A Routing Architecture for MPLS-TE Networks

Imene Chaieb, Jean-Louis Le Roux, Bernard Cousin

A Routing Architecture for MPLS-TE Networks

Imene Chaieb and Jean-Louis Le Roux
France Telecom R&D, 2 Avenue Pierre Marzin
Lannion 22307, France
Email: {imene.chaieb, jeanlouis.leroux}@francetelecom.com

Bernard Cousin
IRISA universite de Rennes 1
Rennes 35042, France
Email: bernard.cousin@irisa.fr

Abstract

This paper deals with Multi-Protocol Label Switching-Traffic Engineering Routing (MPLS-TE Routing) Systems which offer key Traffic Engineering features, such as optimization of resources utilization, Quality-of-Service (QoS) and Fast Recovery. Numerous MPLS-TE Routing systems have been defined in the past with their own advantages and drawbacks. This paper proposes a generic architecture for MPLS-TE Routing Systems, with the main objective of helping in classification, analysis and improvement of these systems or the design of new systems. This architecture includes main functions that may be required in an MPLS-TE Routing System. These functions and their interactions are described. Various approaches and options for the implementation and the distribution of these functions in network elements are qualitatively analyzed. A classification of some well known MPLS-TE routing systems is finally proposed.

KEYWORDS

MPLS, Traffic Engineering, Routing, Quality-of-Service (QoS), Network Architecture

I. INTRODUCTION

The emergence of multi-service IP-centric networks, which transport value added services such as VoIP (Voice/Video Telephony over IP), IP TV, Video on Demand and VPN, leads to the requirement for strict QoS delivery (bounded delay, jitter, packet loss), and high availability. Given the huge increase of last miles capacities, and the reduction of the gap between core and access bandwidth, the over-provisioning approaches followed for years by operators in core and backhaul networks, so as to ensure the QoS, are no longer a panacea today in backhaul networks, and may no longer be a valuable approach in core networks at mid term. Hence Traffic Engineering (TE) mechanisms are required so as to optimize network resources utilization, and maximize the amount of traffic that can be transported, while ensuring the quality of service, with as main objective to reduce network costs and postpone investments. In order to address the traffic increase and satisfy the QoS requirements of multimedia applications, various TE mechanisms are proposed, among those MPLS-TE, a connection oriented mechanism based on the MPLS forwarding paradigm, well suited to TE thanks to its explicit Routing capability. The MPLS-Traffic Engineering approach [1] allows setting up explicitly routed Traffic Engineering-Label Switched Paths (TE-LSPs) whose paths satisfy a set of traffic engineering constraints, including bandwidth. MPLS-TE combines explicit routing capabilities of MPLS with a constraint based routing paradigm based on dynamic resources discovery (ISIS-TE [2], OSPF-TE [3]), constrained path computation, and distributed LSP signalling and resources reservation (RSVP-TE) [4]. MPLS-TE ensures Traffic Engineering functions such as network resources optimization, strict QoS delivery, and fast recovery upon link or node failures. TE-LSPs can be used to route traffic flows between network Edge Routers. In order to efficiently route a set of flows in TE-LSPs, additional mechanisms are required on top of the standard MPLS-TE control plane. This includes essentially a TE-LSP utilization function, responsible for an efficient routing of a set of flows in a set of TE-LSPs, along with an adaptability function responsible for adapting the TE-LSPs topology (LSPs resizing/creation/suppression) according to traffic matrix changes and/or topology changes (failures). These mechanisms (utilization and adaptability) are actually intimately linked to the MPLS-TE path computation function. The combination of the base MPLS-TE control plane building blocks (Routing, Path Computation, Signaling) and these additional functions (adaptability and utilization) form together what we call a MPLS-TE Routing System. In the literature, there are papers that focus on MPLS-TE path computation [5], [6], [7]. They propose efficient algorithms to place TE-LSPs in the network and satisfy a pre-defined set of flow requests. There are also studies which account for TE-LSPs utilization [8] and for adaptation mechanisms [9]. Others, interest in flow admission control and its application in MPLS-TE networks [10]. However, a global study that covers the overall architecture of an MPLS-TE Routing System is not considered. The remainder of this paper is built as follows: In section III, we propose a generic architecture to describe the functions of a MPLS-TE Routing System and their interaction. The objectives of this Generic MPLS-TE Architecture are on the one hand to provide a common framework for classifying and analyzing existing MPLS-TE Routing Systems and on the other hand to help in improving existing systems or define new systems. In section IV, we give a qualitative analysis of some design options. Finally in section V, we use the architecture to classify some well know MPLS-TE routing systems.

II. MPLS-TE Trunks and their Utilization

A TE-Trunk is defined as a set of one or more TE-LSPs used to carry an aggregate traffic demand between two points for a given service class [11]. A TE-Trunk is characterized by its reserved bandwidth and a set of TE parameters (e.g. class of service, delay...). For a load balancing purpose, a set of two or more LSPs may be used to route a given aggregate traffic demand between two end points. The TE-Trunk concept allows accounting for such load balancing. We distinguish two main approaches, deployed in service provider networks, for the utilization of TE-Trunks: The *tactical* approach and the *strategic* approach. The *tactical* approach consists in the deployment of a few TE-Trunks, so as to bypass some congested network segments. The *strategic* approach consists in meshing network Edge Routers with a set of TE-Trunks. These TE-Trunks are used to carry all traffic or some specific traffic classes between Edge Routers. In the following, we focus on the strategic MPLS-TE approach.

In the *strategic* approach, TE-Trunks are initially computed using forecast traffic matrix, and then may be adapted in response to traffic changes or topology changes (e.g. network failures). Once these TE-Trunks are configured in the network, they are used to route traffic flows between Edge Routers. As bandwidth reservation in MPLS-TE is purely logical, additional mechanisms are required to ensure bandwidth guarantees. This requires policing mechanisms on Edge Routers so as to ensure that the actual TE-Trunk load does not exceed the reserved bandwidth. Such policing can be done on a per *flow basis* by performing flow admission control (based on the requested flow bandwidth and the available bandwidth within the TE-Trunk) and then policing each flow individually, or on an *aggregate basis* by limiting the aggregate traffic rate. The adaptation of the TE-Trunk size to traffic load changes requires knowledge of the actual TE-Trunk load. This can rely on the measured traffic load on Edge Routers or on the cumulated amount of bandwidth requested by each flow, in a per flow admission control mode. When a TE-Trunk consists of more than one LSP, a load balancing mechanism is required so as to select the LSP to be used to route a given flow.

III. Functional architecture of MPLS-TE Routing Systems

In this section we propose a generic architecture to describe an MPLS-TE Routing System (Fig. 1). This is a functional architecture that helps in covering a large solution spectrum. It includes a set of functions or building blocks. Some of these building blocks are located in routers, others may be located either in routers or in one or more network servers. We distinguish standard MPLS-TE blocks and implementation specific blocks:

- Standard MPLS-TE functions include the TE Topology Discovery ensured by an IGP-TE protocol (either OSPF-TE or ISIS-TE) and the LSP Signalling function ensured by the RSVP-TE protocol. These standard functions are located in routers.
- Implementation specific functions include the TE-Trunk Path Computation, the TE-Trunk Agent, the TE-Trunk Adaptation and the TE-Trunk Utilization functions. These functions may be located in routers or externalized in one or more network servers. They also include the TE-Manager function which is located in a network server.

In the below sections we focus on the five building blocks in charge of resource optimization: TE-Manager (TM), TE-Trunk AGent (TAG), TE-Trunk Computation (TC), TE-Trunk ADaptation (TAD), and TE-Trunk Utilization (TU).

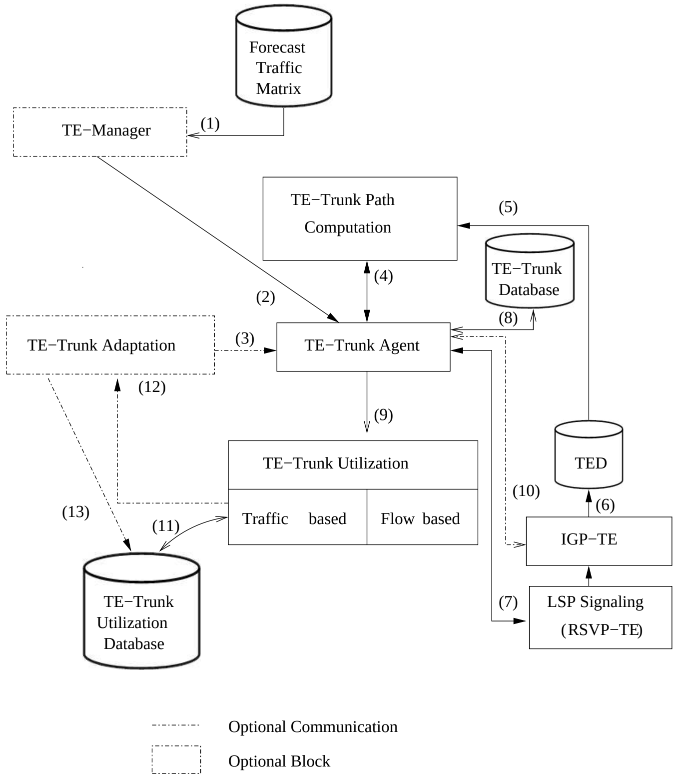

Fig. 1. Generic architecture for MPLS-TE Routing Systems

A. *TE-Manager*

The TE-Manager (TM) is a function which receives the forecast traffic matrix (1), that is the set of aggregate traffic demands between each pair of Edge Routers, and sends TE-Trunk setup/modification/deletion requests to the set of one or more TE-Trunk Agents (2). This function is optional, TE-Trunk definition may be directly done by a network operator and directly configured on Ingress routers.

B. *TE-Trunk Agent*

The TE-Trunk Agent (TAG) is the heart of the architecture. It controls the TE-Trunks establishments/modifications/deletions in the network. It coordinates the actions of the TE-Manager, the TE-Trunk Adaptation, the TE-Trunk Path computation, the TE-Trunk Utilization and the LSP Signalling blocks. It handles TE-Trunk setup/modification/deletion requests sent by the TE-Manager (1), and TE-Trunk modification requests sent by the TE-Trunk Adaptation block (3). It sends TE-Trunk Computation requests to the TE-Trunk Path Computation block (4). Once paths are computed the TE-Trunk Agent sends LSP setup requests to the RSVP-TE module (7) so as to signal the TE-LSPs along the computed paths. Once the TE-LSPs are setup, the TE-Trunk Agent feeds the TE-Trunks Database which contains information related to the established TE-Trunks (TE-Trunk constraints, TE Trunk paths, etc.) (8). It also communicates the established TE-Trunks and their LSPs to the TE-Trunk Utilization block (9). In an Online mode, it may communicate with the IGP-TE (10) and LSP Signalling blocks so as to be notified of network (link/node) and TE-LSPs failures. This communication allows the TE-Trunk Agent to detect failures and call the TE-Trunk Path Computation block so as to reroute the TE-Trunks on paths avoiding failed elements.

C. *TE-Trunk Path Computation*

The TE-Trunk Computation block (TC) is a fundamental building block in MPLS-TE Routing Systems. It has to find TE-Trunks paths by operating on the Traffic Engineering Database (TED) (5) fed up by the IGP (6) and considering the TE-Trunks constraints. It handles Trunk Computation requests sent by the TE-Trunk Agent. A request may correspond to a single TE-Trunk or to a set of TE-Trunks. The request may be a Trunk setup request or a Trunk modification one. The output for a given Trunk is a path or a set of paths whose cumulative bandwidth fits the Trunk(s) request.

D. *TE-Trunk Adaptation*

The TE-Trunk Adaptation block (TAD) is in charge of adapting the TE-Trunk size to the actual traffic load. It increases TE Trunks size (i.e. it increases the amount of bandwidth reserved for the TE-Trunk), so as to anticipate congestion issues, when the load between a pair of nodes increases, and decreases TE Trunks size so as not to waste unused bandwidth when the load between a pair of nodes decreases. Verification of the TE-Trunk load can be done in a *timer driven* manner, in which case the TE-Trunk load in the TE-Trunk utilization databases is periodically checked by the Adaptation block (13) or it can also be done in an *event driven* manner, in which case the TE-Trunk Utilization block notifies the Adaptation block that a TE-Trunk is congested or is going to be congested (12). Note that this block is optional and may not be used in every MPLS-TE Routing Systems.

E. *TE-Trunk Utilization*

The TE-Trunk Utilization block (TU) is in charge of (1)- IP routing within TE-Trunks, (2)- mapping of incoming flows within TE-LSPs (LSP selection among the set of Trunk's LSPs), (3) traffic policing within TE-Trunks (rate limiting), (4)- checking the actual TE-Trunk load (5)- potential flow admission control within TE-Trunks. IP routing within TE-Trunk may rely on the IGP, on BGP or on static routing. Mapping of a flow within a Trunk can be done either in advance, in a flow admission based scenario, or in real time by relying on a hash function that respects flows. Traffic policing, that is, rate limiting on Trunk Edge Routers can be done on a *per-flow basis* or on *an aggregated traffic basis*. The control of the TE-Trunk load consists in updating the TE-Trunk utilization Database with the actual TE-Trunk load. This database contains the current load of all LSPs starting on the routers. Updating the TE-Trunk Utilization database with the actual traffic load within LSPs, can be achieved by measurement on the Edge Routers or by cumulating the requested bandwidth of each admitted flow, when flow admission control is performed. This information can then be used by the TE-Trunk Adaptation block so as to anticipate congestion within a Trunk and ensure adaptation to traffic matrix changes. It can also be used for flow admission control. In a flow aware mode, the TE-Trunk Utilization block can handle flow admission requests; upon reception of a flow admission request, it consults the TE-Trunk Utilization database (11), and if there is an LSP with sufficient resources, the new flow is accepted and the TE-Trunk Utilization database is updated. Else, the flow admission request is rejected. Alternatively when there is no TE-Trunk with enough available bandwidth, the TE-Trunk Utilization block may ask the TE-Trunk Adaptation block to increase a TE-Trunk size in order to provide residual capacity for the new flow (12).

IV. Applying the architecture: Functions distribution and implementation options

In the previous section we proposed a functional architecture for MPLS-TE Systems, which includes in addition to standard MPLS-TE blocks, specific blocks such as TE-Trunk Computation, TE-Trunk Adaptation and TE-Trunk Utilization. This architecture may help in classifying MPLS-TE mechanisms and improve the design of MPLS-TE Systems. A given MPLS-TE Routing System corresponds actually to a specific implementation of this architecture. The blocks of this generic architecture may be located in different elements (Centralized in Network servers or Distributed in Edge Routers). The performances of an MPLS-TE Routing System, in terms of scalability, reactivity and optimality depend on the distribution of these blocks. Before discussing several options for the distribution and implementation of these functions, a description of some classification criteria, which will be used in the discussion, is provided.

A. *MPLS-TE classification criteria*

Several criteria are identified to arrange the various modes and options for MPLS-TE Routing Systems. We distinguish the following:

1) Time Scale:
   - Offline (Off): TE-Trunks are computed and established periodically based on forecast traffic matrices. This mode allows more time for path computation. This implies that there is no TE-Trunk Adaptation

and no LSP re-routing upon network failures. In this mode, recovery can be ensured by pre-computing and pre-establishing backup TE-Trunks.
- Online (On): TE-Trunks are modified (TE-LSPs resizing, LSPs re-routing, LSPs creation/deletion) according to traffic matrix evolution, or network failure. In such mode, path computation time should be minimized.

2) Path Computation Method:
   - Coordinated (Coo): TE-Trunk paths are computed taking into account all TE-Trunks requests.
   - Uncoordinated (Unc): The path(s) of TE-Trunks originated by a given head-end router are computed without taking into account TE-Trunks originated by other head-end LSRs.

3) Function Distribution:
   - Centralized (Cen): The function is located in a single computing element
   - Distributed (Dis): The function is distributed on multiple computing elements.

B. Function Distribution

We discuss in this section the distribution of each architecture's function and its impact on the performances of an MPLS-TE System. Some functions should be only Distributed, others only Centralized and others can be either Distributed or Centralized. When two functions are not located within the same element (e.g. One is located in an Edge Router and the other is located in a TE server) a standard communication protocol is required to manage the relationships and cooperations between the two functions. In contrast, when two functions are located in the same element (e.g. Edge Router) there is no need for any communication protocol, and their relationship may simply rely on a software interface (e.g. Inter Process Communication API).

- The MPLS-TE protocols (RSVP-TE and IGP-TE) are Distributed on the routers (note that the IGP-TE may passively run on the Path Computation Block when it is Centralized, so as to feed the TED).
- The TE-Trunk Utilization block should be Distributed as it is in charge of the routing of incoming flows within the TE-Trunks and of TE-Trunk load measurement on Edge Routers. The centralization of this block would definitely affect the reactivity of the MPLS-TE System due to the amount of information to be communicated between Edge Routers and the TE-server.
- By definition, the TE-manager is always Centralized.
- The TE-Trunk Agent can either be Centralized on a network server or Distributed on Edge Routers. In a Distributed mode, it maintains only TE-Trunks for which it is the head-end. In a Centralized scenario, the TE-Trunk Agent has a global knowledge of all the TE-Trunks, and a communication protocol is required to communicate with Edge Routers (with the LSP signalling process) so as to trigger LSP setup and retrieve failure events. This may rely for instance on a standard management protocol (e.g. SNMP [12]). Note that in the Centralized mode, the notification of network failure events should be event-driven (e.g. SNMP traps) so as to minimize the amount of information to be communicated between the Edges Routers and the TE-server.
- The TE-Trunk Adaptation function should always be linked to the TE-Trunk Agent, that is if the TE-Trunk Agent is Centralized (respectively Distributed), the TE-Trunk adaptation is also Centralized (respectively Distributed). When the TE-Trunk Adaptation is Centralized, a communication protocol is required between the TE-Trunk Utilization block located in Edge Routers and the TE-Trunk Adaptation, so as to retrieve the LSP load. Such dynamic discovery of LSP load should be event-driven so as to minimize the amount of information communicated between Edge Routers and the TE-server (The TE-Trunk Utilization block sends a message to the Adaptation block only whenever a threshold is reached; this avoids the Adaptation block to periodically consult the TE-Trunk Utilization database). The separation of these two functions (the TE-Trunk Adaptation and the TE-Trunk Agent) would not bring any value and would require communicating a lot of information.
- The TE-Trunk Path Computation block may be Distributed or Centralized. (1) If the TE-Trunk Agent is Centralized, the TE-Trunk Path Computation block should also be Centralized (Coordinated mode) because the separation of these two functions would not bring any value and would require the communication of a lot of information (but note that this does not mean that they are collocated). (2) But, if the TE-Trunk Agent is Distributed, the TE-Trunk Path Computation block may either be Distributed or Centralized. When the TE-

Trunk Agent is distributed and the TE-Trunk path Computation is Centralized, the TE-Trunk Path Computation remains Uncoordinated because the set of TE-Trunk Agents send requests independent from each other.

C. *Various implementation options*

There are several approaches for MPLS-TE Routing Systems, which correspond actually to several combinations of the criteria listed above. We discuss below some of these approaches which correspond to particular implementations of our architecture.

1) The Off/Cen/Coo MPLS-TE approach: We consider here an Offline mode based approach, where the TE-Trunks paths, potentially including backup paths, are pre-computed periodically without real time computation constraints, in a Coordinated manner. The Fig. 2 shows the distribution of MPLS-TE functions in such approach . The cursor in position P1, separates the blocks which are Centralized on a TE-server (above the cursor) from those which are Distributed on Edge Routers (under the cursor). There is no TE-Trunk adaptation as we are in an Offline mode. The TE-Trunk Path Computation, the TE-Trunk Agent and the TE-Trunk Adaptation blocks are Centralized (actually the TED is maintained on routers and on the server). Other blocks: TE-Trunk Utilization, IGP-TE, RSVP-TE are located in Edge Routers (actually IGP-TE and RSVP-TE are located in all routers). In this approach, the placement of TE-Trunks can be drastically optimized because the TE-Trunk Path Computation function knows all the requests and can perform a Coordinated path computation, with no time limitation. In return, by definition this Offline approach does not allow reacting upon traffic matrix change or network failure and this may lead to packet loss upon congestion, or service disruption upon failure. Also even if this approach allows the pre-establishment of backup paths, such protection may not work upon a multiple failure cases and hence faces some robustness limitations. Thus, the Off/Cen/Coo approach obviously suffers from a lack of survivability and reactivity as it does not allow adaptation to traffic variations and topology changes.

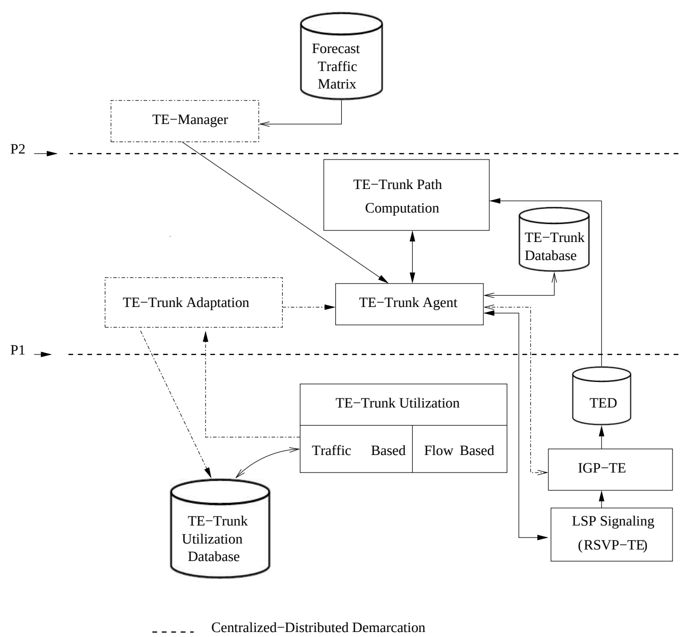

Fig. 2. The Off/Cen/Coo and On/Dis/Unc MPLS-TE mechanisms

2) The On/Cen/Coo MPLS-TE approach: If we add a TE-Trunk Adaptation block and a TE-Trunk re-routing capability to the previous system, it ends up with an Online mode and this results in an On/Cen/Coo MPLS-TE System. Such approach can operate but it may require a lot of information exchanges between Edge Routers and the TE-server and suffers from a limited reactivity. Indeed, the recovery upon a network failure implies the following sequences: (1) Failure discovery on the TE-server, (2) Coordinated path computation and (3) Communication of the new paths to all Edge Routers. Such sequence may take a long time, particularly when a lot of Trunks are impacted.

3) The On/Dis/Unc MPLS-TE approach: In this Online approach, the MPLS-TE Routing System contains an adaptation function. The requests are handled in an Uncoordinated manner. The TE-Manager remains Centralized, however all other architecture's blocks are Distributed on the Edge Routers, which corresponds to the position P2 for the centralized/distributed cursor in the Fig. 2. This approach offers better reactivity and recovery time than the On/Cen/Coo approach thanks to its On/Dis scheme. In fact, in addition to the reactivity offered by the online mode, all functions of a given MPLS-TE System are located in the same element (the Edge Router) and hence this does not require heavy communication between Edge Routers and the TE-server. In return, it affects the optimality because of the Uncoordinated scheme. The higher the level of reactivity and robustness is, the more affected the optimality is. Note that here the optimality may be affected by the order in which the TE-Trunk/LSP paths are computed. If instead of adopting an Uncoordinated scenario, the Coordinated mode was used, this would result in an On/Dis/Coo MPLS-TE System.

4) The On/Dis/Coo MPLS-TE approach: This approach implies that each Edge Router has to be aware of all the TE-Trunk requests and operate in a Coordinated manner with other network Edge Routers. Thus, network resources usage is optimized but not as well as if the Offline mode was used (because of computation time constraints). The limitation of this option is obviously the scalability, Edge Routers are likely to be saturated, because, each Edge Router has to exchange all the information about its own TE-Trunks/LSPs with all other Edge Routers. This cannot scale because of the number of TE-Trunks in the network and their activity (resizing rates,...). So, this approach seems not relevant.
The analysis carried above shows that each approach has its own advantages and drawbacks. Let's us discuss in the following section another approach that brings together conflicting modes to take advantages of each of theme.

5) The hybrid MPLS-TE approach: This approach is based on a hybrid scheme: On/Dis/Unc-Off/Cen/Coo approach. In this case, a TE-Trunk layout is computed and setup periodically (e.g. weekly) in an Offline mode. Between offline reoptimization periods, the Online mode is used and traffic matrices and topology changes are handled dynamically. This mechanism maintains two TE-Trunk Agents: A Centralized TE-Trunk Agent without adaptation function and a Distributed TE-Trunk Agent with an Adaptation function. It maintains also two Path Computation blocks linked to the two TE-Trunk Agents: A Centralized and a Distributed one (Fig. 3). Others blocks are located only in Edge Routers.
This approach can be applied to a protected MPLS-TE layout where primary TE-Trunks are protected by local fast reroute backup TE-Trunks; primary TE-Trunks are established in an Online manner to ensure as much adaptability and reactivity as possible, while Backup TE-Trunks are established in an Offline manner, without real time computation constraints ,so that their path can be optimized as much as possible, in order to avoid backup resources starvation.

Such combination of two MPLS-TE approaches seems really relevant as it takes advantage of the optimality offered by the Offline, Centralized and Coordinated modes and the reactivity, robustness, scalability and survivability assured by the Online, Distributed and Uncoordinated modes. In return, this approach raises issues in terms of cooperation between the two modes.

The key challenge raised by this hybrid scheme relies on how to manage the handover between Offline and Online operations. A difficulty can arise when, in the Offline mode, the resource reservation for TE-LSPs is not explicitly done. When the reservation is not done, the Online mode ignores the resources taken by the Offline mode. A solution can consist in using two bandwidth pools on the same link to separate Online resources from Offline resources. These pools must be dynamically adjusted to avoid blocking the resources of a pool while the other pool is under-utilized. Path computation has to take the bandwidth pool type into consideration. Another problem deals with the morphing of TE-Trunks between an Online and an Offline sequence (e.g. The migration from the Online mode to the Offline one). It consists, in this case, in controlling the transfer of a TE-Trunk management between a Distributed TE-Trunk Agent and a Centralized TE-Trunk Agent. Two scenarios for TE-Trunk management can be distinguished (Fig. 4): (1) Separate the management of the TE-Trunks which are created in Centralized mode (Offline TE-Trunks) from those which are established in Distributed mode (Online TE-Trunks). This approach allows each Agent to manage separately its TE-Trunks. It implies two distinct types of TE-Trunks (Offline TE-

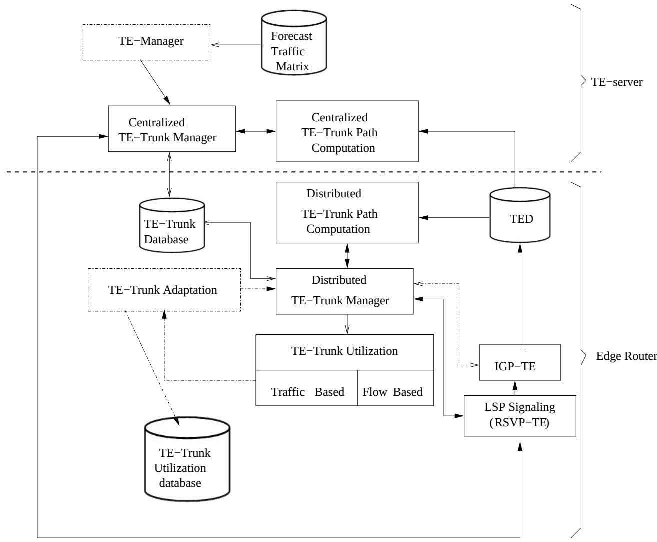

Fig. 3. Hybrid MPLS-TE Approach

Trunks and Online TE-Trunks) and two distinct TE-Trunk databases. (2) Centralized and Distributed control of the same TE-Trunk. In this approach, the management of a given TE-Trunk is shared between the Centralized and the Distributed TE-Trunk Agents. This requires more care in the cooperation between the two TE-Trunks Agents (single database). In both scenarios, an Inter-Agent Communication Protocol (IACP) (3) would be useful, so as to synchronize the two Agents and coordinate the migration between the Offline and Online modes. In case of separate TE-Trunk management the IACP protocol would have to carry information related to the creation, suppression and evolution of Online TE-Trunks, from the Distributed Agent to the Centralized Agent. It would also have to carry notification of an Offline re-optimization, from the Centralized Agent to the Distributed Agent. In case of common TE-Trunk management, the IACP protocol would have to carry information allowing a transfer of TE-Trunk ownership between Centralized and Distributed Agents. For instance this could rely on a master-slave scheme, where the Centralized Agent is the master and notifies the Distributed Agent when it needs to take control of the TE-Trunks, i.e. when an Offline optimization is triggered.

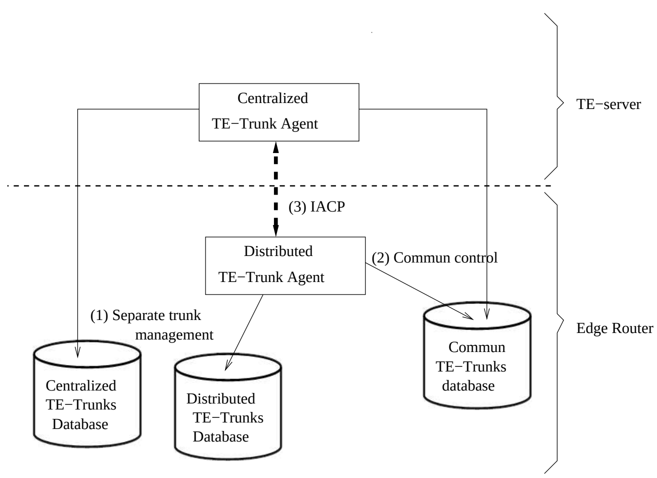

Fig. 4. TE-Trunks management

V. Classification of some existing MPLS-TE mechanisms based on the architecture

Various TE mechanisms have been proposed in the literature. Our generic architecture and our classification criteria can be used to classify these mechanisms. For the sake of illustration, we discuss below five distinct MPLS-TE Mechanisms. The well known ”autobandwidth” mechanism [13] is an On/Dis/Unc TE mechanism, which is actually implemented by a few router vendors. Path computation (TC) relies on a CSPF algorithm performed by the Ingress Router. The MPLS-TE coordination function (TAG) runs on the Ingress router. The Ingress router monitors the traffic rate within TE-Trunks (TU) and periodically runs an adaptation function (TAD) that decides whether TE-Trunk resizing is needed or not. Such decision relies on bandwidth thresholds; if an upper/lower threshold is reached the TE-Trunk bandwidth is increased/decreased accordingly. A hysteresis is used so as to avoid oscillations. Another mechanism based on an Off/Cen/Coo server is proposed in [14]. It is a network planning, design and management system which includes only TM, TAG and TC functions.

MATE [8] is an On/Dis/Unc TE mechanism which aims mainly to avoid network congestion by adaptively balancing the load among multiple paths (TAD function) based on measurement and analysis of path congestion (TU function). **MATE** assumes, that a set of TE-LSPs between each ingress-egress pair are initially established (there is no TC function).

RATES [15] is an On/Cen/Unc MPLS-TE server which can set-up bandwidth guaranteed LSPs between Edge Routers by relying on a minimum-interference routing algorithm (TC function). This algorithm tries to minimize the interference between different routes in a network. This process involves the computation of maximum flow values for all ingress-egress pairs, computation of the so-called critical links (links that are likely to suffer from interference), and the computation of weights to be used for a weighted-shortest path algorithm that chooses the final route. **RATES** includes also a module to communicate with the network and to make decision to compute and to establish LSPs (TAG function).

The **TEAM** architecture [16] defines an On/Cen/Unc automated manager for DiffServ/MPLS networks. **TEAM** is designed for an entirely automated management based on real time measurements of the network state. **TEAM** is an adaptive manager that provides the required Quality of Service to the users by reserving bandwidth resources, and reduces the congestion in the network by distributing the load efficiently. **TEAM** is composed of: (1)- A Traffic Engineering Tool (TET), which adaptively manages the bandwidth and routes in the network. This tool provides LSP routing (TC), LSP preemption and LSP capacity allocation (fine-tuning the LSP capacity allocation to avoid unused reserved bandwidth) (TAD) and traffic measurement in trunks (TU) by proposing an on-line scheme called Estimation and Prediction Algorithm for Bandwidth Brokers (EPABB). An optimal policy to determine and adapt the MPLS network topology based on the traffic load is used (no full meshed network). The TET includes the decision function to establish/fine-tune/preempt a LSP (TAG). (2)- A Measurement and Performance Evaluation Tool (MPET) which measures important parameters in the network and inputs them to the TET, and (3)- a Simulation Tool (ST) which may be used by TET to consolidate its decisions.

The table I summarizes the classification of the five TE mechanisms described above.

TABLE I
Classification of some existing MPLS-TE mechanisms based on the architecture

	TM	TAG	TC	TAD	TU
Autobandwidth(On/Dis/Unc)		x	x	x	x
IP/MPLSView(Off/Cen/Coo)	x	x	x		
MATE (On/Dis/Unc)				x	x
RATES (On/Cen/Unc)	x		x		
TEAM (On/Cen/Unc)		x	x	x	x

VI. Conclusion

MPLS-TE is being deployed by network operators to better optimize their network resources. Routing in MPLS-TE networks is a large and open issue. Studies aimed to improve the MPLS-TE Routing in terms of scalability, stability, robustness, optimality and survivability. In this paper, we have proposed a generic architecture for MPLS-TE Routing Systems, that combines a set of MPLS-TE building blocks such as TE-Trunk Computation, TE-Trunk Utilization and TE-Trunk Adaptation. The goals of this generic architecture are to ease the classification of existing MPLS-TE Routing solutions, and to help in improving existing solutions or designing new solutions.

We have illustrated the interest of such architecture with (1) a qualitative evaluation of various MPLS-TE implementation approaches, (2) the identification of the needs for an Inter-Agent Communication Protocol (IACP) to facilitate the synchronization between Online and Offline control functions in a hybrid mode, and finally (3) a classification of existing approaches.

Further works will focus on a quantitative evaluation of the discussed implementation options with regards to a set of evaluation metrics: optimality, scalability, reactivity, stability and robustness. Also, a functional specification of the IACP protocol will be performed.

Acknowledgement

The authors would like to thank R. Moignard, P. Chemouil, B. Jabbari and Y. Le Louedec for their helpful comments and suggestions.

References

[1] D. Awduche and B. Jabbari, "Internet traffic engineering using multi-protocol label switching (MPLS)," *Computer Networks*, vol. 40, no. 1, pp. 111–129, September 2002.
[2] H. Smit and T. Li, "IS-IS extensions for traffic engineering," RFC 3784, June 2004.
[3] D. Katz, K. Kompella, and D. Yeung, "Traffic engineering (TE) extensions to OSPF version 2," RFC 3630, September 2003.
[4] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP tunnels," RFC 3209, December 2001.
[5] T. L. M. Kodialam, "Minimum interference routing with applications to MPLS traffic engineering," in *INFOCOM*, 2000.
[6] W. Szeto, R. Boutaba, and Y. Iraqi, "Dynamic online routing algorithm for MPLS traffic engineering," *The International Journal of Network and Systems Management*, vol. 10, no. 3, pp. 936–946, 2002.
[7] P. W. S. Suri, M. Waldvogel, "Profile-based routing: A new framework for mpls traffic engineering," *Quality of future Internet Services*, no. 2156, pp. 138–157, 2001.
[8] A. Elwalid, C. Jin, S. H. Low, and I. Widjaja, "MATE : MPLS adaptive traffic engineering," in *INFOCOM*, 2001.
[9] K.-T. K. S. Phuvoravan, "Fast time scale control for MPLS traffic engineering," *CS-TR 4351*, April 2002.
[10] A. Bosco, "Edge distributed admission control in MPLS networks," *IEEE Communications Letters*, vol. 7, no. 2, pp. 88–90, Febrary 2003.
[11] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus, "Requirements for traffic engineering over MPLS," RFC 2702, September 1999.
[12] J. Case, M. Fedor, M. Schoffstall, and J. Davin, "A simple network management protocol (SNMP)," RFC 1098, May 1990.
[13] "Cisco MPLS autobandwidth allocator for MPLS traffic engineering: A unique new feature of cisco IOS software," *White paper*. [Online]. Available: www.cisco.com/
[14] Ip/mplsview. [Online]. Available: www.wandl.com/
[15] P. Aukia, M. Kodialam, P. Koppol, T. Lakshman, H. Sarin, and B. Suter, "RATES: a server for MPLS traffic engineering," *IEEE Network*, vol. 14, p. 34 41, March-April 2000.
[16] J. C. d. O. I. F. A. C. Scoglio, T. Anjali and G. Uhl, "TEAM: A traffic engineering automated manager for diffserv-based MPLS networks," *IEEE Communications Magazine*, vol. 42, p. 34 41, October 2004.