An Automated Traffic Engineering Algorithm for MPLS-Diffserv Domains

By
Dr. Junaid Ahmed Zubairi
SUNY at Fredonia,
Fredonia, New York
Institute of Business Administration,
University of Karachi, Pakistan
August 2002

Outline of Presentation

- Traffic Engineering
- MPLS and Diffserv
- Need for Automated Provisioning
- TELIC: Colorful Traffic Engineering
- TELIC: Focus, Method and Results
- Conclusion

Traffic Engineering

- TE optimizes the network efficiency with the control of the
  - Mapping
  - Distribution
- Of the traffic across the network
- TE tries to balance the load across the network and addresses fault tolerance and congestion avoidance

Traffic Engineering

- Earlier, the routing protocols favored shortest or least cost paths, building up congestion on some paths
- TE was not practiced, leaving the network overloaded in some parts and underutilized in others
IETF’s New Approach

- New applications on the Internet seek timely delivery besides correctness
- IETF has orchestrated the development of new service sensitive protocols for the Internet
- ISP organizations wish to provide most appropriate services for their customers and gain profits
- They want to utilize their bandwidth most efficiently

IETF, ISP’s and Users

- The conventional Internet has been following the “make effort not promises” model, with all the traffic treated in almost the same manner
- The new protocols intend to change it and provide suitable services to various classes of traffic

Motivation for MPLS

- MPLS (Multi Protocol Label Switching) is a very interesting recent development
- Let us see why MPLS was developed
- ATM switches are deployed in the Internet backbones due to their extremely fast switching and provisioning
- All Internet traffic is based on IP. So IP must be carried over the ATM

IP/ATM → MPLS

- Classical IP over ATM (Overlay model) suffers from several problems
- First, all ATM switches are connected in a mesh. A small increase in the number of switches can drastically increase the number of virtual circuits
- The QoS features of ATM are not exploited and all connections are best effort
- IP and ATM have incompatible addressing and control protocols so overlaying is expensive
The industry developed tag switching and label switching to solve the above problems. In label switching, a short fixed length label is encoded into the packet. The intermediate LSR (Label Switched Router) finds the next hop from a table, using the label as an index. If the LSR is an ATM switch, label is just the VPI/VCI identifier. If the LSR is an IP router, the label helps eliminate the destination based routing and reduces the job of the router to label switching.

A label switched path (LSP) must be set up prior to the start of transmission. IP and ATM are tightly integrated with label switching. IP takes over the control path and ATM switches are used only for data transmission. IP can use the ATM switches as label switched nodes (or IP routers).

IP Over ATM With MPLS

Structure of a Label Switching Router

- Layer 1 (IP) performs routing and label distribution
- Layer 2 (ATM) performs fast forwarding
LSP Hierarchy

Tunnel Hierarchy and Label stacks

In ATM, two levels of hierarchy

Label stack allows for arbitrary levels of hierarchy

13

LSP’s in an MPLS Network

14

MPLS

MPLS simplifies the routing problem in an all IP subnet
An MPLS domain has an ingress node that nails down paths through the maze of core routers for every requesting flow until the exit door (egress node)
Thus every router does not have to decide about the path of each packet
In MPLS, the connectionless network is converted into connection oriented network

15

MPLS

Intermediate routers use a “shim header” or a layer 2.5 header to find out the next hop of a packet
This shim header is inserted between the frame header and packet header
It is used by the router to consult a table that tells what path is to be taken for this packet

16
**MPLS**
- Instead of routing, now the routers do label switching
- Since the path is pre-determined, routers can speed up the processing of packets
- Also, the network manager can decide LSP’s (label switched paths) based on load distribution and other administrative goals
- Directing traffic on paths not determined by traditional IGP’s provides flexibility and load balancing

**MPLS**
- MPLS (Multi-Protocol Label Switching) provides LSP (Label Switched Paths) to the requesting traffic trunks
- MPLS runs constrained routing to determine an LSP within an MPLS domain.
- This LSP will run from an ingress node to an egress node of the domain
- LSP may have some QoS features

**MPLS**
- MPLS works well with cell based and packet based networks
- It provides path calculation in advance of the traffic flow, thus introducing connections in the connectionless world!!
- The path could be strictly specified or loosely outlined and backup paths may be specified for handling link failures

**TE With MPLS**
- The LSP setup may follow TE principles thus solving the chronic inefficient utilization problem of the networks
- For example, constrained routing may prefer longer and lightly loaded paths over shortest paths
- MPLS + TE ➔ Balanced and well utilized network
Automated Provisioning

- The networks are growing bigger!!
- The protocols are becoming more complex
- With Diffserv, MPLS, RSVP TE, CR-LDP, COPS and associated protocols, it is impossible to allow manual provisioning
- Therefore, there is a need for automated TE-based path selection algorithms

Constrained Routing

- Constrained routing applies extended IGP parameters to the tree to find a suitable path
- $BW_{\text{avail}}$ and hop count may be used to determine paths
  - Shortest widest path
  - Widest shortest path
  - Shortest distance path ($\text{dist} = 1/BW_{\text{avail}}$)

QoS Traffic Considerations

- If only the available bandwidth is considered, the class of service may not be taken into consideration
- Thus, the best effort traffic may intersect the QoS traffic at several points within the domain
- In Diffserv, this may be a recipe for disaster!!

TELIC

- In this project, an efficient dynamic traffic engineering algorithm is developed for selecting paths across an MPLS-Diffserv domain
- TELIC (Traffic Engineering with Link Coloring) works with a set of traffic requests present at an ingress router of such a domain
- It allocates paths to an egress node using Dijkstra's shortest path algorithm
TELIC

- Each request specifies the amount of bandwidth requested followed by the Diffserv class of service (EF, AF, DF)
- While processing a request, TELIC partitions the network into several monochromatic subgraphs and makes an effort to match the request with an appropriate subgraph

In case a subgraph has no path to the egress node, TELIC merges it with another subgraph as per rules carefully built-in and starts the search all over again
- In case a search is exhausted, rules are available to deallocate a best effort class LSP and start the search again
- TELIC is written as a flexible tool in C++ by student programmer Jason Beuckman

Software Operations

- Traffic requests are read in and placed in a FIFO queue
- The program will then:
  - Look at the type of request
  - Create sub-graphs based on color and available bandwidth to find the best match for a request

If a path is found, the links on the path are updated to reflect the increase in usage
- Higher cost, less bandwidth, different color
- Otherwise, the request is not allocated, and the next traffic request in the queue is processed
Software Features

- Variable bandwidth requests
- Domains and traffic requests can be placed in files so multiple configurations may be tested

Results

- Traffic sets were prepared and applied to TELIC with various topologies of the domain
- The topologies tested include mesh, disjoint multipath, ISP and irregular ISP
- Results reflect a consistent behavior with 100 percent success rate for EF and variable success rates for AF and DF

Questions? Email Track605@aol.com
Results for EF Traffic

Results for AF Traffic

Results for DF Traffic

Discussion of Results

- Using class based priority, EF received 100 percent of the requested bandwidth
- AF's request and allocation remained identical until a fixed maximum beyond which AF could be not allocated additional bandwidth
- DF's trend is also similar but with a wider gap between demand and supply
Discussion of Results

- The request and allocation gap between AF and DF is due to the specific topology of the network being tested.
- The network has two bottleneck links: Node 8 to egress and Ingress to node 1.
- Both links become congested quickly, thus resulting in denial of additional bandwidth requests.

ISP and Disjoint Multipath Results

- We have also configured and tested networks that are based on ISP, disjoint multipath and mesh topologies.
- In the ISP and disjoint multipath topologies, the results are similar to the ones shown.
- In mesh topology, the allocated bandwidth falls short of demand considerably for AF and DF.
- It is because of the reduced number of links from the ingress towards the egress node.
Current Work

- Currently, one student is working to implement a colorless shortest distance algorithm that is based on \( \text{dist} = \frac{1}{\text{BW}_{\text{avail}}} \)
- Thus we will be able to compare the performance of TELIC vs. shortest distance routing

Future Work

- Future work includes modeling the request arrival and LSP holding time with different probability distributions
- Enhancements proposed for this algorithm include fault tolerance and prevention of excessive delays in best effort traffic
- We would like to extensively test the improved algorithm on OPNET’s MPLS platform