2024 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS) | 979-8-3503-9172-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ANTS63515.2024.10898264

Increasing Information-Carrying Capacity by Exploiting Diverse Traffic Characteristics in Multi-Band Optical Networks

Ramanuja Kalkunte
University of California, Davis, USA
rkalkunte@ucdavis.edu

Forough Shirin Abkenar
University of California, Davis, USA
fshirina@ucdavis.edu

Sifat Ferdousi
University of California, Davis, USA
sferdousi@ucdavis.edu

Rana Kumar Jana
Indraprastha Institute of Information Technology, Delhi, India
ranaj@iiitd.ac.in

Anand Srivastava
Netaji Subhas University of Technology Delhi, India
anand@nsut.ac.in

Abhijit Mitra
Indraprastha Institute of Information Technology, Delhi, India
abhijit@iiitd.ac.in

Massimo Tornatore
Politecnico di Milano, Italy
massimo.tornatore@polimi.it

Biswanath Mukherjee
University of California, Davis, USA
bmukherjee@ucdavis.edu

***Abstract*—Efficient network management in optical backbone networks is crucial for handling continuous traffic growth. In this work, we address the challenges of managing dynamic traffic in C- and C+L-band optical backbone networks while exploring application flexibility, namely the compressibility and delayability metrics. We propose a strategy, named Delay-Aware and Compression-Aware (DACA) provisioning algorithm, which reduces blocking probability, thereby increasing information-carrying capacity of the network compared to baseline strategies.**

***Index Terms*—Optical networks, multi-band, dynamic traffic, compressibility, delayability, blocking probability, information-carrying capacity.**

I. Introduction

The growth of heterogeneous and bandwidth-hungry 5G/6G applications [1] calls for new efficient on-demand bandwidth-provisioning strategies in backbone networks. To support these diverse applications, optical backbone networks, which typically support quasi-static traffic, may soon need to evolve to adapt to increasingly dynamic traffic patterns [2]. The emergence of Elastic Optical Networks (EONs) has allowed operators to maximize utilization of the available C-band spectrum in Single-Mode Fibers (SMFs) [3]. However, to overcome the limited capacity of C band, optical backbone networks are migrating toward Multi-band (MB) transmission [4], starting with the L band due to matured technology. As such, operators are already expanding their infrastructure to support MB transmission for ever-growing service requirements [5].

Examining the characteristics of novel dynamic network services, two key flexibilities can be observed, *Compressibility* and *Delayability*. Delayability means that some services can be deferred in time, while compressibility means that the amount of bandwidth required by a connection can be reduced while still offering an acceptable level of service. Both delayability and compressibility can be leveraged by service providers and operators to accommodate more traffic at a given time. While some prior research has studied delayability [6], others have focused on compressibility metrics [7]. In this work, we differentiate traffic based on their application flexibility (see Section II-A) and highlight the *joint impact* of both *delayability and compressibility*. We propose a novel Delay-Aware and Compression-Aware (DACA) provisioning algorithm, which exploits delayability and compressibility characteristics of dynamic traffic to enhance the information-carrying capacity of the network. This enhancement is demonstrated by showing a reduction in Blocking Probability (BP).

II. System Model

A. Diverse Traffic Types

In this work, we categorize traffic into three major types based on its delayability and compressibility metrics:

- *Type 1*: Non-Delayable and Non-Compressible, e.g., Google search. Statistics show that there are about 99,000 global search queries every second [8].
- *Type 2*: Compressible, e.g., video streaming.
  - *2a*: Compressible and Delayable, e.g., on-demand videos provided by YouTube, Netflix, etc. Statistics reveal that 40% of YouTube users in the UK access the platform daily [9].
  - *2b*: Compressible and Non-Delayable, e.g., live stream on platforms such as Zoom and Microsoft Teams. Statistics show that about 27% of Internet users watch live streaming content on a weekly basis [10].
- *Type 3*: Delayable and Non-Compressible, e.g., data backup.
  - *3a*: User backup, e.g., Google Drive and iCloud.
  - *3b*: Enterprise backup, e.g., AWS and Microsoft Azure. On average, medium-sized companies benefit from backing up data every 24 hours [11].

B. Network Model

We consider an elastic optical backbone network topology, $G(V,E)$, comprising $|V|$ nodes and $|E|$ links, where V and E represent set of nodes and set of links, respectively. Our study evaluates two scenarios: network operating in only C band and in C+L bands. Each band is composed of 133 channels with a frequency spacing of 37.5 GHz [12]. We consider that the incoming traffic is dynamic in nature and remains in the network for a given period of time. Considering various traffic types discussed in Section II-A, there is a total of R requests, where each request $r \in R$ is described using a tuple $(s, d, \gamma, \overline{\gamma}, q, \phi, \delta, \tau)$. Here, s and d denote source and destination nodes, respectively, and they are generated using a gravity model consisting of traffic generation probability of each node (see Fig. 1). Parameter γ represents the required data rate and $\overline{\gamma}$ is the minimum acceptable data rate defined by the Service Level Agreement (SLA) between the users and the service providers. Also, q represents the traffic type (refer to Section II-A), ϕ is the scaling factor to compress the data rate (γ), δ is the maximum tolerable delay, and τ indicates the holding time of the request. Note that parameters are adjusted for different traffic types; e.g., as shown in Table I, type 2b traffic has $\phi = 50\%$ and δ marked as Not Applicable (N/A), indicating that it is compressible up to 50% but non-delayable.

In this study, each request is assumed to occupy one lightpath; and, to provision these lightpaths, their SLA must be satisfied. SLA compliance is ensured by providing adequate Quality of Transmission (QoT), which is often quantified in optical networks using Generalized Signal-to-Noise Ratio (GSNR). The data rate and modulation format of each lightpath is obtained using a GSNR window, as specified in [12]. To calculate GSNR based on the current network state, we employ two Machine Learning (ML)-based QoT estimators: one for C-band-only and one for C+L bands. The physical-layer model and the ML models for QoT estimation are adopted from [12]. Note that dynamic traffic can cause frequent changes in the network state resulting in variations in the GSNR of the lightpaths. This may lead to SLA violations causing some connections to be dropped. This study, however, focuses solely on the number of blocked connections (BP).

III. Delay- and Compression-Aware Provisioning

Delayability and Compressibility have significant impact on traffic provisioning. We propose an algorithm, namely Delay-Aware and Compression-Aware (DACA), to jointly exploit these characteristics with the objective to reduce BP, and hence increase the information-carrying capacity of the network. This strategy utilizes the delayability and compressibility of certain incoming traffic to postpone their provisioning time and/or compress their bandwidth to accommodate them efficiently when spectral resources, i.e., Frequency Slots (FSs), are not available. Initially, DACA checks if it can provision a request r in its uncompressed form; if the request cannot be provisioned due to unavailable slots, it is delayed. When the maximum tolerable delay (δ) of r has been exhausted and no slots are available, DACA compresses the request. Compression lowers the required data rate γ by a factor of ϕ, thereby reducing the number of FSs required to provision the request. If compression does not help, the request is blocked.

Algorithm 1 outlines DACA, which takes network topology (G) and a set of requests (R) as input. The number of blocked requests (N_b) is initialized to zero. We assume that, for a 24-hour period, peak hours of operation is between p_s and p_e, and the rest of the period is considered off-peak. GSNR of all active lightpaths is re-estimated every t^p time unit during peak hours and every t^o time unit during off-peak hours. After re-estimation, lightpaths for which the data rates (γ) are within their SLA ($\overline{\gamma}$) are retained in the network. Routing and Spectrum Allocation (RSA) of the requests employ k-Shortest Path and First-Fit (FF) strategies, respectively. For each time unit t and each incoming traffic $r \in R$, RSA identifies the candidate path(s) and the FSs. Each request is then checked and decisions are made based on the data rate (Γ_{FS}) of the available slot(s). If Γ_{FS} is greater than the required data rate (γ), the path and FSs are assigned to the request. If not, request r, based on its traffic type q, is either delayed by 1 time unit or compressed to $\phi * \gamma$. Otherwise, r is blocked. Once a request completes its service, i.e., its holding time (τ), the lightpath is released. All type 3b traffic, for which [arrival time + delayability factor (δ)] falls within $(p_s, p\prime_e)$, where $p\prime_e = p_e + m$, will be postponed until $p\prime_e$ with $\delta = 0$.

Algorithm 1 DACA Provisioning Algorithm

Input: $G(V, E)$, R;
Output: N_b;
1: **Initialize:** $N_b = 0$, p_s, p_e, $p\prime_e$, t^p, t^o;
2: **for** each time t **do**
3:   **if** $p_s < t < p_e$ **then**, re-estimate GSNR every t^p time unit
4:   **else**
5:     Re-estimate GSNR every t^o time unit
6:   **for** incoming request $r \in R$ **do**
7:     Perform corresponding RSA;
8:     **if** slot(s) available and $\Gamma_{FS} \geq \gamma$ **then**
9:       Assign slot(s) to request r;
10:     **else**
11:       **if** $q == 2a$ **then**
12:         **if** $\delta > 0$ **then** delay provisioning; $\delta -= 1$;
13:         **else** $\gamma = \phi * \gamma$; perform lines 7-9;
14:       **else if** $q == 2b$ **then**
15:         $\gamma = \phi * \gamma$; perform lines 7-9;
16:       **else if** $q == 3a$ **then**
17:         **if** $\delta > 0$ **then** delay provisioning; $\delta -= 1$;
18:         **else** perform lines 7-9;
19:       **else if** $q == 3b$ **then**
20:         **if** $p_s < (t + \delta) < p\prime_e$ **then** delay r till $p\prime_e$; $\delta = 0$;
21:         **else if** $\delta > 0$ **then** delay provisioning; $\delta -= 1$;
22:         **else** perform lines 7-9;
23:       Block r; $N_b += 1$;
24: **Return** N_b

IV. Numerical Evaluation

A. Modeling and Simulation Setup

An event-driven, custom-built Python simulator is used to model (i.e., emulate) a dynamic traffic environment in a C-/C+L-band framework, while incorporating physical-layer model from [12]. We use, as reference topology, the BT-UK network (see Fig. 1), which comprises 22 nodes and 35 bi-directional links, where the average link length is about 147 km [3]. We repeat and average the simulations for 10 seeds, each with 15,000 demands. Network traffic follows a Poisson distribution with arrival rates of λ_{peak} and $\lambda_{off-peak}$, which varies across each traffic type, during peak and off-peak hours, respectively. Table I presents all traffic parameters, where

TABLE I

TRAFFIC PARAMETERS: TYPE (q), HOLDING TIME (τ), ARRIVAL RATE (λ), COMPRESSION FACTOR (ϕ), DELAYABILITY (δ), AND DATA RATE (γ)

q	Percentage of Total Requests	τ (min)	λ_{peak} (requests/s)	$\lambda_{off-peak}$ (requests/s)	ϕ	δ (min)	γ (Gbps)
1	20%	5	8	2	N/A	N/A	$\{100, 200\}$
2a	45%	$\mathcal{U}[30, 90]$	100	25	50%	$\mathcal{U}[3, 5]$	$\{200, 400\}$
2b	25%	$\mathcal{U}[20, 40]$	48	12	50%	N/A	$\{200, 400\}$
3a	6%	$\mathcal{U}[8, 12]$	8	2	N/A	$\mathcal{U}[2, 4]$	$\{100, 200\}$
3b	4%	$\mathcal{U}[360, 600]$	4	1	N/A	$\mathcal{U}[360, 720]$	400

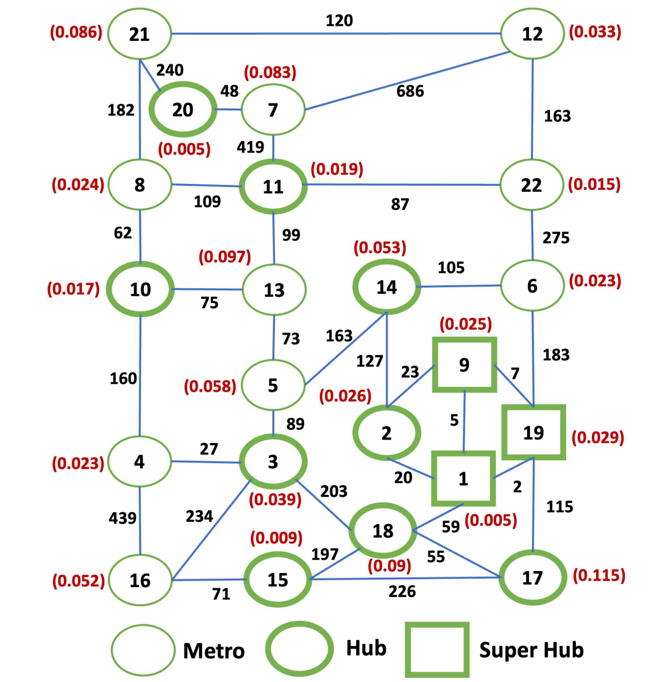

Fig. 1. BT-UK network: link lengths in kilometer (km) and traffic generation probabilities in parentheses [12], [13].

$\mathcal{U}[\cdot,\cdot]$ denotes a uniform distribution. For γ, $\{\cdot\}$ indicates that an element in a set is selected uniformly. Traffic is generated over multiple days. Figure 4 shows a 24-hour snapshot with peak hours from $p_s = $ 8am to $p_e = $ 8pm. For traffic type 3b, $m = 2$, such that $p\prime_e = $ 10pm. We set $t^p = 10$ and $t^o = 100$.

B. *Baseline Strategies*

To demonstrate the efficiency of DACA (see Section III), we compare it with a set of baseline strategies described below:

- **No Delay, No Compression (NDNC)** processes all traffic uniformly with no delay or compression. If incoming traffic cannot be provisioned, it is blocked.
- **Delay-Aware (DA)** only delays the provisioning of requests when resources are unavailable. If incoming traffic cannot be immediately provisioned and is delayable, DA postpones its provisioning. Note that provisioning of traffic type 3b follows the approach mentioned in DACA.
- **Compression-Aware (CA)** only compresses the traffic when resources are unavailable. If incoming traffic cannot be provisioned and is compressible, CA implements compression and provisions the request in its compressed form.

C. *Performance Evaluation in C vs. C+L*

We now evaluate the performance of DA, CA, and DACA w.r.t. NDNC, as shown in Fig. 2. For context, NDNC resulted

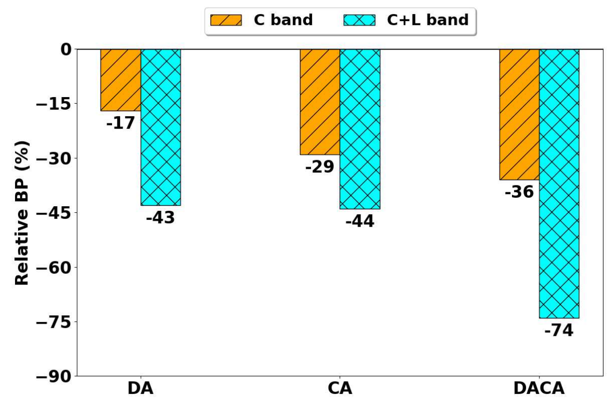

Fig. 2. Relative BP induced by all strategies w.r.t. NDNC in C and C+L.

in BP of 13.4% of total network traffic for the C-band-only scenario. Relative to this, DA and CA lower the BP by 17% and 29%, respectively. This can be attributed to the fact that, while NDNC accommodates incoming requests by only checking for available bandwidth, DA and CA employ delayability and compressibility, respectively, which allows the network to operate with greater flexibility. However, with DACA, we observe a 36% reduction in BP. In DACA, if a request cannot be provisioned after delaying, it may still be possible to provision in its compressed form. This allows the network to serve more connection requests compared to NDNC, DA, and CA. Similarly, activation of L band led to reduced BP of 4.6% of total traffic in NDNC. Due to additional bandwidth in L band, note that DA, CA, and DACA show a significant improvement of 43%, 44%, and 74%, respectively, in BP. The rest of this study will focus on the performance of these strategies in the C+L-band framework.

D. *Performance Evaluation based on Traffic Types*

In this section, we assess the impact of each strategy on different traffic types w.r.t. NDNC. In Fig. 3, we see that DA reduces the BP of types 2a, 2b, and 3a, by 57%, 12%, and 60%, respectively, and maintains the same BP as NDNC for 3b. Here, DA delays all requests that cannot be provisioned immediately, and hence provisions a higher volume of types 2a, 2b, and 3a, compared to NDNC. Note that traffic from the previous day, particularly type 3b, occupies bandwidth overnight for a significantly longer duration. This results in lack of resources for type 1, leading to its higher BP of 57%.

In CA, we observe that the BP for types 2a and 2b is reduced by 46% and 60%, respectively. While CA can compress and provision types 2a and 2b, which constitutes 70% of the total traffic, such provisioning has an adverse effect on other traffic types. Consequently, there is a rise in the BP of traffic types 1,

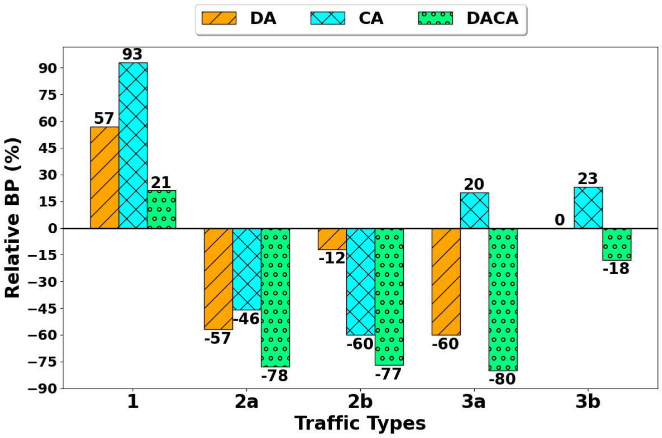

Fig. 3. Relative BP induced by all strategies w.r.t. NDNC per traffic type.

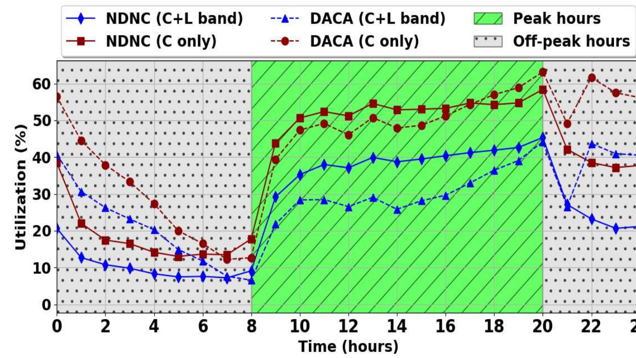

Fig. 4. Average resource utilization per hour (NDNC vs. DACA).

3a, and 3b by 93%, 20%, and 23%, respectively. Additionally, unlike DA and DACA, type 3b is never delayed, which along with existing traffic from the previous day, contributes to its increased BP and also affects other types. On the other hand, since CA blocks a higher volume of types 2a, 3a, and 3b than DA, it has more resources available for 2b. Thus, the relative BP of type 2b is significantly higher than DA.

In DACA, requests can be provisioned in a compressed form even after their maximum tolerable delay (δ) is exhausted, resulting in 78% drop in BP for type 2a w.r.t. NDNC. Compressibility also reduces BP for type 2b (non-delayable and compressible) by 77%. Delayability allows an operator to delay types 2a, 3a, and 3b to ensure spectrum availability, while compressibility reduces the number of channels occupied by types 2a and 2b. The combination of these two factors reduces the relative BP for types 3a and 3b w.r.t. NDNC, while traffic from the previous day leads to a rise in BP of type 1.

To gain a better understanding of resource consumption in the network, Fig. 4 shows the weighted average spectrum utilization per hour for both NDNC and DACA across C-band-only and C+L-band frameworks. We notice that, due to limited bandwidth in C band, utilization is between 50% and 60% during peak hours. Note that the uptick observed at 22 hours (10pm) in DACA indicates the bandwidth consumed by type 3b, most of which was delayed until 10pm. However, addition of L band reduces this usage to about 40% in NDNC. DACA further reduces utilization to about 30%. While both NDNC and DACA have access to additional bandwidth in C+L, DACA utilizes this bandwidth more effectively by deferring most of type 3b traffic to a later time, allowing other traffic types with much shorter holding times to be served in the network. Overall, these experiments indicate a significant increase in the network's information-carrying capacity.

V. Conclusion

We considered a diverse-dynamic traffic scenario in an optical backbone network and proposed a provisioning strategy, called DACA, which exploits delayability and compressibility metrics. Numerical results show that the proposed strategy significantly reduces the blocking probability and hence increases the information-carrying capacity of the network. It also shows that DACA is more effective in C+L framework, offering a promising solution for managing traffic growth in future multi-band networks.

Acknowledgment

This work was supported by National Science Foundation (NSF) Grant No. 2226042. We appreciate reviewers' feedback which helped us improve this paper.

References

[1] Cisco, "Cisco annual report (2018–2023) white paper," 2020.

[2] R. Wang, S. Bidkar, R. Nejabati, and D. Simeonidou, "Load-aware nonlinearity estimation for efficient resource allocation in elastic optical networks," *Proc. International Conference on Optical Network Design and Modeling*, Budapest, Hungary, May 2017.

[3] A. Mitra, D. Semrau, N. Gahlawat, A. Srivastava, P. Bayvel, and A. Lord, "Effect of reduced link margins on C + L band elastic optical networks," *Journal of Optical Communications and Networking*, vol. 11, no. 10, pp. C86–C93, Oct. 2019.

[4] N. Sambo, A. Ferrari, A. Napoli, N. Costa, J. Pedro, B. Sommerkorn-Krombholz, P. Castoldi, and V. Curri, "Provisioning in Multi-Band Optical Networks," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2598–2605, May 2020.

[5] Light Reading Inc., "Chile Entel is building the world's first 800G mesh backbone network, entering the F5.5G Era," Feb. 2024. [Online]. Available: https://www.lightreading.com/network-technology/chile-entel-is-building-the-world-s-first-800g-mesh-backbone-network-entering-the-f5-5g-era

[6] J. Comellas, F. Nourmohammadi, and G. Junyent, "Impact of Traffic Delay Tolerance on Elastic Optical Networks Performance," *Proc. International Conference on Transparent Optical Networks*, Angers, France, July 2019.

[7] C. S. K. Vadrevu, R. Wang, M. Tornatore, C. U. Martel, and B. Mukherjee, "Degraded Service Provisioning in Mixed-Line-Rate WDM Backbone Networks Using Multipath Routing," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 840–849, June 2014.

[8] SEO.ai, "How Many People Use Google?" April. 2024. [Online]. Available: https://seo.ai/blog/how-many-people-use-google

[9] Cybercrew, "YouTube Usage Statistics UK," Feb. 2024. [Online]. Available: https://cybercrew.uk/blog/youtube-usage-statistics-uk/

[10] Restream Inc., "48 live streaming stats you should know in 2024," Dec. 2023. [Online]. Available: https://restream.io/blog/live-streaming-statistics/

[11] EnvisionIT Solutions, "How Often Should Your Business Perform a Data Backup?" Oct. 2022. [Online]. Available: https://blog.envisionitsolutions.com/how-often-should-your-business-perform-a-data-backup

[12] R. Kalkunte, R. K. Jana, S. Ferdousi, A. Srivastava, A. Mitra, M. Tornatore, A. Lord, and B. Mukherjee, "GSNR-aware resource re-provisioning for C to C+L-bands upgrade in optical backbone networks," *Photonic Network Communications*, vol. 47, no. 3, pp. 139–153, June 2024.

[13] T. Ahmed, A. Mitra, S. Rahman, M. Tornatore, A. Lord, and B. Mukherjee, "C+L-Band Upgrade Strategies to Sustain Traffic Growth in Optical Backbone Networks," *Journal of Optical Communications and Networking*, vol. 13, no. 7, pp. 193–203, July 2021.