An Efficient Partial Link Monitoring Scheme for Inter-Domain Routing Under Dynamic Traffic Scenarios

Pengfei Zhang[1], Yaohui Jin[1], Weiqiang Sun[1], Wei Guo[1], Weisheng Hu[1], N. Ghani[2]

[1]*State Key Lab of Advanced Communication Systems and Networks, Shanghai Jiao Tong University*
[2]*Electrical & Computer Engineering Department, University of New Mexico*

Abstract: We investigate link-state advertising in multi-domain networks with dynamic traffic. A novel triggering scheme is proposed to monitor a subset of domain links, and thereby achieve a balance between salability and accuracy.

1. Introduction

The proliferation of carrier grade optical transport/switching technologies has gained significant momentum in recent years. As these deployments increase, the topic of inter-domain *traffic engineering* (TE) to achieve *quality of service* (QoS) support across domain has become a key focus [1]. It is here that inter-domain routing protocols play a vital role in disseminating critical topological and resource state between domains to assist with provisioning needs. Now the most prevalent inter-domain protocol today is the *border gateway protocol* (BGP) [2]. However, this offering only provides end-point reachability exchange for packet-routing networks and is generally insufficient for inter-domain circuit and/or QoS TE routing. Alternatively, the *open shortest path first* (OSPF-TE) [3] protocol is much more amenable for *generalized multi-protocol label switching* (GMPLS) TE support [4] owing to its inherent link-state routing design. For example, this protocol can be used in conjunction with topology abstraction techniques to propagate condensed domain-level resource state in a hierarchical manner. OSPF-TE also provides improved state accuracy and faster convergence behaviors as compared to BGP. However, a key challenge here is scalability, particular when advertising abstracted state between large numbers of domains.

Now in terms of scalability, triggering policies play a vital role in inter-domain routing as they control the amount of disseminated state. Along these lines, two main classes of triggering policies have been proposed, namely *period-based advertising* (PA) and *threshold-based advertising* (TA). However, in large-scale settings both of these strategies face a tradeoff between accuracy and scalability. Along these lines, recent work in [5] has proposed an improved *partial link-state monitoring and advertising* (PLMA) scheme which only monitors a subset of domain-level links when computing aggregate link-state information for full mesh abstraction. Namely, here only the resource-critical links (i.e., those which are deemed as bottleneck links which will contribute the most to the abstract link attributes) are tracked and used to trigger the advertising updates. However, since PLMA only considers stable traffic conditions with uniformly-distributed *traffic matrices* (TM) [5], it is deemed as a static/fixed strategy, termed herein as F-PLMA. Given that generalized traffic profiles with a domain can be highly non-uniform and also exhibit dynamic time-varying characteristics, the application of this scheme would require constant changes to the subset of bottlenecked links being monitored. As a result, one solution approach could be to predict changes in domain-level traffic matrices using appropriate estimation methods [6], and then use these values to vary the monitored link sets for dynamic demands.

Along these lines, in this paper considers inter-domain routing for realistic network settings with dynamic (time-varying) traffic profiles. Specifically, an enhanced *variable partial link-state monitoring and advertising* (V-PLMA) scheme is proposed to achieve an acceptable performance tradeoff between scalability and state accuracy at the inter-domain level. The paper is organized as follows. Section 2 presents the detailed overview of the proposed multi-domain advertising model. This is followed by detailed simulation results in Section 3 which gauge the performance of the proposed scheme versus other triggering policies (including PLMA) under dynamic traffic loads. Section 4 then presents conclusions and future work directions.

2. PLMA Under Dynamic Traffic Demands

The proposed V-PLMA scheme adapts the base F-PLMA approach by dynamically changing a domain's *critical link set* (CLS) to adapt to changes in real-time link-state statistics. For example, Figure 1 shows a 4-node domain with three border nodes (A, C and D), along with two sets of normalized traffic matrices. Specifically, Figure 1a shows a uniform traffic matrix distribution for which the (F-PLMA) CLS is defined as links A-B and C-D. Conversely, Figure 1b shows a non-uniform traffic case, which exhibits significant change in link loadings. As a result, the corresponding CLS is now comprised of links A-B, A-

C, and C-D, respectively. From this illustration, it is clear that the base F-PLMA scheme will not incorporate all congested links in the second case, thereby yielding reduced accuracies in abstracted link-state. In turn, this will result in higher levels of inter-domain request blocking.

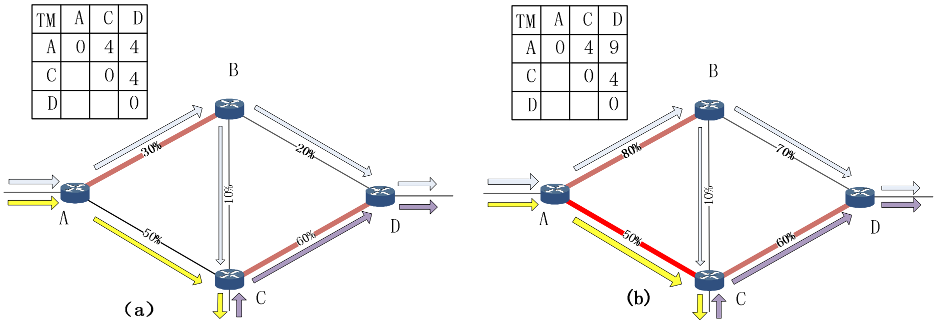

Figure 1: Simple example of changing CLS under dynamic traffic matrix: (a) Initial CLS (brown links) for fixed PLMA under uniform traffic matrix; (b) CLS change (red link) with traffic matrix variation

Now consider the design of the proposed variable PLMA scheme, i.e., V-PLMA. Here, the key concern is how to select the CLS from the link-state statistics under dynamic traffic profiles. Along these lines, first consider an abstract link $e_{(s,d)}$ between border nodes s and d within a given domain. The aggregate bandwidth of this link, $Be_{(s,d)}$, is defined as the maximum available bandwidth along all paths between the border nodes, i.e., termed by the set of routes $\pi \in \varphi_{(s,d)}$. In particular, here the bandwidth of each path is simply defined as the minimum in available bandwidth, b_e, of all links along the route π. Now akin to the basic PLMA scheme in [5], the bandwidth state variation of an abstract link $e_{(s,d)}$ is only determined by monitoring a few (i.e., critical) physical links, denoted by the set $E^p_{(s,d)}$. Using this subset, the aggregate bandwidth of an abstract link is given as:

$$Be_{(s,d)} = \max_{\pi\in\varphi(s,d)} \min_{e\in\pi} b_e = \max_{e\in E^p_{(s,d)}} b_e \qquad (1)$$

where b_e denotes the *instantaneous* available bandwidth, as opposed to the average link bandwidth used in [5]. Now under dynamic traffic matrices, the value b_e can be further defined as a time-varying function, i.e., denoted as $b_e(t)$. Using this function, a simple thresholding-mechanism is proposed to check for changes in the CLS. Namely, the relative change in instantaneous bandwidth in the time interval $[t_1,t_2]$ is first computed, termed as $\rho^e_{(t1,t2)}$, and the CLS recalculated to see if this value exceeds a given threshold τ, i.e.,

$$\rho^e_{(t_1,t_2)} = \frac{|b_e(t_2) - b_e(t_1)|}{b_e(t_1)} > \tau \qquad (2)$$

In general, t_1 and t_2 can represent any update checking/monitoring times. The overall psuedocode for the proposed V-PLMA triggering algorithm (CLS re-computation) is shown in Table 1. Here time is discretized into integral intervals, denoted by k, and the input traffic demand matrix is defined as a function of these intervals, i.e., $\boldsymbol{T}(k)$, Table 1. Hence the V-PLMA algorithm basically loops through all border node pair combinations, i.e., full-mesh abstraction links, and performs the threshold check in Eq. (2). If any threshold change is detected, the appropriate CLS set for that abstract link is revised as per the latest traffic matrix values, otherwise it is left unchanged.

3. Performance Evaluation

The performance of the proposed dynamic triggering policy is analyzed using discrete event simulation with the 16-node NSF topology. In particular, 3 nodes are chosen as domain border nodes, and traffic is injected into the domain following dynamic weighted Poisson distributions. In particular, the traffic weights are regularly varied over time following a preset pattern, yielding good dynamic characteristics. Now in order to gauge the performance of the V-PLMA scheme under dynamic demands, a novel *link monitoring ratio* (LMR), ρ_M, is also introduced, i.e., defined as the number of CLS links, η_{CLS}, divided by the total number of domain links η_{total}, i.e., $\rho_M=\eta_{CLS}/\eta_{total}$. In addition, the blocking probability of inter-domain requests is also used as another metric, defined as the number of rejected request over all incoming requests. The overall results for the F-PLMA and V-PLMA schemes are now presented.

First, Figure 2 plots the LMR values for the two schemes. Here it is seen that the V-PLMA algorithm appropriately increases (decreases) the percentage of monitored links under high (low) traffic loads. However, the scalability of this solution remains good, as the number of monitored links is still bounded to under 20% of links even under high loads (and remains within 5% of the F-PLMA scheme). Next, the blocking probabilities for the PLMA scheme are plotted in Figure 3, along with those for *periodic advertising* (PA) and *threshold-based advertising* (TA). Here the proposed V-PLMA scheme gives very competitive blocking performance, within 5% of that for the non-partial TA strategy (i.e., which advertises all links). This indicates a very high level of domain state accuracy. Moreover, the proposed solution also does significantly better than the F-PLMA scheme at higher traffic loads (with increased demand variations). By contrast, the basic PA scheme gives the highest blocking.

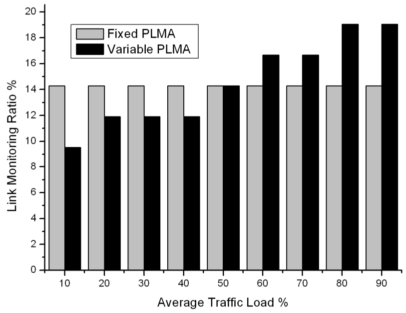

Figure 2: Link Monitoring Ratio comparison

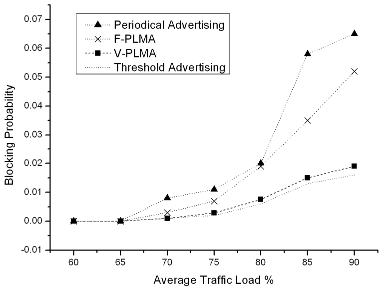

Figure 3: Blocking probability comparisons

```
Algorithm 1: CLS modification threshold
Input: Directed graph G=(V,B,E),
        Discrete traffic matrix T(k),
        Border node pair (s,d)
Output: A subset $E^p_{(s,d)}(k)$ of CLS

begin:
  Get available bandwidth $b_e(1)$ of all links at k=1
  $E^p_{(s,d)}(1) \leftarrow$ PLMA: $b_e(1)$

  while T(k) is not empty
      For every link e in $E^p_{(s,d)}(k-1)$:
          $b_e(k) \leftarrow T(k)$
          $\rho^e(k) \leftarrow | b_e(k) - b_e(k-1) |/ b_e(k-1)$
          if any $|\rho^e(k)| > \tau$
              $E^p_{(s,d)}(k) \leftarrow$ PLMA: $b_e(k)$
          else
              $E^p_{(s,d)}(k) \leftarrow E^p_{(s,d)}(k-1)$
          end if
  end while
end
```

Table 1: CLS modification threshold algorithm

4. Conclusions

A novel inter-domain link-state advertising (triggering policy) scheme for dynamic traffic demand scenarios is proposed, termed as V-PLMA. The solution is fully-compatible with the GMPLS-TE routing framework and can be readily deployed using existing protocols such as OSPF-TE. Although the scheme entails more link monitoring overheads, simulation results show very good scalability and accuracy (i.e., lower blocking). Future efforts will look at extending this solution for operation with other topology aggregation methods (e.g., such as star, bus) and other link-state propagation models, e.g., diversity state.

5. Acknowledgements

This work has been supported by the 973 program under grant number 2010CB328205 as well as by the 863 program and NSFC. The authors are very grateful for this support.

6. References

[1] N. Wang, K. Ho, G. Pavlou, M. Howarth, "An Overview of Routing Optimization for Internet Traffic Engineering," *IEEE Communications Surveys and Tutorials*, 2008.
[2] J. Stewart III, *BGP4: Inter-Domain Routing in the Internet*, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1998.
[3] D. Katz, K. Kompella, D. Yeung, "Traffic engineering (TE) Extensions to OSPF Version 2," IETF RFC 3630, September 2003.
[4] S. Sanchez-Lopez, *et al*, "A Hierarchical Routing Approach for GMPLS-Based Control Plane for ASON," *IEEE ICC 2005*, Seoul, Korea, June 2005.
[5] Y. Yu, Y. Jin, *et al*, "On the Efficiency of Inter-Domain State Advertising in Multi-Domain Networks," *IEEE Globecom 2009*, Hawaii, HI, November 2009.
[6] A. Medina, *et al*, "Traffic Matrix Estimation: Existing Techniques and New Directions", *ACM SIGCOMM 2002*, Pittsburgh, PA, August 2002.