RADWAN: Rate Adaptive Wide Area Network

Rachee Singh
rachee@cs.umass.edu
University of Massachusetts, Amherst

Monia Ghobadi
mgh@microsoft.com
Microsoft Research

Klaus-Tycho Foerster
ktfoerster@cs.aau.dk
Aalborg University

Mark Filer
mafiler@microsoft.com
Microsoft

Phillipa Gill
phillipa@cs.umass.edu
University of Massachusetts, Amherst

ABSTRACT

Fiber optic cables connecting data centers are an expensive resource acquired by large organizations with significant monetary investment. Their importance has driven a conservative deployment approach with redundancy and reliability baked in at multiple layers. In this work, we take a more aggressive approach and argue for adapting the capacity of fiber optic links based on their signal-to-noise ratio (SNR). We investigate this idea by analyzing the SNR of over 2,000 links in an optical backbone for a period of 3 years. We show that the capacity of 64% of IP links can be augmented by at least 75 Gbps, leading to an overall capacity gain of over 134 Tbps. Moreover, adapting link capacity to a lower rate can prevent 25% of link failures. This means using the same links, we get higher capacity and better availability.

We propose RADWAN, a traffic engineering system that allows optical links to adapt their rate based on the observed SNR to achieve higher throughput and availability while minimizing the churn during capacity reconfigurations. We evaluate RADWAN using a testbed consisting of 1,540 km fiber with 16 amplifiers and attenuators. We then simulate the throughput gains of RADWAN at scale compared to the state-of-the-art. Our results show that with realistic traffic matrices and conservative churn control, RADWAN improves the overall network throughput by 40%. The service provider we study has invested in this idea and is rolling out the necessary infrastructure to deploy the first capacity variable link between Canada and Europe this year.

1 INTRODUCTION

Optical backbones are million dollar assets, with fiber comprising their most expensive component. Companies like Google, Microsoft and Facebook purchase or lease fiber to support wide-area connectivity between distant data center locations but have not been able to fully leverage this investment due to conservative provisioning of the optical network. We show that wide area fiber links exhibit significantly better signal quality (measured by the signal-to-noise-ratio or SNR) than the minimum required to support transmission at 100 Gbps, leaving money on the table in terms of link capacities.

In other words, there is potential to operate fiber links at higher capacity, thereby increasing the throughput of existing optical networks. We analyze historical SNR from 2,000 optical channels in a backbone network and find that the capacity of 64% links can be augmented by 75 Gbps or more, leading to a capacity gain of over 134 Tbps in the network. Doing this simplistically, operators can statically raise link capacities to a higher value (*e.g.,* 150 Gbps or 200 Gbps) supported by the link's SNR. However, aggressively using the fiber in such a manner increases the rate of link failures as operating near the SNR makes links susceptible to failures when signal quality fluctuates.

On the other hand, degradation in the SNR of a link below a threshold is treated as a *link failure* and the link is effectively "down." We show that this is wasteful, as at least 25% of current failures can be mitigated by reducing the rate of transmission from 100 Gbps to 50 Gbps.

At the core of these issues is a fundamental orthodoxy in the operation of wired networks: a fiber link is either up with a fixed capacity or down, largely oblivious to changes in the quality of the underlying optical signal. With this school of thought, operators are forced to account for large margins between the actual SNR and the operating capacity to avoid frequent link failures. In contrast, wireless networks employ a variety of schemes to adapt the transmission rate in response to changing signal quality [3, 23, 27]. However, the adoption of adaptive bit-rate schemes has been slow in wireless communication because of practical challenges. Specifically, adapting transmission rates to the wireless channel quality is hard as the quality can vary at time-scales shorter than a single packet transmission time [27]. In addition, obtaining accurate measurements of received signal strength indication (RSSI) of wireless media (a proxy for true SNR) is hard in practice [14] because of issues like mis-calibration and packet corruption.

We argue that optical links are well positioned to be rate-adaptive. First, signal quality varies at a much coarser time granularity in fiber than in wireless media (hours as

opposed to milliseconds). This stability can be leveraged in wide-area networks to amortize the cost of infrequently shifting between multiple discrete modulation schemes as signal quality changes. Second, unlike wireless signals, optical signal quality is easily inferred from bit-error rate (BER) reported post forward error correction (FEC). Leveraging these benefits, we present RADWAN (Rate Adaptive WAN), a system that adapts channel bit-rates in WANs to improve the overall throughout and availability of the network.

RADWAN consists of a centralized rate-adaptive WAN controller that gathers SNR from all fiber channels in the network for adjusting the modulation format of the channels to achieve higher or lower data rates. In traditional wide area setting, QPSK modulation format supports data rates of 100 Gbps for distances upto 3,000 km, 8QAM allows 150 Gbps for distances up to 2,100 km, and 16QAM allows 200 Gbps for distances up to 800 km (see §7 for a discussion on distance). By switching links to a lower modulation format (BPSK with data rates of 50 Gbps), RADWAN allows critical WAN links to function at lower data rates instead of failing altogether. We refer to these variable capacity links in RADWAN as *dynamic capacity links*. By building on top of existing software based WAN controllers [15], RADWAN allows existing traffic engineering schemes to exploit dynamic capacity to improve network throughput. We make two key contributions to make rate adaptive WANs practical:

Optimal WAN traffic engineering. A major challenge associated with dynamically adapting link capacities in WANs is the latency incurred by network hardware to change a link's modulation format. To reconcile the latency of capacity changes and benefits of adapting link capacities in WANs, the RADWAN controller re-formulates the centralized traffic engineering optimization problem to avoid unnecessary capacity reconfiguration (Section 4). We evaluate the RADWAN controller by comparing the throughput gains of employing RADWAN at scale over the state-of-the-art controller. Our results show that in a real-world network topology with realistic traffic demands and conservative traffic churn settings, RADWAN improves the overall network throughput by 40% (Section 6).

Avoiding high latency of modulation reconfiguration. We build a testbed emulating a four datacenter WAN connected via 1,540 km of fiber. Using this testbed, we confirm the viability of modulation reconfiguration to achieve greater network throughput. We benchmark the behavior of the RADWAN controller as it reacts to SNR degradation by switching to a lower modulation format. During the modulation change, the line-rate traffic on the affected link is migrated to a backup path until the modulation change can complete (Section 5). Our experiments show that reconfiguring modulation formats on commodity hardware incurs a latency of 68 seconds, on average. We develop a prototype that demonstrates the feasibility of decreasing this reconfiguration time by a factor of 1,000 (Section 7.1).

RADWAN opens the door to revisit several classic networking problems in light of dynamic capacity links and opens new lines of research. For instance, are there graph abstractions that capture networks with dynamic capacity links? How do classical networking algorithms (such as maximum-flow problem [10]) change in the presence of variable link capacities? Are there smart capacity planning, failure-recovery, load-balancing, or on-demand bandwidth allocation algorithms that can benefit from rate adaptive links? RADWAN provides a foundation for beginning to think about these problems. In addition to new avenues of research, RADWAN has also been successful at convincing network operators about the potential for dynamic capacity links with the first operational dynamic capacity link between Canada and Europe set to roll out in 2018.

2 QUANTIFYING THE OPPORTUNITY

We first investigate the signal quality in a large optical backbone, comprised of 2,000 optical channels. Our dataset consists of the average, minimum, and maximum SNR per channel, aggregated over 15 minute intervals for 3 years. We characterize the SNR of these channels and quantify its variations. In contrast to wireless networks [27], where signal quality may vary on short time intervals and where estimating SNR is complicated by signal interference, signals in fiber optical media do not suffer from these challenges.

2.1 Characterizing SNR in our WAN

First, we note that the average SNR can support much more than the configured link rate of 100 Gbps. In Figure 1(a) we plot the distribution of the average SNR of all channels in the network and observe that over 64% of the channels have an SNR that can support data rates of 175 Gbps.This represents a significant opportunity to improve the throughput of optical links by operating closer to the actual SNR of the signal.

But what about stability? While the average may be well above what is needed to run the links at 100 Gbps, adjusting link rates requires that SNR be stable to avoid frequent capacity fluctuations. Figure 1(b) shows the SNR over time of 40 channels on the same WAN fiber cable. We observe that the SNR of these channels is largely stable. Occasional dips, suggest impairments in fiber or other optical hardware. Despite these occasional dips, the SNR of all channels remains well above 6.5 dB which is required to carry 100 Gbps of traffic at fixed capacity, highlighting over-provisioning by network

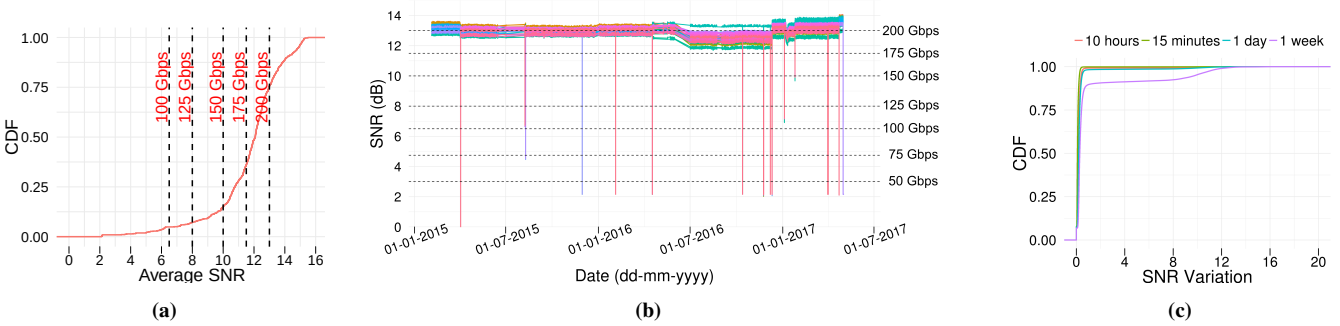


Figure 1: (a) Distribution of the average SNR of over 2,000 channels in a backbone network for 3 years. Note that the SNR of optical channels is much higher than the required SNR for 100 Gbps bit rate (6.5 dB). (b) SNR variations in 40 optical channels (i.e., IP links) on a wide area fiber cable. Dotted lines represent the feasible link capacity for a particular SNR. (c) Variations in the channel SNR in intervals of different durations. Observe that most links do not observe significant variation in SNR for several hours.

operators in an attempt to prevent link failures. Other fibers in the network show similar behavior.

We further consider variability of SNR across all links on different time-scales. For each time interval of size 15 minutes, 10 hours, 1 day and 1 week, we calculate the variability of SNR (the difference between maximum SNR and minimum SNR) for all optical channels in the backbone network. Figure 1(c) shows the distribution of SNR variation in time intervals of different sizes. We confirm that SNR remains stable over several hours at a time. A small fraction ($\leq 5\%$) of links observe variation of over 1 dB in the 10 hour interval. Moreover, while our SNR measurements are aggregated over 15 minute intervals, we argue that our conclusions are sound, as Figure 1(c) shows the variations in SNR in fiber over 15 minutes are negligible. This is in contrast with wireless media where significant SNR changes can happen within a few milliseconds.

Why do we need variable bandwidth links? Based on our observation of stable but over-provisioned SNR of links, one might be tempted to operate links closer to the actual SNR by simply making a one-time decision to increase the transmission rate of all links. However, we find that the frequency of link failures increases if we cannot dynamically adapt to changes of SNR. This is because infrequent but sizeable variations in SNR occur in fiber links. While the SNR of a small fraction of links changes significantly in a few hours, 10% of all links undergo 2 dB of change in SNR within a week (Figure 1(c)). For example, we select a fiber where each link on this fiber (i.e., optical channel) has a high enough SNR to make all capacity denominations feasible over the period of 3 years. We then analyze the number of failures that links would undergo if they were modulated with higher but static capacities. Figure 2(a) shows that links on this fiber do not see significant increase in the number of failures as the capacity is increased up to 175 Gbps. But some of the links would observe a 100 failures if driven at 200 Gbps. This behavior is repeated with other fibers but depending on the number of links, fiber length, technology, and age of equipment, the point at which the failures start to increase is different for each fiber and each channel on the fiber. Hence, it is impossible to select a one-size-fits-all static capacity that is higher than 100 Gbps.

Furthermore, we characterize the duration of link failures. Figure 2(b) plots the duration of link failures for the different modulated capacities (based on the link's average SNR). We observe that such failure events last several hours which is unacceptable. This impact on reliability underscores the need to adapt capacity in the face of SNR drops, if we want to tighten the gap between required and observed SNR values.

2.2 Complete loss-of-light is uncommon

Today, when the SNR of a link's optical signal drops below its pre-determined threshold, the link is declared down. However, we show that not all failures are complete loss-of-light. SNR drops may be caused by planned maintenance work (e.g., a line card replacement) or unplanned events (e.g., fiber cut, hardware failure, human error). While some of these impairments make the link unusable (e.g. fiber cuts), others may simply lower the signal quality (e.g. degradation of an amplifier) without completely shutting off the signal. Links undergoing failures due to lowered signal quality can still be used to send traffic at a reduced rate, highlighting another opportunity to improve link availability.

We quantify this opportunity by manually analyzing seven months of failure tickets (250 events) reported by WAN field operators, to understand their root causes. We identify five types of root cause: (i) **Power failure:** This indicates a power line failure causing a complete loss-of-light; (ii) **Fiber cuts:**

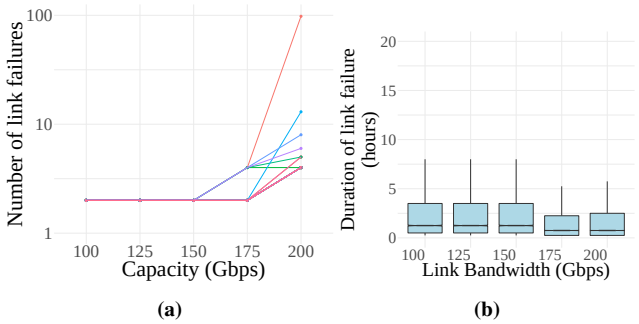


Figure 2: (a) Number of link failures for 40 links (one color per link) for a given capacity. For this particular fiber, while increasing capacity up to 175 Gbps does not increase link failure events, achieving 200 Gbps capacity comes at the cost of increased link failures. (b) Duration of failures if WAN links operate at a given capacity.

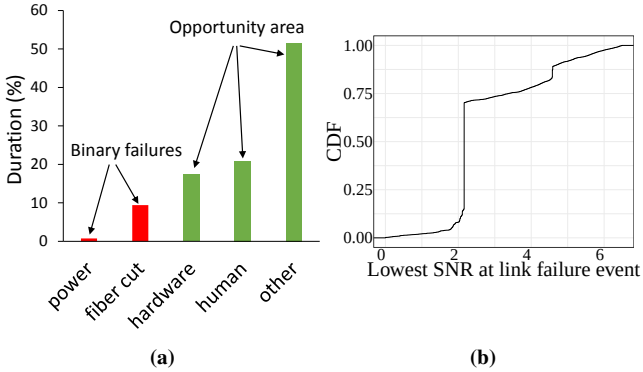


Figure 3: (a) Categorization of failure root causes in terms of duration. Contrary to common belief, fiber cuts are not the major root cause of failures in WANs. Unplanned events during planned maintenance or hardware failures are more probable. (b) Distribution of the lowest SNR values when a link failure event happened. The lowest SNR is above 3.0dB (which is sufficient to drive a link at 50 Gbps) 25% of the time.

This is an accidental break in an optical fiber (mostly due to construction projects) that leads to a complete loss-of-light; (*iii*) **Hardware failure:** This refers to a malfunction of optical hardware, such as degrading amplifiers, transponders, or optical cross connects; (*iv*) **Human error:** In this category, failures are mostly due to humans inadvertently disturbing an operational fiber while working on another; (*v*) **Other:** This category refers to all other cases with an unidentified root cause.

We analyze the contribution of each root cause to total outage duration (Figure 3(a)). Fiber cuts and power failures make up 10% of the total outage durations, highlighting that complete outage of fibers is relatively rare. Collateral damage by human operators represents 20% of the total outage time. Similar observations have been made in a previous study [13]. The remaining categories are hardware failure and other undocumented failures. Failures with undocumented causes occur when technicians do not log the exact action taken, but we know they were not instances of fiber cuts.

Even current failures could present an opportunity to run links at reduced rates. To narrow down the opportunity area, we record the lowest SNR of failure events (when the SNR falls below the 100 Gbps threshold which is 6.5 dB). Figure 3(b) shows the distribution of SNR values at link failures. A minimum SNR of above 3.0dB, which is enough to drive a link at 50 Gbps capacity, is observed in 25% of the failures. Therefore, even with conservative 100 Gbps capacity, 25% of link failures could have been avoided by driving the impacted links at 50 Gbps, highlighting the improvement in availability offered by dynamic capacity links.

3 DYNAMIC CAPACITY LINKS

Our characterization of optical links suggests that links are currently operating well below their potential transmission rates, based on SNR values. However, operating links at constant transmission rates closer to the observed SNR increases the likelihood of link failures. To balance this trade-off, we propose dynamic adjustment of physical link capacities in *centrally controlled wide area networks* by changing the *modulation format* of optical signals. These choices are motivated by the latest hardware and software developments in the industry:

Adapting bit-rates by modulation change. Recent advances in the development of bandwidth variable transceivers (BVTs) enable our goal of reclaiming lost network throughput due to over-provisioning and improving availability by decreasing transmission rates in the face of low SNR (vs. incurring a link failure). State-of-the-art BVTs are capable of modulating signals on the fiber with three different formats: 16QAM, 8QAM and QPSK. All other factors being constant, signals in 16QAM format can carry traffic at 200 Gbps, 8QAM can carry 150 Gbps and QPSK can carry 100 Gbps. However, these transceivers were designed with the assumption that operators would make a one-time choice of modulation format.

This is reflected in the latency incurred in changing the modulation of ports on modern Arista 7504 switches. In our experiments, we find that on average, changing the modulation of a port incurs a latency of over 1 minute. During this time, the link undergoing the modulation change is down and cannot carry traffic. This is because of the assumption that the modulation change is a one-time event. To benchmark the reconfiguration latency, we experiment with a transceiver evaluation board and investigate ways of reducing capacity reconfiguration time (Section 7.1). However, we note that it will take some engineering effort to make hitless capacity

change production ready. We discuss intermediate routing states to handle link flaps with low disruption in Section 7.1.

Software Driven WANs. Effective utilization of network infrastructure in modern WANs is enabled by software-driven centralized traffic engineering (TE) [15, 16, 20] that maximizes the network flow for changing demand matrices. Therefore, we consider the implementation of dynamic capacity links in such networks. TE controllers are consumers of network link capacities, as they make decisions about routing flows along the best paths with available capacity. Had link capacity reconfiguration been a hitless phenomenon, existing TE controllers could function largely unmodified with dynamic capacity links. However, capacity reconfiguration is expensive, as it causes a link outage lasting for over a minute. We discuss the impact of this additional constraint on TE controllers in the next section.

4 TRAFFIC ENGINEERING WITH DYNAMIC CAPACITY LINKS

In a network with dynamic capacity links, the state of the network in each run of the TE optimization algorithm is dependent on the links' underlying SNR. Therefore, TE controllers must be modified to gather the SNR of all links in the network and treat the link capacities as variables. We propose the RADWAN centralized TE controller which can leverage dynamic capacity links to achieve higher network throughput and availability. RADWAN handles a spike in the demand matrix by upgrading the capacities of one or more links. However, state-of-the-art bandwidth variable transceivers (BVTs) require over a minute to change the capacity of a link (Section 7.1), rendering the link unusable for that period. In response to this link flap, existing traffic flows must be migrated away from the link undergoing capacity reconfiguration. Such flow migrations can cause transient congestion in the network and, hence, must be done minimally.

Therefore, we argue that in a network composed of dynamic capacity links, the objective of traffic engineering changes from simply maximizing the network throughput to maximizing throughput *while minimizing churn caused by link capacity reconfigurations*. In Section 4.1 we discuss how network churn can be quantified to achieve low disruption while meeting traffic demands with link capacity reconfiguration.

4.1 Quantifying network churn

Present day hardware does not support hitless capacity changes; therefore, we propose to deal with churn induced by link capacity changes in software. As a first step towards this, we introduce a definition of churn induced by a link capacity change in terms of the rate of traffic on the link. The capacity change (either an increase to meet demands or a decrease due to lowered signal quality) of link l carrying f_l units of traffic will displace the f_l units. Displacement of large flows is more likely to cause transient congestion as opposed to smaller flows. Therefore, we define churn induced by the capacity change of link l as:

$$churn(l) = f_l \tag{1}$$

The overall churn induced by capacity changes in a network, C, is a summation of the churn from each link undergoing capacity change:

$$C = \sum_{\text{links}} churn(l) \tag{2}$$

We note that this is one of the many possible ways to define the churn caused by link flaps in the network. We encourage practitioners to consider other definitions which deter inducing churn into preferred traffic classes (e.g., interactive traffic over background traffic).

4.2 Computing flow allocations

When computing allocations of flows along different paths in a network composed of dynamic capacity links, the goal of RADWAN is to maximize the network utilization (as was the case with earlier work [15, 16, 20]) while keeping churn due to capacity reconfigurations minimal. In this section, we formulate this goal as a constrained optimization problem using the definition of churn from Section 4.1. RADWAN periodically evaluates the optimization goal to assign traffic flows along network paths. In each round of its operation, RADWAN has access to attributes of the network state which serve as input to the optimization problem. We now describe various elements of the RADWAN controller.

Inputs. Traditional TE controllers take as input the *network topology* and *traffic demand matrix* to compute allocations of flows along label switched network paths. In addition to these, our controller requires SNR measurements for all physical links in the network. Using this information, the controller derives the *potential* capacity of each link, over its existing capacity.[1] Implicitly, the controller is also aware of the existing flow on all links in the network, assigned in the previous round of controller operation.

Allocation Objective. Algorithm 1 describes the optimization goal of RADWAN. At its core, the optimization is a modified multi-commodity flow that maximizes overall throughput of the network while augmenting link capacities minimally. The optimization variables $b_{i,j}$ specify the allocation of flow i along path j in the network. Allocation of

[1] Even if there is potential to increase a link's capacity by, say, 50 Gbps, the controller must do an upgrade only if this extra capacity is needed to meet traffic demands.

Algorithm 1: Traffic Engineering Optimization

Inputs:

d_i: flow demands for source destination pair i

c_l: capacity of each link l

p_l: potential capacity increase of each link l

$I_{j,l}$: 1 if tunnel j uses link l and 0 otherwise

f_l: existing flow on link l ($f_l \leq c_l$)

T_i: set of tunnels that are set up for flow i

Outputs:

$b_i = \sum_j b_{i,j}$: b_i is allocation to flow i

$b_{i,j}$ is allocation to flow i along tunnel j

Maximize: $\sum_i b_i$ - $\epsilon(\sum_l \text{churn}(l))$

subject to:

$\forall i, 0 \leq b_i \leq d_i$

$\forall i, j, b_{i,j} \geq 0$

$\forall l, \sum_{i,j} I(j,l) b_{i,j} \leq c_l + p_l$

$\forall i, \sum_{j \in T_i} b_{i,j} \geq b_i$

$$\text{churn}(l) = \begin{cases} 0, & \sum_{i,j} b_{i,j} I(j,l) \leq c_l \\ f_l, & \text{else} \end{cases}$$

flow along link l in the network is constrained by the sum of the link capacity (c_l) and the potential increase in capacity (p_l) depending on the link's SNR. ϵ is a small positive constant that denotes the relative importance of the two aspects of the objective function: maximizing throughput and minimizing churn. Finally, in a given round, network churn caused by the capacity change of link l is 0 if the optimal flow assigned to the link is less than or equal to the link's capacity (c_l). However, if the link has more flow assigned to it than its current capacity, it induces network churn equal to the amount of traffic on it (f_l), as assigned in the previous round of flow allocation. The nature of network churn makes the objective function of the optimization piece-wise linear.

Approximation to Linear Program. To efficiently solve the optimization objective described in Algorithm 1, we approximate the definition of churn as:

$$\text{churn}(l) = \max(0, (\sum_{i,j} b_{i,j} I(j,l) - f_l)) \quad (3)$$

This monotonically increasing value of churn, depending on the flow assignments $b_{i,j}$, is different from the actual churn value which is essentially a step function; however, this reasonable approximation allows us to convert Algorithm 1 to an efficiently solvable linear program.

Managing Churn. For the duration of a link flap, no traffic can be routed along this link. As the impacted links will be offline for just one minute, the affected traffic (churn) has to be managed efficiently for low disruption. We thus compute a single intermediate flow allocation, where the churn is distributed along routes without link flaps. We show in Section 6 that a single intermediate step suffices, as the number of link flaps per reconfiguration is low in practice (see Figure 9). Methods for networks with highly unstable SNR are described in Section 7.1. Once hitless capacity reconfiguration is production ready, the intermediate flow allocation step described in this paragraph can be omitted.

4.3 Controller Implementation

We implement RADWAN, the traffic engineering controller based on the goals outlined in the previous subsection. The controller implements Algorithm 1 using the popular optimization library CVXPY [5] in Python 2.7.

RADWAN computes flow allocations for the input demand matrix in each round of its operation. Before solving the optimization, RADWAN uses the link-level SNR information to determine: (*i*) Links for which the total capacity must now be reduced since the new SNR is too low to support the existing capacity. These *capacity downgrades* must be performed even though they will cause the impacted links to be down for roughly a minute. (*ii*) The potential capacity of other links, above their current capacity, depending on the SNR of the link. For instance a link could currently be operating at 100 Gbps but if it has an SNR of 10.2 dB, it has a potential capacity increase of 50 Gbps as its capacity can be augmented to 150 Gbps.

In what follows, we present the results of an extensive testbed evaluation of RADWAN (§ 5), benchmarking the effect of optically changing links' capacities on the IP layer. Next, we simulate RADWAN and compare it with our implementation of the SWAN controller as described in [15]. We perform a data-driven evaluation of the behavior and performance of these two controllers in § 6 and show the gains of capacity variable links on the overall network throughput.

5 TESTBED EVALUATION

In this section, we build a testbed consisting of 1,540 km of fiber and 16 optical amplifiers to evaluate the feasibility of deploying RADWAN in a moderate sized WAN. Our goal is to highlight the impact of modulation changes on realistic traffic flows. In addition, we provide insights to both researchers and practitioners into the state-of-the-art hardware components required to realize a rate-adaptive wide area network.

5.1 Testbed Implementation Details

We build a moderate sized testbed which emulates a 4-datacenter WAN as shown in Figure 4(a) to evaluate RADWAN. Each datacenter consists of a router connected to

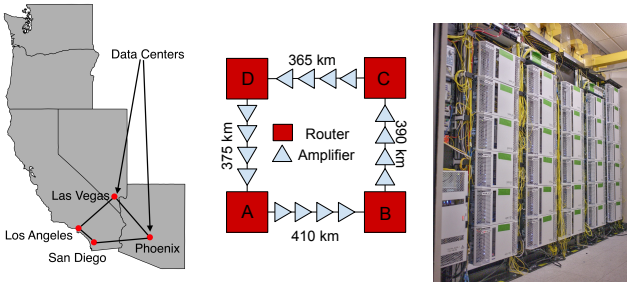


Figure 4: (a) Geographic scale of the testbed built to demonstrate the operation of RADWAN**. Our testbed emulates a WAN connecting four major cities on the west coast of the United States. (b) Logical view of the testbed where four routers (logically split from a modular chassis switch) emulate four datacenters. These routers are connected via hundreds of kilometers of optical fiber and regularly spaced amplifiers. (c) Photograph of the electrical and optical equipment in our testbed.**

its neighbors using hundreds of kilometers of optical fiber. To prevent signal deterioration, we connect Erbium Doped Fiber Amplifiers (EDFAs) at approximately every 65-120 kilometers of fiber length. For simplicity, Figure 4(b) represents the logical view of the WAN.

Note that we had access to only one Arista 7504 modular chassis; therefore, we used Virtual Routing and Forwarding (VRF) [6] to logically split the same physical switch into four routers (named *A*, *B*, *C* and *D* in Figure 4(b)). Each VRF has a separate routing table and routing protocol instances. By configuring relevant physical interfaces to be in separate VRFs and connecting the interfaces via optical components (fiber, amplifiers), we achieved a logical topology whereby traffic sent between ports on the switch was sent out on the wire. We verified bi-direction connectivity between each pair of nodes *A*, *B*, *C* and *D*. The Arista 7504 has integrated bandwidth variable transceivers manufactured by Acacia Inc. (the BVT module, AC 400, is described in detail in Section 7.1). These allow us to configure three modulation formats (QPSK, 8QAM and 16QAM) on the switch ports. The complete testbed, including optical and electrical equipment is shown in Figure 4(c).

The part of the RADWAN controller responsible for configuring the switch was implemented using Arista's PyEAPI [2] framework. With this, we could programmatically configure the modulation formats of different ports, program routes and query status of our commands.

To generate line rate traffic flows in the topology, we used a Spirent traffic generator [25]. With the help of the Spirent device we programmed 400 Gbps of TCP traffic flows to test the dynamic capacity links of the testbed.

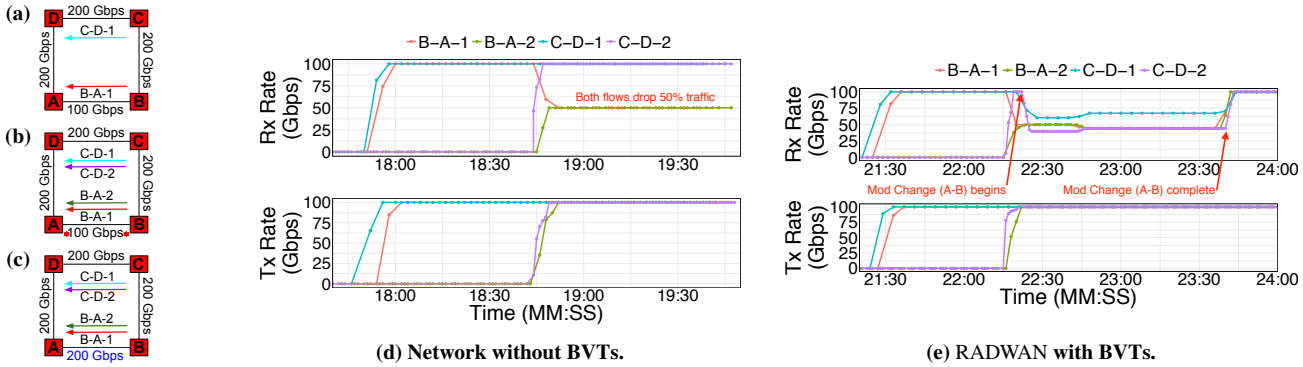
5.2 Benchmarking the WAN testbed

Reacting to SNR degradation. Optical signals in fiber can become attenuated because of ill-functioning amplifiers, disturbances caused during maintenance windows or even ambient temperature conditions. We now demonstrate that RADWAN reacts to signal attenuation by switching to a lower order modulation format that can be supported by the degraded SNR. In the laboratory setting, we use a Variable Optical Attenuator (VOA) device to add configurable amounts of noise (measured in dB) so that we can demonstrate signal attenuation. We connect the VOA between routers *A* and *B* in the test topology. On the underlying switch, this connection is implemented by connecting *Ethernet*4/1/1 to *Ethernet*3/1/1 with 410km of optical fiber. The Ethernet ports are in separate VRFs (not directly connected), so we setup static routing such that traffic sent from one to the other is sent over the fiber connection. Every five seconds, we increase the noise from the VOA by 1 dBm.

We measure the SNR of the signal on each end of the connection and observe that the SNR of the received signal on *Ethernet*3/1/1 steadily deteriorates as the level of noise increases (Figure 5). Once the added noise reaches 16 dBm, the transceiver can no longer recover from the increased errors,[2] and the port goes down. At this point, the controller reduces the modulation format of the port from 16QAM to 8QAM. The modulation change takes approximately 70 seconds to complete. We then resume incrementing the noise level using the VOA. When the noise level reaches 18 dBm, the transceiver can no longer recover from the errors to support 8QAM format and the port goes down again. Our controller reacts by reducing the modulation format yet again, this time from 8QAM to QPSK. After roughly 70 seconds of down time, the ports come back up with QPSK modulation format. The addition of noise of 23dBm or more renders the link unusable, even in the lowest supported modulation format. At this point, the link has failed and such a failure is irrecoverable with the current set of hardware.

Modulation Change Latency. In the previous benchmarking experiment, we changed the modulation format of a link in the testbed in response to SNR degradation. We observe that each change in modulation format changes the status of the ports involved to *down*, making them unavailable for sending and receiving traffic. In Figure 5, we observe that modulation change operations take approximately 70 seconds. This aspect of the latency of modulation change guides the design of the RADWAN controller (Section 4).

[2] Acacia BVTs have 15% soft decision FEC enabled by default.



(d) Network without BVTs. **(e)** RADWAN **with BVTs.**

Figure 6: (a) describes the network and link capacities. At the start, all links except link $A-B$ are in 16QAM modulation format, capable of carrying 200 Gbps. $A-B$ being in QPSK format can carry 100 Gbps. In the beginning, there are two flows in the network, each of 100 Gbps from $B \rightarrow A$ and $C \rightarrow D$. As additional demand of 100 Gbps gets added ($B-A-2$) and ($C-D-2$) described in (b), the link $A-B$ gets congested, leading to 50% traffic drops in flows $B-A-1$, $B-A-2$ in the absence of RADWAN, as seen in the Rx Rate in Figure(d). However, in a RADWAN deployment, the controller reacts to the increased demand by increasing the capacity of $A-B$ link to 200 Gbps (seen in (c) by changing the modulation format to 16QAM. While this causes temporary disruption due to rerouting of $B-A$ flows along the $C-D$ link, once the modulation change is complete, the network can carry the flows of 400Gbps without any drops, as seen in the Rx Rate of (e).

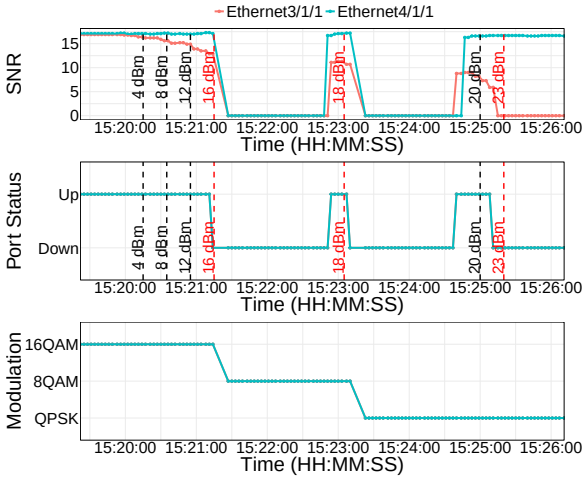


Figure 5: Impact of attenuation on link SNR, port status and modulation format as the amount of signal attenuation increases (shown with dotted vertical lines).

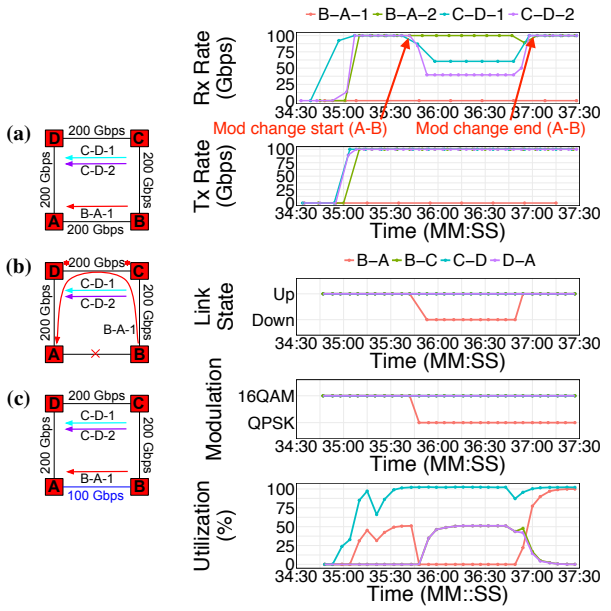
5.3 Evaluating Modulation Change

We demonstrate the capability of RADWAN to react to SNR degradation by reducing the modulation format of ports, allowing links with reduced signal quality to function at lower rates. In this section we provide an end-to-end evaluation of RADWAN as it attempts to meet changing demand matrices by upgrading the capacities of links in the WAN. Additionally, we show that RADWAN migrates flows from a link undergoing capacity up-/downgrade (due to improved/poor SNR) to alternate paths until the modulation change is complete.

In each of the following experiments, we show the transmission rate (Tx Rate) of the traffic we attempt to send between nodes in the topology. An overwhelmed node responds to high traffic volume by dropping a portion of the flows. We capture the net traffic received by the sink node of a flow as the receive rate (Rx rate). In the ideal case, the Tx and Rx rates should match, implying that all the traffic sent by the source is reaching the sink node.

Link capacity upgrade. Figure 6(a) shows the starting state of a network where there are two flows of 100 Gbps each, one from Node B to Node A (flow $B-A-1$) and the other from Node C to Node D (flow $C-D-1$). With the introduction of two additional 100 Gbps flows ($B-A-2$) and ($C-D-2$) as shown in Figure 6(b), the network becomes congested, because link $A-B$ can only carry 100 Gbps of traffic. As seen in the Rx rate in Figure 6(d), both $B-A-1$ and $B-A-2$ share the $A-B$ link fairly and drop 50% of their traffic. However, RADWAN can salvage this congestion by increasing the capacity of the $A-B$ link (as seen in Figure 6(c)). To do this, the RADWAN controller reacts to the increased demand by changing the modulation of the $A-B$ link, causing it to be down for roughly one minute. This temporarily congests the $C-D$ link (observe the Rx rate of all flows drops in Figure 6(e)), due to rerouting of the $B-A$ flows. However, once the modulation change is complete, all flows can be transmitted successfully with no packet drops. We note that without augmenting the capacity of link $A-B$, the network could not satisfy 400 Gbps of demand but dynamic capacity links with RADWAN enable us to meet the increased demand.

Link capacity downgrade. Figure 7(a) shows the starting state of our testbed where the network is carrying three flows of 100 Gbps, two from Node C to D ($C{-}D{-}1$, $C{-}D{-}1$) and one from Node B to A ($B{-}A{-}2$). All links in the network can carry 200 Gbps of traffic. Observe that the Rx rate in Figure 7(d) matches the Tx rate, implying there is no packet loss. Now, we attenuate the signal between Node A and B using a VOA device such that the switch ports can no longer sustain transmission at 200 Gbps. Therefore, the link goes down (Figure 7(b)), causing ($B{-}A{-}2$) to get routed over the longer path $B{\rightarrow}C{\rightarrow}D{\rightarrow}A$ which is configured as the backup route. This transition of the $B{-}A{-}2$ flow along the longer path is visible in the utilization of links in the network (Figure 7(d)). Links $B{-}C$ and $D{-}A$ now are carrying 100 Gbps of the $B{-}A{-}2$ flow (therefore are 50% utilized). Note that this leads to congestion on link $C{-}D$ which can only carry 200 Gbps of traffic and thus drops 100 Gbps of traffic from $C{-}D$ flows. The RADWAN controller can mitigate this congestion by reducing the modulation format of $A{-}B$ link to QPSK from 16QAM. It takes roughly one minute for the modulation change to take effect as observed in the *down* status of link $A{-}B$ in Figure 7(d). Once the modulation change is complete, link $A{-}B$ is back up and carries the $B{-}A{-}2$ flow without any congestion in the network (Tx/Rx rates match again). The new network state is shown in Figure 7(c). Therefore, our experiments show that RADWAN can react to traffic demands and signal quality by adapting the capacity of links in the WAN.



(d) Reaction of RADWAN to poor SNR on $B{-}A$

Figure 7: (a) describes the starting network state and link capacities. At the start, all links are in 16QAM modulation format, capable of carrying 200 Gbps. There are three flows in the network, each of 100 Gbps. One from $B{\rightarrow}A$ and two from $C{\rightarrow}D$. Due to signal attenuation, the link $A{-}B$ fails as seen in (b), causing the $B{-}A{-}2$ flow to get routed over the longer path $B{\rightarrow}C{\rightarrow}D{\rightarrow}A$. Observe that the utiilization of links $B{-}C$ and $D{-}A$ increases in (d). This causes link $C{-}D$ to get congested and it drops parts of the $C{-}D{-}1$, $C{-}D{-}2$ flows (Tx rate falls below the Rx rate in (d)). RADWAN reacts to this situation by reducing the modulation format of link $A{-}B$ to QPSK as is allowed by the lowered SNR of the link (see (c)). Once the modulation change completes, all flows are routed along direct paths without any packet loss as confirmed by the Tx/Rx rates and link utlizations in (d).

6 LARGE SCALE EVALUATION

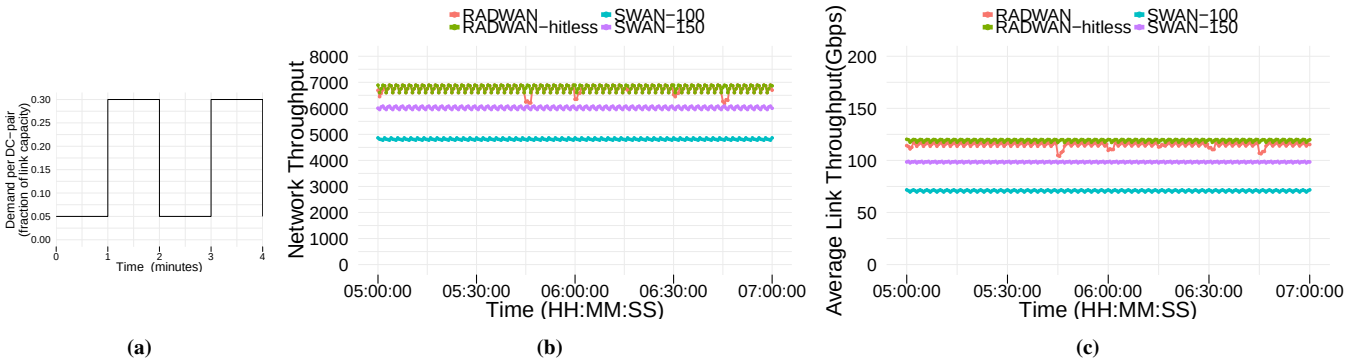
In section §2, we used 3 years of SNR measurements to demonstrate that an overall *capacity gain* of 67% is possible by augmenting the capacity of links from 100 Gbps to 125, 150, 175, or 200 Gbps depending on their average SNR. This represents the upper bound of the *throughput gain* achievable using RADWAN. The actual network throughput depends not only on the network state (link capacities, tunnels *etc.*) but also the traffic demand and acceptable churn (defined in §4). In this section, we simulate the operation of RADWAN in a large backbone network with realistic traffic demands to compute the network throughput achieved by RADWAN. We compare the throughput and availability of the network under RADWAN and state-of-the-art SWAN controller.

Both controllers are aware of the underlying signal quality of links. But unlike SWAN, RADWAN uses the SNR to update link capacities, choosing amongst discrete choices of 50, 100, 125, 150, 175 and 200 Gbps. As outlined in the previous section, RADWAN only upgrades the capacity of a link to meet increased traffic demand that cannot be met otherwise. Capacity downgrades are done to prevent link failures such that the lower quality link can continue to function at a reduced rate.

6.1 Simulation Setup

We consider the network topology of a large commercial WAN and gather SNR measurements from the optical fiber connecting the nodes in the topology for four randomly chosen days in the years 2016 and 2017. Both RADWAN and SWAN compute flow allocations along various network paths to meet an elastic demand between each pair of nodes in the network.

Since WANs operate links at 100 Gbps currently, we consider the performance of SWAN in a fixed capacity network where each link operates at 100 Gbps if the SNR is above the threshold of 100 Gbps modulation and otherwise the link is down. We refer to this scheme as SWAN-100 in the analysis. However, operators can be more aggressive with respect to



(a) (b) (c)

Figure 8: (a) Traffic demand pattern between each pair of nodes in the network similar to prior work [15]. (b) Optimal network flow achieved by different traffic engineering schemes. For better visibility, (b) zooms into two hours of the simulation period. RADWAN **achieves 40% higher network throughput compared to the state-of-the-art mechanism, SWAN-100** (RADWAN **and** RADWAN**-hitless are overlapping curves on top of the graph). We also compare SWAN's performance with fixed capacity links operating with static 150 Gbps higher modulation format. While SWAN-150 provides an improvement over SWAN-100,** RADWAN **achieves 12% higher throughput than SWAN-150. (c) Average per link throughput. We observe that** RADWAN **achieves 68% higher per link throughput compared to SWAN-100.**

utilizing the network by operating links at fixed but higher capacity of 150 Gbps. We refer to SWAN operating in such a network as SWAN-150. SWAN-150 is used to compare the benefit of using rate adaptive schemes like RADWAN over a network with higher but fixed link capacities. While hardware limitations prevent hitless capacity changes presently, we simulate the performance of RADWAN under both hitless (RADWAN-HITLESS) and non-hitless (RADWAN) link capacity change behavior.

The traffic demand between each node pair varies periodically every two minutes (demand pattern shown in Figure 8(a)). Our choice of network demands is similar to previous work [15] since rapid changes in demand matrices stress test the TE controllers. We also offset the traffic demand between each pair of nodes by a randomized value to ensure that at any given point of time, there is sufficient variety of demands in the network.

Simulation Parameters. Unless otherwise stated, the control loop of both controllers is executed every 30 seconds as stated in [15]. In addition, we assume that the demand between each pair of nodes can be split across $k = 2$ shortest paths between the nodes. For RADWAN, we set the churn trade-off parameter ϵ (defined in §4) to a conservative value of 0.001. We perform several runs of this experiment with each run lasting for one day. We find that across four randomly chosen days, our results are similar. Hence, for the sake of brevity, the figures show results from one experimental run.

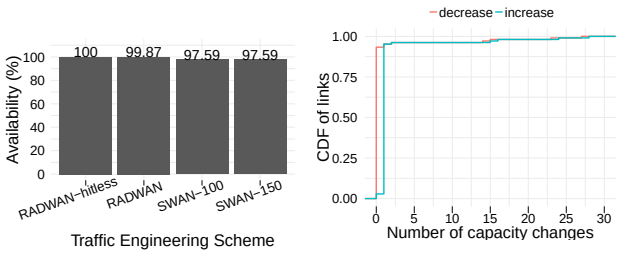
6.2 Evaluation Metrics

We focus on the following three key aspects of effective, cost-efficient network design to evaluate RADWAN.

Network Throughput. First, we compute the optimal network flow that RADWAN can achieve in each run of the controller and compare it with the optimal flow that SWAN achieves for the same network conditions. This provides the network throughput enabled by both controllers for each run of their control loops for the duration of a day. Figure 8(b) shows the network flow for both RADWAN and SWAN for two hours of a day (zooming into two consecutive hours, picked randomly for the sake of better visibility in the figure). We observe that RADWAN manages to push 40% more traffic than SWAN-100 in the same network. The same observation holds consistently with other hours and days we simulated. We note that while RADWAN

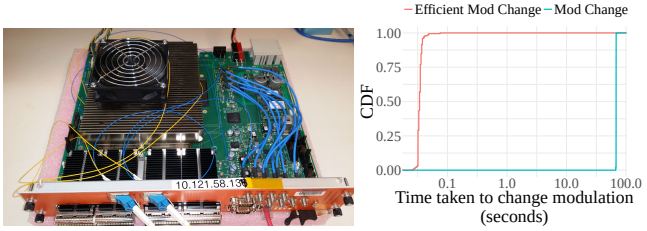
Link Throughput. Next, we compare RADWAN and SWAN's per link throughput. For each run of the TE control loop, we compute the total traffic carried by each link and average it over all links in the network. Figure 8(c) shows the distribution of average link throughput over time (again zoomed over two hours for better visibility). We find that, on average, RADWAN increases the utilization of network links by 68% compared to SWAN-100, getting more utility from each link in the network.

Link Availability. We compute the availability of links under each TE scheme as the fraction of total simulation time for which the link was up for carrying traffic. Figure 9(a) compares the average availability of links in the simulation period. As expected, RADWAN-HITLESS allows links to be 100% available since it instantly adapts the link capacity to the lower or higher SNR. RADWAN has slightly lower link availability since capacity changes are not hitless and cause the links undergoing capacity changes to be unavailable for



(a) (b)

Figure 9: (a) Comparison of the average availability of all links in the network as they operate under different TE schemes. (b) Distribution of the number of capacity reconfigurations occurring per link in the network. We note that only 6% of the links changed their capacities more than once in the simulation period.



(a) Evaluation board (b)

Figure 10: (a) AC400 BVT to analyze modulation change latency. (b) CDF of the time taken to change modulation (capacity) of a fiber link using the BVT. Link capacity changes take 68 seconds, on average. But we demonstrate ways to change the modulation efficiently, such that it takes only 35 milliseconds, on average.

roughly one minute. However, we note that RADWAN provides better availability over the state-of-the-art SWAN-100 scheme. This is because RADWAN adapts links to lower capacities in place of failing them when the signal quality degrades. Even though capacity reconfigurations done by RADWAN are not hitless, we note that the link availability under RADWAN does not suffer much since very few links undergo rapid changes in capacity (as seen in Figure 9b).

7 DISCUSSION

In this section, we provide insights on future directions of hitless capacity change as well as a discussion on cost and the impact of capacity change on distances the light can travel.

7.1 Hitless Capacity Change

BVTs and dependency graphs. Dependency graphs [19, 21] are one of the seminal techniques used for consistent network updates [9]. By specifying an old and a new network state, each individual routing change is performed only when safe to do so. However, when a link e is to be modulated, with a flow f utilizing e before and after the link flap, dependency graphs perform poorly: no alternative path is specified for f.

RADWAN manages link flaps by computing an intermediate routing state for flows during reconfiguration. As such, RADWAN specifies a two-step dependency graph: in order for a scheduled link flap to be activated, the affected traffic is rerouted beforehand. [3] Due to the benevolent nature of SNR in our 3 year data set of 2,000 links, coupled with the churn minimization of Algorithm 1, RADWAN jointly activates all link flaps. In more volatile SNR scenarios, RADWAN can be set to activate link flaps over multiple dependent iterations.

We conjecture that such intermediate consistency methods can eventually be phased out once hitless capacity changes become production ready, as discussed in the next section.

Towards hitless capacity change. BVTs are not yet optimized to handle the latency of a modulation change. State-of-the-art BVTs can only change the link modulation after bringing the module to a lower power state. This translates to a link flap for higher layer protocols. The duration of such link failures is a challenge in the deployment of dynamic capacity links in production networks. To quantify this, we obtain an evaluation board to the Acacia AC400 bandwidth variable transceiver [1].. This module is essentially the same as the one integrated in the switch linecard used as part of our testbed in Section 5. Since the evaluation board exposes an API to program the transceiver, we use it for understanding the modulation change procedure. We change the link's modulation 200 times from QPSK to 16QAM and analyze the time taken.

Figure 10b shows the AC400 bandwidth variable transceiver module. We observe that the average downtime of the link undergoing capacity change is 68 seconds, similar to the observation made in Section 5. We investigate the cause of latency in capacity reconfiguration and find that majority of this time is associated with turning the laser back on after reprogramming the transceiver module. We plot the distribution of time taken to change modulations without turning off the laser and find that it only takes approximately 35 ms on average. This suggests an opportunity to strive towards hitless capacity changes in the fiber.

7.2 Cost and Distance

One of the key benefits of deploying bandwidth variable links is their cost savings. While we are sadly unable to report the exact discounted retail price that we have been quoted, our

[3] OWAN [18] also deals with consistent cross-layer reconfiguration in WANs. However, OWAN is designed for Reconfigurable Optical Add-Drop Multiplexers, where wavelengths are exclusively *either* activated or deactivated: link flaps due to BVTs are not considered.

device manufacturers reveal that the cost of BVTs is on par with the cost of 100 Gbps static transceiver. This comparable cost is in fact one of the motivating factors for operators to start deploying BVTs even if they are being programmed a few times in their life time.

A disadvantage of using higher order modulations is that they limit the distance that the light can travel. As mentioned in §1 current QPSK modulation format supports data rates of 100 Gbps for distances upto 3,000 km, 8QAM allows 150 Gbps for distances up to 2,100 km, and 16QAM allows 200 Gbps for distances up to 800 km. We analyzed the fiber distances in our WAN and found that 50% of our fiber paths are less than 800 km (thus capable of supporting 16-QAM), the remaining 45% are between 800 to 2,100 km (thus capable of supporting 8-QAM), and only 5% of paths are longer than 2,100 km. While in our analysis, we did not take the distances into account, we believe it will not dramatically impact our results. Moreover, device manufacturers are working towards improving the distances for higher order modulation.

8 RELATED WORK

Our work builds on several lines of related research as categorized below.

Optical and IP layer orchestration. Singh *et al.* [24] recently analyzed the SNR of links in a large North American backbone over a period of 2.5 years and proposed adapting link capacities to the SNR optical channels. We extend their study period to 3 years, also broadening their initial measurement and testbed quantifications. We also propose a centralized TE controller system RADWAN and evaluate the interaction between dynamic capacity links and IP layer flows with simulations at scale and in a realistic testbed. The study by Jin *et al.* [18] on cross-layer optimization between IP and optical layers wavelengths is similar in spirit to our motivation of bridging the gap between optical and IP layers. In their work, Jin *et al.* show reconfiguration of wavelengths provides latency gains for deadline driven bulk transfers, also providing a competitive analysis of scheduling single-hop transfers in [17]. But their work still keeps the capacity of each wavelength static. In contrast, our work focuses on the reconfiguration of the *capacity* of wavelengths, without the migration of wavelengths across links. We also provide measurements from an operational backbone and argue for changing link capacities with a focus on throughput and reliability. An interesting future direction is to study the throughput and latency gains of the combination of the two proposals: a fully programmable WAN topology were both capacities and placement of wavelengths on fiber is informed by the centralized TE.

WAN measurements. Govindan *et al.* [13] study 100 failure events across two WANs and data center networks, offering insights into the challenges of maintaining high levels of availability for content providers. Although they do not isolate optical layer failures, they report on root causes of failures, including optical transmitters. We complement their work by focusing on optical layer failures. Ghobadi *et al.* study Q-factor data from Microsoft's optical backbone [11, 12] and provide insights on the data. Our work complements their analysis on several fronts. First, we take a deep dive into the impact of temporal changes of SNR on link capacities in terms of capacity gain, availability gains, and realistic throughput gains. Second, we propose and build the system infrastructure required to achieve capacity variable links and benchmark the throughput gains using realistic IP level data. Third, we build a comprehensive testbed and evaluate the impact of capacity reconfiguration, as well as amplifiers on the path. Our work closes the loop for enabling capacity variable links. Similarly, Filer *et al.* [7] studied the deployed optical infrastructure of Microsoft's backbone; they discuss the benefits of optical elasticity, express a long-term goal of unifying the optical control plane with routers under a single Software Defined Network controller and recognize YANG [4] and SNMP as potential starting points for a standard data model and control interface between the optical layer and the WAN traffic controller. In this work, we explore how programmability in the optical layer can bring throughput gains and present a cross-layer WAN traffic controller for dynamic capacity links. Marian *et al.* [22] focused on IP and TCP layer measurements such as packet loss and packet inter-arrival times on fiber optics spans. In contrast, we capture failures in the optical layer using failure tickets.

Hardware feasibility studies. Yoshida *et al.* [28, 29] studied the use of 12.5 GHz spectrum slices for allocating bandwidth variable connections to improve the spectrum usage. Although their works did not consider real-time adjustment of the capacity, they provided the foundation for the feasibility of building the necessary hardware with variable bandwidth capabilities that is the enabler of our work. We use real world measurements and build the system that fills the gap between optical and IP layers. Fischer *et al.* [8] and Teipen *et al.* [26] efforts towards commercializing higher-speed optical transmission have demonstrated the need for advanced modulation formats, several of which require similar transceiver hardware architecture. Their work showed that adaptive transceivers can be built to support a number of possible operational configurations but without real-time reconfiguration mechanism. In this paper, we discuss advantages of reconfigurable capacities in real-time based on live SNR measurements.

9 CONCLUSION

In this work, we quantify the throughput and reliability benefits of rate adaptive wide area networks. Our analysis of the SNR of over 2,000 links in an optical backbone for a period of 3 years shows that the capacity of 64% of IP links can be increased by ≥75Gbps, yielding an overall throughput gain of 134 Tbps. Furthermore, 25% of link failures can be avoided by reducing the transmission rate to 50 Gbps from 100 Gbps. To leverage these benefits, we present RADWAN, a traffic engineering system that dynamically adapts link rates to enhance network throughput and availability. We evaluate RADWAN in a testbed with 1,540 km optical fiber and also simulate throughput and availability gains at scale. In comparison to state-of-the-art software WANs, RADWAN has 40% higher network throughput in a realistic network with elastic traffic demands. We also address the challenge of the hardware delay in modifying a link's capacity. We analyze the cause of this delay in current optical transceivers and propose a potential solution to reduce this delay from over a minute to a few milliseconds.

REFERENCES

[1] Acacia Communications. 2015. Acacia Bandwidth Variable Transceiver Module. http://ir.acacia-inc.com/phoenix.zhtml?c=254242&p=irol-newsArticle&ID=2103147. (March 2015).

[2] Arista Networks. 2017. Python client for Arista eAPI. https://github.com/arista-eosplus/pyeapi. (Dec. 2017).

[3] John C. Bicket. 2005. *Bit-rate Selection in Wireless Networks*. Master's thesis. Massachusetts Institute of Technology.

[4] Martin Bjorklund. 2010. YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF). RFC 6020. (Oct. 2010).

[5] Steven Diamond and Stephen P. Boyd. 2016. CVXPY: A Python-Embedded Modeling Language for Convex Optimization. *Journal of Machine Learning Research* 17 (2016), 83:1–83:5.

[6] E. Rosen, Y. Rekhter. 2006. BGP/MPLS IP Virtual Private Networks (VPNs). RFC 4364. (Feb. 2006).

[7] 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 (July 2016), A45–A54.

[8] J. K. Fischer, S. Alreesh, R. Elschner, F. Frey, M. Nölle, C. Schmidt-Langhorst, and C. Schubert. 2014. Bandwidth-Variable Transceivers based on Four-Dimensional Modulation Formats. *Journal of Lightwave Technology* 32, 16 (Aug 2014), 2886–2895.

[9] Klaus-Tycho Foerster, Stefan Schmid, and Stefano Vissicchio. 2016. Survey of Consistent Network Updates. *CoRR* abs/1609.02305 (Sept. 2016).

[10] Saul I. Gass and Arjang A. Assad. 2006. *An Annotated Timeline of Operations Research: An Informal History*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

[11] 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 *Optical Fiber Communication Conference*. Optical Society of America, M2J.2.

[12] Monia Ghobadi and Ratul Mahajan. 2016. Optical Layer Failures in a Large Backbone. In *Internet Measurement Conference*. ACM.

[13] Ramesh Govindan, Ina Minei, Mahesh Kallahalla, Bikash Koley, and Amin Vahdat. 2016. Evolve or Die: High-Availability Design Principles Drawn from Googles Network Infrastructure. In *SIGCOMM Conference*. ACM.

[14] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2010. Predictable 802.11 packet delivery from wireless channel measurements. *SIGCOMM Comput. Commun. Rev.* 41, 4 (Aug. 2010), 12.

[15] Chi-Yao Hong, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Vijay Gill, Mohan Nanduri, and Roger Wattenhofer. 2013. Achieving High Utilization with Software-driven WAN. *SIGCOMM Comput. Commun. Rev.* 43, 4 (Aug. 2013), 15–26.

[16] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hölzle, Stephen Stuart, and Amin Vahdat. 2013. B4: Experience with a Globally-deployed Software Defined Wan. *SIGCOMM Comput. Commun. Rev.* 43, 4 (Aug. 2013), 3–14.

[17] Su Jia, Xin Jin, Golnaz Ghasemiesfeh, Jiaxin Ding, and Jie Gao. 2017. Competitive analysis for online scheduling in software-defined optical WAN. In *INFOCOM*. IEEE.

[18] 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 *SIGCOMM Conference*. ACM.

[19] Xin Jin, Hongqiang Harry Liu, Rohan Gandhi, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Jennifer Rexford, and Roger Wattenhofer. 2014. Dynamic Scheduling of Network Updates. *SIGCOMM Comput. Commun. Rev.* 44, 4 (Aug. 2014), 539–550.

[20] Hongqiang Harry Liu, Srikanth Kandula, Ratul Mahajan, Ming Zhang, and David Gelernter. 2014. Traffic Engineering with Forward Fault Correction. *SIGCOMM Comput. Commun. Rev.* 44, 4 (Aug. 2014), 527–538.

[21] Ratul Mahajan and Roger Wattenhofer. 2013. On consistent updates in software defined networks. In *HotNets*. ACM.

[22] T. Marian, D.A. Freedman, K. Birman, and H. Weatherspoon. 2010. Empirical characterization of uncongested optical lambda networks and 10GbE commodity endpoints, In DSN. *DSN*.

[23] Andrew Mcgregor and Derek Smithies. 2010. Rate Adaptation for 802.11 Wireless Networks: Minstrel. http://blog.cerowrt.org/papers/minstrel-sigcomm-final.pdf.

[24] Rachee Singh, Monia Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. 2017. Run, Walk, Crawl: Towards Dynamic Link Capacities. In *HotNets*. ACM.

[25] Spirent Communications. 2018. Spirent TestCenter. https://www.spirent.com/Products/TestCenter. (Jan. 2018).

[26] Brian Thomas Teipen, Michael Eiselt, Klaus Grobe, and Jörg-Peter Elbers. 2012. Adaptive Data Rates for Flexible Transceivers in Optical Networks. 7 (05 2012).

[27] Mythili Vutukuru, Hari Balakrishnan, and Kyle Jamieson. 2009. Cross-layer Wireless Bit Rate Adaptation. *SIGCOMM Comput. Commun. Rev.* 39, 4 (Aug. 2009), 3–14.

[28] Y. Yoshida, A. Maruta, K. i. Kitayama, M. Nishihara, T. Tanaka, T. Takahara, J. C. Rasmussen, N. Yoshikane, T. Tsuritani, I. Morita, S. Yan, Y. Shu, Y. Yan, R. Nejabati, G. Zervas, D. Simeonidou, R. Vilalta, R. Muñoz, R. Casellas, R. Martinez, A. Aguado, V. Lopez, and J. Marhuenda. 2015. SDN-Based Network Orchestration of Variable-Capacity Optical Packet Switching Network Over Programmable Flexi-Grid Elastic Optical Path Network. *Journal of Lightwave Technology* 33, 3 (Feb 2015), 609–617.

[29] Y. Yoshida, A. Maruta, K. Kitayama, M. Nishihara, T. Tanaka, T. Takahara, J. C. Rasmussen, N. Yoshikane, T. Tsuritani, I. Morita, S. Yan, Y. Shu, M. Channegowda, Y. Yan, B. R. Rofoee, E. Hugues-Salas, G.

Saridis, G. Zervas, R. Nejabati, D. Simeonidou, R. Vilalta, R. Muñoz, R. Casellas, R. Martinez, M. Svaluto, J. M. Fabrega, A. Aguado, V. Lopez, J. Marhuenda, O. G. de Dios, and J. P. Fernandez-Palacios. 2014. First international SDN-based network orchestration of variable-capacity OPS over programmable flexi-grid EON. In *OFC 2014*. 1–3.