Making Intra-domain Traffic Engineering Resistant to Failures

Baobao Zhang zbb@netarchlab.tsinghua.edu.cn
Jun Bi junbi@tsinghua.edu.cn
Jianping Wu jianping@cernet.edu.cn
Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing 100084, China
Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084, China

Categories and Subject Descriptors
C.2.2 **[Computer-Communication Networks]:** Network Protocols

General Terms
Algorithms, Design.

Keywords
Traffic engineering, Failure recovery

1. INTRODUCTION
The Internet has been changing from transferring text data to multimedia data since the past years. The multimedia applications are sensitive to packet losses. Packet losses are mainly due to link overload and failures. For safeguarding against failures, the backup routing is needed. Accordingly, we call the routing for the normal data transferring the primary routing. For avoiding link overload, TE (Traffic Engineering) is needed. The solution to TE is typically formulated to minimize the maximum link utilization (MLU). We call the MLU before a failure the normal-state MLU. If a failure occurs, the backup routes against the failure will be invoked. The invoked backup routes have been working until the new primary routes are re-computed on the new topology. We call the MLU during the backup routes work the transition-state MLU. To our knowledge, the existing predict-based TE methods optimize the primary routing and the backup routing independently and in this paper it is the first time to optimize both the primary routing and the backup routing based on predicted TE.

The primary routing is intrinsically linked with the backup routing, so in this paper we propose a unified approach, called Unified TE (UTE), to together optimizing the primary routing and the backup routing for both the normal-state MLU and the transition-state MLU as a whole. The emerging Openflow-MPLS [1] provides convenient deployment conditions for UTE. The difficulty of UTE is that the optimal normal-state MLU and the optimal transition-state MLU cannot be always simultaneously achieved. We now take an example. Figure 1 is a topology where the bandwidth of each link is 100. Table 1 shows the traffic demands and the routes on the topology in Figure 1. 'S1' and 'S2' in Table 1 denote two routing schemes. In Scheme 1, the normal-state MLU is 0.6, which is a minimum, but the transition-state MLU when Link C-E fails is 1.2, leading to link overload. In Scheme 2,

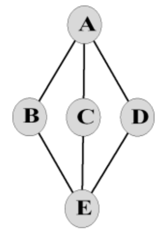


Figure 1. Topology for Table 1

Table 1. Traffic demands and routes

Traffic demands: A→E 180; C→A 60		
	Primary routes:	Backup routes for Link C-E
S1	A-B-E:60; A-C-E:60 A-D-E:60; C-A: 60	C-A-B-E: 30 C-A-D-E: 30
S2	A-B-E:80; A-C-E:20 A-D-E:80; C-A: 60	C-A-B-E: 10 C-A-D-E: 10

the transition-state MLU is decreased to 0.9, which is a minimum, but the normal-state MLU is increased to 0.8. The both states of TE performance are important as follows. Although the sum time during the transition-state period is much shorter than that during the normal-state period, the good transition-state TE performance is very important because the probability of link overload during the transition-state period is much higher than that during the normal-state period. Even though that, the good normal-state TE performance is still very important because if a network often maintains a low MLU, the network operators can delay to upgrade the network infrastructure. How to achieve a reasonable tradeoff is a promising but challenge problem. In this paper, we study the mathematical relationship between the normal-state TE performance and the transition-state TE performance. As an early stage of this work, we only sketch up the ideas of doing tradeoffs. In the future, we will develop the practical and flexible tradeoff system. For better understanding the impact of the primary routing on the backup routing, we compare UTE with the separate TE. The separate TE works as follows. The optimal primary routing is first computed and the optimal backup routing based on the computed optimal primary routing is then computed. For comparison fairness, we let UTE use such a strategy, which is to make the normal-state TE performance optimal. Using simulations, we find that UTE using such a strategy can achieve better transition-state TE performance than the separate TE. Because the number of the optimal primary routings is enormous, the separate TE will arbitrarily select an optimal primary routing while UTE will select the optimal primary routing that makes the backup routing best among all the optimal primary routings. It implies that the primary routing does have a significant impact on the backup routing in the aspect of TE.

In this paper, we only consider the protection against single link failures because [2] shows that 70% failures are single link failures. In the future, we will consider the protection against the single router failures and multiple failures. At present, UTE uses the pro-active TE as the TE module, i.e. computing routes based on a given topology and a given traffic matrix, which can be obtained from the past history. UTE uses the bypass restoration as the failure restoration module, i.e. once a failure is locally detected, the affected traffic will be detoured around the failure.

SIGCOMM'13, Aug 12-16 2013, Hong Kong, China
ACM 978-1-4503-2056-6/13/08.

Besides the pro-active TE and the bypass restoration, there are reactive TE and path-level restoration. In the future, we will explore other combinations of unified approaches. We suppose that the capacity of each link is symmetric in this paper.

2. APPROACH

In this section, we will describe our unified TE approach, UTE. In addition, we study the mathematical relationship between the normal-state TE performance and the transition-state TE performance. In this section, G denotes a topology; D denotes a traffic matrix; X denotes a primary routing; Y denotes a backup routing.

The normal-state TE performance ratio of X on G and D is denoted by $PP(G, D, X)$, which is equal to the ratio of the MLU achieved by X on G and D to the optimal MLU on G and D.

We let $\mathrm{FP}(G, D, X, Y, F)$ be the ratio of the MLU achieved by X and Y on G and D in the transition-state period of the single link failure F to the optimal MLU on the new topology removing the failed link F from G and D.

The transition-state TE performance ratio of a primary routing X and a backup routing Y on G and D is:

$$BP(G, D, X, Y) = \max_{F \text{ is a link failure}} FP(G, D, X, Y, F)$$

The objective of UTE is to minimize $BP(G, D, X, Y)$, where G and D are given constants, X and Y are variables. The problem can be formulated as a linear programming (LP) model as shown in Table 2. The LP problem can be solved in polynomial time. We refer readers to the technical report [3] for the specified formulation of the LP model. In our LP model, we suppose that the link failure is bidirectional.

Table 2. The LP model of UTE

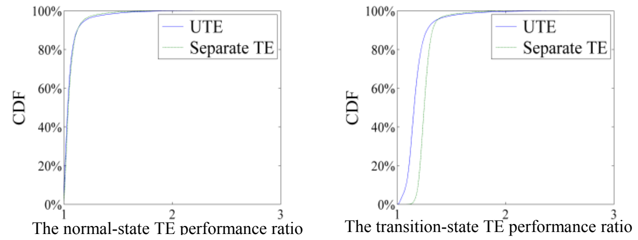
min $BP(G, D, X, Y)$
subject to: X is a primary routing
Y is a backup routing
$PP(G, D, X) \leq t$

We let s be the optimal value of the objective in the LP model of UTE. Obviously, given G and D, s can be represented as a function of t in Table 2. We let the function be $s = f_{G,D}(t)$, where $t \geq 1$. We let $s' = f'_{G,D}(t) = \frac{ds}{dt}$, which is the derived function of $s = f_{G,D}(t)$. We can get some theories and findings on the two functions as follows.

Theorem 1: Function $s = f_{G,D}(t)$ *is a non-strictly monotone decreasing function.*
Proof: see the technical report [3].

Theorem 2: For function $s = f_{G,D}(t)$*, the range of s is:* $1 \leq s \leq 3$.
Proof: see the technical report [3].

Using simulations, we find that $|s'| = |f'_{G,D}(t)|$ is non-strictly monotone decreasing as t increases in most of cases. We refer readers to the technical report [3] for the details. We now sketch up the ideas of doing tradeoff. First of all, we need to study the acceptable performance requirements for the normal-state and transition-state TE. If the both requirements cannot be simultaneously satisfied, the priority level should be given according to the network state and the desire of the network operators. If the normal-state performance needs to be sacrificed for improving the transition-state performance, the mathematical property of $|s'| = |f'_{G,D}(t)$ should be paid careful attention to.



(a) Normal-state comparison (b)Transition-state comparison

Figure 2. The comparisons between UTE and separate TE.

3. EVALUATION

In this section, we will compare UTE with the separate TE. We obtained the real topology and real traffic matrices from the Abilene network [5]. The topology has 12 nodes and 30 links. The traffic matrices are collected every five minutes from 2003-03-01 to 2003-09-10. We use the traffic matrix in the last 5-minite interval to compute the primary and backup routing for the current 5-minite interval. The ratio of the MLU achieved by using the last past traffic matrix to the theoretical limit achieved by using the current traffic matrix is called the TE performance ratio. We computed the normal-state and the transition-state performance ratios for UTE and the separate TE. The comparison results are shown in Figure 2 (a) and (b). The vertical coordinates in Figure 2 denote the CDF (Cumulative Distribution Function) of the performance ratios. We find that the UTE and the separate TE has nearly the same results in the aspect of the normal-state performance. In addition, most of the normal-state performance ratios are very close to 1, which implies that pro-active TE is effective. In the aspect of the transition-state performance, UTE has much better performance than the separate TE.

4. CONCLUSION AND FUTURE WORK

In this paper, we build a LP model to optimizing the primary routing and the backup routing as a whole for TE. In addition, we explore the mathematical relationship between the normal-state TE performance and the transition-state TE performance. In the future, we will develop a reasonable tradeoff system for the normal-state and transition-state TE performance based on their practical significances and their mathematical characteristics.

5. ACKNOWLEDGMENTS

This research is supported by the National High-tech R&D Program ("863" Program) of China(No.2013AA010605), the National Science Foundation of China (No.61073172), and National Key Basic Research Program ("973" Program) of China (No.2009CB320501).. We would like to thank Yanyang Wang and Binyang Liu for their valuable comments on this work. Jun Bi is the corresponding author.

6. REFERENCES

[1] Ali Reza Sharafat, Saurav Das, Guru Parulkar, Nick McKeown, “MPLS-TE and MPLS VPNs with OpenFlow”, In Proc. ACM SIGCOMM, August 2011.

[2] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C. Chuah, and C. Diot, “Characterization of failures in an IP backbone,” in Proc. IEEE INFOCOM, Mar. 2004

[3] Technical Report: http://netarchlab.tsinghua.edu.cn/~zbb/THU-NetArchLab-RTG-TR-UTE-20130517.pdf

[4] The traffic matrices and topology of the Abilene network: http://totem.run.montefiore.ulg.ac.be/datatools.html