Achieving IP Routing Stability with Optical Bypass

Mohit Chamania, Marcel Caria and Admela Jukan
Technische Universität Carolo-Wilhelmina zu Braunschweig
Email: chamania, caria, jukan@ida.ing.tu-bs.de

***Abstract*—The evolution of next generation services has led to significant increase in Internet backbone traffic, and multi-layer (hybrid) IP over circuit/optical layer solutions are being explored to cope up with the growing demands for capacity. Optical circuit bypass is typically used to increase capacity in the IP layer without need for over-provisioning, which in turn reduces OPEX of the IP network. Current proposals for IP topology reconfigurations in multi-layer networks do not take into consideration the effect of modifying the network topology on routing stability. We present a new bypass-based IP topology upgrade mechanism which can be used with high frequency without significantly affecting routing in the network. We present an ILP based approach to compute the optimal bypasses in the IP layer in case of congestion and numerical results show that our proposed solution is scalable and efficient.**

I. Introduction

The emergence of next generation services has led to significant increase in the Internet backbone traffic in recent years. Traditional network design paradigms use over-provisioning of resources in the network to make it future-proof as can be seen in the Internet-2 backbone network [1], where every link in the network is 80% over-provisioned, i.e., ensuring that the maximum utilization in any IP link is no more than 20%. However, the monumental growth of Internet services, sudden traffic churns in response to high-bandwith applications, along with growing concerns to reduce network CAPEX and OPEX is indicative that simple over-provisioning paradigms cannot support future backbone networks. Network providers are therefore looking towards solutions which can reduce over-provisioning while avoiding frequent hardware upgrades.

To this end, the network community has been exploring the possibility of using dynamic Layer 1/Layer 2 circuits to increase the capacity in the IP layer. Significant improvements in the area of optical networks and control planes have reduced the setup time for lightpaths/layer-2 circuits from days to mere hours and these improvements have led to the development of IP-over-WDM network concepts, where the IP layer is responsible for the routing of network flows, and the WDM layer can increase/decrease link capacity as required and even provision new links in the network in case of sudden increase in network traffic [2]. The WDM layer therefore is responsible for the design of the *virtual topology* used by the IP layer to route traffic. While a vast amount of research has been dedicated to the cost optimal virtual topology at the IP layer, little attention has been paid to the issues of routing stability in the IP layer, which directly depends on the frequency of optical bypass reconfiguration.

Early proposals for virtual topology reconfiguration assume the knowledge of the traffic between source-destination (s,d) pairs and proposed ILP/heuristic mechanisms [3], [4]. A genetic algorithm to reconfigure virtual topologies has been presented in [5]. Other approaches assume incomplete/inaccurate knowledge about the traffic matrix in the network. [6] presents an approach for topology reconfiguration in conjunction with traffic estimation by modifying the network topology in multiple steps, and using traffic measurements after each step to account for traffic estimation errors. Other approaches use mechanisms such as the gravity model [7] to estimate the traffic matrices in large scale networks before performing topology reconfiguration. A heuristic approach for reconfiguring multi-domain virtual network topologies has also been proposed in [8]. Most of these approaches lead to significant routing changes at the IP layer, and are therefore suitable for application over large time scales; achieving IP routing stability with optical bypass is still an open challenge.

In this paper, we propose a new approach for increasing capacity in the virtual IP layer by adding optical bypasses across congested links in the IP network. These bypasses are not advertised in the IP routing protocols, and only the ingress node for each bypass is configured separately to route certain flows on these bypasses, thus maintaining routing stability. The bypasses are established across congested links, and specific flows are rerouted over the bypass to reduce traffic on congested links. The ingress and egress routers of the bypass are on the original routing path of the re-routed flows, thus ensuring that link traffic in other areas is not adversely affected by the creation of a bypass. An Integer Linear Programming (ILP) based approach is presented which can compute the optimal set of bypasses for congestion free operation. The results shown demonstrate the feasibility of the proposed method.

The rest of the paper is organized as follows: Section II describes the concept and ILP formulation, numerical results are presented in section III and section IV concludes the paper.

II. The Concept

The proposed multi-layer virtual topology reconfiguration paradigm is designed to ensure that congestion in the IP network is avoided with minimal routing changes. Minimum routing changes imply stable routing in the IP layer and can be coupled with fast reconfiguring transport networks to quickly adapt to traffic changes without the need for over-provisioning of IP links. The basic premise of our proposal is shown in Fig. 1, where traffic is sent from sources S_i to destinations D_j and links $2-3$ and $3-4$ are congested. The dashed lines indicate bypasses which are established to remove congestion in the network. In this scenario, the bypass between nodes $2-3$ reduces congestion in link 2-3. As the delay characteristics of the bypass may be different from that of a link, we ensure that flows are not routed over the bypass and the original links simultaneously, and thus all traffic between specific (s, d) pairs only is routed over a bypass. In this example, traffic from all sources S_i to D_1 is routed on the shortcut $2-3$. Similarly, the shortcut $2-4$ is established to alleviate the congestion on both links $2-3$ and $3-4$ and traffic from all sources S_i to D_4 is routed over this shortcut. Note that an alternate solution could have been the establishment of bypasses between $2-3$ and $3-4$ to reduce the congestion on both links. The optimal solution is however dependent on the capacity of the shortcuts and the observed overflow on each link. For example, in Fig. 1, a single shortcut may not have enough capacity to alleviate the congestion on link $2-3$. In this case, the shortcut set ($2-3$, $3-4$) cannot reduce congestion in the network.

Another notable fact is that we do not advertise the shortcuts in the IP layer. Router 2 is configured to divert the flows between specific (S_i, D_j) pairs on to the corresponding shortcuts, while other routers such as 1 and 5 are unaware of the existence of these shortcuts. Routers 3 and 4 route the IP packets coming in from the shortcuts using the same routing rules as that for other packets and therefore do not need to reconfigure their routing rules. This ensures that there is minimal routing reconfiguration in the IP layer: in this case only router 2 observes a change in routing. Also note that the egress of a bypass must be on the original routing path, and therefore bypass $2-4$ cannot be used to route traffic to D_1.

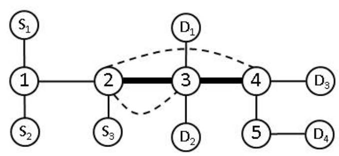

Fig. 1. A topology with congested links demonstrating the use of our proposed shortcut mechanism

In our model, we assume that the traffic matrix is known, and the routing configuration of the IP network is known. The network is represented as a graph $G(V, E)$ with nodes $v_i \in V$ and links $e_{ij} \in E$ with capacity of a link represented as C_{ij}. The traffic from nodes s to d is given by λ_{sd}, and the maximum allowable utilization of any link/bypass is given by α ($\alpha \in (0, 1]$). We assume that there are T different classes of bypasses corresponding to different granularities, each with capacity C^t_{BP} and cost $Cost_t$. Two constants are introduced to characterize routing, namely:

- ψ^{sd}_{xy} - boolean value set to 1 if the traffic from node s to node d is routed over the path $s \rightarrow x \rightarrow y \rightarrow d$, 0 otherwise. Note that in the constraint, s may be equal to x, d may be equal to y, and the path from x to y may contain multiple hops, and $x \neq y$.
- $\psi^{sd}_{xy,ij}$ - boolean value set to 1 if the traffic from node s to node d is routed over the path $s \rightarrow x \rightarrow i \rightarrow j \rightarrow y \rightarrow d$ and the link $e_{ij} \in E$, 0 otherwise. Note that s may be equal to x, x may be equal to i, j may be equal to y, and y may be equal to d.

The constant ψ^{sd}_{xy} indicates if any two (possibly non-neighbouring) nodes x and y lie on the path from s to d can therefore be used to bypass traffic from s to d, while the constant $\psi^{sd}_{xy,ij}$ indicates if the bypass from x to y for traffic from s to d would bypass this traffic from link i to j. As an example, in Fig. 1 $\psi^{S_1D_4}_{24,23} = \psi^{S_1D_4}_{24,34} = 1$ indicating that the bypass from router 2 to 4 bypasses the links $2-3$ and $3-4$ for traffic from S_1 to D_4.

The variables used in the ILP are :

- X^t_{xy} - Boolean variable with value 1 to indicate a bypass of type t is established between the nodes x and y
- f^{sd}_{xy} - Boolean variable with value 1 to indicate if traffic from s to d is routed via the bypass from x to y

The ILP attempts to minimize the total cost of the bypasses setup to relieve congestion in the network, and the objective function is expressed as

$$Minimize: \sum_{xy}\sum_{t}\left(X^t_{xy}\cdot Cost_t\right) \tag{1}$$

$$\forall i,j\colon\ \sum_{sd}\lambda_{sd}\cdot\psi^{sd}_{ij}\left[1-\sum_{xy}\psi^{sd}_{xy,ij}\cdot f^{sd}_{xy}\right]\leq\alpha\cdot C_{ij} \tag{2}$$

$$\forall x,y\colon\ \sum_{t}X^t_{xy}\leq 1 \tag{3}$$

$$\forall s,d\ \ \forall x,y\colon\ f^{sd}_{xy}\leq\sum_{t}X^t_{xy} \tag{4}$$

$$\forall x,y\colon\ \sum_{sd}f^{sd}_{xy}\cdot\lambda_{sd}\leq\alpha\cdot\sum_{t}\left(C^t_{BP}\cdot X^t_{xy}\right) \tag{5}$$

$$\forall s,d\ \ \forall x,y\colon\ f^{sd}_{xy}\leq\psi^{sd}_{xy} \tag{6}$$

$$\forall s,d\ \ \forall i,j\colon\ \sum_{xy}\psi^{sd}_{xy,ij}\cdot f^{sd}_{xy}\leq 1 \tag{7}$$

Eq. 2 is the link capacity constraint and the term $\left[1-\sum_{xy}\psi^{sd}_{xy,ij}\cdot f^{sd}_{xy}\right]$ ensures that traffic from s to d using any bypass which covers the links e_{ij} do not contribute to the link capacity usage. The constraints on the bypasses are presented in Eq. 3, 4, 5. Eq. 3 ensures that there can only be one bypass between any pair of nodes, Eq. 4 ensures that flows can only be routed over a bypass if one exists and Eq. 5 ensures that the capacity utilization on any bypass in less than α. The routing constraints are presented in Eq. 6, 7 with Eq. 6 indicating that a flow can only be routed over a valid bypass, while Eq. 7 ensures that any flow between a given (s, d) pair is not routed over multiple overlapping bypasses. For example, in Fig. 1, the flow from S_1 to D_4 cannot be routed over the bypasses $2-3$ and $2-4$ simultaneously. The solution of the ILP gives the optimal set of bypasses with their granularities that must be established to ensure congestion-free network for a minimum cost. Note that the routing costs in the transport layer have not been taken into consideration in this ILP.

III. Numerical Results

We study the performance of our scheme on the Atlanta reference network using the traffic matrix given in [9]. We assume that shortest path first (SPF) routing is used in the IP layer, and the link capacities are assigned so that the initial link utilization is 0.71, while the maximum allowed utilization $\alpha = 0.9$. We have four different types of bypasses with capacity and normalized costs as: (50, 1), (200, 2), (1000, 4) and (5000, 8), with all capacities in Mbps. To test the performance of our scheme, we randomly select a number of (s, d) pairs and increase the traffic on these pairs by 150%. We then use the ILP to compute the optimal set of bypasses, and the results of the same are presented in Fig. 2. For each results set, the values are averaged over 20 runs.

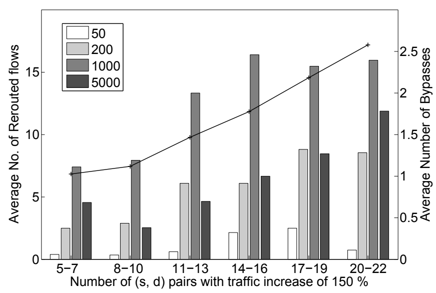

Fig. 2. No. of Bypasses(bar graph) and average no. of rerouted (s, d) flows (line graph) with increasing traffic. Traffic is increased by randomly selecting x number of (s, d) pairs in the range $a - b$ ($a \leq x \leq b$) as shown on the X axis, and increasing the traffic on these (s, d) pairs by 150%

Fig. 2 shows that the average number of bypasses required increases with the number of (s, d) pairs subject to increased load. However, rate of increase of total number of bypasses is very low, suggesting that the system can remain stable even when large number of (s, d) pairs exhibit traffic increase. We can also see that the total number of bypasses required is very low, with only 6 bypasses required on an average in cases when 17 to 19 (s, d) pairs show a significant traffic increase, suggesting that the approach is viable in current networks. The increase in the number of (s, d) pairs also leads to the increase in the number of higher capacity bypasses, while bypasses with lower capacity first increase and then decrease in number. This is observed as the traffic on congested links also increases with the increase in the number of chosen (s, d) pairs, and low capacity bypasses can not relieve congestion on these links. Note that the number of (s, d) flows rerouted onto bypasses increase linearly with the increase in number of (s, d) pairs with high traffic, and are fewer in number than the number of (s, d) pairs with high traffic, indicating that the solution does not disrupt the routing significantly.

Fig. 3 shows that the normalized number of bypasses per congested link is approximately constant with the increase in number of (s, d) pairs, while the fraction of high capacity bypasses increase with increasing number of (s, d) pairs. Fig. 3 also shows the average number of hops traversed by bypasses with capacity 200, 1000, and 5000 Mbps are in the range of 1.5-2.5 hops. We observed link congestion between adjoining links (one common router) because of the characteristics of

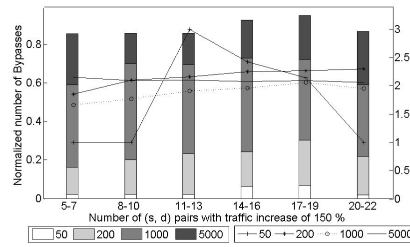

Fig. 3. Normalized no. of bypasses(bar graph) and average hop count(line graph) with increasing traffic. Traffic is increased by randomly selecting x number of (s, d) pairs in the range $a - b$ ($a \leq x \leq b$) shown on the X axis, and increasing the traffic on these (s, d) pairs by 150%

shortest path routing, and therefore high capacity bypasses typically had a span of 2 or more hops. On the other hand, the bypasses with capacity 50 Mbps was used mostly to supplement capacity on congested links at very low or very high loads while the occurrence of longer length bypasses can be attributed to routing of small number of flows across multiple congestion sites.

IV. Conclusion and Future Work

We have presented a novel approach to facilitate bypass usage while ensuring routing stability in the IP layer. An ILP based formulation was presented and the numerical results indicate that the proposed formulation can effectively tackle increase in traffic flow between a large number of (s, d) pairs in a scalable fashion. The average number of hops per bypass and the average number of bypasses per congested link in the network were seen to remain constant, suggesting that the mechanism is ideally suited to handle traffic fluctuations in a network at short time scales. Future research in this area would involve taking into consideration routing at the transport layer, heuristic mechanisms to compute the bypass requirement in polynomial time as well as mechanisms to compute bypasses without traffic matrix information.

References

[1] Internet2 Headroom Practice, https://wiki.Internet2.edu/confluence/download/attachments/17383/Internet2+Headroom+Practice+8-14-08.pdf?version=1

[2] L. Cheng, J. Ellson, A. Jukan, P. Lamy, E. Varma, "*Network Engineering - Control of Dynamic Link Topology in User Networks,*" Bell Labs Technical Journal (BLTJ), Vol. 8, No 1, pp. 207-218, July 2004.

[3] S.F. Gieselman, N.K. Singhal, B. Mukherjee, "*Minimum-cost virtual-topology adaptation for optical WDM mesh networks,*" IEEE ICC, vol.3, no., pp. 1787-1791 Vol. 3, 16-20 May 2005

[4] B. Mukherjee, et al. "*Some Principles for Designing a Wide-Area WDM Optical Network,*" IEEE/ACM Transactions on Networking, vol. 4, no. 5, pp. 684-696 ,Oct. 1996.

[5] R.J. Durn, et al. "*Genetic algorithm to design logical topologies in reconfigurable WDM networks,*" Photonic Network Communications, vol. 17, no. 1, pp. 21-33, Feb. 2009

[6] Y. Ohsita, et al. "*Gradually Reconfiguring Virtual Network Topologies based on Estimated Traffic Matrices,*" IEEE INFOCOM, May 2007

[7] Y. Zhang, et al. "*Fast Accurate Computation of Large-Scale IP Traffic Matrices from Link Loads,*" ACM SIGMETRICS, June 2003

[8] K. Shiomoto, et al. "*Distributed Virtual Network Topology Control Mechanism in GMPLS-Based Multiregion Networks,*" IEEE JSAC, vol. 21, no. 8, pp. 1254-1262, Oct. 2003

[9] Atlanta reference Network, http://sndlib.zib.de/coredata.download.action?objectName=atlanta&format=native&objectType=network