# On the Efficacy of WDM Virtual Topology Design Strategies

Xuezhou Ma
Department of Computer Science
North Carolina State University
Email: xuezhou_ma@ncsu.edu

Khaled Harfoush
Department of Computer Science
North Carolina State University
Email: harfoush@cs.ncsu.edu

***Abstract*—Existing WDM virtual topology models mainly aim at maximizing the network throughput by optimizing predetermined objective functions. While the literature is rich in variants of such objective functions, they share a few deficiencies. Specifically, they abstract the problem with one *fixed* objective assuming that the throughput hindrance is *uniform* across the network, and do not consider node structure nor router utilization. These factors, when considered, affect network bottlenecks limiting a network throughput. As a result, *none* of the existing models fits all ISP networks. In this paper, we introduce a novel algorithm to determine a network bottleneck based on the 1) physical topology, 2) traffic demand and 3) technology constraints, and a topology model leading to optimized network throughput.**

***Index Terms*—WDM optical networks, Internet backbone, virtual topology design, optimization**

## I. Introduction

Existing WDM virtual topology models mainly aim at maximizing the network throughput. Maximizing network throughput is reasonable given that ISPs are profit driven and a higher throughput translates into a higher profit. Still, distinct models optimize distinct objectives, which result in different virtual graphs. For example, heuristics in [1] primarily connect distant nodes with large demand; while approaches in [2] tend to place lightpaths between neighboring nodes. Refer to Figure 1 for two published ISP graphs: National LambdaRail (NLR) and Sprint network. In Sprint, many *end-to-end* lightpaths such as (SJ-RLY) are set up, resulting in a fairly convoluted virtual topology. NLR, on the other hand, exhibits purely *point-to-point* pattern with no lightpath bypassing any intermediate node. This observation, however, creates much confusion as optimizations aimed at the same goal lead to virtual graphs of totally different topological features.

In this paper we aim at explaining this paradox from a new perspective rather than focusing on individual design strategies. As network throughput is limited by bottleneck elements, whether at routers or links, it is essential for an effective model to identify these elements. We show that none of the existing models identify bottleneck elements in all cases. They abstract the problem with one fixed objective assuming that the throughput hindrance is uniform across the network, and do not consider node structure nor router utilization. As a result, a model suitable for one network (or one branch) may not be suitable for others. Our results reveal that the distinctness of ISP network bottlenecks is the main cause for the distinctness in the corresponding virtual graphs. Furthermore, we introduce a novel algorithm to determine a network bottleneck based on the 1) physical topology, 2) traffic demand and 3) technology constraints, and a topology model leading to optimized network throughput.

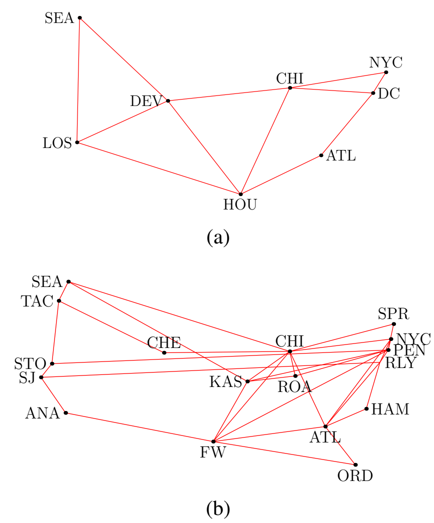



Fig. 1. Virtual topology maps for (a) NLR [3] and (b) Sprint [4]. Note that each edge in the maps may represent multiple parallel lightpaths.

The rest of this paper is organized as follows. In Section II, we survey related research work. In Section III, we discuss the virtual topology design factors. In Section IV, we propose our approach. In Section V, we present the simulation results on published and synthetic ISP networks. We finally conclude in Section VI.

## II. Related Work

For years, many popular design metrics have been considered and each of them was shown individually to maximize the throughput for only certain networks or scenarios. In [1], the *packet hop distance* was minimized since the average hop distance is inversely proportional to the network throughput, when traffic demand is balanced across the network. In [5], the *maximum link utilization* was minimized because an overall throughput *scale-up* was often limited by over-congested links. In [6], a virtual layout with the *minimum usage of network resources* (e.g., wavelengths) was selected to fight against traffic

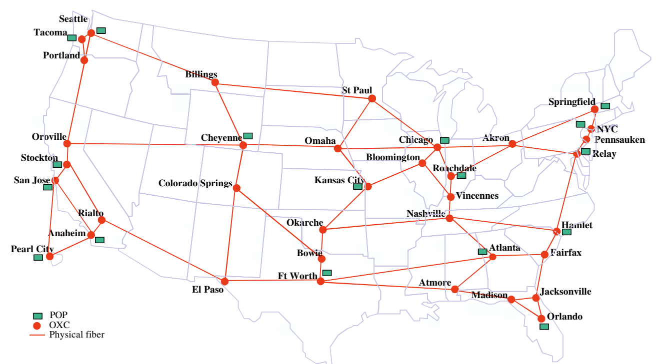



Fig. 2. Physical topology for Sprint network [4]. Each edge in the graph represents optical fibers. The PoP nodes are marked with green box.

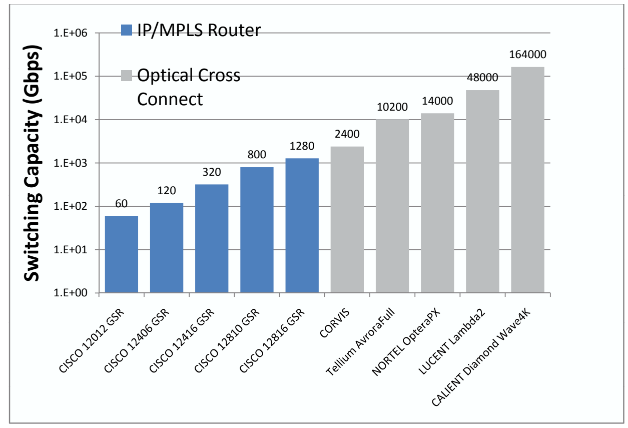



Fig. 3. Comparison of OXCs and routers in term of switching capacity. Information in this figure is collected from [1] and manufacturer web sites based on data from 2006.

fluctuations. In [1], the *demand delay product*, the product of traffic rate and end-to-end queuing delay, was considered equivalent to the amount of data "on the line". In the case router ports are scarce resources, maximizing the *channel load* by grooming low-rate flows into one channel was shown more appropriate [2]. Despite the rich literature in this area, relying on one objective function in all cases is expected *not* to lead to the maximum throughput in all networks independent of bottleneck elements in these networks.

The optimization problem is typically casted as a Linear Programming (LP) problem. The optimal solution is computed using software packages such as *Cplex* for small sized cases. To make it tractable for real networks, heuristic algorithms were introduced to obtain the approximate or near-optimal results. For example, MMT [1] primarily connects distant nodes with large demand. MLDA [5] establishes lightpaths between node pairs in descending order of their traffic intensity. MRU [2] first places lightpaths between neighboring nodes. Two other greedy approaches, *simulated annealing (SA)* [7] and LPLDA [8], apply an iterative algorithm by starting with an initial random virtual graph and adjusting the configuration of adjacent nodes during each iteration. SA and LPLDA are computationally expensive. The lack of knowledge about the bottleneck location hinders the efficacy of the adjustments, creating a large number of less-relevant intermediate steps.

## III. WDM Topology Design Factors

In this section, we make a case that in order to obtain realistic and optimized virtual topologies the following factors need to be considered: 1) node structure, 2) link/router utilization, 3) traffic demand, and 4) technology constraints.

In general, backbone networks consist of two types of nodes: *PoP nodes* which are equipped with both Optical Cross-Connects (OXCs) and routers, and *OXC nodes* which contain only optical switches. The OXCs switch lightpaths in optical domain from input links to output links, and routers switch data packets (when converted into electronic domain) over the lightpaths. Note that OXC nodes can not originate/terminate lightpaths, and while they should show in a network physical graph, they are absent from the virtual graph. For example, by comparing Sprint physical topology, Figure 2, with the corresponding virtual map, Figure 1 (b), we find that the outgoing traffic at Atmore is not forwarded directly. Instead, it is first aggregated at Orlando PoP through regional fabric (not shown), then transmitted to the destination by taking one of the two output lightpaths (ORD-FW) and (ORD-ATL). Existing WDM virtual topology models assume that all nodes were created *equal*, as PoP nodes, which leads to unrealistic node connectivity and unoptimized topologies.

In early networking systems, both the number and the capacity of wavelengths were limited. For example, the RACE project [9] in 1998 had only four wavelengths on one link with each running at 2.5 Gbps. That is why researchers have repeatedly minimized the lightpath utilization to increase the throughput. Recently, however, optics technologies have improved tremendously. Newly deployed infrastructures support as many as 64 wavelengths at 40 Gbps each [9], which is far more than the capacity of core routers in the market. As an illustration, Figure 3 compares the switching capacity of optical switches and IP routers. We use the Cisco 12000 GSR series to represent typical router specifications in 2006, compared to OXCs from Lucent, Tellium, *et al.* in the same time period. It is clear from the figure the optical capacity is several order of magnitude larger than the electronic processing rate, implying that the routers are more likely to be the bottleneck in modern networks. Unfortunately, existing topology models rarely take router utilization into consideration.

Also, traffic demands are not uniform across the network. PoPs located at major cities with large population can exchange traffic at a much higher rate than others [10]. There are always some nodes and/or links that carry excessive load, creating bottlenecks. The traffic demands, together with the layout of physical topology, affect the bottleneck location and intensity. For instance, a "hot spot" NYC in Figure 2 is expected to employ mostly direct connections to other PoPs in order to avoid congestion at intermediate nodes, while establishing too many end-to-end lightpaths may exhaust the

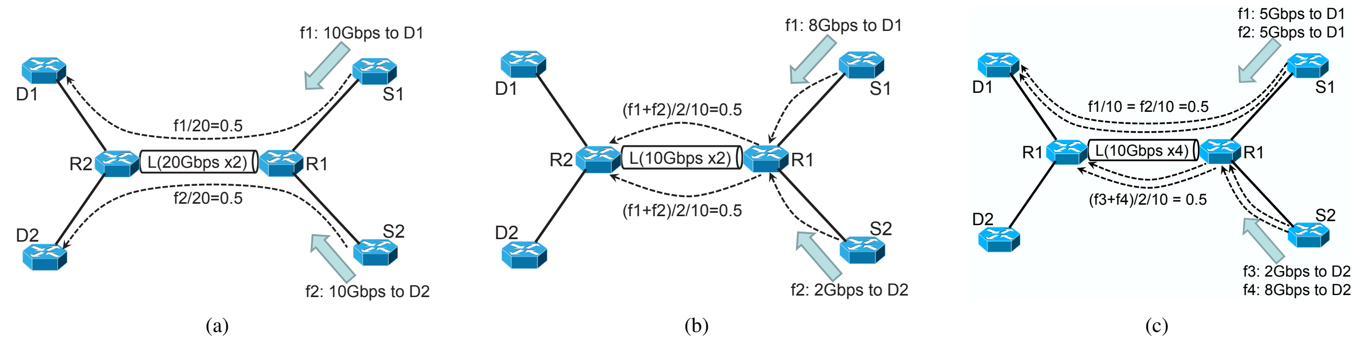



Fig. 4. An illustration of preferred lightpath layout. (a) router capacity is the bottleneck; (b) wavelength capacity is the bottleneck; (c) both router and wavelength capacity are bottlenecks.

scarce wavelengths and result in new bottlenecks at the output links. Another big city, CHI, with five output links will not have such problem though. A realistic virtual topology model thus needs to take into consideration the traffic distribution over physical infrastructure.

Moreover, the design of a WDM virtual topology is affected by technology constraints. The choice of the supporting technology is an investment decision balancing the cost and the expected revenue based on traffic volume. While major ISPs can deploy costly switches with large port count to accommodate millions of users, a small ISP with limited number of subscribers has to rely on less expensive, hence less capable equipment. The bottleneck elements will shift easily as different technologies are deployed. Previous literature that evaluated their models only on few Tier-1 ISP networks can hardly be effective for others.

## IV. A Bottleneck-Oriented Design

### A. Problem Definition and Assumptions

Refer to Table I for the terminology used. Network throughput, $T$, is the sum of the accommodated traffic between all node pairs. With given traffic demands, $T$ is also known as the maximum scale-up factor. The WDM virtual topology design problem aims at identifying a virtual topology $E_v$ that maximizes the throughput $T$, given physical topology $G_p$, PoP node placement $V_v$, traffic matrix $\Lambda$, and technology constraints $w$, $c$, and $c_r$.

TABLE I
Notations

| Notation | Definition |
|---|---|
| $G_p = (V_p, E_p)$ | Physical topology with the set of nodes $V_p$ and the set of fiber links $E_p$ |
| $G_v = (V_v, E_v)$ | Virtual topology with the set of PoP nodes $V_v$ and the set of lightpaths $E_v$ |
| $\Lambda = \{\lambda^{sd}\}$ | Long-term traffic demand between nodes $s$ and $d$ |
| $\rho^L/\rho^N$ | Maximum lightpath/node utilization |
| $c_r$ | Electronic switching capacity at a node |
| $c$ | Capacity of each wavelength |
| $w$ | Number of wavelengths per fiber |
| $T$ | Network throughput |

We make the following assumptions: 1) each physical link is bi-directional and composed of a single fiber supporting $w$ wavelengths; 2) the wavelength-continuity constraint is not considered since a design with no or limited converting capability has been addressed in [11]; 3) traffic flows go through the smallest possible number of lightpaths, i.e., shortest-path-first (SPF) routing protocol. SPF algorithm may not be optimal for all cases, but it holds well for *static* network design problems [12].

### B. Key Idea

In order to maximize the throughput, WDM virtual topology design needs to minimize the utilization of bottleneck elements, whether routers or links. The utilization of a lightpath is defined as the percentage of the wavelength capacity that is used by traffic crossing the lightpath. The utilization of a node is defined as the percentage of the router capacity that is used by traffic crossing the router. When routers are the bottlenecks, it is desirable to *bypass* electronic processing at routers by creating end-to-end lightpaths. On the other hand, when wavelength capacity is the bottleneck, it is desirable to deploy traffic grooming techniques, thus not bypassing electronic processing, in order to: 1) distribute traffic evenly among parallel lightpaths, and/or 2) spread the load over alternative link paths.

Let $\rho^L$ denote the *maximum lightpath utilization* in the network, and $\rho^N$ denote the *maximum node utilization*. A virtual topology which relies on end-to-end lightpaths (referred to as a *transparent* topology) is thus mostly trying to avoid bottlenecks at routers, and a topology which relies on hop-by-hop lightpaths (referred to as an *opaque* topology) is trying to avoid bottlenecks at links. An optimal topology design should balance the use of these two schemes to minimize $\rho = \max\{\rho^L, \rho^N\}$.

Refer to Figure 4 for an example. In this example, two sources $S1$ and $S2$ send two flows with rates $f1$ and $f2$ towards destinations $D1$ and $D2$. The flows cross routers $R1$ and $R2$ and the link $L$ connecting them. In Figure 4(a), link $L$ has two 20 Gbps wavelengths and the routers, $R1$ and $R2$, each has 20 Gbps capacity, i.e., $w$=2, $c$=20 and $c_r$=20. If flows $f1$ and $f2$ (10 Gbps each in this case) are processed at $R1$, they will exhaust the router capacity and create node congestion as $\rho^N = (f1 + f2)/20 = 100\%$. Thus, it is desirable to bypass the routers in Figure 4(a) with

**Algorithm 1** The BOA Algorithm

```
1: INPUT: Virtual topology $G_v$ (2-connected)
2: repeat
3:   Implement SPF routing on $G_v$ and compute:
4:     $T \leftarrow$ the throughput
5:     $k \leftarrow$ the node with maximum utilization $\rho^N$
6:     $(i,j) \leftarrow$ the lightpath with maximum utilization $\rho^L$
7:   if $\rho^N \geq \rho^L$ then
8:     $S \leftarrow$ the list of traffic flows being routed at $k$
9:     while $S \neq \emptyset$ do
10:      $\lambda^{sd} \leftarrow$ pop the largest flow in $S$
11:      try to create a new lightpath $(s,d)$. If fails,do
12:        $c \leftarrow$ connectivity of $G_v - \{\forall(u,v)|p_{uv}^{sd}=1\}$♠
13:        if $c > 1$ then
14:          Replace $\{\forall(u,v)|p_{uv}^{sd}=1\}$ by a channel $(s,d)$
15:        end if
16:    end while
17:  else
18:    try to add one more parallel $(i,j)$. If fails, do
19:      $(p,q) \leftarrow$ search for the qualified channel♣
20:      if $(p,q)$ exists then
21:        Divide $(p,q)$ into segments containing $(i,j)$
22:      else
23:        Divide $(i,j)$ into pt-to-pt channels
24:      end if
25:  end if
26:  Recompute the new throughput $T'$
27: until $T' < T$
28: OUTPUT: new virtual topology $G_v^{'}$ (2-connected)
```

a utilization $\rho = \rho^L = \max\{f1/20, f2/20\} = 50\%$. In Figure 4 (b), we reduce the wavelength capacity to 10 Gbps (i.e., $w$=2, $c$=10 and $c_r$=20) and let $f1$ and $f2$ be 8 Gbps and 2 Gbps, respectively. Bypassing routers in this case leads to $\rho = \rho^L = \max\{f1/10, f2/10\} = 80\%$, while applying traffic grooming at $R_1$, as shown in Figure 4(b), provides a lower $\rho = \max\{(f1+f2)/2/10, (f1+f2)/20\} = 50\%$. Finally, in Figure 4(c), end-to-end lightpaths lead to a $\rho = \rho^L = 80\%$, and hop-by-hop lightpaths lead to $\rho = \rho^N = 100\%$. Neither design is optimal. A layout combining both, as shown in the figure, leads to the best $\rho = 50\%$.

### C. Algorithmic Details

We provide an iterative heuristic algorithm, called bottleneck-oriented approach (BOA), that does not make any assumption about the location of network bottlenecks. In each iteration, $\rho^L$ and $\rho^N$ are computed and maintained. We identify the most congested element and set up the lightpaths accordingly. Refer to Algorithm 1. If the bottleneck is a router, lines 8-11 establish a new end-to-end lightpath to carry the largest flow routing through this router. In case of insufficient wavelengths, line 14 removes the series of intermediate channels currently used by $\lambda_{sd}$ and replaces them by a direct connection $(s,d)$, with subject to network connectivity constraints shown in lines 12-13. Otherwise, if the bottleneck is a channel, say $(i,j)$, a parallel channel will first be attempted so that the load of $(i,j)$ can be split with the new one (see line 18). In case the creation of the new parallel channel fails due to the shortage of wavelengths, we search for one existing "longer" channel that can be fragmented into $(i,j)$ and other segments (see lines 19-21). As marked by ♣, a qualified channel must traverse all links traversed by $(i,j)$ and we select the one which carries the least traffic load. The initial input $G_v$ can be either a real virtual graph in use, or a baseline configuration such as MMT [1] and MLDA [5].

The convergence of Algorithm 1 to a lower $T$ is guaranteed since $T' < T$ is checked after each iteration. Furthermore, the algorithm cannot lead to a disconnected graph since we always start with a *2-connected* graph (line 1) and enforce *2-connectivity* in each iteration (as marked by ♠ in Algorithm 1). 2-connectivity is a common property of core networks to ensure resilience [10]. Each iteration of BOA takes $O(|V_v|^3 lg|V_v|)$ of time and $O(|V_v|^2)$ of space in a worst case. The total number of iterations performed depends on input graph $G_v$ and is bounded by $O(|V_v|^2)$.

TABLE II
ISP NETWORK CONFIGURATIONS

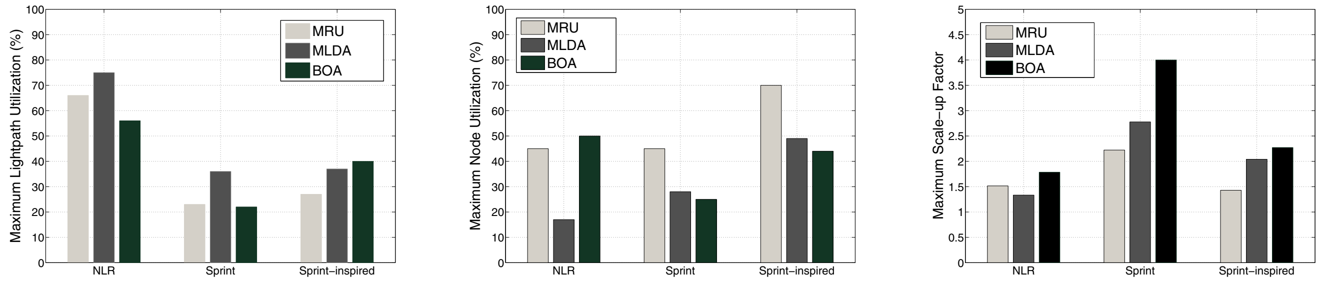
| **Network** | $\Lambda$ (Tbps) | $w$ | $c$ (Gbps) | $c_r$ (Gbps) |
|---|---|---|---|---|
| NLR | 2 | 8 | 5 | 160 |
| NLR-inspired | 2 | 8 | 5 | 160 |
| Sprint | 10 | 16 | 10 | 320 |
| Sprint-inspired | 50 | 64 | 10 | 320 |

## V. PERFORMANCE EVALUATION

Our experiments are based on two real networks: NLR (Figure 1) and Sprint (Figure 2), and two fictitious networks: *NLR-inspired* and *Sprint-inspired*. NLR-inspired network is created by adding three *extra* links, (SEA-CHI), (LOS-DC), and (HOU-NYC), which reflect NLR's next-generation infrastructure proposal in 2009. To represent emerging DWDM technologies, Sprint-inspired network has the number of wavelengths on each link upgraded to 64. The detailed network configurations are summarized in Table II.

Since ISP traffic demand is rarely published, we consider the following traffic model. The required bandwidth between two nodes is proportional to the product of their populations (refer to traffic model in [10]), and traffic has to be aggregated at a PoP in the vicinity before being transmitted. This scenario reflects the demand irregularity and hierarchical structure of Internet backbone networks.

Each simulation starts with a simple MMT configuration (see Section II) and then BOA is applied till all traffic demands are accommodated. Depending on the setup, the resulting virtual topology could be transparent, opaque or a combination thereof. The results are listed in Table III. In all cases, BOA balances $\rho^L$ and $\rho^N$, thus reduces the bottleneck utilization. Since NLR has initial link congestion much larger than its node congestion (i.e., $\rho^L = 68\%$ and $\rho^N = 42\%$), the bottlencks are on the links. The resulted NLR virtual graph is




(a) (b) (c)

Fig. 5. Comparison of BOA with two other design approaches in terms of (a) the maximum lightpath utilization, (b) the maximum node utilization, (c) the network throughput.

TABLE III
NETWORK BOTTLENECK AND DESIGN STRATEGY

| Network | Desired Design | Network Bottleneck | |
|---|---|---|---|
| | | $\rho^L$ (%) | $\rho^N$ (%) |
| NLR | opaque | $68 \rightarrow 55$ | $42 \rightarrow 50$ |
| NLR-inspired | combination of two patterns | $50 \rightarrow 44$ | $40 \rightarrow 45$ |
| Sprint | combination of two patterns | $26 \rightarrow 22$ | $20 \rightarrow 23$ |
| Sprint-inspired | transparent | $27 \rightarrow 40$ | $68 \rightarrow 44$ |

laid out with all point-to-point connections. Surprisingly, even with an opaque graph, NLR still suffers from link congestion more than from the nodes (i.e., $55\% > 50\%$). This observation suggests that as traffic volume increases, NLR is likely to set up more parallel channels along the congested links to treat the bottleneck. Also, one observes that by adding three extra links, the bottleneck utilization in NLR-inspired network decreases significantly. Particularly, $\rho = \rho^N = 45\%$ in NLR-inspired network compared to $\rho = \rho^L = 55\%$ in NLR network. The optimized NLR-inspired network has both transparent and opaque layouts.

The Sprint network yields comparable link and node congestions in the initial configuration (i.e., $\rho^L = 26\%$ and $\rho^N = 20\%$). The resulting virtual graph has both point-to-point and end-to-end connections and matches well with the published Sprint map. For the Sprint-inspired network, it is obvious that a significant shortage of switching capability at nodes (i.e., $\rho^N = 68\%$ compared to $\rho^L = 27\%$) pushes the resulting graph towards a transparent layout. Our results imply that the Sprint network is likely to evolve towards a transparent virtual graph in order to maximize the throughput.

We next compare the performance of BOA to two other popular heuristics MLDA and MRU in terms of the maximum lightpath utilization, maximum node utilizations, and network throughput. As shown in Figure 5, BOA outperforms the other heuristics, leading to better network throughput. It achieves that by balancing the maximum link utilization and the maximum node utilization. These results reinforce our arguments that none of the existing heuristics perform best in all setups. BOA on the other hand can pinpoint the bottleneck elements, which is crucial to maximize the network throughput.

## VI. CONCLUSIONS AND FUTURE WORK

In this paper, we investigate the existing WDM virtual topology models and make a case that none of them fits all ISP networks. This is mainly due to the fact that different networks have different bottleneck elements as they vary in traffic demands and the technology deployed. We further propose an adaptive design model which iteratively identifies the bottleneck elements and sets up suitable lightpaths accordingly. The efficacy of our model is tested on published and synthetic ISP networks. Simulation results highlight the superiority over existing designs. Our approach targets a static design problem and may not be applicable to dynamic switching technologies such as optical burst switching (OBS) and optical packet switching (OPS).

## REFERENCES

[1] B. Mukherjee, *Optical WDM networks*. Springer Science, 2006.
[2] K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 122–133, 2002.
[3] "National LambdaRail network," 2009, http://noc.nlr.net.
[4] A. Nucci, N. Taft, P. Thiran, H. Zang, and C. Diot, "Increasing the link utilization in IP over WDM networks using availability as QoS," *Photonic Network Communications*, vol. 9, pp. 55–75, 2005.
[5] R. M. Krishnaswamy and K. N. Sivarajan, "Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers," *IEEE/ACM Transactions on Networking*, vol. 9, no. 2, 2001.
[6] M. Tornatore, G. Maier, and A. Pattavina, "WDM network design by ILP models based on flow aggregation," *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 709–720, 2007.
[7] S. Xu, K. Sezaki, and Y. Tanaka, "A two-stage simulated annealing logical topology reconfiguration in IP over WDM networks," in *Proceedings 11th IEEE International conference on Telecommunications Network Strategy and Planning Symposium*, 2004, pp. 327–332.
[8] N. Skorin-Kapov, *Greedy Algorithms*. InTech, 2008, vol. 30.
[9] K. Liu, *IP over WDM*. John Wiley & Sons, 2002.
[10] X. Ma, K. S., and K. Harfoush, "Towards realistic physical topology models for internet backbone networks," in *Proceedings of 6th IEEE International Symposium on High-Capacity Optical Networks and Enabling Technologies (HONET)*, 2009, pp. 36–42.
[11] J. Teng and G. N. Rouskas, "Wavelength selection in OBS networks using traffic engineering and priority-based concepts," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1658–1669, 2005.
[12] G. Iannaccone, C. Chen-Nee, S. Bhattacharyya, and C. Diot, "Feasibility of IP restoration in a Tier 1 backbone," *IEEE Journal on Networking*, vol. 18, no. 2, pp. 13–19, 2004.