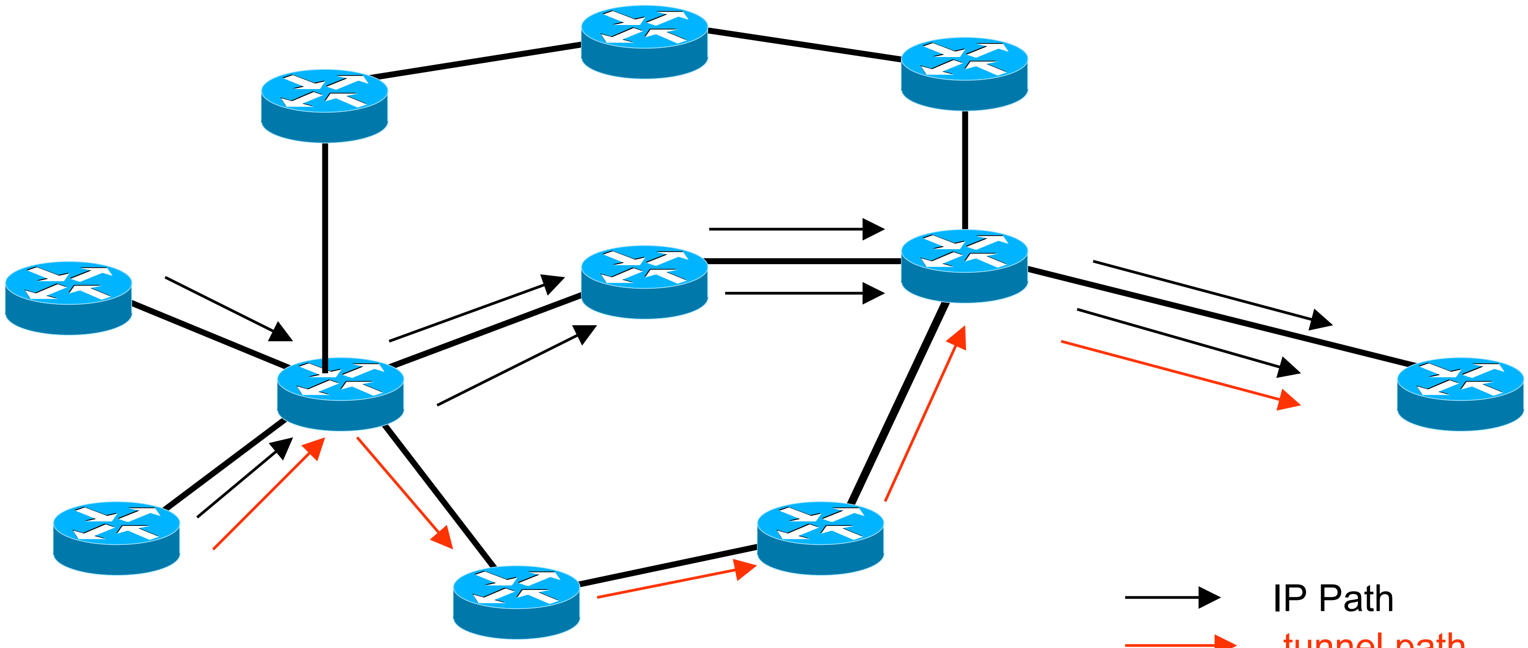
On the Performance of MPLS Traffic Engineering (TE) Queues for QoS Routing

C.J. (Charlie) Liu
AT&T Laboratories

Yihan Li, Shivendra S. Panwar
Dept. of ECE, Polytechnic University

Deficiency in IP Routing

- IP traffic takes shortest path destination based routing.
- Shortest path may not be the only path, or the best path.
- Alternate paths may be under-utilized while the shortest path is over-utilized.

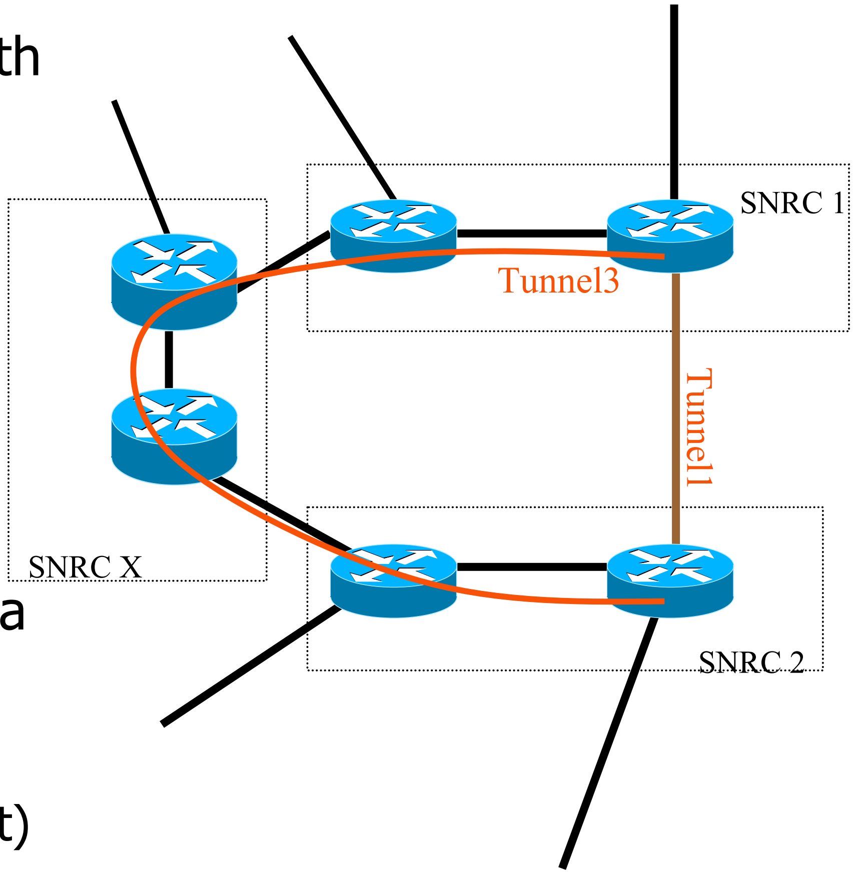


Application of MPLS TE Tunnels in a SP's Network

- Provide alternative to potential prolonged link congestion due to traffic growth
  - long lead time for facility order, and capital constraint
  - **TE allows path selection** without adjusting link OSPF cost
    - routing flexibility for unexpected traffic growth
  - Unequal cost load sharing
- Provide fast restoration in case of link failure to reduce packet loss
- Provide a means to integrate voice and data network
  - Voice carried by connection oriented entities on top of IP backbone network which transports data traffic on a hop-by-hop basis

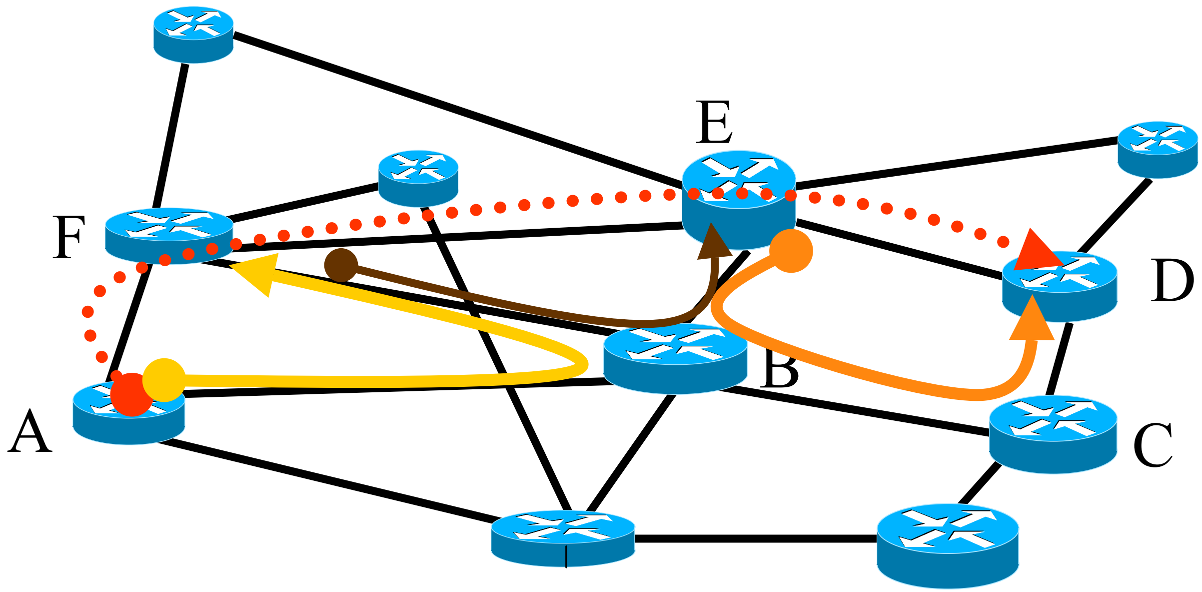
An Example of TE Application
- Re-route from the Congested Link

- Tunnel head and tail ends coincide with congested link.
- Not necessary to know culprit traffic source(s)
- Two tunnels :
  - short (the congested physical link configured as tunnel)
  - long (multiple hops)
- Traffic trunk: long tunnel effective as a separate physical link
- Tunnel cost assignment
  - equal cost (same as original OSPF cost) for all the tunnels
  - assign bandwidth based on desired load splitting ratio



Backup Tunnel for Fast Reroute

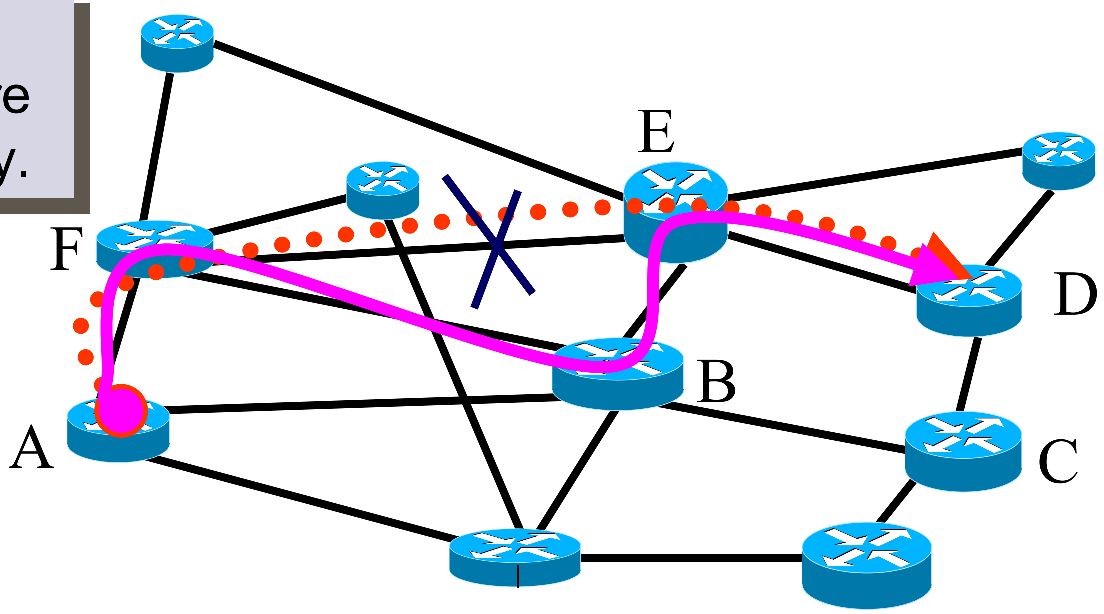
- Suppose we have mission critical traffic from A to D as **AFED**
  - Tunnel **ABF** protect AF link
  - Tunnel **FBE** protects FE link
  - Tunnel **EBCD** protects ED link



Fast Reroute (FRR) Example - Link Protection

- Suppose we have mission critical traffic from A to D as **AFED**
  - Tunnel ABF protect AF link
  - Tunnel FBE protects FE link
  - Tunnel EBCD protects ED link

When link FE fails,
- F will signal A about failure
- FRR reroutes immediately.

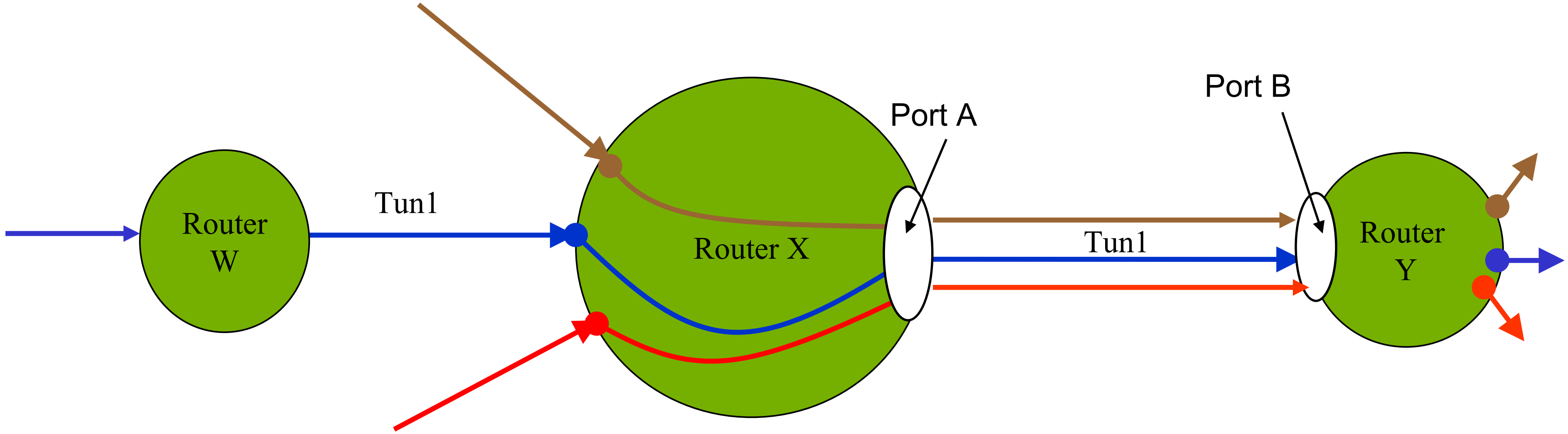


Problem Statement

- Admission control mechanism is available only at tunnel setup time, but not at packet sending time.
  - Bandwidth reservation is policed only at tunnel setup time to limit the number of tunnels traversing a given link.
  - At packets sending time, tunnel traffic has to compete with all other traffic either in another tunnel or non- tunnel traffic

MPLS TE Queue

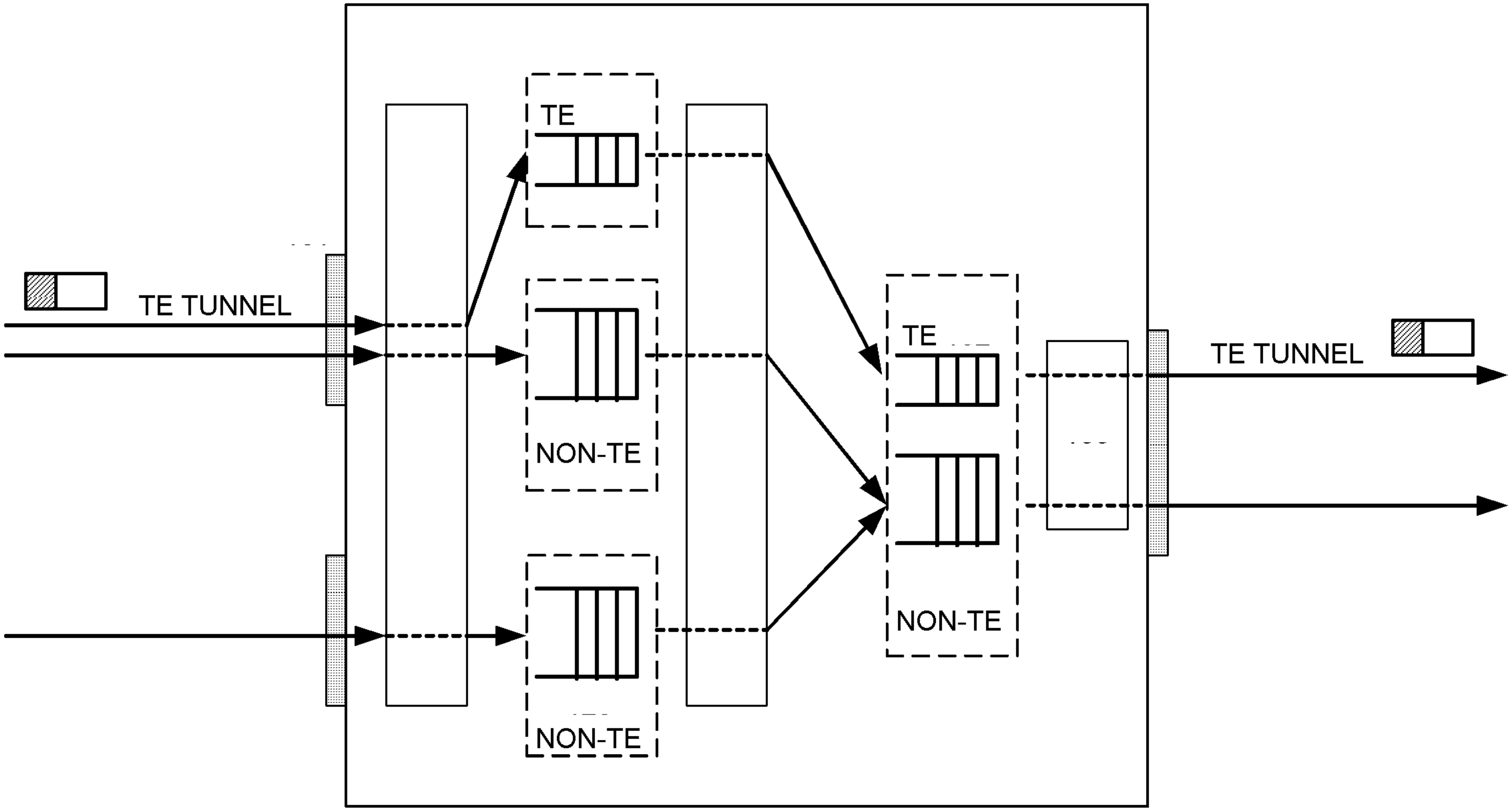
- Tun1 traffic competes with non-tunnel traffic for BW available in X-Y link.
- If RSVP reserved BW does not have receive preferential treatment than other traffic, mixing tunnel traffic and non tunnel traffic.
- **If there is preferential treatment,**
  - Router X recognizes tunnel packets by Label associated with the packets.
  - Create MPLS TE Queue to for tunnel traffic.
  - Packets in MPLS TE queue take priority over packets in any other non- TE queues.



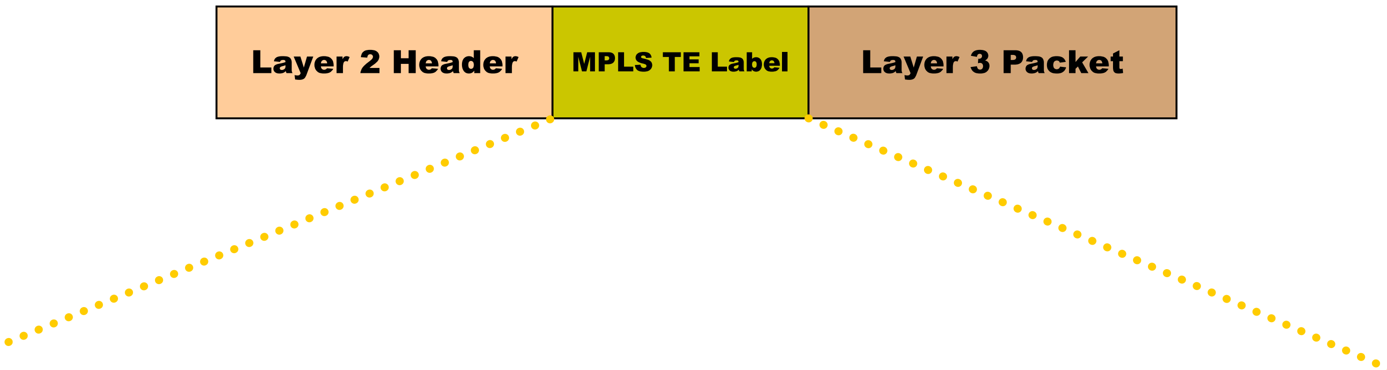
MPLS TE Queue

- MPLS TE Queues are created to ensure tunnel traffic's priority over non-tunnel traffic.
  - **Only real time traffic will be sent into MPLS TE tunnels via policy routing.**
  - **Input TE queues are shared by tunnels with the same head end.**
  - **Output TE queues are shared by tunnels with the same tail end.**
- Tunnel packets are identified by the label associated with the packets, and sent to a TE queue based on its label.
- It is a new idea to enable scalable MPLS TE Tunnels deployment with QoS guarantee for real time traffic in Service Provider's network.

Abstract Diagram



MPLS TE Label



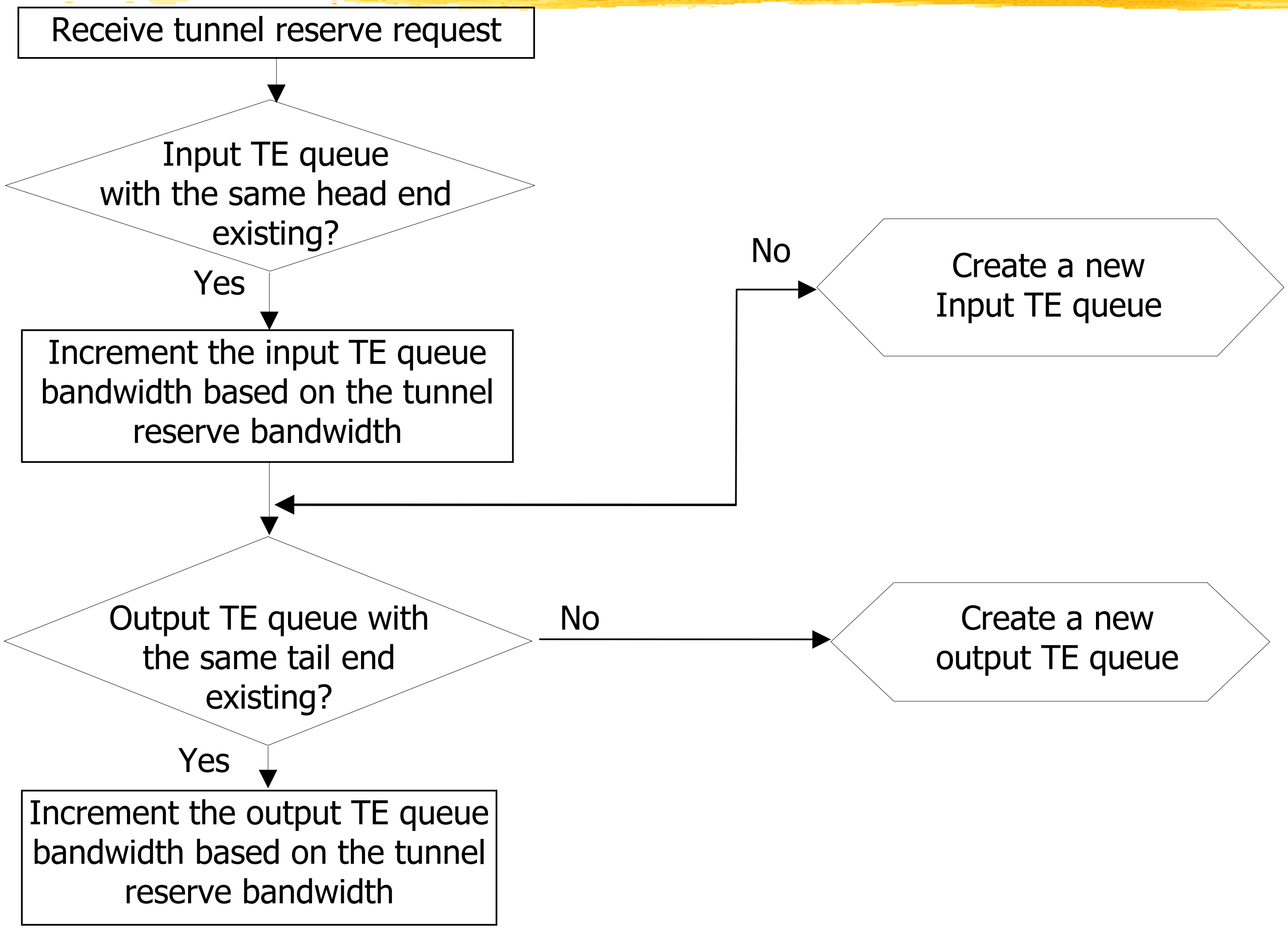

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Label                  | Exp |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Often called a "shim" header

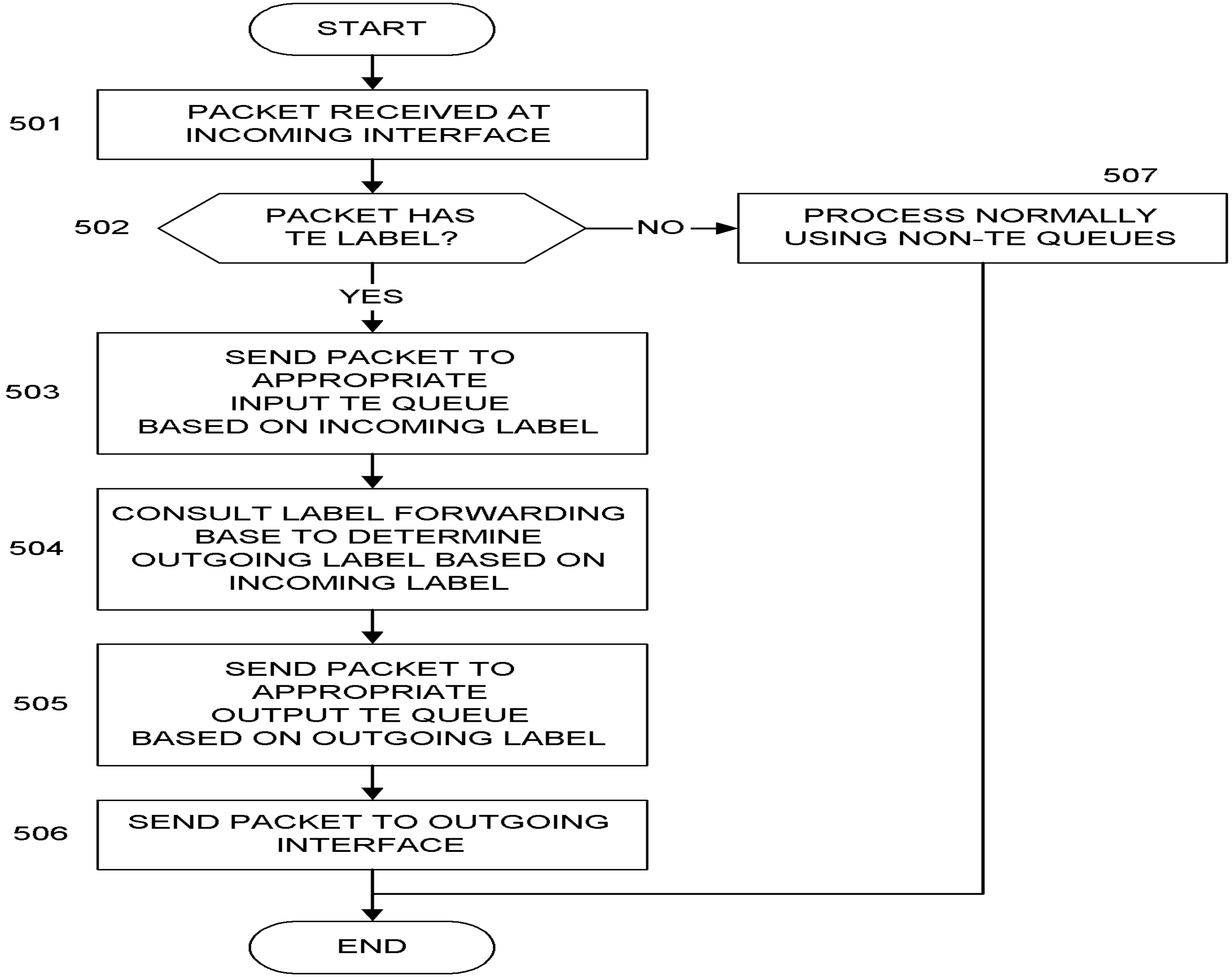
- **Label: Label Value, 20 bits**
- **Exp: Experimental, 3 bits**
- **S: Bottom of Stack, 1 bit**
- **TTL: Time to Live, 8 bits**

RFC 3032. MPLS Label Stack Encoding

MPLE TE Queues Creation



Switching Process for Tunnel Packets

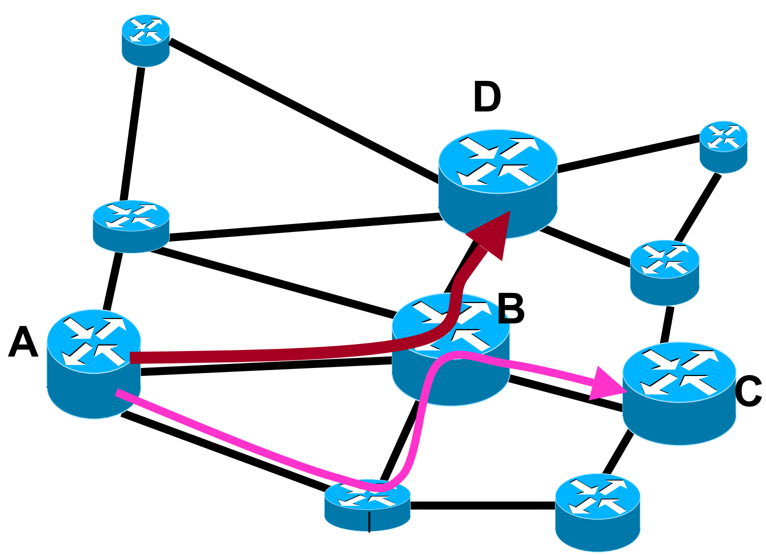


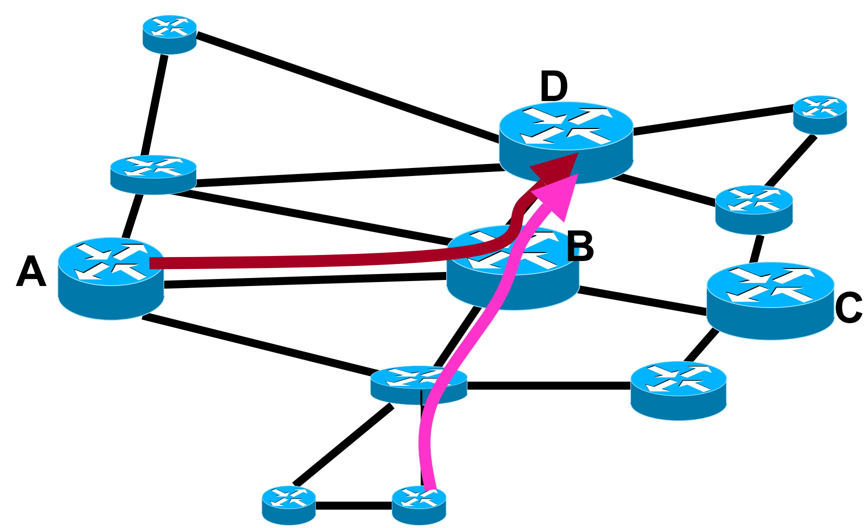
Number of MPLS TE Queues Per Router

- How many queues can a router effectively support?
  - For a large network with full meshed TE tunnels, the number of tunnels can easily go to many thousands.
- Limit MPLS TE tunnels at Backbone only
  - For a typical IP core network with 36 backbone routers, there will be 1260 tunnels.
  - Assuming each tunnel, on the average, will traverse five routers, including its head end and tail end.
  - Each router, on the average, will have to accommodate 175 tunnels.
    - head end of 35 tunnels, tail end of 35 tunnels, and mid point of 105 tunnels.

Number of MPLS TE Queues Per Router

- Allow multiple TE tunnels share the same
  - input TE queue if they originate from the same head end LSR
  - output TE queue if they terminate at the same tail end LSR





Brown tunnel and **pink tunnel** can share the same input TE queue in **Router B**.

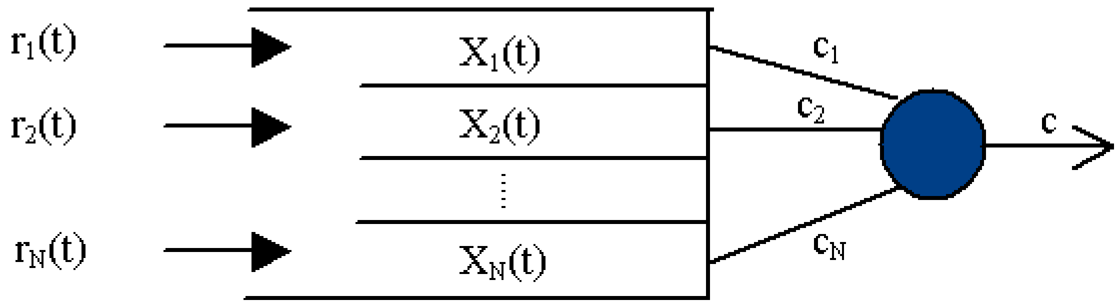
Brown tunnel and **pink tunnel** can share the same output TE queue at **Router B.**

Performance Analysis

- System model
  - An output switch with output TE queues is considered.
  - The process from the time packets enter output TE queues to the time they are forwarded to the next hop is analyzed.
  - Assume that each traffic source can be modeled as a **continuous-time Markov process**.
  - The system is analyzed and simulated as a Generalized **Processor Sharing (GPS)** system.

The System Model

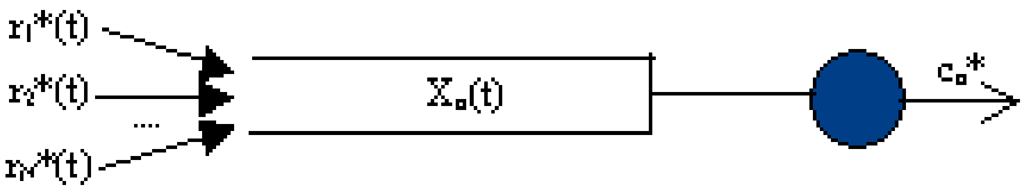
- Assume that each input maintains ***N*** (output) TE queues and ***K*** non-TE queues.
- All TE queues have the same priority, which is higher than the priorities of non-TE queues.
- When all TE queues are served, the residual service is distributed to non-TE queues.
- The buffer is infinite For each TE queue. Need to find out **the overflow probability** with threshold ***B***.
- Each queue is modeled as a **Markov Modulated Fluid Process (MMFP).**



Analyze a Queue with GPS Scheduling

- When one queue is analyzed, the system can be simplified as below.
- Resolve the problem with Fluid-Flow Model.
- Details can be found in the reference paper

S. Mao and S. Panwar, "The Effective Bandwidth of Markov Modulated Fluid Process Sources with a Generalized Processor Sharing Server," Globecom 2001, pp. 2341-2346.

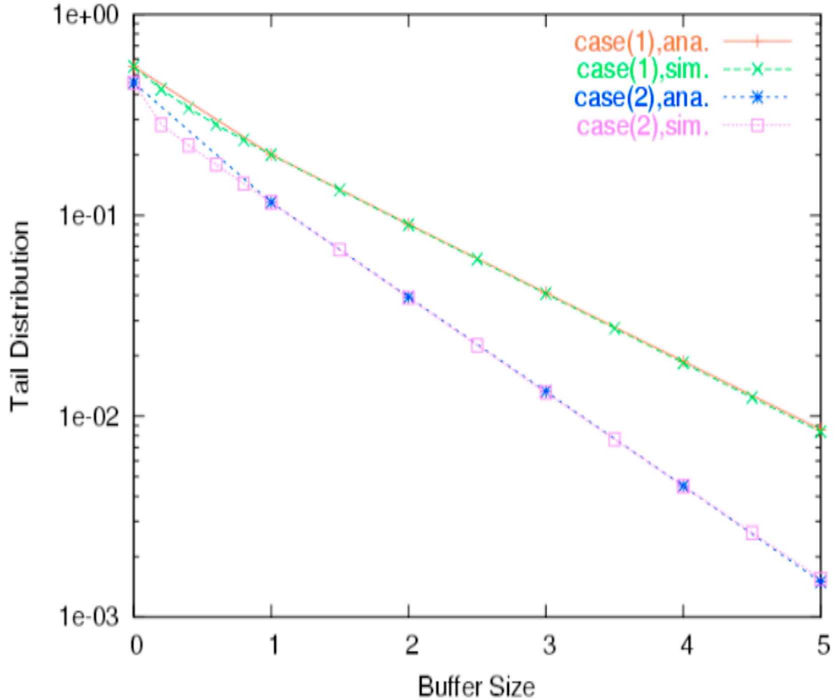


Analysis and Simulation Results

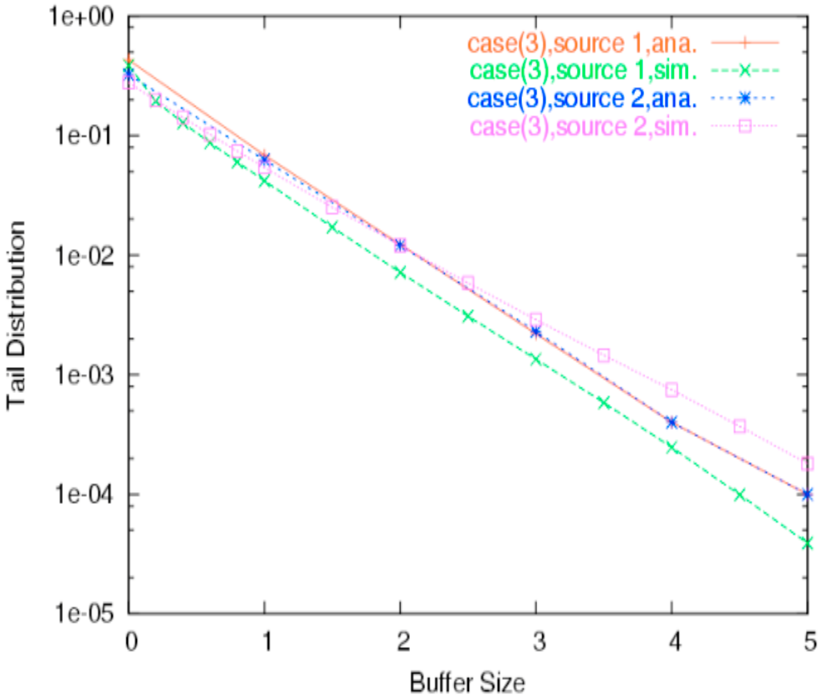
- Assume that there are three on-off sources,
  - two for TE traffic, and
  - one for non-TE traffic.
- Three cases are considered:
  - One queue: all traffic share one queue.
  - Two queues: all TE traffic share one TE queue and non-TE traffic goes to the non-TE queue
  - Three queues: each TE source traffic goes to its own TE queue and non-TE traffic goes to the non-TE queue
- Parameters:
  - α: the transition rates from off to on
  - β: the transition rates from on to off
  - **p**: is the input rate when the source is on.
  - The guaranteed service rate for source 1 and source 2 are 0.7 and 0.4, respectively, and for source 3 in case 1 is 0.6.

	α	β	p
Source 1	0.4	1.0	1.2
Source 2	0.4	1.0	1.0
Source 3	1.0	1.0	1.2

Analysis and Simulation Results



Tail distributions of case 1 and 2.



Tail distributions of case 3.

- With the selected system parameters,
  - using TE queue leads to lower overflow probabilities for TE tunnel traffic, and
  - using multiple TE queues can further improve the service of TE tunnel traffic.