

# An Integrated Network Management Framework for Inter-domain Outbound Traffic Engineering\*

Mina Amin, Kin-Hon Ho, Michael Howarth, and George Pavlou

Centre for Communication Systems Research, University of Surrey, UK  
{M.Amin, K.Ho, M.Howarth, G.Pavlou}@surrey.ac.uk

**Abstract.** This paper proposes an integrated network management framework for inter-domain outbound traffic engineering. The framework consists of three functional blocks (monitoring, optimization and implementation) to make the outbound traffic engineering adaptive to network condition changes such as inter-domain traffic demand variation, inter-domain routing changes and link failures. The objective is to keep the inter-domain link utilization balanced under any of these changes while reducing service disruptions and reconfiguration overheads. Simulation results demonstrate that the proposed framework can achieve better load balancing with less service disruptions and re-configuration overheads in comparison to alternative approaches.

## 1 Introduction

Outbound Traffic Engineering (TE) [1,2,3,4] which has become increasingly important and has been well studied, is a set of techniques for controlling traffic exiting a domain by assigning the traffic to the best egress points (i.e. routers or links). The general problem formulation of outbound TE is: given the network topology, BGP routing information and inter-domain Traffic Matrix (TM), determine the best Egress Point (EP) for each traffic demand so as to optimize the overall network performance [2]. Since inter-domain links are the most common bottlenecks in the Internet [2], optimizing their resource utilization becomes a key objective of outbound TE.

In practice, network conditions change dynamically, which can make fixed outbound TE solutions obsolete and subsequently cause some inter-domain links to become congested over time. One such dynamic change is inter-domain traffic variation, which is typically caused by changes in user or application behavior, adaptations from the TCP congestion control or even routing changes from other domains [5]. In addition to these traffic variations, transient and non-transient inter-domain peering link failures might occur. According to [7] transient inter-domain link failures are common events and their duration is less than a few minutes. Upon failure of a peering link, there may be a large amount of traffic shifted to other available EPs, potentially leading to congestion on these new serving EPs if they are not carefully chosen. In

---

\* This work was undertaken in the context of FP6 Information Society Technologies AGAVE (IST-027609) and EMANICS Network of Excellence on the Management of Next Generation Networks (IST-026854) projects, which are partially funded by the Commission of the European Union.

theory, although it is possible to perform outbound TE based on the other proposals in the literature [2,3,4] whenever any of those changes occur, it may require huge computational overheads and a large number of EP re-configurations given that previous proposals have not considered the reduction of reconfiguration changes and overheads. This can lead to excessive service disruptions and is not practical. As a consequence, lack of TE solutions that react to those dynamic changes rapidly enough will leave the network *unmanaged*. It is thus the focus of this paper to make outbound TE more adaptive to fast-changing IP networks by taking into consideration practical network operation and management constraints such as time-efficiency, reconfiguration overheads and service disruptions.

In this paper, we propose an Inter-domain Outbound Traffic Engineering (IOTE) framework that consists of two re-optimization components: (1) Primary Egress Point (PEP) re-optimizer that is designed to manage dynamic traffic variation and routing changes. This component handles primary outbound TE which determines EP selection under Normal State (NS, i.e. no inter-domain link failure); (2) Secondary Egress Point (SEP) re-optimizer that is designed to manage inter-domain link failures. This component handles secondary outbound TE which determines EP selection under Failure States (FS, i.e. transient and non-transient inter-domain link failure). A time-efficient heuristic algorithm is proposed for each optimization component. The overall objective of the **IOTE FRAMEWORK** is, *in spite of dynamic changes in network conditions, to balance the loads among inter-domain links under both NS and FSs, while reducing reconfiguration overheads and service disruptions.*

To the best of our knowledge, there is no such an integrated network management approach like the **IOTE FRAMEWORK** that addresses primary and secondary outbound TE simultaneously. The authors in [6] propose a multi-objective outbound inter-domain TE re-optimization that handles changes of the traffic pattern or routing failures with a minimal burden on BGP. However, they do not consider network performance under transient inter-domain link failures. On the other hand, the authors in [8] propose an intra-domain TE solution that is robust to transient intra-domain link failures and argue that relying on reactive robust solutions may not be appropriate or even feasible, since quickly computing and deploying a new robust solution can be challenging especially in today's large networks. In a similar fashion, changing EP configuration dynamically to avoid a transient failure may not be a practical solution since there is not sufficient time for network operators to configure their networks before recovering from the transient failure. Instead, in order to avoid human configuration and achieve fast recovery from inter-domain link failures, we pursue a proactive robust approach to manage the transient inter-domain link failure through the pre-computation of SEPs.

We compare the performance of **IOTE FRAMEWORK** with two alternative strategies. The first strategy does not consider any PEP or SEP re-optimization at all, while the second only considers PEP re-optimization. In our simulation model, we generate a series of random events to be handled by the different strategies, attempting to emulate realistic changes in network conditions. These events include traffic variation, routing changes and transient and non-transient inter-domain link failures. Simulation results demonstrate that the **IOTE FRAMEWORK** has the following key advantages over the other two alternatives: (a) in spite of network condition changes, maintains a better load balancing on inter-domain links under both NS and FSs; (b) limits the service disruptions and reconfiguration overheads.

This paper is organized as follows. In section 2, we present the proposed **IOTE FRAMEWORK** in detail. Section 3 presents the optimization problem handled by the PEP and SEP re-optimization components. We detail the proposed heuristic algorithms in Section 4. Section 5 presents two alternative strategies for performance comparison. Then, we present our evaluation methodology and simulation results in Sections 6 and 7 respectively. Finally we conclude the paper in Section 8.

## 2 Inter-domain Outbound Traffic Engineering Framework

The proposed **IOTE FRAMEWORK** is illustrated in Figure 1. The key idea of the framework is to continuously monitor<sup>1</sup> the network conditions and, if some optimization triggering policies are met, initiate the PEP and SEP re-optimization components based on the latest network conditions. The PEP and SEP solutions are then finally configured in the network if some implementation policies are met. The framework comprises three functional blocks which we explain in detail:

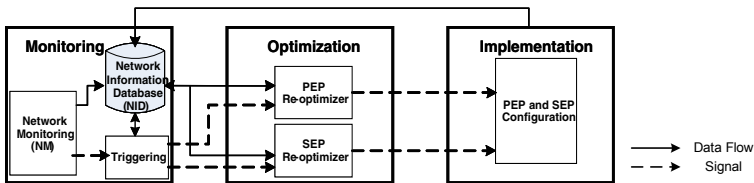


Fig. 1. Inter-domain Outbound Traffic Engineering Framework

**1) Monitoring Block:** It consists of Network Monitoring (NM) module, Network Information Database (NID) and triggering module. The NM module continuously monitors the network conditions to establish a global view of the network. The network information, which will be stored in the NID, includes inter-domain link utilization, overall traffic demand and BGP routing information. The authors in [9] presented a distributed management infrastructure that enables real-time views of network traffic to be generated. The key concept of their approach is that each router monitors its local resources (e.g. utilization of the attached links) and then stores the monitored data in local databases. When a real-time global view of network is needed for network management, the console system that is controlled by the network operator retrieves and processes the information from the databases at each router through specific query languages. To apply this distributed monitoring infrastructure to outbound TE, each egress router monitors the utilization of inter-domain links attached to it and collects the updated BGP routing information from the local Routing Information Base (RIB). On the other hand, each ingress router monitors the updated traffic

<sup>1</sup> In this paper, continuous monitoring can refer to 10 minutes interval according to [5,6]. However, there is a trade off between the accuracy of network conditions and monitoring overheads. In fact, the higher the accuracy of network conditions then the higher the monitoring overheads. The network operators may choose their best strategy to perform the network monitoring.

demand. Note that there are currently several hundred thousand prefixes in the Internet and collecting real-time changes for all the prefixes is challenging. As suggested in [1], TE can consider only a small number of prefixes that is responsible for large volume of traffic. As such, this monitoring block only needs to pay attention to these prefixes in order to significantly reduce the monitoring complexity as well as to make real-time data generation more efficient.

When the latest network conditions are known from monitoring, the NM module signals the triggering module. The triggering module invokes the re-optimization modules in the optimization block if some optimization triggering policies are met. The policy can be event-driven: re-optimization is invoked if an event occurs. In this paper, we use this event-driven policy for triggering the PEP and SEP re-optimizers as follows: (i) The PEP re-optimizer is invoked if the latest network utilization obtained by the monitoring exceeds a tolerance threshold  $\alpha$ . This is a common policy since network providers often take actions to minimize congestion in their networks. Without loss of generality, in this paper we assume  $\alpha=50\%$  to be the borderline of congestion. In summary, the PEP re-optimizer aims to keep the network utilization under NS below  $\alpha$ . (ii) The SEP re-optimizer is invoked when the network information database is updated by the NM. Note that, since the network may suffer from dramatically poor performance under FSs, keeping the SEP solution updated according to the changes is very important.

**2) Optimization Block:** it consists of PEP and SEP re-optimizers and requires as input the latest network information from the NID. The task of PEP re-optimizer is to re-assign the primary egress points to traffic flows under NS. The key objective is to achieve inter-domain load balancing while reducing reconfiguration overheads and service disruptions. The PEP re-optimizer is designed for managing dynamic traffic variation and routing changes. On the other hand, the task of SEP re-optimizer is to pre-compute a set of optimal secondary (i.e. backup) egress points for the traffic. Upon failure of an inter-domain link, the traffic affected by the failure will be shifted to the secondary egress points. The key objective is to achieve inter-domain load balancing under any single inter-domain link failure while reducing backup reconfigurations. The SEP re-optimizer is designed for managing inter-domain link failures. It is worth to mention that changing secondary egress points does not cause service disruption as the primary BGP routes remain intact. Details of the PEP and the SEP re-optimizers will be presented in Sections 3.1 and 3.2 respectively.

**3) Implementation Block:** it enforces the solutions produced by the PEP and SEP re-optimizer into the network according to the implementation policies. For the PEP re-optimization, the solution is enforced if it leads to better inter-domain load balancing than the previous configuration. On the other hand, a benefit-based implementation policy is used for the SEP re-optimization. The SEP solution is enforced if there is a significant gain in reducing the network utilization under FS compared to the previous attempt. The rationale of using this policy is to maximize the lifetime of the previous SEP solution in order to reduce frequent SEP reconfigurations. In fact, we avoid SEP reconfiguration until the latest solution provides a significant performance gain to the network. In this paper, we consider 10% performance gain as the significant gain for the SEP implementation policy. The current PEP and SEP configurations are updated in the NID in order to maintain the latest network information.

One way to implement the PEP and SEP solutions is to assign, for each prefix, the largest and the second largest value of BGP *local-pref* for the selected primary and secondary egress point respectively. To achieve faster failure recovery, the SEPs can be implemented by the proposal in [7] in which an IP tunnel is established to move traffic from the failed PEP to the pre-computed SEP.

### 3 Problem Formulation

Here, we present the optimization problem to be tackled by the PEP and SEP re-optimizer in the IOTE framework<sup>2</sup>. Table 1 shows the notation used in this paper.

**Table 1.** Notation used in this paper

NOTATION	DESCRIPTION
$K$	A set of destination prefixes, indexed by $k$
$L$	A set of egress points, indexed by $l$
$S$	A set of states $S = \{\emptyset \cup (\forall l \in L) \}$ , indexed by $s$
$I$	A set of ingress points, indexed by $i$
$t(k, i)$	Bandwidth demand of traffic flows destined to destination prefix $k \in K$ at ingress point $i \in I$
$Out(k)$	A set of egress points that have reachability to destination prefix $k$
$c_{inter}^l$	Capacity of the egress point $l$
$x_{sk}^l$	A binary variable indicating whether prefix $k$ is assigned to the egress point $l$ in state $s$
$u_s^l$	Utilization on non-failed egress point $l$ in state $s$ . Its value is zero when $s=l$
$U_{max}(s)$	Maximum egress point utilization in state $s$
$U_{worst}$	Worst case maximum egress point utilization across all states
$R, R'$	Primary and secondary egress point reconfiguration limits
$r_{PEP}, r_{SEP}$	The number of actual primary and secondary egress point reconfigurations per re-optimization

#### 3.1 Outbound TE PEP Re-optimization Problem Formulation

The PEP re-optimizer requires the following two items as input: (1) Network utilization: the latest utilization of each inter-domain link together with the maximum and the minimum; (2) The current possibly suboptimal PEP configuration: this includes the latest traffic matrix and BGP routing information. Note that the best EP for each destination prefix according to inter-domain BGP routing policy is known from the BGP routing information.

The task of the PEP re-optimizer is to re-assign the best EPs for destination prefixes, with the objective of balancing the utilization among inter-domain links under normal state ( $s=\emptyset$ ) while reducing reconfiguration overheads and service disruptions. More specifically, the objective of inter-domain load balancing can be achieved by minimizing the maximum inter-domain link utilization. Moreover, inter-domain load

<sup>2</sup> In this paper, we focus the TE re-optimization objective on inter-domain resources due to the reason that capacity over-provisioning is usually employed by ISPs within their IP backbones [10]. In addition, since the objective of this paper is to demonstrate the principle of the outbound TE re-optimization, we consider only the single egress selection [2] as the outbound TE optimization problem. We leave the multiple egress selection as future work.

balancing and reducing EP changes (i.e. reconfigurations) are contradictory objectives: increasing the number of EP changes can improve inter-domain load balancing. In addition, balancing their trade-off is non-trivial. We therefore resort to using the  $\epsilon$ -constraint method [11], which is one of the most favored methods of generating compromising bi-objective solutions. According to the  $\epsilon$ -constraint method, the performance of an objective is optimized, while the other one is constrained so as not to exceed a tolerance value. Hence, we choose to place a constraint on the number of EPs changes that may be attained by the PEP re-optimization while minimizing the maximum inter-domain link utilization. Therefore, the optimization problem to be tackled by the PEP-re-optimizer can be formulated with the objective:

$$\text{Minimize } U_{\max}(\mathcal{O}) = \text{Minimize } \text{Max}_{l \in L} (u_{\mathcal{O}}^l) = \text{Minimize } \text{Max}_{l \in L} \left( \frac{\sum_{k \in K} \sum_{i \in I} x_{\mathcal{O}k}^l t(k, i)}{c_{inter}^l} \right) \quad (1)$$

subject to the following constraints:

$$r_{PEP} \leq R \quad (2)$$

$$\forall k \in K: \sum_{l \in \text{Out}(k)} x_{\mathcal{O}k}^l = 1 \quad (3)$$

$$\forall l \in L, k \in K: x_{\mathcal{O}k}^l \in \{0, 1\} \quad (4)$$

Constraint (2) ensures that the number of EP changes does not exceed the limit  $R$ . A method used in this paper to determine  $R$  is presented in section 6.4. Constraints (3) and (4) ensure that only one EP is selected for each destination prefix as the PEP.

### 3.2 Outbound TE SEP Re-optimization Problem Formulation

The SEP re-optimizer requires as input the current SEP configuration as well as those inputs required by the PEP re-optimizer. The task of the SEP re-optimizer is to re-assign secondary egress points for destination prefixes, with the objectives of minimizing the worst case maximum inter-domain link utilization across all FSs (we assume single inter-domain link failure) while reducing secondary egress point changes. Similar to the PEP re-optimizer, we place a constraint on the number of secondary egress point changes while minimizing the worst case maximum inter-domain link utilization. Therefore, the optimization problem in the SEP re-optimizer can be formulated with the objective:

$$\text{Minimize } U_{\text{worst}} = \text{Minimize } \text{Max}_{s \in S} U_{\max}(s) \quad (5)$$

$$\text{where } \forall s \in S: U_{\max}(s) = \text{Max}_{l \neq s} (u_s^l) = \text{Max}_{l \neq s} \left( \frac{\sum_{k \in K} \sum_{i \in I} x_{sk}^l t(k, i)}{c_{inter}^l} \right) \quad (6)$$

subject to the following constraints:

$$r_{SEP} \leq R' \quad (7)$$

$$\forall k \in K, s \in S: \sum_{l \in \text{Out}(k)} x_{sk}^l = 1 \quad (8)$$

$$\forall l \in L, k \in K, s \in S : x_{sk}^l \in \{0, 1\} \quad (9)$$

$$\forall l \in L, k \in K \quad \text{if } x_{\emptyset k}^l = 1 \quad \text{then } \begin{cases} x_{sk}^l = 1 & \forall s \in S \setminus \{l\} \\ x_{sk}^l = 0 & \forall s = l \end{cases} \quad (10)$$

The term  $x_{sk}^l t(k, i)$  consists of flows that are assigned to EP  $l$  as their PEP and also flows that are assigned to EP  $l$  as their SEP. Constraint (7) ensures that the number of SEP changes does not exceed the limit  $R'$ . A method used in this paper to determine  $R'$  is presented in section 6.4. Constraints (8) and (9) are equivalent to constraints (3) and (4), ensuring that only one EP is selected for each destination prefix as the SEP under each FS. Constraint (10) ensures that if prefix  $k$  is assigned to EP  $l$  under NS, then this prefix remains on  $l$  for all FSs except when the current FS is the failure on  $l$ . Note that, in comparison to the PEP re-optimization that minimizes the maximum link utilization only under NS, the SEP re-optimization optimizes the worst case maximum link utilization across all the FSs as expressed by the objective function (5).

## 4 Proposed Heuristics

**PEP Re-optimization Heuristic:** Local search algorithms have been shown to produce good results for many combinatorial optimization algorithms. We therefore propose an iterative local search algorithm for our heuristic as the following steps:

**Step 1.** Set  $r_{PEP}$  to zero and identify EPs with the maximum and minimum utilization ( $U_{max}(\emptyset), U_{min}(\emptyset)$ ).

**Step 2.** Among all the prefixes whose PEP is the EP with maximum utilization ( $U_{max}(\emptyset)$ ), search for the prefix that by reassigning it to the EP with minimum utilization ( $U_{min}(\emptyset)$ ) would minimize the maximum EP utilization according to objective function (1). Re-assign the prefix to that EP, update both values of  $U_{max}(\emptyset)$  and  $U_{min}(\emptyset)$ , and set  $r_{PEP} = r_{PEP} + 1$ .

**Step 3.** Repeat step 2 until either  $r_{PEP}$  reaches the limit  $R$  or there is no obvious performance improvement for  $U_{max}(\emptyset)$  compared to the previous iteration. We define the threshold of obvious performance improvement to be 5%.

**SEP Re-optimization Heuristic:** Similar to the PEP re-optimization heuristic, we also propose an iterative local search algorithm for our SEP heuristic as follows:

**Step 1.** Set  $r_{SEP}$  to zero, calculate the maximum EP utilization under each potential FS

**Step 2.** Identify the EP  $l'$  with the worst case maximum link utilization  $U_{worst}$  under all FSs (i.e. the link with the highest  $U_{max}(s)$  for all FSs). Calculate the utilization of EP  $l'$  with the minimum link utilization ( $U_{min}(s)$ ) for the state when  $l'$  has the maximum utilization.

**Step 3.** Among all the prefixes whose SEP is  $l'$ , search for the prefix that by re-assigning it to  $l'$  within that state would minimize the worst case maximum EP utilization according to objective function (5). Re-assign the prefix to  $l'$ , update both values of  $U_{max}(s)$  and  $U_{min}(s)$ , and set  $r_{SEP} = r_{SEP} + 1$ .

**Step 4.** Repeat step 2 to 3 until either  $r_{SEP}$  reaches the limit  $R'$  or there is no obvious performance improvement for the worst case performance compared to the previous iteration. We define the threshold of obvious performance improvement to be 5%.

## 5 Alternative Strategies

In this section, we present two alternative outbound TE strategies.

**NO-REOPT:** In this strategy neither the PEP nor the SEP re-optimization is considered. Therefore, in spite of any changes, the current configurations are always used.

**PEP-REOPT-ONLY:** This strategy only considers the PEP re-optimization. Therefore, in case of an EP failure (transient or non-transient), the affected traffic will be shifted according to the current SEP configuration. In comparison to the **NO-REOPT**, this strategy attempts to reactively improve the network performance under non-transient FSs, if the latest network performance obtained by the monitoring violates the threshold criterion. In fact, the PEP re-optimization is triggered to minimize the maximum EP utilization under the particular FS (i.e. in this special case that EP  $l$  has failed we have  $s=l$  instead of  $s=\emptyset$ ). Obviously, this strategy cannot improve the network performance in case of a transient failure due to the very short duration of the failure.

## 6 Evaluation Methodology

### 6.1 Network Topology and Inter-domain Traffic Matrix

Our experiment is performed on a topology with 30 egress routers, each being associated with an inter-domain link. We assume the capacity of all the inter-domain links to be identical. As suggested in [1], TE can focus only on a small fraction of destination prefixes that is responsible for a large fraction of traffic. Hence, we consider 4000 such prefixes in this paper. In fact, each of them may not merely represent an individual prefix but also a group of distinct destination prefixes that have the same set of candidate EPs [12] in order to improve network and TE algorithm scalability. Hence, the number of prefixes we consider could actually represent an even larger value of actual prefixes. We assume that, in the initial network condition, each EP acknowledges reachability to all the considered destination prefixes.

We generate a synthetic inter-domain traffic matrix for our evaluation. The traffic matrix consists of a set of inter-domain traffic flows that originates from each ingress point towards each of the considered destination prefixes. Each inter-domain traffic flow is associated with a randomly generated bandwidth demand according to uniform distribution. We remark that our traffic matrix generation is just our best attempt to model inter-domain traffic, as no synthetic model for actual behavior of traffic in real networks can be found in the literature.

### 6.2 Performance Metrics

The following metrics are used in our evaluation. For all these metrics, lower values are better than high values. The last two metrics are the re-optimization cost metrics.

- **The maximum EP utilization:** This refers to both  $U_{\max}(\emptyset)$  under NS and the  $U_{\max}(s)$  under FS  $s$  in objective functions (1) and (6) respectively.
- **Service Disruption per re-optimization:** A traffic flow (service) is disrupted if it is shifted to another EP due to re-optimization. We represent this



metric by  $SD$  and calculate it by adding the volume of all traffic flows disrupted for the PEP re-optimization.

- **The number of actual PEP and SEP reconfigurations per re-optimization:** These refer to  $r_{PEP}$  in (2) and  $r_{SEP}$  in (7) respectively.

### 6.3 Generated Events

Since no realistic model has been investigated for changes in network conditions, such as traffic variations, routing changes, inter-domain transient failures (TF) and non-transient failures (NTF), we generate a series of random events that attempts to emulate those realistic changes by assigning an occurrence probability to each event. By summarizing several relevant findings in [1,5,7,8], we consider TFs to be the most common event [7]. Hence, we assign the highest probability for TFs. The second highest probability is then assigned for traffic demand variations due to possibly frequent changes of user demands [5]. The lowest probabilities are assigned for routing changes and NTFs due to the stable nature of popular prefixes [1] and rare possibility of fiber-cut which is responsible for NTFs [8]. The performance of all the strategies under these events is investigated in section 7.

### 6.4 Determination of PEP and SEP Reconfiguration Limits

As mentioned earlier, minimizing the maximum EP utilization and reducing EP reconfigurations are contradictory objectives: the larger the number of EP reconfigurations, the better the expected value of the objective functions (1) and (5). To balance this tradeoff we calculate the PEP and the SEP re-configuration limits to restrict the negative effects brought from reconfiguration such as service disruptions and overheads while keeping the objective functions as low as possible.

Figures 2(a)-(b) illustrate the maximum EP utilization under NS and the worst case maximum EP utilization across all FSs respectively as a function of the number of actual EP reconfigurations. In Figure 2(a), the leftmost point on the curve represents a suboptimal PEP selection with maximum EP utilization under NS equal to 50%. In fact, since we chose the PEP re-optimization threshold to be 50% ( $\alpha = 50\%$ ), we generated a suboptimal PEP selection solution with 50% maximum EP utilization under NS. Then this suboptimal PEP selection solution is improved by the **IOTE FRAMEWORK** using the PEP re-optimization heuristic without considering any reconfiguration constraint. The knee on the curve at point (41.2, 100) shows that only a very small load balancing improvement can be attained beyond 100 PEP reconfigurations. Hence, we set the PEP reconfiguration limit ( $R$ ) to 100 and use the current PEP setting as an input for the SEP reconfiguration limit. In Figure 2(b), the leftmost point on the curve represents a suboptimal SEP selection with the worst maximum EP utilization across all FSs equal to 82.4%. Then this suboptimal SEP selection solution is improved by the **IOTE FRAMEWORK** using the SEP re-optimization heuristic without considering any reconfiguration constraint. The steepness of the left part of the curve indicates that large improvements in load balancing under FSs can be attained without large increase in the SEP reconfiguration. The knee on the curve at 400 SEP reconfigurations shows that only a very small improvement can be attained beyond this value. Hence, we set the SEP reconfiguration limit ( $R'$ ) to 400.

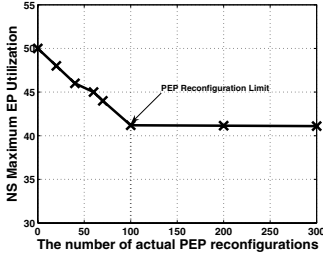


Fig. 2(a). PEP Reconfiguration limit

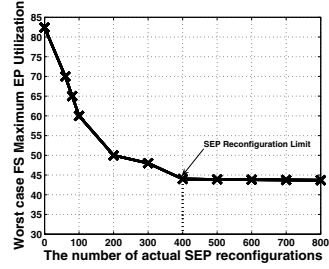


Fig. 2(b). SEP Reconfiguration limit

## 7 Simulation Results

### 7.1 Evaluation of the Maximum EP Utilization

We investigate the performance of all the strategies under several random events generated based on their occurrence probabilities. The randomly generated events are shown in Figure 3 and have the following order: (1) a sub-interval of small traffic fluctuations together with 5 TF. This sub-interval corresponds to positions [0 3]<sup>3</sup>; (2) a sub-interval of gradual traffic increase together with 3 TF, 1 NTF and 1 TF. It corresponds to positions [3 10]; (3) a sudden downward traffic surge, corresponds to position 10; (4) a sub-interval of small traffic fluctuations together with 5 TF. It corresponds to positions (10 13.7); (5) sudden routing changes, corresponds to position 13.7; (6) a sub-interval of small traffic fluctuations together with 7 TF. It corresponds to positions (13.7 18); (7) a sudden downward traffic surge, corresponds to position 18; (8) a sub-interval of gradual traffic increase with 1 TF, 1 NTF and 4 TF. It corresponds to positions (18 23.1); (9) sudden routing changes, corresponds to position 23.1; (10) a sub-interval of gradual traffic decrease with 6 TF. It corresponds to positions (23.1 27.1); (11) a sudden downward traffic surge corresponds to position 27.1 and finally (12) a sub-interval of small traffic fluctuations together with 6 TF. It corresponds to positions (27.1 30]. In addition, Figures 4(a), (b) and (c) show the maximum EP utilization under NS and FS *s* achieved by **No-REOPT**, **PEP-REOPT-ONLY**

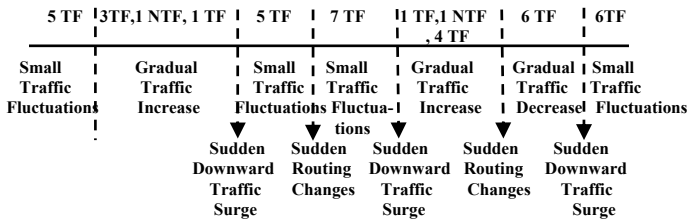


Fig. 3. Randomly generated events

<sup>3</sup> Note that, we will be using interval notations for the remainder of this section. In this notation, a “[” or “]” indicates that the number is inclusive, while a “(” or “)” indicates that the number is exclusive.

and **IOTE FRAMEWORK** respectively. The x axis represents the positions of the random events. All the simulation results presented in this paper are the average of 20 trials.

Figures 4(a)-(c) show that all the strategies perform identically both under NS and FSs until the first time their latest measured performance reaches the PEP re-optimization threshold value (i.e. 50% maximum EP utilization). This is due to our assumption that all the strategies start with the same initial solutions for fair comparison. However, once their measured performance violates the threshold value they start to react differently.

Figure 4(a) shows that **NO-REOPT** is the worst performer under all the events and cannot keep the maximum EP utilization under NS below the threshold value and its maximum EP utilization under FSs has dramatically poor performance. This phenomenon was expected due to the fact that neither PEP nor SEP re-optimization is considered in this strategy and the initial PEP and SEP solutions become less appropriate for the subsequent changes in the network conditions such as accumulation of traffic matrix variations and routing changes.

In contrast, Figure 4(b) shows that **PEP-REOPT-ONLY** can keep the maximum EP utilization under NS below the threshold value by the PEP re-optimization heuristic<sup>4</sup>. However, since no SEP re-optimization is considered in this strategy, its maximum EP utilization under FSs becomes poor and gets worse after subsequent events. Nevertheless, the overall FS network performance degradation in **PEP-REOPT-ONLY** is less severe compared to **NO-REOPT**. This result was expected since **NO-REOPT** does not apply any optimization heuristic as a result the failure of a congested EP and the assignment of its traffic flows over the non-optimized SEP may result in the assignment of a large number of traffic flows over already congested EPs which can cause a huge performance degradation. Whereas, in **PEP-REOPT-ONLY** as a result of an EP failure and the assignment of its flows over the non-optimized SEP does not lead to that much performance degradation due to the fact that the EPs are balanced under NS by PEP re-optimization. Moreover, **PEP-REOPT-ONLY** improves the maximum EP utilization when it exceeds the threshold value after NTFs. In total, Figure 4(b) shows 4 PEP re-optimizations to improve the maximum EP utilization after the traffic variations and routing changes and 2 PEP re-optimizations to improve the maximum EP utilization after the 2 NTFs which results to overall 6 PEP re-optimizations.

However, Figure 4(c) shows that the **IOTE FRAMEWORK** can keep the maximum EP utilization under NS below the threshold value by PEP re-optimization and moreover, can improve the maximum EP utilization both for TFs and NTFs by SEP re-optimization. In fact, its FS worst case performance is respectively 44% and 20% better than the FS worst case performance of the **NO-REOPT** and the **PEP-REOPT-ONLY**. Note that in the **IOTE FRAMEWORK** the maximum EP utilization in FSs is proactively re-optimized for both TFs and NTFs, whereas in the **PEP-REOPT-ONLY**, there is no

<sup>4</sup> Note that in **PEP-REOPT-ONLY** and **IOTE FRAMEWORK**, the maximum EP utilization under NS might exceed the tolerance threshold due to sudden changes. Nevertheless, both strategies are able to minimize the utilization below the tolerance threshold after the PEP re-optimization under the condition where there exist sufficient capacity to accommodate the latest overall traffic demands.

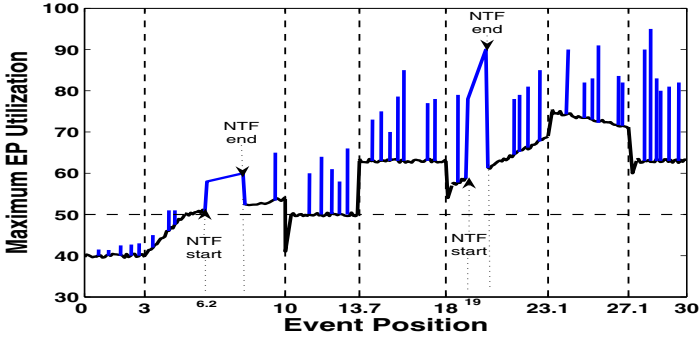


Fig. 4(a). Maximum EP Utilization of **No-REOPT** over event position

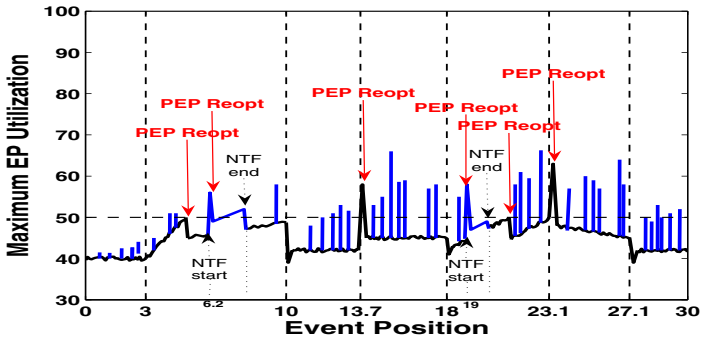


Fig. 4(b). Maximum EP Utilization of **PEP-REOPT-ONLY** over event position<sup>5</sup>

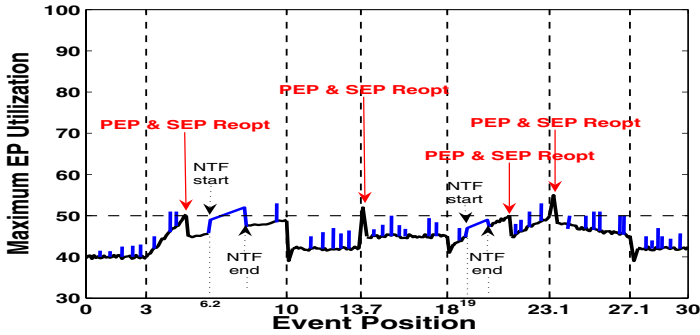


Fig. 4(c). Maximum EP Utilization of **IOTE-FRAMEWORK** over event position<sup>5</sup>

<sup>5</sup> Note that, in some cases, even though the maximum EP utilization violates the tolerance threshold, there is no re-optimization due to the reason that for those cases their re-configuration policy are not met.

re-optimization for TFs due to their very short duration<sup>6</sup> but there are reactive re-optimizations for NTFs. As a result, the significant performance degradation shown in Figure 4(b) due to TFs and NTFs do not occur in Figure 4(c). Furthermore, in the **IOTE FRAMEWORK** the network performance degradation under sudden routing changes is not as serious as the one in the **PEP-REOPT-ONLY**. This result was expected since SEP re-optimization performed in the **IOTE FRAMEWORK** at the earlier stages alleviates the performance degradation compared to no SEP re-optimization in the **PEP-REOPT-ONLY**.

In total, Figure 4(c) shows 4 PEP together with 4 SEP re-optimizations. Overall, the **IOTE FRAMEWORK** has achieved (1) much better performance compared to the **No-REOPT** and almost the same performance as the **PEP-REOPT-ONLY** regarding the maximum EP utilization under NS, and (2) significantly better performance compared to alternative strategies regarding maximum EP utilization under FSs and routing changes.

## 7.2 Evaluation of Re-optimization Cost Metrics

In this section, we compare the re-optimization cost metrics (i.e.:  $SD$ ,  $r_{PEP}$ ,  $r_{SEP}$ ) of the **PEP-REOPT-ONLY** and the **IOTE FRAMEWORK**. Obviously, for the **No-REOPT**, all these cost metrics are zero since this strategy does not perform any re-optimization.

In Table 2, each column represents a re-optimization cost metric while each row corresponds to the  $N^{th}$  re-optimization. In each metric column, the first value (a) corresponds to the **PEP-REOPT-ONLY** and the second value (b) corresponds to the **IOTE FRAMEWORK**. The table shows that, in total, the **PEP-REOPT-ONLY** has higher service disruption and PEP reconfigurations in comparison to the **IOTE FRAMEWORK**. This result was expected since **PEP-REOPT-ONLY** attempts to re-optimize the network performance degradation due to NTFs by PEP re-optimization after the failure, resulting in two more PEP re-optimizations that corresponds to the 2<sup>nd</sup> and the 4<sup>th</sup> re-optimizations. Whereas in the **IOTE FRAMEWORK** the proactive SEP re-optimizations that correspond to the 1<sup>st</sup> and 3<sup>rd</sup> rows take care of both the TFs and the NTFs and result to zero re-optimizations on the 2<sup>nd</sup> and 4<sup>th</sup> rows. Moreover, since **PEP-REOPT-ONLY** does not perform any SEP re-optimization, it requires more PEP reconfiguration for re-optimizing the network performance after sudden routing changes which corresponds to the 3<sup>rd</sup> and the 6<sup>th</sup> re-optimizations in Table 2. Note that at these two re-optimizations the  $r_{PEP}$  have exceeded the PEP limit calculated in section 6.4. The reason is that after the first PEP re-optimization process, the maximum EP utilization under NS is still over the threshold, as a result the re-optimization is triggered again<sup>7</sup>. Whereas in the **IOTE FRAMEWORK** the proactive SEP re-optimizations take care of routing changes and result to less service disruption and re-configurations on the 3<sup>rd</sup> and 6<sup>th</sup> rows.

<sup>6</sup> If TF happens at the time of network conditions monitoring and violates the network performance threshold criterion, the PEP re-optimization is triggered. However, since the TF has a very short duration, it is recovered earlier than the configuration takes place. At this point network operator could simply ignore such re-optimization. In this paper, we assume that the network operator takes care of this task and therefore no re-optimization due to TFs are shown in the graphs.

<sup>7</sup> In fact, in Table 2 on the 3<sup>rd</sup> row the first value of  $r_{PEP}$  is the sum of two numbers of actual PEP re-configurations that correspond to the two consecutive PEP re-optimizations ( $r_{PEP}=100+70=170$ ). Similarly, for the first value of  $r_{PEP}$  on the 6<sup>th</sup> row we have  $r_{PEP}=100+40=140$ .

In summary, the **IOTE FRAMEWORK** incurs almost 40% less service disruptions/PEP reconfigurations in comparison to the **PEP-REOPT-ONLY** at the cost of 4 SEP re-optimizations which result in 710 SEP reconfigurations, to keep the network performance under FSs well balanced. We recall that the SEP reconfiguration does not cause service disruption. In addition, less service disruptions/PEP reconfigurations in our framework may imply better network stability compared to the **PEP-REOPT-ONLY**. Also, in our framework the numbers of actual PEP and SEP reconfigurations per re-optimization have never exceeded their limits.

**Table 2.** Re-optimization cost metrics for (a) **PEP-Reopt-Only** and (b) **IOTE framework**

RE-OPTIMIZATION	SD		$r_{PEP}$		$r_{SEP}$	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
1	9533	9533	70	70	0	150
2	5958	0	40	0	0	0
3	22641	11916	170	90	0	200
4	7150	0	40	0	0	0
5	9533	9533	60	60	0	160
6	19066	13108	140	100	0	200
Total	73881	44090	520	320	0	710

## 8 Conclusion

In this paper, we have addressed the problem of existing outbound TE solutions in case of dynamic changes in network conditions such as traffic variations, routing changes and inter-domain link failures. Hence, we have proposed an Inter-domain Outbound Traffic Engineering (IOTE) framework that aims to balance the load on inter-domain links under both normal and failure states, while reducing service disruptions and reconfiguration overheads. We developed time-efficient heuristics to achieve the framework objectives and compared its performance to two alternative strategies. Our simulation results show that our proposed framework performs better compared to the alternative strategies regarding our objectives. We believe that our work provides insights to network operators on how to keep a balanced network especially under transient and non-transient inter-domain failures in spite of traffic variations and inevitable routing changes by limiting egress point changes.

## Reference

- [1] N. Feamster et al., "Guidelines for Interdomain Traffic Engineering," *ACM CCR*, October 2003.
- [2] B. Bressound et al., "Optimal Configuration for BGP Route Selection," *IEEE INFOCOM*, 2003.
- [3] S. Uhlig et al., "Interdomain Traffic Engineering with Minimal BGP Configurations," *ITC*, 2003.
- [4] K. Ho et al., "Multi-objective Egress Router Selection Policies for Interdomain Traffic with Bandwidth Guarantees," *Proc. IFIP Networking*, 2004.

- [5] R.Teixeira et al, "Traffic Matrix Reloaded: Impact of Routing Changes," Proc. *PAM*, 2005.
- [6] S. Uhlig et al, "Designing BGP-based outbound traffic engineering techniques for stub ASes" *ACM SIGCOMM CCR*, October 2004.
- [7] O. Bonaventure et al., "Achieving Sub-50 Milliseconds Recovery Upon BGP Peering Link Failures," Proc. *ACM CONEXT*, 2005.
- [8] A. Sridharan et al., "Making IGP Routing Robust to Link Failures," Proc. *IFIP Networking*, 2005.
- [9] K.S. Lim et al, "Real-time Views of Network Traffic using Decentralized Management," Proc. *IFIP/IEEE Integrated Network Management (IM)*, 2005
- [10] T. Telkamp, "Traffic Characteristics and Network Planning," *NANOG 2002*.
- [11] V. Chankong et al, *Multiobjective Decision Making-Theory and Methodology*, Elsevier, 1983
- [12] A. Broido et al., "Their Share: Diversity and Disparity in IP Traffic," Proc. *PAM*, 2004.