xianneng Zou, Jilong Wang. 2023. FlexWAN: Software Hardware Co-design for Cost-Effective and Resilient Optical Backbones. In *ACM SIGCOMM 2023 Conference (ACM SIGCOMM '23), September 10, 2023, New York, NY, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3603269.3604824

1 INTRODUCTION

The demand for Wide Area Networks (WANs) has surged during COVID due to increased online interactions like virtual meetings and cloud gaming. Cloud service providers, including Tencent, Meta, and Microsoft, respond by augmenting WAN capacity. This capacity expansion involves adding wavelengths, each requiring optical transponders and occupying a specific spectrum width in the fiber. However, the high cost of optical hardware, such as transponders, fiber cable, and equipment in the optical line system (OLS), makes long-haul WAN capacity a costly proposition [46].

Current cloud providers utilize fixed optical backbones for long-haul WAN capacity, employing a rigid spectrum grid such as 50 GHz or 75 GHz [5] to simplify optical layer management. Traditionally, data rates per wavelength occupy a 50 GHz spectrum width with 100 Gbps bandwidth capacity for distances up to 3000 km. However, recent advances propose the use of bandwidth-variable transponders (BVT) to carry higher data rates per wavelength [46, 47, 49]. In our scenario, we assume the wavelengths occupy a larger channel spacing of 75 GHz [49] [1], with 300 Gbps, 200 Gbps, and 100 Gbps data rates modulated at 8QAM, QPSK, and BPSK, respectively, for distances up to 1100 km, 2000 km, and 5000 km. Notably, different modulation formats of the wavelength occupy the same channel spacing of 75 GHz.

The inflexible nature of the optical backbone leads to the inefficient use of expensive optical resources. Our measurements on a large-scale production WAN, consisting of thousands of IP links and hundreds of optical paths, reveal that over 50% of optical paths are shorter than 200 km (§ 3). Signals traveling shorter paths experience less attenuation and result in higher signal-to-noise ratios (SNR), enabling modulation at higher-order formats like 256QAM to support data rates of up to 800 Gbps. This increase in data rate per wavelength translates to a reduced need for transponders to provision the same WAN capacity. However, there are limitations to aggressively increasing the order of modulation format. Extremely high-order modulation formats require precise signal generation and are susceptible to optical impairments [13, 31, 38, 41]. Additionally, the Shannon-Hartley theorem [45] establishes a maximum data

[1]The technology of optical transponders has evolved to support higher data rates per wavelength by increasing the channel spacing from 50 GHz [47] to 75 GHz [49].

rate limit for a given channel spacing [2]. Consequently, wavelengths cannot support extremely high data rates at existing channel spacing (e.g., 75 GHz), even with modulation at high-order formats.

The Shannon-Hartley theorem presents a promising opportunity for flexible adjustment of wavelength channel spacing to accommodate potentially high data rates. However, achieving this flexibility is not straightforward, as existing optical hardware devices operate on a rigid spectrum grid. To address this limitation, we propose the design and implementation of a novel flexible WAN infrastructure called FlexWAN. This infrastructure aims to provision cost-effective WAN capacity and enhance resilience to optical failures. FlexWAN addresses both *system-level* challenges, such as physically supporting flexibility at the optical layer and overcoming spectrum-related issues in a multi-vendor optical backbone, as well as *algorithmic* challenges, such as optimizing network planning strategies and optical restoration decisions. Through our proposed approach, we aim to revolutionize the design and operation of optical backbones, enabling them to efficiently adapt to evolving traffic demands.

To achieve a flexible WAN infrastructure, we introduce *spacing-variable hardware* within FlexWAN. This involves incorporating adjustable components into spacing-variable transponders (SVTs) to generate wavelengths that are optimally modulated for the required data rate and channel spacing based on the transmission distance. Additionally, the spectrum-sliced optical line system (OLS) dynamically adjusts the passband width to adapt to the channel spacing of each wavelength. This approach allows each wavelength to efficiently utilize the limited spectrum resources to carry the WAN capacity. However, the implementation of spacing-variable hardware at the optical layer presents practical challenges in a multi-vendor optical backbone. Each vendor employs a customized controller to interface with the devices they manage. Configuring thousands of IP links and hundreds of optical paths from multiple vendors using these individual controllers increases the likelihood of spectrum-related issues. To address this, FlexWAN leverages a centralized controller interfacing with heterogeneous devices in a vendor-agnostic manner to coordinate network-wide optical devices, effectively avoiding potential problems related to multi-vendor configurations (§ 4).

The flexibility at the optical layer presents new algorithmic challenges for cloud providers when choosing an optimal and cost-effective network planning strategy. In addressing this, FlexWAN formulates the problem of provisioning WAN capacity with a focus on minimizing hardware costs (§ 5). To evaluate transponder performance, we conduct production-level testbed experiments in collaboration with a vendor (§ 6). Leveraging the testbed's transponder specifications, FlexWAN's proposed planning strategy for WAN capacity provisioning achieves impressive results. Compared to existing optical backbones, FlexWAN saves 57% of transponders and reduces spectrum usage by 36%, while continuing to meet up to 8 × the present-day demands using existing optical line system (OLS) and fiber deployments (§ 7). Furthermore, we extend FlexWAN to incorporate the goal of failure resilience by introducing optical restoration. In the overloaded network at 5 × present-day demands, FlexWAN improves revived bandwidth capacity by 15% (§ 8). Additionally, through two years of production experience, we have gained valuable operational insights that have further enhanced the system's performance (§ 9).

This work **does not** raise any ethical issues.

[2]The Shannon-Hartley theorem defines the maximum data rates based on channel spacing and signal-to-noise ratio (SNR), with the formulation $C = W \log_2(1 + S/N)$, where C is the data rate of an optical channel, W is the channel spacing, and S and N are the signal and noise power, respectively.

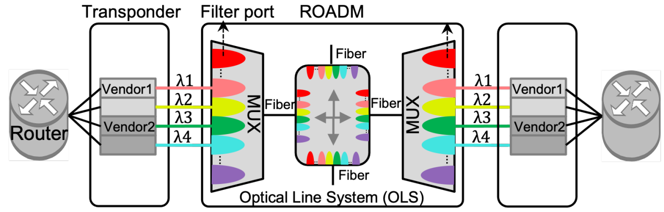


Figure 1: Overview of the fixed optical backbones.

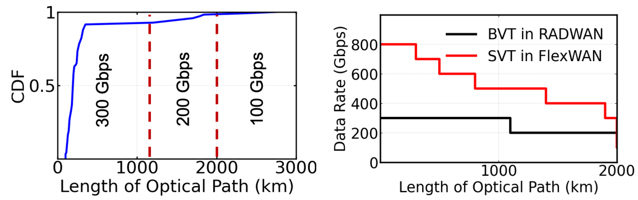
2 BACKGROUND

Optical backbone is a fundamental infrastructure to provision the long-haul WAN capacity. Figure 1 zooms-in into a specific IP link with a focus on aspects relevant to our work. Each IP link contains multiple optical transponders and an optical line system (OLS) to carry several Tbps of traffic. To enable data transmission, optical transponders receive electrical signals from connected router ports and convert them into wavelengths called *optical channel*. These wavelengths operate at data rates of hundreds of Gbps and occupy specific spectrum widths known as *channel spacing*. Multiple optical channels are combined using an arrayed waveguide grating (AWG) [42] multiplexer (MUX), where each filter port in the MUX corresponds to a passband defined by upper and lower frequency bounds. These channels then travel through the limited spectrum available for long-haul transmission in the fiber, typically within the C-band [6]. To configure the direction of each wavelength, the Reconfigurable Optical Add-Drop Multiplexer (ROADM) [3] comes into play. The ROADM allows for dynamic reconfiguration of wavelengths, enabling them to map to various surrogate fibers for efficient transmission.

Today's cloud providers operate fixed optical backbones where all the optical hardware devices run on the same rigid spectrum grid (e.g., 50 GHz or 75 GHz [5]). This fixed grid approach is adopted to maintain simplicity and ease of management within the optical layer. Specifically, each wavelength generated by optical transponders is bound to a fixed channel spacing, while the equipment in the OLS is designed based on a fixed grid, ensuring that the passband maintains the same rigid spectrum width throughout the network.

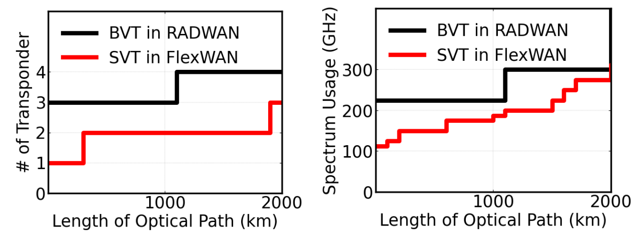
Fixed-rate WAN. In traditional wide-area settings, the fixed-rate WAN approach prevails [27, 28]. In this configuration, the wavelengths generated by optical transponders are limited to a single modulation format. For instance, each wavelength occupies a 50 GHz spectrum width and is modulated at a data rate of 100 Gbps, with a maximum optical reach of 3000 km. Here, the optical reach refers to the farthest distance a signal can travel while still allowing error-free decoding at the destination. The passband width in the OLS remains fixed at 50 GHz in this context.

Rate-adaptive WAN (RADWAN). As the signal traveling distance is diverse, signals traveling short distances undergo less attenuation, leading to a high SNR. The SNR ultimately decides the data



(a) Optical path length (b) Maximum data rate

Figure 2: (a) The distribution of the optical path lengths in a production WAN and (b) The maximum data rates supported by BVT in RADWAN [47] and SVT in FlexWAN according to the traveling distance required.



(a) Minimum transponder number (b) Minimum spectrum usage

Figure 3: A typical example of hardware costs for provisioning 800 Gbps WAN capacity at different optical path lengths.

rate of the signal - higher the SNR, higher order modulation format and consequently the data rate of the signal. Recent advances [46, 47, 49] called RADWAN proposed using state-of-the-art optical transponder called bandwidth-variable transponder (BVT) that provides multiple modulation formats to support different data rates per wavelength according to the traveling distance required. As the technology of transponders evolves to occupy a channel spacing of 75 GHz [49], we adapt the BVT in [47] into our scene: 300 Gbps, 200 Gbps, and 100 Gbps by modulating the signals in 8QAM, QPSK, and BPSK for distances up to 1100 km, 2000 km, and 5000 km respectively. The passband width in the OLS is equal to 75 GHz.

3 MOTIVATION

3.1 Far from Shannon limits

We conducted extensive measurements on a large-scale production WAN, examining hundreds of optical paths that traverse thousands of IP links. The term "optical path" refers to the wavelength's journey from the source region to the destination region, carrying bandwidth capacity for a specific IP link. Figure 2(a) provides a visual representation of the distribution of optical path lengths across all IP links. Notably, the optical path lengths exhibit significant variation, ranging from as short as 100 km to well over 2000 km.

The data reveals an essential insight – approximately 50% of the optical paths are less than 200 km in length, which is considerably shorter than the optical reach at 300Gbps in RADWAN. According to the Shannon-Hartley theorem [45], the maximum data rates depend on the channel spacing and signal-to-noise ratio, with the formulation $C = W \log_2(1 + S/N)$. In this context, C represents the data rate of an optical channel, W is the channel spacing, and S/N denotes the signal-to-noise ratio (SNR). The short distances traveled by these signals result in lower attenuation and a higher SNR. The high-quality signal indicates that the data rate of 300Gbps is significantly below the Shannon limits. By employing higher-order modulation formats like 256QAM, the signal's data rate can be increased to support up to 800Gbps. However, it is crucial to consider the limitations of aggressively increasing the order of modulation formats. Extremely high-order modulation formats necessitate precise signal generation and are more susceptible to optical impairments, including chromatic dispersion and fiber nonlinearity [13, 31, 38, 41].

The Shannon-Hartley theorem sets a definite limit on the maximum data rate for a given channel spacing, which means that a wavelength cannot support high data rates like 800 Gbps at a low channel spacing (e.g., 75 GHz), even with high-order modulation formats. However, this theorem opens up a promising opportunity for adapting wavelength channel spacing to accommodate potentially higher data rates. To achieve this flexibility, we introduce spacing-variable transponders (SVTs) in FlexWAN, enabling the provision of high data rates tailored to the required traveling distance. We conducted production-level testbed experiments to evaluate the performance of SVTs, and the specifications are detailed in Table 2 in Appendix A.2. Figure 2(b) illustrates the comparison between the maximum data rates supported by bandwidth-variable transponders (BVTs) and SVTs at various optical path lengths. Notably, we observe a significant gap between SVT and BVT, particularly for short optical paths. This observation serves as a strong motivation for delving deeper into the optical hardware costs associated with provisioning WAN capacity.

3.2 Waste of optical hardware

The cost of long-haul WAN capacity is mainly influenced by transponders, fiber, and equipment in the OLS. On the one hand, providers can effectively save on hardware costs by reducing spectrum usage in the fiber, considering the limited spectrum available for long-haul transmission, typically within the C-band [6]. This optimization allows for the allocation of the remaining spectrum to provision additional WAN capacity and address optical failures without the need to purchase or lease new fiber. On the other hand, the marginal cost of provisioning WAN capacity is dominated by the extra hardware resources, namely transponders. To illustrate this cost analysis, we consider a representative example that studies the hardware costs for provisioning 800Gbps WAN capacity, with the specifications of the spacing-variable transponders (SVTs) detailed in AppendixA.2. By understanding the key cost drivers, we can better optimize the deployment of WAN capacity and ensure cost-effective solutions for future network expansions.

Waste of transponders. In Figure 3(a), we observe a significant difference in the minimum pairs of transponders used between the SVTs and the bandwidth-variable transponders (BVTs) as the optical path length varies. When utilizing SVTs, we require notably fewer transponders compared to BVTs for the same path length. For instance, when the optical path is shorter than 300 km, we only need one pair of SVTs to provision the desired WAN capacity of 800 Gbps, whereas three pairs of BVTs are necessary for the same task. This is because SVTs can support high data rates of 800 Gbps at short optical paths, resulting in a reduced need for additional transponders. Conversely, for long transmission distances such as

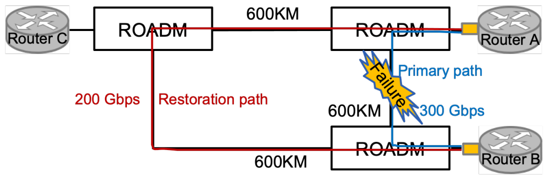
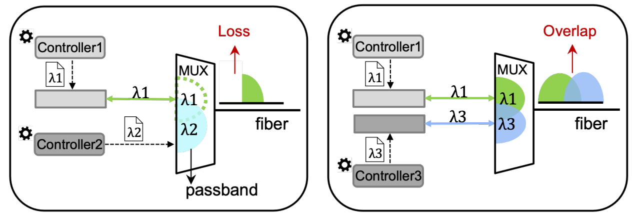


Figure 4: Optical restoration when fiber cuts occur.



(a) Channel inconsistency (b) Channel conflict

Figure 5: Spectrum-related issues.

1800 km, the number of transponders used in SVT is only half of that used in BVT, further highlighting the cost-saving advantages of using spacing-variable transponders.

Waste of spectrum in the fiber. In Figure 3(b), we observe a consistent trend, where there is a substantial disparity in spectrum usage between the BVTs and SVTs. Specifically, when considering optical paths shorter than 300 km, the spectrum width occupied by three pairs of BVTs amounts to 225 GHz, with each optical channel utilizing 75 GHz. In contrast, the maximum spectrum usage by a single pair of SVTs is only 150 GHz, significantly reducing the overall spectrum consumption. This discrepancy illustrates how SVTs enable more efficient utilization of the limited spectrum available in the fiber, particularly for shorter optical paths.

The representative results demonstrate that the traditional fixed optical backbone is uneconomical because it wastes both the number of transponders and the valuable spectrum in the fiber.

3.3 Poor performance addressing fiber cuts

Fiber cuts are a common cause of failures in the optical layer, and their impact can be severe, resulting in data losses of several terabits per second [7, 49]. To address these fiber cuts and mitigate their effects, optical restoration has emerged as an advanced solution [35–37]. Optical restoration involves reconfiguring the wavelengths from the cut fiber onto healthy fibers, effectively restoring network connectivity. This underscores the remarkable effectiveness and significance of optical restoration in safeguarding network resilience and minimizing disruptions caused by fiber cuts. Our measurements of a production WAN have revealed valuable insights. Specifically, we found that 90% of the restored optical paths were longer than their original configurations (as elaborated in § 8). In some extreme cases, the restored path length exceeded that of the original by over ten times.

We illustrate a representative optical restoration case in Figure 4. As the length of the primary path is 600 km, the wavelength generated by BVT in RADWAN is initially modulated at 300 Gbps bandwidth capacity with an optical reach of 1100 km. When the fiber in the primary path fails, the wavelength should be reconfigured to a restoration path to revive the lost capacity. However, the restoration path is longer up to 1200 km, exceeding the optical reach at 300 Gbps. The wavelength should lower its achievable data rate to 200 Gbps at a low-order modulation format, resulting in the bandwidth capacity not being fully revived. The loss of bandwidth capacity in turn hampers the network's ability to meet traffic demands. A more attractive approach is to revive all the bandwidth capacity even if the restoration path is longer. We achieve it by using our proposed SVT which supports flexible adjustment of wavelength channel spacing. For example, we could increase the channel spacing of the wavelength to 87.5 GHz, utilizing an additional 12.5 GHz spectrum to cope with noise introduced by longer optical path.

3.4 Challenges

The limitations of fixed optical backbones have prompted cloud providers to explore a new approach: a flexible WAN infrastructure that provides cost-effective WAN capacity while ensuring resilience to optical failures. However, achieving this goal is challenging, as it involves addressing both *system-level* challenges, such as physically supporting flexibility at the optical layer and handling spectrum-related issues in a multi-vendor optical backbone, and *algorithmic* challenges, including finding optimal network planning strategies and optical restoration plans. Overcoming these obstacles requires innovative solutions and a comprehensive understanding of the complexities involved in creating a flexible WAN infrastructure.

Challenge1 (system-level challenges): existing hardware and software do not support the flexible WAN infrastructure. Existing optical hardware devices operate on a rigid spectrum grid. This is attributed to the inherent characteristics inside the device. For example, the electro-optic modulator (EOM) inside the BVT only inputs/outputs wavelengths at a fixed channel spacing. However, the implementation of spacing-variable hardware at the optical layer presents practical challenges in a multi-vendor optical backbone. As vendor diversity is common in production, each vendor uses a vendor-customized controller to interface the optical devices they manage. Configuring thousands of IP links and hundreds of optical paths from multiple vendors using these individual controllers increases the likelihood of spectrum-related issues, such as channel inconsistency and channel conflict. The channel inconsistency indicates that the spectrum used by the wavelength and the passband provided by the equipment in OLS along the wavelength's optical path is not fully identical. This phenomenon will cause the loss of signals at the destination and finally lead to the loss of capacity (see Figure 5(a)). The channel conflict indicates that the wavelengths whose optical paths intersect use the same spectrum frequency in the fiber. The signals are hard to merit error-free decoding at the destination due to the signal overlap (see Figure 5(b)).

Challenge2 (algorithmic challenges): the flexibility at the optical layer offers new algorithmic problems. As the flexibility at the optical layer offers a vast number of possibilities, it is challenging for cloud providers to choose optimal decisions. On the one hand, with the goal of providing a cost-effective optical backbone, it is challenging to choose the best planning strategy (e.g., transponder number, modulation format of each wavelength) to provision the WAN capacity with physical constraints (e.g., optical reach and limited spectrum). On the other hand, we extend FlexWAN to incorporate the goal of failure resilience by introducing optical restoration. As fiber cut affects multiple IP links, it is challenging to provide an optimal restoration plan with the maximized

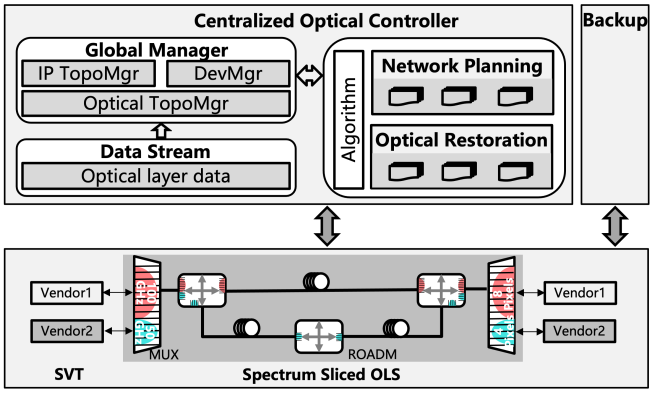
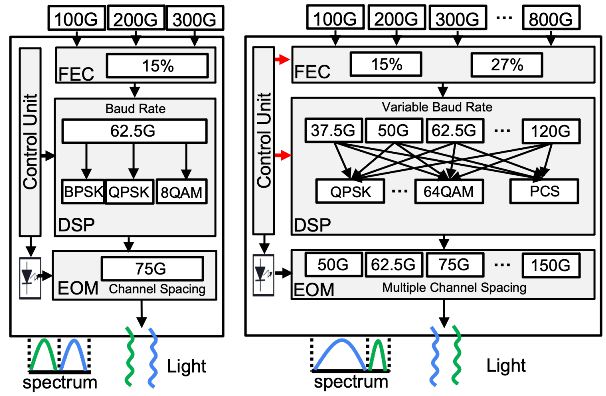


Figure 6: FlexWAN's high-level design.



(a) BVT in RADWAN (b) SVT in FlexWAN

Figure 7: Detailed workflow of optical transponder.

restoration capacity because the restoration paths of these affected IP links intersect and compete for the limited spectrum in the fiber.

4 FLEXWAN DESIGN AND IMPLEMENTATION

4.1 High-Level Design

FlexWAN's goal is to design a novel flexible WAN infrastructure to provision cost-effective WAN capacity and enhance resilience to optical failures. FlexWAN addresses the challenges of the existing optical backbone to achieve this goal. More concretely, to address the rigid issue of optical hardware and achieve a flexible WAN infrastructure, FlexWAN introduces spacing-variable hardware at the optical layer to generate wavelengths that are optimally modulated for the required data rate and channel spacing based on the transmission distance (§ 4.2). The key differences in the optical backbone infrastructure between the traditional WANs and FlexWAN are shown in Appendix A.1. To address spectrum-related issues in a multi-vendor optical backbone, FlexWAN leverages a centralized controller interfacing with devices in a vendor-agnostic manner to coordinate network-wide devices. To address algorithmic problems, FlexWAN provides two algorithms that find the best network planning strategy and optical restoration decision respectively (§ 4.4). Figure 6 presents FlexWAN's high-level design.

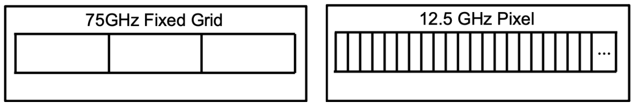
4.2 Spacing-variable Hardware

Existing optical hardware devices operate on a rigid spectrum grid, leading to the inefficient use of expensive optical resources. FlexWAN introduces *spacing-variable hardware* at the optical layer where spacing-variable transponder and spectrum-sliced optical line system (OLS) physically support the flexibility inside the device. We present several key designs of optical hardware devices below.

Spacing-variable transponder (SVT). Figure 7(a) zooms in into a typical bandwidth-variable transponder (BVT) that provides multiple modulation formats proposed in RADWAN [47]. We simplify the transponder into a control unit and several key components, including forward error correction (FEC) module, digital signal processor (DSP), and electro-optic modulator (EOM), which are closely related to our work. There are three parallel workflows in the DSP, each of which modulates the signals at a specific order format (e.g., BPSK, QPSK, and 8QAM). The control unit controls the DSP to determine the data rate of the wavelength by adjusting the modulation format. However, these workflows share the rigid components providing fixed capabilities such as fixed FEC in the FEC module and fixed channel spacing in the EOM. The inherent rigid characteristics of these components in the BVT make it hard to support high data rates at short optical paths.

To break this limit, we partially corporate with vendors to design the spacing-variable transponder (SVT), as shown in Figure 7(b). Compared to the BVT, the key components inside SVT are adjustable. Specifically, the FEC module provides several choices of FEC capability by adding different ratios of redundant data (e.g., 15%, 27%). The higher the ratio of redundant data in the signal, the higher the capability of meriting error-free decoding at the destination in long traveling distances. The variable baud rates (e.g., 37.5 GBd, 50 GBd) and different modulation formats (e.g., QPSK, 8QAM, PCS) have been almost fully meshed in the DSP. The different combinations of baud rate and modulation format promote the signals to support multiple data rates. Here, the probabilistic constellation shape (PCS) [20] is an advanced modulation technique that supports finer-granularity data rates. The multiple combinations of FEC, baud rate, and modulation format in these adjustable components allow the wavelengths to be optimally modulated for the required data rate and channel spacing based on the transmission distance. Note, the flexibility inside the transponder does not bring much complexity to the control unit because the control unit only needs to receive the configuration parameters from the controller to configure each module.

Spectrum-sliced optical line system (OLS). As the channel spacing of a wavelength generated by the SVT may change dynamically, the passband provided by the fixed-grid equipment in OLS can not adapt to such dynamics (see Figure 8(a)). Providing a larger passband for a signal will result in the waste of valuable spectrum in the fiber while providing a smaller passband will cause the loss of signal. Therefore, we introduce the spectrum-sliced optical line system (OLS) where the passband width can dynamically adapt to the channel spacing of the wavelength. Specifically, we utilize the Liquid Crystal on Silicon (LCoS) [26] based wavelength selective switch (WSS) at each optical site (e.g., MUX, ROADM) in the OLS. As shown in Figure 8(b), the LCoS-based WSS reduces the granularity of the grid from 75GHz down to 12.5 GHz pixel or lower (if



(a) Traditional fixed passband (b) Pixel-wise flexible passband

Figure 8: Comparision of the passband in the OLS.

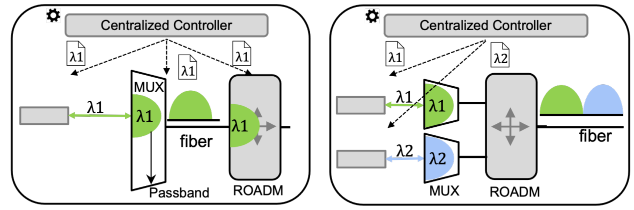
required) [2], called *pixel-wise WSS*. The passband can be dynamically adjusted to a higher or lower width according to the number of pixels selected. Note the selected pixels should be continuous. In this way, the passband width could be precisely aligned with the spectrum width occupied by the wavelength. For example, if a wavelength occupies a 100 GHz channel spacing, the passband is configured with eight continuous pixels. For these wavelengths requiring a large channel spacing to carry higher data rates, the passband could be configured with more continuous pixels.

The spacing-variable hardware supports a flexible WAN infrastructure physically and allows each wavelength to efficiently utilize the limited spectrum resources to carry the WAN capacity.

4.3 Network-wide Device Coordination

The implementation of spacing-variable hardware at the optical layer presents practical challenges in a multi-vendor optical backbone. Configuring hundreds of optical paths from multiple vendors using these individual controllers increases the likelihood of spectrum-related issues. Therefore, we need a centralized controller to coordinate network-wide optical devices to overcome these issues. To achieve this, similar to previous work [7], we utilize a standard device model for each type of device so that the heterogeneous devices across vendors are uniformly abstracted into a group of logic components. Then, the device model provides the mapping of these abstracted logic components to specify the detailed workflow between them. At this point, these heterogeneous optical devices can be abstracted the same at a high level as each vendor follows the same device model to abstract its devices. By utilizing a standard device model, the network-wide devices can be interfaced by a centralized controller in a vendor-agnostic manner.

Coordinating optical device. The centralized control of optical devices enables the controller to hold a holistic view of optical devices, which promotes the coordination of these optical devices. Specifically, each optical device is allocated with an IP address and the centralized controller uses this IP address to locate the optical device in the optical backbone. For these optical devices along a wavelength's optical path, the centralized controller uses the same configuration parameters as the wavelength's spectrum to configure the passband of these devices. Figure 9(a) depicts an example of achieving channel consistency. The centralized controller configures the transponder to generate a wavelength occupying the spectrum of λ_1. When the optical path of the wavelength is determined, the controller configures the equipment at each optical site (such as MUX, ROADM) along the wavelength's optical path to provide the passband of λ_1. In this way, the spectrum required by the wavelength and the spectrum of the passband in the OLS is fully identical and the loss of the signal caused by channel inconsistency can be avoided. On the other hand, for these wavelengths generated by transponders passing through the same fiber, the centralized controller should configure the non-overlapping spectrum for them.



(a) Consistent channel (b) Non-overlapping channel

Figure 9: Two typical examples.

As shown in Figure 9(b), the optical paths of the two wavelengths intersect in a fiber. The centralized controller configures the two wavelengths with λ_1 and λ_2 respectively. Note, the λ_1 and λ_2 do not overlap in the spectrum. Therefore, the two wavelengths will not interfere with each other during the transmission.

Our experience has shown that there is *zero* spectrum inconsistency and conflict issue by leveraging a centralized control of the optical backbone, demonstrating its effectiveness.

4.4 Centralized Optical Controller

The centralized optical controller consists of four key modules, i.e., global manager, network planning, optical restoration, and data stream. The global manager module maintains a global view of network topology at both IP and optical layers and contains the standard device model. The network planning module contains a cost-effective network planning algorithm to generate the optimal optical layer configuration strategy. As FlexWAN incorporates the goal of failure resilience, the optical restoration module contains a restoration algorithm to provide the best restoration plan. The data stream module collects and stores optical layer data used for detecting optical events, such as fiber cuts.

Global manager. It mainly consists of three parts: IP topology manager (IP TopoMgr), optical topology manager (Optical TopoMgr), and device manager (DevMgr).

- IP TopoMgr: It holds a network-wide view of an IP-layer topology and stores the demands of bandwidth capacity of each pair of two IP nodes (i.e., IP links). Determining bandwidth capacity for an IP link is challenging. Providing a capacity much larger than the traffic demands will cause the waste of optical hardware while providing a capacity close to the traffic demands will potentially lead to traffic loss when traffic surges or network failures happen. Several works have studied efficient bandwidth capacity provision [10, 46]. Since it is not the core of our work, we use the bandwidth capacity of each IP link provided by network operators according to their experience.
- Optical TopoMgr: It holds a physical topology of optical hardware devices, connected by optical fibers. For each IP link, the optical TopoMgr provides several optical paths which consist of a set of ROADMs and fibers. As optical hardware devices and fibers may suffer from optical failures, the optical TopoMgr ingests the optical layer data from the data stream module to update the physical topology. Once an optical failure happens, the optical TopoMgr will notify the optical restoration module to generate the optimal restoration plan. Here, we focus on fiber cuts, the most common and disruptive cause of failures in the optical layer [7, 49].
- DevMgr: It contains a standardized device model and configuration parameters of each optical device. After receiving the detailed

configurations from the network planning module and the optical restoration module, it utilizes the IP address to locate the device in the network and issues a Yang file [16] containing detailed configuration parameters to configure the device through the Netconf protocol [24]. The configuration parameters of each type of optical device may differ from each other. For example, transponders will be configured with non-overlapping spectrums to avoid spectrum conflicts if their generated wavelengths' optical paths intersect in the same fiber.

Network planning. It contains an optimal and cost-effective network planning algorithm. Specifically, it obtains both the IP and optical topology and bandwidth capacity demand of each IP link from the global manager and runs the network planning algorithm *offline* with the goal of minimizing hardware costs. Once the optimal configuration plan is generated, it returns the configuration parameters of all devices to the DevMgr in the global manager. Since bandwidth capacity changes monthly or yearly, the network planning module serves as a long-term strategy and is operated *infrequently*. We will present the detailed algorithm in § 5.

Optical restoration. It contains an optimal optical restoration algorithm with the goal of maximizing restored capacity. Specifically, it receives the device configuration details of the existing optical backbone and current optical topology from the global manager and replies with the detailed restoration plan including modulation formats of influenced transponders and restored paths of these influenced wavelengths. As we have a set of failure scenarios based on the link failure model [17, 40], the restoration plan for each fiber cut scenario can be produced *offline*. This algorithm is executed *occasionally*. We will present the detailed algorithm in § 8.

Data stream. It *periodically* collects optical layer data, including key performance indicators of each optical device. The optical layer data contains meaningful information about the optical layer status. For example, the transmitted and received power of two terminal devices at each end of a fiber cable could be used to identify the status of the fiber cable. The data stream module is an online database (e.g., Kalfa system) that is equipped with a scalable collector to support one-second granularity data collection [7]. Such high-precision data can be used to detect optical failures in real time.

Fault tolerance. The centralized optical controller in FlexWAN is cloud-based and designed to be robust to unpredictable control-plane failures, such as software bugs, network disconnections, and even natural disasters (e.g., earthquakes, hurricanes, floods). The main controller is deployed as close to the devices to reduce the control delay. By utilizing the containerization technique in the cloud, multiple backups of the controller are deployed in multiple geo-disjoint areas that provide redundancy at multiple levels. This setup allows the controller to survive under network failures or natural disasters.

5 NETWORK PLANNING

The flexibility at the optical layer presents new algorithmic challenges for cloud providers when choosing an optimal and cost-effective network planning strategy. In addressing this, FlexWAN formulates the problem of WAN capacity provisioning with the goal of minimizing hardware costs. The formulations are shown in Algorithm 1 and we will describe them below in detail.

Algorithm 1 : Optimal Network Planning

Inputs:

$G(V,E)$: The IP topology with routers V and IP links E

$G_o(V_o,E_o)$: The optical topology G_o with ROADMs V_o and fibers E_o

c_e : The bandwidth capacity of link e

$P_{e,k}$: The k-th optical path of link e

$w \in W$: The spectrum slot w

(d,l,Y) : A set of modulation formats in SVT with data rate d, optical reach l, and channel spacing Y, where j-th format is represented as d_j, l_j, Y_j

$\pi_{\phi}^{e,k}$: 1 if fiber ϕ belongs to the path k of link e, and 0 otherwise

$s_w^{j,q}$: 1 if SVT at the j-th format uses slot w at q-th order, and 0 otherwise

Outputs:

$\xi_{\phi,w}^{e,k}$: 1 if path k of link e uses slot w of fiber ϕ, and 0 otherwise

$Y_{j,q}^{e,k}$: 1 if path k of link e uses SVT at j-th format and q-th order, and 0 otherwise

$\lambda_j^{e,k}$: The number of SVTs at j-th format used by path k of link e

Minimize: $\sum_e \sum_k \sum_j \lambda_j^{e,k} + \epsilon \sum_e \sum_k \sum_j \lambda_j^{e,k} Y_j$

subject to:

(1) $\sum_k \sum_j d_j \lambda_j^{e,k} \geq c_e;$ $\forall e$

(2) $(l_j - |P_{e,k}|)\lambda_j^{e,k} \geq 0;$ $\forall e,j,k$

(3) $\sum_e \sum_k \xi_{\phi,w}^{e,k} \leq 1;$ $\forall \phi, w$

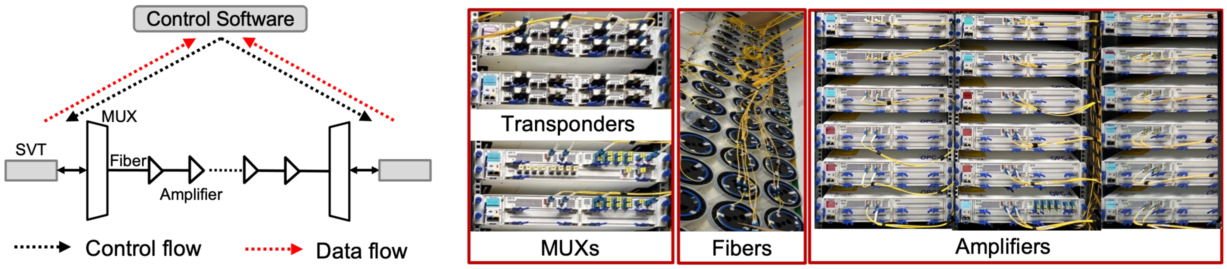
(4) $\xi_{\phi,w}^{e,k} = \xi_{\phi',w}^{e,k} \quad if\ \pi_{\phi}^{e,k} = \pi_{\phi'}^{e,k} = 1;$ $\forall e,k,w$

(5) $\xi_{\phi,w}^{e,k} = \sum_j \sum_q Y_{j,q}^{e,k} s_w^{j,q} \times \pi_{\phi}^{e,k};$ $\forall e,k,\phi,w$

(6) $\lambda_j^{e,k} = \sum_q Y_{j,q}^{e,k};$ $\forall e,k,j$

Inputs: The IP topology and optical topology of the WAN are modeled as the graph $G = (V,E)$ and $G_o(V_o,E_o)$ respectively. The demand of bandwidth capacity of an IP link $e \in E$ is c_e. As the IP layer bandwidth capacity can be aggregated using the link aggregation protocol [34], we use K shortest path (KSP) algorithm to find the K optimal optical paths in the optical topology $G_o(V_o,E_o)$ for each IP link, where the k-th optical path of link e is $P_{e,k}$. The pixel-wise WSS in the OLS slices the spectrum in the fiber into a finer granularity (i.e., 12.5 GHz) with a total number of $|W|$ pixels, where the slot w represents the index of the pixel. Each SVT provides multiple formats with different combinations of data rate d, optical reach l, and channel spacing Y (see Table 2), where the j-th format is represented as d_j, l_j, and Y_j respectively. Here, the channel spacing Y can be translated to the number of pixels. As the wavelength generated by the SVT should use the continuous pixels in the fiber, the total number of choices for the spectrum that a wavelength modulated at j-th format can occupy is determined. We make the order of these choices based on the starting pixel q. $s_w^{j,q}$ is a binary variable indicating if the SVT modulated at the j-th format adopting the q-th order uses slot w, and 0 otherwise. For example, SVT modulated at the 1-st (j=1) format (i.e., wavelength modulated at 100 Gbps occupying a 50 GHz spectrum width) occupies 4 pixels and the 5-th order (q=5) occupies slots w of 5, 6, 7, and 8, i.e., $s_5^{1,5} = 1$, $s_6^{1,5} = 1$, $s_7^{1,5} = 1$, and $s_8^{1,5} = 1$. $\pi_{\phi}^{e,k}$ is a binary variable indicating if fiber $\phi \in E_o$ belongs to the path k of link e.

Outputs: FlexWAN generates the optimal planning strategy for WAN capacity provisioning. $\xi_{\phi,w}^{e,k}$ is a binary variable indicating if the pixel w of fiber ϕ is used by the path k of link e. The decision of



(a) (b)

Figure 10: Production-level testbed. (a) Detailed workflow of the testbed. (b) Photo of the testbed.

$\gamma_{j,q}^{e,k}$ is a binary variable indicating if the path k of link e uses SVT at the j-th format and q-th order. The total number of SVTs at the j-th format for the optical path k of link e is represented as $\lambda_j^{e,k}$. We note that $\lambda_j^{e,k}$ is an integer.

Objective function: The goal of FlexWAN in network planning is to minimize hardware costs. The hardware costs contain direct and indirect costs. The direct cost is dominated by the extra transponders, represented by the sum of SVTs at all modulation formats for all IP links and optical paths, i.e., $\sum_e \sum_k \sum_j \lambda_j^{e,k}$. The indirect cost is the spectrum usage in the fiber. Although cloud providers have existing fiber deployments, reducing spectrum usage allows for the allocation of the remaining spectrum to provision additional WAN capacity and address optical failures without purchasing or leasing a new fiber. The total spectrum usage by all IP links is represented as $\sum_e \sum_k \sum_j \lambda_j^{e,k} Y_j$. The parameter ϵ defines the balance between FlexWAN's tendency to minimize the direct or indirect cost.

Bandwidth capacity constraints: Each IP link between two edge routers has a demand of bandwidth capacity c_e to meet traffic demand amount between two regions. Under the hood, there are a set of optical paths, each of which provides a portion of bandwidth capacity. Optical paths are a pre-computed set using the KSP algorithm on the optical topology $G_o(V_o, E_o)$ between the IP link source and destination region. Each path k contains multiple wavelengths, each of which supports a specific data rate. The sum of the data rates of wavelengths from all optical paths for an IP link should meet the demand of the bandwidth capacity.

$$\sum_k \sum_j d_j \lambda_j^{e,k} \geq c_e \tag{1}$$

Optical reach constraints: The length of an optical path k of link e is $|P_{e,k}|$. Each SVT provides multiple modulation formats with different optical reaches. For the transmission distance of $|P_{e,k}|$, the SVT at j-th format can not be adopted if the optical reach l_j is less than the required transmission distance, i.e., $\lambda_j^{e,k} = 0$. As the transmission distances longer than the optical reach of the signal make the SNR too low to merit error-free decoding at the destination, we only use the modulation format whose optical reach l_j is longer than the transmission distances. The transponder number $\lambda_j^{e,k}$ should be an integer. We summarize the above constraints.

$$(l_j - |P_{e,k}|)\lambda_j^{e,k} \geq 0 \tag{2}$$

Spectrum conflict constraints: The spectrum overlap among wavelengths in the fiber $\phi \in E_o$ affects the transmitted '1' and '0' signals to be correctly decoded. We demonstrate how we address spectrum conflict in the formulation. Since we have sliced the spectrum in the fiber into W pixels, the spectrum overlap can be restated as the repeat uses of the pixel. Each slot w in the fiber can only be used at most once. Therefore, the sum of slot w usage in fiber ϕ from all IP links and optical paths should be no more than 1.

$$\sum_e \sum_k \xi_{\phi,w}^{e,k} \leq 1 \tag{3}$$

Spectrum consistency constraints: Each optical path contains several wavelengths that traverse one or several fibers to provide the bandwidth capacity. For these fibers that a wavelength passes through, the occupied spectrum by this wavelength should be the same. For each optical path k of an IP link e, the status (occupied or spare) of each slot along the fiber path should be the same.

$$\xi_{\phi,w}^{e,k} = \xi_{\phi',w}^{e,k} \quad if\ \pi_{\phi}^{e,k} = \pi_{\phi'}^{e,k} = 1 \tag{4}$$

Spectrum status constraints: There are a number of wavelengths that provide the bandwidth capacity for optical path k of link e. At the optical layer, each wavelength is modulated at the j-th format occupying the spectrum starting at q-th order, i.e., $\gamma_{j,q}^{e,k}$ in the fiber ϕ that belongs to the optical path k. The occupied spectrum of that wavelength can be represented by $s_w^{j,q}$. Therefore, the status of slot w in fiber ϕ can be represented by the sum of all modulation formats of wavelengths and slots occupied by these wavelengths. On the other hand, the status of slot w in the fiber ϕ by optical path k of link e can be directly represented by the variable $\xi_{\phi,w}^{e,k}$.

$$\xi_{\phi,w}^{e,k} = \sum_j \sum_q \gamma_{j,q}^{e,k} s_w^{j,q} \times \pi_{\phi}^{e,k} \tag{5}$$

Transponder number constraints: As each wavelength will occupy a specific spectrum, starting at q-th order. For each optical path k of link e, the total number of SVT modulated at the j-th format can be counted by the number of orders that these wavelengths modulated at the same j-th format use.

$$\lambda_j^{e,k} = \sum_q \gamma_{j,q}^{e,k} \tag{6}$$

6 TESTBED EVALUATION.

In this section, we evaluate the performance of our proposed spacing variable transponder (SVT) by partly collaborating with *vendor A* using a production-level testbed experiment. Here, we omit the vendor name for confidentiality reasons.

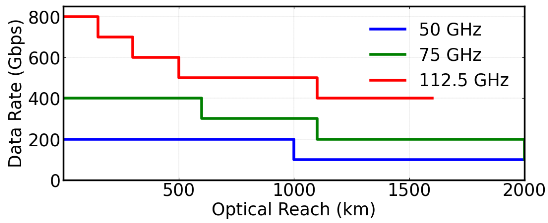


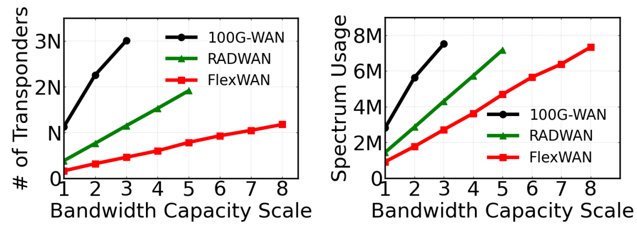
Figure 11: Specifications of the SVT

Setup. As shown in Figure 10, our testbed is combined of a pair of SVTs, MUXs, multiple bundles of fibers, tens of amplifiers, and a centralized controller. In our testbed, all optical hardware devices and software are identical to our production optical backbone. The centralized controller is used to adjust the modulation format of the generated wavelength by sending the control flow including configuration parameters to the SVT. Meanwhile, it receives the performance data including the post-FEC bit error rate (BER) from the SVT. Here, the post-FEC BER indicates whether the signal can be correctly decoded [7]. The positive values of the post-FEC BER show that the SNR of the signal is too low to merit error-free decoding at the destination. There are multiple bundles of fibers to control the optical path length. We introduce an amplifier for each 50∼100 km fiber which is consistent with the production network. In this experiment, we control the format of SVT and gradually increase the fiber length. If the post-FEC BER increases from 0 to a positive number, we obtain the maximum transmission distance at the current format.

SVT Specifications. Figure 11 depicts the data rates and optical reaches of the wavelength when it is modulated to occupy a typical channel spacing. Note that the channel spacing of 75 GHz is widely used in the existing fixed optical backbone. The detailed specifications are shown in Table 2 (see Appendix A.2). We observe that wavelengths occupying a fixed channel spacing should be modulated at lower data rates if the transmission distance becomes longer. A longer transmission distance always leads to a lower SNR, and thus the signal should lower its achievable capacity. If the wavelength transmits a longer optical path while supporting the same data rate, it should be modulated at occupying more channel spacing. For example, the wavelength is modulated at 400 Gbps occupying the channel spacing of 75 GHz for distances up to 600 km. However, the wavelength can support distances up to 1600 km if it is modulated to occupy the channel spacing of 112.5 GHz. Furthermore, the wavelength can be modulated at a high data rate (800 Gbps) if the optical path is short. We use the specifications of the SVT measured in the testbed to conduct cost-effective network planning and optimal optical restoration.

7 COST SAVING WITH FLEXWAN

We implement FlexWAN 's optimization algorithm using Julia programming language [14] and the Gurobi solver [1]. Note that Algorithm 1 solves a minimization problem with a mixed integer program (MIP). In practice, the Gurobi solver utilizes linear programming (LP) relaxation and provides the lower bound for the minimization problem with a gap of less than 0.1%. All topologies of Algorithm 1 we formulate were solved within hours of runtime, which is acceptable because the network planning algorithm is running *offline* and *infrequently*.



(a) Transponder number (b) Spectrum usage

Figure 12: (a) Transponder number and (b) Spectrum usage vs. Bandwidth capacity scales for benchmark schemes and our proposed FlexWAN.

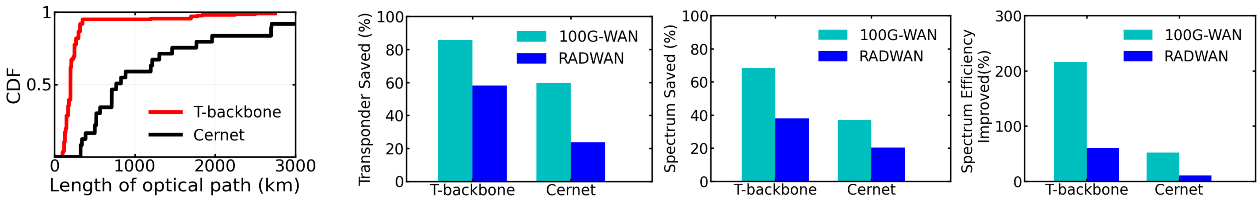
7.1 Reducing hardware costs of capacity.

As FlexWAN utilizes a cost-effective network planning algorithm to reduce hardware costs, we study the performance of FlexWAN in our production optical backbone.

Benchmark schemes. Benchmark schemes operate fixed optical backbones where all the optical hardware devices run on the same rigid spectrum grid. **100G-WAN**: The fixed-rate WAN approach prevails in the traditional wide-area settings. We used the 100G-WAN proposed in Microsoft [27, 28] that each wavelength occupies a 50 GHz spectrum width modulated at a data rate of 100 Gbps, with a maximum optical reach of 3000 km. **RADWAN**: Recent advance [46, 47, 49] called RADWAN proposed BVT that supports multiple data rates according to the traveling distance required. In our settings, the wavelength can be modulated in BPSK, QPSK, and 8QAM format to carry 100 Gbps, 200 Gbps, and 300 Gbps with 5000 km, 2000 km, and 1100 km optical reach respectively. All these modulation formats occupy the same channel spacing of 75 GHz.

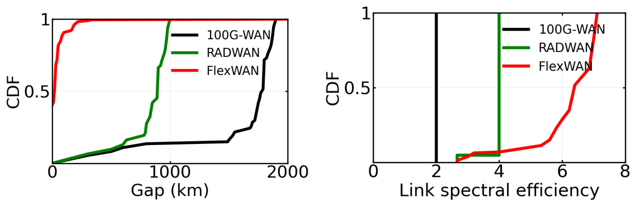
Hardware costs vs. bandwidth capacity scale. The optical backbone is often over-provisioned in an attempt to future-proof the network for the potential increase in demands. We obtain the demand of bandwidth capacity of each IP link in the production and evaluate the performance of FlexWAN by increasing the bandwidth capacity scale. Figure 12 shows the transponder number and spectrum usage by different settings of the optical backbone. We omit detailed values and represent them by *'N'* and *'M'* for confidentiality reasons. Figure 12(a) shows that the number of transponders is approximately linear to the scale of bandwidth capacity. This phenomenon is expected because providing twice the capacity always requires twice the number of transponders. FlexWAN is the most cost-effective because it requires the least number of transponders to provide the same bandwidth capacity. For example, FlexWAN saves transponders by 85% and 57% at the bandwidth capacity scale of 1, as compared to 100G-WAN and RADWAN respectively. We observe a similar trend of spectrum usage in Figure 12(b). FlexWAN reduces the spectrum usage by 67% and 36%, as compared to 100G-WAN and RADWAN respectively. The main reason is that wavelengths generated by SVTs can make the best use of the spectrum resources to carry the WAN capacity. An important observation is that FlexWAN supports the bandwidth capacity scale up to 8×, while 100G-WAN and RADWAN only support the scale up to 3× and 5× respectively. FlexWAN can support 2.7 × and 1.6× capacity scale, as compared to 100G-WAN and RADWAN respectively.

Optical reach vs. fiber path length. We then study the gap between the optical reach of a wavelength and the actual fiber path



(a) (b)

Figure 13: (a) The distribution of the optical path lengths on T-backbone and Cernet topologies weighted by the bandwidth capacity. (b) Percent of reduced costs and improved link spectrum efficiency by FlexWAN over 100G-WAN and RADWAN.



(a) Gap (b)

Figure 14: (a) The CDF of the gap and (b) The CDF of the link spectral efficiency.

length, i.e., *gap = optical reach - fiber path length.* The gap indicates if the optical reach of a wavelength is close to its optical path length. The lower the gap, the closer to the optimality of the format the wavelength is modulated. Figure 14(a) plots the distribution of the gaps. We observe that about 90% of the gaps in FlexWAN are less than 100 km, indicating that the optical reach of the wavelengths is quite close to their optical path length. However, 80% of the gaps in 100G-WAN are more than 1000 km, indicating that the actual fiber path length is much shorter than the optical reach of the signal. Signals traveling short distances undergo less attenuation, leading to a high SNR. The 100G-WAN has not efficiently utilized the high signal quality to enhance the signals to support more data rates.

Link Spectral Efficiency. The link spectral efficiency refers to the amount of data transmitted over a certain spectrum width, defined as $\frac{data\,rate}{spectrum\,width}$. Figure 14(b) plots the CDF of the link spectral efficiency of all wavelengths. The link spectral efficiency in 100G-WAN is fixed at 2 because all wavelengths occupy a 50 GHz spectrum width to carry 100 Gbps data rate. RADWAN improves the link spectral efficiency by supporting more data rates at the same channel spacing when the optical path is short. FlexWAN is the most spectrally efficient because SVT provides flexible combinations of data rates and channel spacing, enabling it to make the best use of spectrum resources to carry the WAN capacity for the distance required.

7.2 Impact of Network Topology.

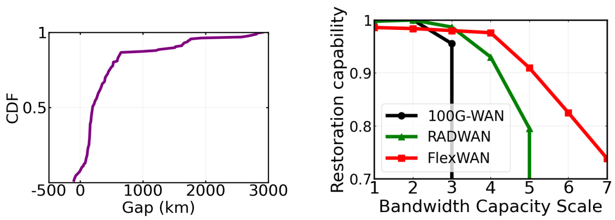
We then evaluate our proposed FlexWAN on different network topologies. We already have detailed information about the optical backbone we study, referred to as T-backbone. We additionally evaluate FlexWAN using the network topology for a prominent production network called Cernet released in [4] and regard it as the optical topology. We assume Cernet operates a point-to-point optical backbone and use distributions in [49] to generate the IP topology and bandwidth capacity. Figure 13(a) shows the optical path length distribution on the T-backbone and Cernet topology weighted by the bandwidth capacity. We observe that the median optical path length of the T-backbone topology is much shorter than that of the Cernet topology. We then study the reduced cost and average improved spectrum efficiency of FlexWAN over 100G-WAN and RADWAN in Figure 13(b). We find that FlexWAN shows a consistent potential of saving hardware costs with at most 85% transponders and 67% spectrum and improving spectrum efficiency by at most 215% in two network topologies. An important observation is that the performance gains on the T-backbone topology are larger than that on the Cernet topology. Note that the path length of the T-backbone topology is much shorter than that of the Cernet. This phenomenon shows that our FlexWAN always leads to more cost savings on the topology whose optical path length is always short. We believe our proposed FlexWAN can be extended to other network topologies and will show similar performance gains.

8 FAILURE RESILIENCE IN FLEXWAN

Optical restoration [35–37] is the advanced solution to address fiber cuts [49] by reconfiguring the cut fiber's wavelengths to healthy fibers. Our measurements of a production WAN in Figure 15(a) show 90% of the restored paths are longer than their original paths. In some extreme cases, the restored path length exceeds that of the original by over ten times. In this section, we must revisit the optical restoration ability of FlexWAN since it fundamentally re-architects present-day optical backbone. We assume the optical restoration algorithm is running on the configured optical backbone with the optimal network planning strategy. Our goal of optical restoration is to *maximize the restored capacity with the existing optical hardware.*

Failure scenario. There are many studies in the context of traffic engineering in the WAN to address link failures [17, 40, 49]. In these works, the link failure model was deterministic (k-failures) [40] or probabilistic [17]. FlexWAN provides the restoration plan for a failure scenario set that contains both 1-failures and probabilistic failures. For each failure scenario i, the formulations are shown below in detail.

Additional inputs and outputs. For each failure scenario i, the new optical topology is represented as $G'_o(V_o, E_o)$. c'_e represents the affected bandwidth capacity of the affected IP link e. We use the KSP algorithm in the optical topology $G'_o(V_o, E_o)$ to find the K optimal optical paths for the affected IP link where the k-th restoration path is represented as $P'_{e,k}$. These spare transponders whose original wavelengths are passing through the cut fiber are available to carry wavelength at the restoration path and the transponder number of link e that can be used for optical restoration is N_e. ϕ_w represents the status of slot w in fiber ϕ after network planning where '1' represents that the slot is spare and can be used for optical



(a) (b)

Figure 15: (a) The distribution of the gaps between the restored and original paths. (b) Restoration capability with different scales.

restoration. These outputs of optical restoration are similar to the outputs in Algorithm 1 that output variables $\xi'^{e,k}_{\phi,w}$, $\gamma'^{e,k}_{j,q}$, and $\lambda'^{e,k}_{j}$ represent the spectrum occupancy in the fiber, the specification of each SVT for restoration, and the number of SVT at j-th format.

Optimization goal and constraints: The goal of FlexWAN in optical restoration is to maximize the total restored capacity. The higher restored capacity always reduces the loss of network traffic and the network can achieve higher network availability under failures. The total restored capacity is the sum of the data rates of wavelengths generated by these spare transponders from all affected IP links. We present the formulation below.

Maximize: $\sum_e \sum_k \sum_j d_j \lambda'^{e,k}_{j}$

subject to:

(7) $\sum_k \sum_j d_j \lambda'^{e,k}_{j} \le c'_e$; $\forall e$

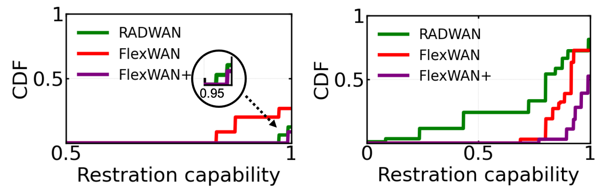
(8) $\sum_k \sum_j \lambda'^{e,k}_{j} \le N_e$; $\forall e$

(9) $\sum_e \sum_k \xi'^{e,k}_{\phi,w} \le \phi_w$; $\forall \phi, w$

(10)-(13) $Constraints(2,4,5,6)\ with\ (P'_{e,k}, \xi'^{e,k}_{\phi,w}, \gamma'^{e,k}_{j,q}, \lambda'^{e,k}_{j})$

Constraint (7) ensures that the total restoration capacity from all optical paths should be less than the affected capacity of IP link e, i.e., c'_e. Constraint (8) ensures that the total number of transponders of IP link e used for optical restoration should be less than its spare transponders. The optical restoration can only use the existing optical hardware. Constraint (9) ensures the spectrum occupancy of the restored optical paths for all IP links should be less than the available spectrum w in each fiber. Constraints (10)-(13) formulate the optical reach constraints, spectrum consistency constraints, spectrum status constraints, and transponder number constraints mentioned in the constraints (2,4,5,6).

Evaluations: We implement and solve the optical restoration algorithm in the T-backbone topology with the goal of maximizing restored capacity. The optical restoration algorithm solves a maximization problem with a MIP. We use the similar optimization solver that has been used in Algorithm 1. Figure 15(b) shows the average restoration capability in all failure scenarios at different capacity scales. The restoration capability is the ratio of the restored capacity to the affected capacity. We observe that both 100G-WAN and RADWAN can restore almost all the affected capacity when the network is underloaded (e.g., scale = 1). This phenomenon is reasonable because the two benchmark schemes provide too much redundancy at each wavelength that the optical reach of the signal is much longer than its optical path. For example, assume that the original optical path of an IP link is 1200 km. The wavelength in RADWAN is modulated at 200 Gbps bandwidth capacity with an optical reach of 2000 km while the wavelength in FlexWAN is modulated at a high data rate of 500 Gbps occupying a 125 GHz spectrum width with an optical reach of 1200 km. The wavelengths in FlexWAN are optimally modulated for the required data rate and channel spacing based on the transmission distance. When there is a fiber cut, the affected wavelength will be reconfigured to the healthy fiber for optical restoration. If the restored path is longer up to 2000 km, the wavelength in RADWAN can still support a 200 Gbps data rate while the wavelength in FlexWAN should be remodulated at a lower data rate of 300 Gbps (See Table 2), resulting in part of capacity loss. However, when the network is overloaded at 5× the present-day demands, both 100G-WAN and RADWAN perform worse than FlexWAN because they provide more redundancy at each wavelength at the cost of more transponders and spectrum usage. Therefore, there are not enough spectrum resources for the optical restoration. In contrast, FlexWAN still has available spectrum and outperforms RADWAN by 15% because wavelengths in FlexWAN are always more spectrally efficient.



(a) Underload (scale=1) (b) Overload (scale=5)

Figure 16: The restoration capability in the underloaded and overloaded optical backbones.

A natural question we ask is how can FlexWAN perform better than RADWAN even if the network is underloaded. We further introduce FlexWAN+, which provides extra half of the saved transponders by the current IP link for each IP link. Figure 16 shows the distribution of restoration capability under all failure scenarios. Figure 16(a) shows that FlexWAN+ performs better than RADWAN in the underloaded network, indicating that these extra transponders can well handle performance degradation caused by the longer restoration paths. We observe a similar performance in Figure 16(b) when the network is overloaded. The network operators should make the balance between the percentage of saved transponders and the optical restoration performance.

9 OPERATIONAL EXPERIENCE

FlexWAN has been running in production for two years and we have learned important operational lessons.

Smooth optical backbone evolution. Cloud providers should evolve their optical backbone to accommodate the increase of traffic demands. Recently, more aggressive transponders have been widely adopted in the industry to support higher data rates per wavelength by increasing the channel spacing from 50 GHz [47] to 75 GHz [49]. However, the inflexible nature of the existing optical backbone requires the evolvement of almost all the equipment in the OLS to ensure that the passband maintains the same rigid spectrum width of 75 GHz throughout the network. This operation requires the replacement of almost all the equipment that is costly and may introduce unpredictable failures. FlexWAN is a novel flexible WAN infrastructure that utilizes the spectrum-sliced OLS where the passband width can dynamically adapt to the channel spacing of the wavelength. If the wavelength occupies larger channel spacing, cloud providers only need to reconfigure the equipment in the spectrum-sliced OLS to provide the passband with more continuous pixels to accommodate the advanced transponders.

Zero-touch misconnection recovery. The physical connection between transponders and the equipment in the OLS is crucial for establishing a proper network link. However, in a multi-vendor optical backbone, misconnections between these devices occur frequently based on our previous operational experience. For instance, network operators may unintentionally use transponders from different vendors connected to the same filter port in the MUXs at both ends of the fiber, causing issues with wavelength transmission and reception. Rectifying these misconnections necessitates on-site manual operations, leading to additional overhead and potential disruptions in the network. However, FlexWAN introduces the spectrum-sliced OLS where the passband of each filter port in the MUX supports all spectrum frequencies. When there is a misconnection, network operators only need to reconfigure the spectrum of the filter port to provide the passband with the same spectrum as the wavelength generated by the transponder.

Vendor-agnostic optical backbone. Cloud providers always introduce new vendors during their long-term operations. Vendor diversity is essential to prevent monopolies and mitigate concurrent optical failures. However, this diversity introduces challenges as each vendor implements a vendor-specific subsystem with a customized controller that controls their optical devices. As a result, the optical backbone becomes fragmented, making it difficult to manage and configure the entire network efficiently. The workload of configuring the network increases linearly with the number of vendors, adding complexity to the management process. In contrast, FlexWAN is vendor-agnostic because FlexWAN utilizes a standard device model to achieve the centralized control of the optical backbone. If the optical devices from new vendors follow the same device model, these devices can be seamlessly introduced and directly interfaced in FlexWAN. The complexity of FlexWAN will not increase no matter how many vendors are introduced.

System reliability. In a multi-vendor optical backbone, each vendor typically deploys its vendor-customized controller on a single physical server. However, control-plane failures can occur for various reasons, such as server malfunctions, software bugs, or network disconnections, leading to reduced system reliability. The presence of multiple vendors in the network increases the likelihood of such failures. As the number of vendors in the network rises, the probability of encountering at least one vendor failure also increases, potentially compromising the overall stability of the system. FlexWAN uses a centralized controller to interface with network-wide optical devices. The centralized controller is deployed in the cloud with multiple copies to guarantee system reliability.

10 RELATED WORK

Optical network design. Cloud providers have studied the design of optical networks in depth [10, 11, 15, 18, 19, 30, 46]. A recent work by Facebook [10] presented the Hose model that abstracted the aggregated traffic demand per site to save capacity. Microsoft [46] proposed a network design tool, called Shoofly, that optically bypassed network hops to reduce the OEO conversion and saved hardware cost. FlexWAN is orthogonal to previous work. FlexWAN is a novel flexible infrastructure that solves both system-level and algorithmic challenges to provide a cost-effective optical backbone.

Flexible optical technology. Despite prior works have studied the flexible optical technology [8, 12, 21, 22, 29, 43, 44, 48] and shown promising benefits, it was based on laboratory or simulation experiments and far from reaching the stage of large-scale deployment. To the best of our knowledge, FlexWAN is the first work to deploy flexible technology in the large-scale optical backbone. FlexWAN makes two novel contributions. First, FlexWAN presents the detailed design and implementation of deploying the flexible optical backbones (§ 4) by introducing spacing-variable hardware and a centralized controller to coordinate these devices. Second, FlexWAN uses the specifications of transponders in practice to generate the optimal network planning and optical restoration strategies.

Optical layer control. Several works have studied the control interface of optical networks. Cox [23] and Filer et al. [25] expressed a long-term goal of using a centralized SDN controller to control the open optical line system in Microsoft's optical backbone. Miao [7] demonstrated the feasibility of controlling the optical network with a centralized controller to collect optical data. In contrast to previous works, the goal of achieving centralized control of the optical backbone in FlexWAN is to coordinate network-wide optical devices from multiple vendors to overcome spectrum-related issues.

Optical failure recovery. Traffic engineering (TE) [9, 17, 32, 33, 39, 40] has been widely studied to recover the network traffic when links failed. FlexWAN proposes the optical restoration algorithm to revive the lost capacity in the optical layer. Recently proposed restoration-aware TE solution, ARROW [49] is the most relevant work to this paper. In contrast, FlexWAN proposes the design and implementation of a novel flexible WAN infrastructure and demonstrates the failure resilience under this flexible infrastructure. Furthermore, FlexWAN has been running in production for years.

11 CONCLUSION

This paper presents a novel flexible WAN infrastructure called FlexWAN to provision cost-effective WAN capacity. FlexWAN solves system-level challenges by introducing spacing-variable hardware and leveraging a centralized controller to coordinate network-wide optical devices. FlexWAN solves the algorithmic challenges by formulating the problem of provisioning WAN capacity with the goal of minimizing hardware costs. Our experimental results show that FlexWAN can save at least 57% of transponders and reduce 36% of spectrum usage while continuing to meet up to 8× the present-day demands using existing hardware and fiber deployments. FlexWAN further incorporates failure resilience that revives 15% more capacity in the overloaded optical backbone.

ACKNOWLEDGEMENTS

We sincerely thank the anonymous reviewers for their valuable feedback on earlier versions of this paper. We also thank the teams at Tencent for their contributions to the work. Jilong Wang and Xianneng Zou are the corresponding authors. Jilong Wang was supported by the National Key Research and Development Program of China under Grant No. 2020YFE0200500.

REFERENCES

[1] [n. d.]. Gurobi optimizer. http://www.gurobi.com. ([n. d.]).

[2] 2020. Spectral Grids for WDM Applications: DWDM Frequency Grid, document ITU-T G.694.1. Available:https://www.itu.int/rec/T-REC-G.694.1/en. (2020).

[3] 2021. ADVA ROADM. https://www.adva.com/en/products/technology/roadm. (2021).

[4] 2021. Cernet. https://www.edu.cn/cernet_fu_wu/about_cernet/. (2021).

[5] 2021. Spectral grids for WDM applications: DWDM frequency grid. https://www.itu.int/rec/R-REC-P.2001/en. (2021).

[6] 2022. C-band. https://www.fiberlabs.com/glossary/conventional-band/. (2022).

[7] 2022. Detecting Ephemeral Optical Events with OpTel. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*. USENIX Association, Renton, WA. https://www.usenix.org/conference/nsdi22/presentation/miao

[8] Forough Shirin Abkenar and Akbar Ghaffarpour Rahbar. 2017. Study and analysis of routing and spectrum allocation (RSA) and routing, modulation and spectrum allocation (RMSA) algorithms in elastic optical networks (EONs). *Optical Switching and Networking* 23 (2017), 5–39.

[9] Firas Abuzaid, Srikanth Kandula, Behnaz Arzani, Ishai Menache, Matei Zaharia, and Peter Bailis. 2021. Contracting Wide-area Network Topologies to Solve Flow Problems Quickly. In *18th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 21)*. 175–200.

[10] Satyajeet Singh Ahuja, Varun Gupta, Vinayak Dangui, Soshant Bali, Abishek Gopalan, Hao Zhong, Petr Lapukhov, Yiting Xia, and Ying Zhang. 2021. Capacity-efficient and uncertainty-resilient backbone network planning with hose. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*. 547–559.

[11] Ajay Kumar Bangla, Alireza Ghaffarkhah, Ben Preskill, Bikash Koley, Christoph Albrecht, Emilie Danna, Joe Jiang, and Xiaoxue Zhao. 2015. Capacity planning for the Google backbone network. (2015).

[12] Sadananda Behera, Anshuman Deb, Goutam Das, and Biswanath Mukherjee. 2019. Impairment aware routing, bit loading, and spectrum allocation in elastic optical networks. *Journal of Lightwave Technology* 37, 13 (2019), 3009–3020.

[13] Pablo Wilke Berenguer, Markus Nölle, Lutz Molle, Talha Raman, Antonio Napoli, Colja Schubert, and Johannes Karl Fischer. 2015. Nonlinear digital pre-distortion of transmitter components. *Journal of lightwave technology* 34, 8 (2015), 1739–1745.

[14] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. 2017. Julia: A fresh approach to numerical computing. *SIAM review* 59, 1 (2017), 65–98.

[15] Martin Birk, Gagan Choudhury, Bruce Cortez, Alvin Goddard, Narayan Padi, Aswatnarayan Raghuram, Kathy Tse, Simon Tse, Andrew Wallace, and Kang Xi. 2016. Evolving to an SDN-enabled ISP backbone: Key technologies and applications. *IEEE Communications Magazine* 54, 10 (2016), 129–135.

[16] Martin Bjorklund. 2010. *YANG-a data modeling language for the network configuration protocol (NETCONF)*. Technical Report.

[17] Jeremy Bogle, Nikhil Bhatia, Manya Ghobadi, Ishai Menache, Nikolaj Bjørner, Asaf Valadarsky, and Michael Schapira. 2019. TEAVAR: striking the right utilization-availability balance in WAN traffic engineering. In *Proceedings of the ACM Special Interest Group on Data Communication*. 29–43.

[18] Andrew Brzezinski and Eytan Modiano. 2005. Dynamic reconfiguration and routing algorithms for IP-over-WDM networks with stochastic traffic. *Journal of Lightwave Technology* 23, 10 (2005), 3188.

[19] Angela L Chiu, Gagan Choudhury, George Clapp, Robert Doverspike, Mark Feuer, Joel W Gannett, Janet Jackel, Gi Tae Kim, John G Klincewicz, Taek Jin Kwon, et al. 2012. Architectures and protocols for capacity efficient, highly dynamic and highly resilient core networks. *Journal of Optical Communications and Networking* 4, 1 (2012), 1–14.

[20] Junho Cho and Peter J Winzer. 2019. Probabilistic constellation shaping for optical fiber communications. *Journal of Lightwave Technology* 37, 6 (2019), 1590–1607.

[21] K Christodoulopoulos, P Soumplis, and E Varvarigos. 2013. Planning flexible optical networks under physical layer constraints. *Journal of Optical Communications and Networking* 5, 11 (2013), 1296–1312.

[22] Konstantinos Christodoulopoulos, Ioannis Tomkos, and Emmanuel A Varvarigos. 2011. Elastic bandwidth allocation in flexible OFDM-based optical networks. *Journal of Lightwave Technology* 29, 9 (2011), 1354–1366.

[23] Jeffrey Cox. 2015. SDN control of a coherent open line system. In *Optical Fiber Communication Conference*. Optical Society of America, M3H–4.

[24] Rob Enns, Martin Bjorklund, Juergen Schoenwaelder, and Andy Bierman. 2011. *Network configuration protocol (NETCONF)*. Technical Report.

[25] Mark Filer, Jamie Gaudette, Monia Ghobadi, Ratul Mahajan, Tom Issenhuth, Buddy Klinkers, and Jeff Cox. 2016. Elastic optical networking in the Microsoft cloud. *Journal of Optical Communications and Networking* 8, 7 (2016), A45–A54.

[26] Steve Frisken, Glenn Baxter, Dmitri Abakoumov, Hao Zhou, Ian Clarke, and Simon Poole. 2011. Flexible and grid-less wavelength selective switch using LCOS technology. In *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*. IEEE, 1–3.

[27] Monia Ghobadi, Jamie Gaudette, Ratul Mahajan, Amar Phanishayee, Buddy Klinkers, and Daniel Kilper. 2016. Evaluation of elastic modulation gains in microsoft's optical backbone in North America. In *2016 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE, 1–3.

[28] Monia Ghobadi and Ratul Mahajan. 2016. Optical layer failures in a large backbone. In *Proceedings of the 2016 Internet Measurement Conference*. 461–467.

[29] Vasileios Gkamas, Konstantinos Christodoulopoulos, and Emmanouel Varvarigos. 2015. A joint multi-layer planning algorithm for IP over flexible optical networks. *Journal of lightwave technology* 33, 14 (2015), 2965–2977.

[30] Jennifer Gossels, Gagan Choudhury, and Jennifer Rexford. 2019. Robust network design for IP/optical backbones. *Journal of Optical Communications and Networking* 11, 8 (2019), 478–490.

[31] Zonglong He, Jinxiang Song, Kovendhan Vijayan, Christian Häger, Alexandre Graell i Amat, Henk Wymeersch, Peter A Andrekson, Magnus Karlsson, and Jochen Schröder. 2022. Periodicity-enabled size reduction of symbol based predistortion for high-order QAM. *Journal of Lightwave Technology* 40, 18 (2022), 6168–6178.

[32] Chi-Yao Hong, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Vijay Gill, Mohan Nanduri, and Roger Wattenhofer. 2013. Achieving high utilization with software-driven WAN. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*. 15–26.

[33] Chi-Yao Hong, Subhasree Mandal, Mohammad Al-Fares, Min Zhu, Richard Alimi, Chandan Bhagat, Sourabh Jain, Jay Kaimal, Shiyu Liang, Kirill Mendelev, et al. 2018. B4 and after: managing hierarchy, partitioning, and asymmetry for availability and scale in google's software-defined WAN. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. 74–87.

[34] IEEE. [n. d.]. IEEE 802.3ad Link Aggregation. https://www.ieee802.org/3/ad/. ([n. d.]).

[35] Rainer R Iraschko and Wayne D Grover. 2000. A highly efficient path-restoration protocol for management of optical network transport integrity. *IEEE Journal on Selected Areas in Communications* 18, 5 (2000), 779–794.

[36] Xin Jin, Yiran Li, Da Wei, Siming Li, Jie Gao, Lei Xu, Guangzhi Li, Wei Xu, and Jennifer Rexford. 2016. Optimizing bulk transfers with software-defined optical WAN. In *Proceedings of the 2016 ACM SIGCOMM Conference*. 87–100.

[37] Takuya Kanai, Yumiko Senoo, Kota Asaka, Jun Sugawa, Hideaki Tamai, Hiroyuki Saito, Naoki Minato, Atsushi Oguri, Seiya Sumita, Takehiro Sato, et al. 2018. Novel automatic service restoration technique by using self-reconfiguration of network resources for a disaster-struck metro-access network. *Journal of Lightwave Technology* 36, 8 (2018), 1516–1523.

[38] Ginni Khanna, Stefano Calabrò, Bernhard Spinnler, Erik De Man, and Norbert Hanik. 2015. Joint adaptive pre-compensation of transmitter I/Q skew and frequency response for high order modulation formats and high baud rates. In *Optical Fiber Communication Conference*. Optical Society of America, M2G–4.

[39] Praveen Kumar, Yang Yuan, Chris Yu, Nate Foster, Robert Kleinberg, Petr Lapukhov, Chiun Lin Lim, and Robert Soulé. 2018. Semi-oblivious traffic engineering: The road not taken. In *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*. 157–170.

[40] Hongqiang Harry Liu, Srikanth Kandula, Ratul Mahajan, Ming Zhang, and David Gelernter. 2014. Traffic engineering with forward fault correction. In *Proceedings of the 2014 ACM Conference on SIGCOMM*. 527–538.

[41] Antonio Napoli, Pablo Wilke Berenguer, Talha Rahman, Ginni Khanna, Mahdi M Mezghanni, Lennart Gardian, Emilio Riccardi, Anna Chiadò Piat, Stefano Calabrò, Stefanos Dris, et al. 2018. Digital pre-compensation techniques enabling high-capacity bandwidth variable transponders. *Optics Communications* 409 (2018), 52–65.

[42] K Okamoto and H Yamada. 1995. Arrayed-waveguide grating multiplexer with flat spectral response. *Optics letters* 20, 1 (1995), 43–45.

[43] Panos Papanikolaou, Konstantinos Christodoulopoulos, and Emmanouel Varvarigos. 2017. Joint multi-layer survivability techniques for IP-over-elastic-optical-networks. *Journal of optical communications and networking* 9, 1 (2017), A85–A98.

[44] Panos Papanikolaou, Konstantinos Christodoulopoulos, and E Varvarigos. 2018. Optimization techniques for incremental planning of multilayer elastic optical networks. *Journal of Optical Communications and Networking* 10, 3 (2018), 183–194.

[45] Claude E Shannon. 1949. Communication in the presence of noise. *Proceedings of the IRE* 37, 1 (1949), 10–21.

[46] Rachee Singh, Nikolaj Bjorner, Sharon Shoham, Yawei Yin, John Arnold, and Jamie Gaudette. 2021. Cost-effective capacity provisioning in wide area networks with Shoofly. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*. 534–546.

[47] Rachee Singh, Manya Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. 2018. RADWAN: rate adaptive wide area network. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. 547–560.

[48] Luis Velasco, Miroslaw Klinkowski, Marc Ruiz, and Jaume Comellas. 2012. Modeling the routing and spectrum allocation problem for flexgrid optical networks. *Photonic Netw. Commun.* 24, 3 (2012), 177–186.

[49] Zhizhen Zhong, Manya Ghobadi, Alaa Khaddaj, Jonathan Leach, Yiting Xia, and Ying Zhang. 2021. ARROW: restoration-aware traffic engineering. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*. 560–579.

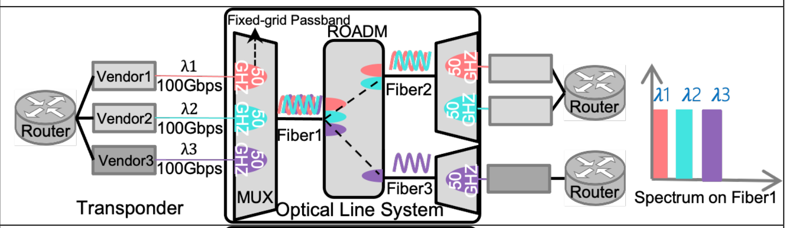
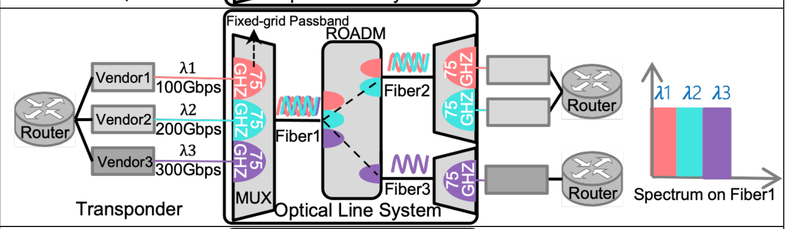
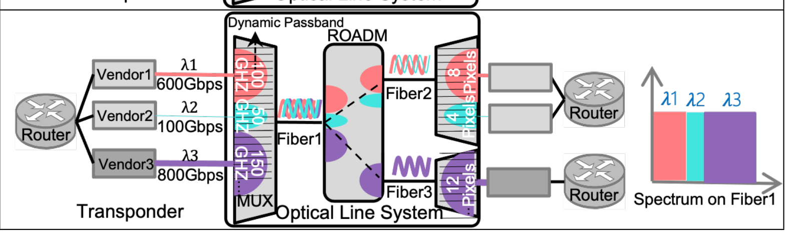
Approach	**Transponder**		**Passband in OLS**	**Illustration**
	Data rate	Channel spacing	Spectrum width	
Fixed-rate WAN (100G-WAN)	Fixed	Fixed	Fix-grid	
Rate-adaptive WAN (RADWAN)	Variable	Fixed	Fix-grid	
FlexWAN	Variable	Variable	Dynamic	

Table 1: Optical backbone infrastructure comparison with prior work.

A APPENDIX

Appendices are supporting material that has not been peer-reviewed.

A.1 Optical Backbone Infrastructure Comparison

Cloud providers mainly develop two main optical backbone infrastructures to provide long-haul WAN capacity: (*i*) rate-fixed WAN; (*ii*) rate-adaptive WAN. Table 1 illustrates the key differences between these infrastructures and FlexWAN.

In the fixed-rate WAN (first row in Table 1), all transponders only support a single modulation format (e.g., QPSK) that generated wavelengths support a fixed data rate occupying a fixed data rate. For example, each wavelength occupies a 50 GHz spectrum width modulated at 100 Gbps bandwidth capacity. The equipment in the OLS is based on a fixed grid that provides the passband with the same spectrum width of 50 GHz.

In the rate-adaptive WAN (second row in Table 1), the wavelength can carry multiple data rates according to the traveling distances required. As the transponders evolves to occupy channel spacing of 75 GHz [49], we adapt the BVT in [47] into our scene: 300 Gbps, 200 Gbps, and 100 Gbps by modulating the signals in 8QAM, QPSK, and BPSK for distances up to 1100 km, 2000 km, and 5000 km. Despite there being variable data rate options, all the optical hardware devices run on the same rigid spectrum grid of 75 GHz.

FlexWAN is a flexible optical backbone infrastructure (third row in Table 1). FlexWAN introduces spacing-variable hardware physically. The spacing-variable transponder (SVT) has adjustable components built in to generate wavelengths that are optimally modulated for required data rate and channel spacing based on the transmission distance. The spectrum-sliced OLS dynamically changes the passband width to adapt to the channel spacing of the wavelength. This approach allows each wavelength to efficiently utilize the limited spectrum resources to carry the WAN capacity.

A.2 Spacing-variable transponder (SVT))

The specifications of the spacing variable transponder (SVT) proposed by FlexWAN are shown in Table 2. The values are measured by partially collaborating with vendors using a production-level testbed experiment (§ 6).

	100 Gbps	200 Gbps	300 Gbps	400 Gbps	500 Gbps	600 Gbps	700 Gbps	800 Gbps
Channel Spacing	Optical Reach	Optical Reach	Optical Reach	Optical Reach	Optical Reach	Optical Reach	Optical Reach	Optical Reach
50 GHz	3000	1000	/	/	/	/	/	/
62.5 GHz	/	1500	/	/	/	/	/	/
75 GHz	5000	2000	1100	600	/	/	/	/
87.5 GHz	/	/	1500	1000	600	300	/	/
100 GHz	/	/	2000	1500	900	400	200	/
112.5 GHz	/	/	/	1600	1100	500	300	150
125 GHz	/	/	/	1700	1200	600	350	200
137.5 GHz	/	/	/	1800	1300	700	450	250
150 GHz	/	/	/	1900	1400	800	500	300

Table 2: Data rates and optical reaches (km) of signals of our proposed SVT. Symbol / means that this format is not recommended.