Network Design with MPLS TE

Source:

Traffic Engineering with MPLS

Nortel's cgNet and OPi Projects

Motivation

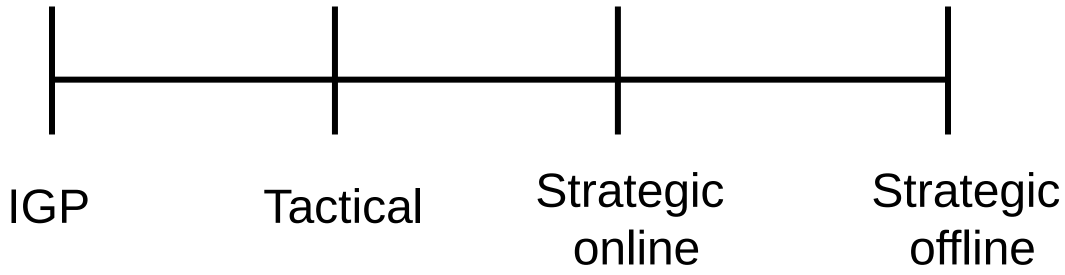
- Have covered how MPLS TE works

- How do you use MPLS TE in the network design?

Types of MPLS TE Design

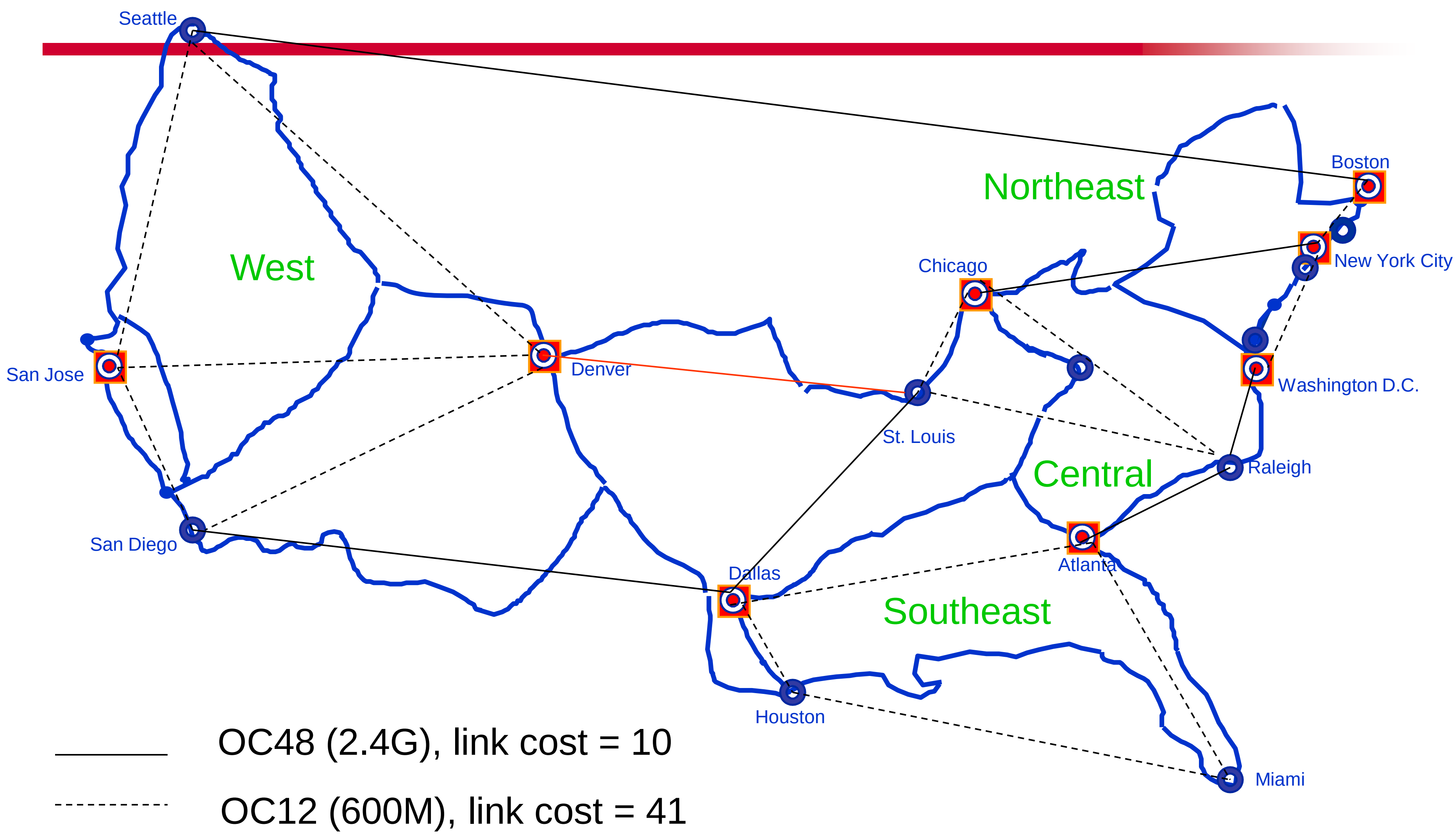
- Tactical MPLS TE
  - ✓ Build MPLS TE tunnels as needed to work around congestion or failures
- Strategic MPLS TE
  - ✓ Full mesh of TE tunnels or partial mesh in some part of the network
    - ➢ Online path calculation conducted by the routers
    - ➢ Offline path calculation done by an external device

Network Design Spectrum



- IGP: determined by routing protocols based on link metrics
- **Tactical**: traffic mostly forwarded along the IGP-calculated path, but build occasional TE LSPs
  - ✓ Avoid congestion
  - ✓ Unequal-cost forwarding
  - ✓ Support special QoS/SLAs requirements
  - ✓ Failure protection
- **Strategic online**: build a full (or partial) mesh of TE LSPs; reserve bandwidth based on SLAs or traffic profiles for efficiency; headends run CSPF and make decisions locally
- **Strategic offline**: build a full (or partial) mesh of TE LSPs; path calculation done offline; central TE server makes decisions “globally”

Case Study



Seattle
West
San Jose
Denver
San Diego
Chicago
St. Louis
Dallas
Houston
Northeast
Boston
New York City
Washington D.C.
Raleigh
Central
Atlanta
Southeast
Miami
OC48 (2.4G), link cost = 10
OC12 (600M), link cost = 41

Sample Network

- 4 regions: Northeast, Central, Southeast, West
- Each region has 4-5 Points of Presence (POPs)
- All POPs consist of
  - ✓ 2 WRs (WAN Routers)
  - ✓ 0 or more DRs (Distribution Routers)
  - ✓ 3 to 20 CRs (Customer Routers)
  - ✓ If a POP has > 11 CRs, it has 3 DRs; otherwise, it has 0 DRs.
  - ✓ Total: 28 WRs, 12 DRs, and 128 CRs

Tactical TE Design

- Goal: build TE LSPs to work around congestion
- Key things to deal with:
  - ✓ When to build TE LSPs?
  - ✓ Where to put them?
  - ✓ When to take them down?
  - ✓ What other features can be used?

When to Build TE LSPs?

- Suppose OC-12 link between St. Louis and Denver is congested

- If congestion is detected, how long do you wait before working around it with TE LSPs?

Where to Put TE LSPs?

- If putting TE LSPs is decided, the next question becomes "Where to place TE-LSPs so that it doesn't create **another problem**?"
- Suppose 900Mbps on link St. Louis → Denver for a significant amount of time (network monitoring)
- How to split the traffic?
  - ✓ 2 LSPs?
    (1. St. Louis→Denver; 2. St. Louis→Dallas→San Diego→Denver)
  - ✓ 3 LSPs?
    (3. St. Louis→Chicago→NY→Boston→Seattle→Denver)
  - ✓ What's the traffic amount on each?(600M, 300M) or (450M, 450M)?
  - ✓ Others considerations?
    1. Equal cost forwarding without bandwidth reservation
    2. Reserve bw based on actual traffic, e.g., 450M each
    3. Reserve bw based on the ratio of traffic
    4. No reservation, but set the load-share ratio

Tunnel Placement Optimization – Termination Region

- Laying out LSPs of 300M and 450M in a network whose bottleneck is 600M links may still be risky. Why?
- Optimization (in this context): ensuring that the network traffic takes the path with the lowest possible delay while maximizing the available bandwidth along a path.
- Suboptimal forwarding exists in the previous case.
  - ✓ Path 3: traffic from St. Louis → Seattle may go through a long path.
  - ✓ Similarly, path 2
- How to improve this?
  - ✓ Consider the **termination region**.
  - ✓ Terminate the suboptimal LSPs at the entry point to the Western region.
  - ✓ Route the packet appropriately from the tunnel tail using IP in the region.

Tunnel Placement Optimization - Sources

- Problem solved, i.e., still have suboptimal packet paths?
- **Sources** of traffic from St. Louis to Denver?
- Consider: Which one or ones is or are not likely to send traffic to the West region through Denver?
- If evenly distributed, 300M each from
  - ✓ St. Louis, Chicago, Raleigh
  - ✓ Chicago→ St. Louis→ Chicago →NY → Boston → Seattle?
  - ✓ Suboptimal
- How to get around the suboptimality?
  - ✓ Build TE LSPs farther away from the source of congestion.
  - ✓ Examples?
  - ✓ Ultimately, a full mesh of TE LSPs → full control

When to Remove Tactical TE Tunnels

- TE LSPs are temporary solution to work around the problem that was caused by a temporary event.
- Clear the network when the event is gone.
- Two common cases:
  - ✓ When LSPs no longer needed
    - ➢ How do you tell?
  - ✓ When they’re causing problems
    - ➢ Congestion elsewhere
    - ➢ 1st thing: check if any TE LSPs are crossing that link
- One of the things MPLS TE buys you is the capacity to **prolong the life of the existing capacity**, thus putting off the amount of time until you have to buy new equipments.
- Cooperation between TE service providers and content providers for special events
  - ✓ More ISPs are looking at content distribution networks (CDNs).

Online Strategic TE Design

- The logical progression of the tactical TE design is to have a partial or full mesh of TE LSPs, typically the WAN core routers. It gives most control and better network utilization.

Tactical	Strategic
Reserves bw as necessary	Reserve bw that matches the actual traffic sent (SLAs and traffic estimation)
Small no. of tunnels	Large no. of tunnels
Difficult to track what tunnels are where and whey they're there	Easier to track, because you know the no. and where they go

In practice, MPLS network topology discovery and management is challenging for tactical approach after some time.

How Well Does it Scale?

- **Scalability** is the most important issue.
- Specific issues:
  - ✓ Scope of the TE cloud
    - ➢ Full mesh between WRs, CRs or DRs
  - ✓ How many LSPs can a router be a headend, a midpoint, or a tailend for?
  - ✓ How many is too many?
    - ➢ Convergence time → scalability limit
    - ➢ Vary form network to network: router CPU utilization, size of the network
    - ➢ Cisco guidelines (~ 5 years old, need to be updated):
      - 600 LSPs as a headend
      - 10,000 LSPs as a midpoint
      - 5000 LSPs as a tail
  - ✓ Example: Full mesh with 28 WRs, 12 DRs, and 128 CRs, how many LSPs?

Scalability – Full Mesh between WRs

- How many LSPs can a router be a headend, a midpoint, or a tailend for?
  - ✓ Ln: total number of LSPs in the network
  - ✓ Wn: total number of WRs
  - ✓ M: number of LSPs for which a router is a midpoint
  - ✓ Headend for:
    - ➢ A WR is connected to every other WR: 27 LSPs
  - ✓ Tailend for:
    - ➢ Every WR has an LSP to a WR: 27 LSPs
  - ✓ Midpoint for:
    - ➢ For 28 WRs: Ln = Wn * (Wn-1) = 28 * 27 = 756 LSPs
    - ➢ Max LSPs that a WR is a midpoint for M = Ln – 2*(Wn – 1) (headend plus tailend).
    - ➢ Average LSPs that a WR can be a midpoint for = M / Wn = (756 – 2*27) / 28 ~ 25
- How about between CRs and DRs?
  - ✓ Midpoint LSPs per WR is ~2000 to 4000 for the topology

Other Growth Factors

- In tactical method, periodically monitor and clean up manually.
- Strategic architecture does not really have housekeeping, but need to consider for online path calculation:
  - ✓ Tunnel reoptimization
  - ✓ Tunnel resizing
  - ✓ Multiple parallel LSPs

Tunnel Reoptimization

- Reoptimization: recalculation of CSPF for (all) existing LSPs to see if there is more optimal path.
- Example:
- Different kinds of reoptimization:
  - ✓ **Periodic reoptimization**
    - ➢ How frequent? 1 hour? 60 seconds? Scalability?
  - ✓ **Event-driven reoptimization**
    - ➢ The result could trigger another event which may trigger reoptimization …
  - ✓ **Cooperation between TE and CDNs**
  - ✓ **Manual reoptimizatio**n
- Tunnel resizing: changing the bandwidth requirements on tunnels, based on demands. Might change the path a tunnel takes.
  - ✓ Frequency (how often, what size)
  - ✓ Traffic patterns
  - ✓ Traffic forecast
- Multiple parallel LSPs

Offline Strategic TE Design

- Packing problem: dealing with optimal use of resources
- Example:
  - Max-fill
  - Min-fill

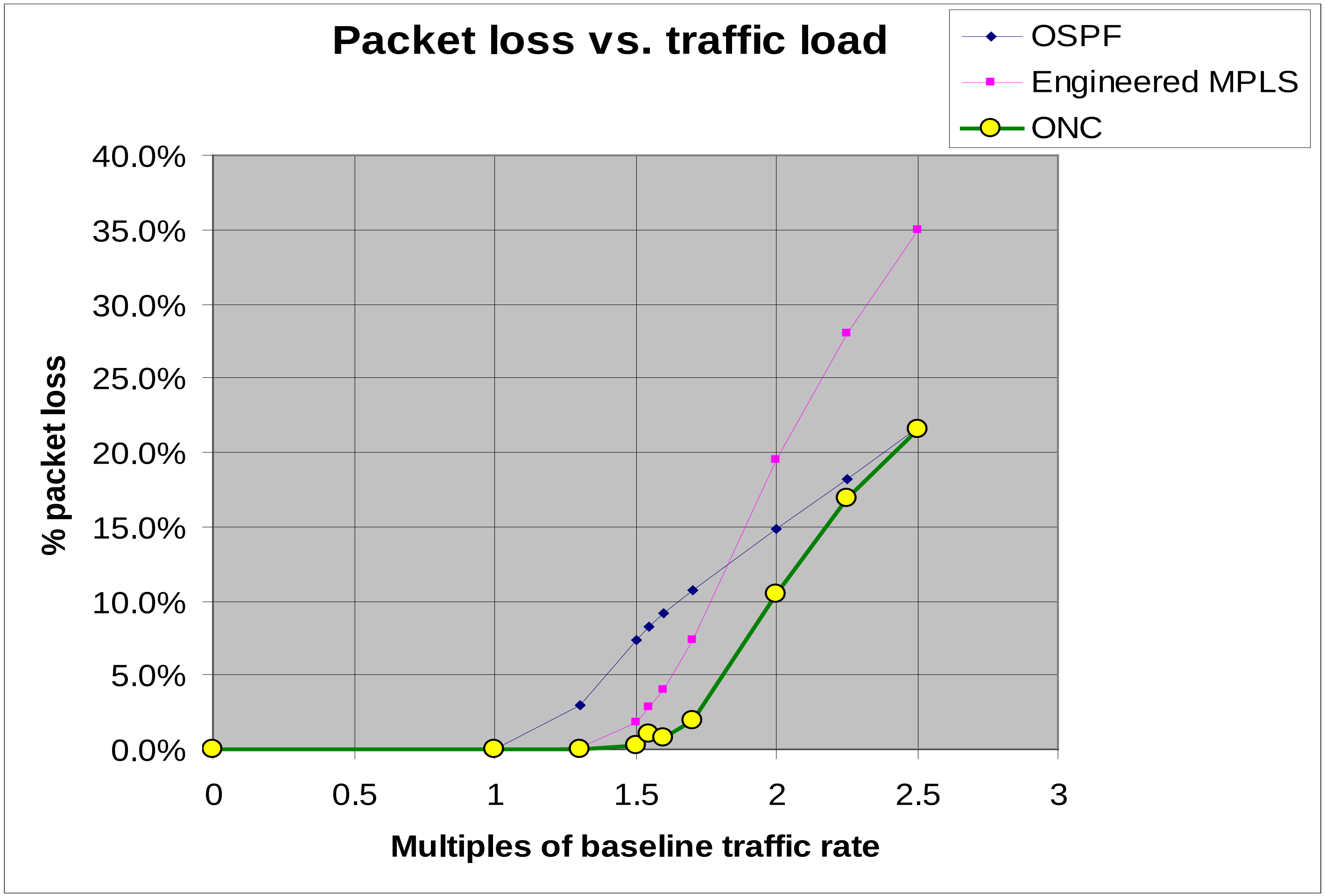
Example on min-fill and max-fill

	Start	After the First LSP Is Placed	After the Second LSP Is Placed	After the Third LSP Is Placed
min (20,30,40)	{40,60}	{20,60}	{20,30}	{F}
max (20,30,40)	{40,60}	{40,40}	{10,40}	{10,0}
min (20,40,30)	{40,60}	{20,60}	{20,20}	{F}
max (20,40,30)	{40,60}	{40,40}	{0,40}	{0,10}
min (30,20,40)	{40,60}	{10,60}	{10,40}	{10,0}
max (30,20,40)	{40,60}	{40,30}	{20,30}	{F}
min (30,40,20)	{40,60}	{10,60}	{10,20}	{10,0}
max (30,40,20)	{40,60}	{40,30}	{0,30}	{0,10}
min (40,20,30)	{40,60}	{0,60}	{0,40}	{0,10}
max (40,20,30)	{40,60}	{40,20}	{20,20}	{F}
min (40,30,20)	{40,60}	{0,60}	{0,30}	{0,10}
max (40,30,20)	{40,60}	{40,20}	{10,20}	{10,0}

Offline Strategic TE Design

- Other techniques:
  - ✓ Use tunnel priority to give larger trunks better priority.
  - ✓ The larger the network, the worse the packing problem may be. Split larger LSP requirements into smaller ones.
  - ✓ Use a **centralized offline tool** to find optimal paths. Most efficient in terms of network utilization, but most complicated as well.
    - ➢ Global vs. local view
    - ➢ Adopt profile-based routing or traffic forecast /estimation methods
  - ✓ Centralized tool → **SDN**

An Example of Automatic Bandwidth Adjustment – a Tactical Approach



OSPF: IGP

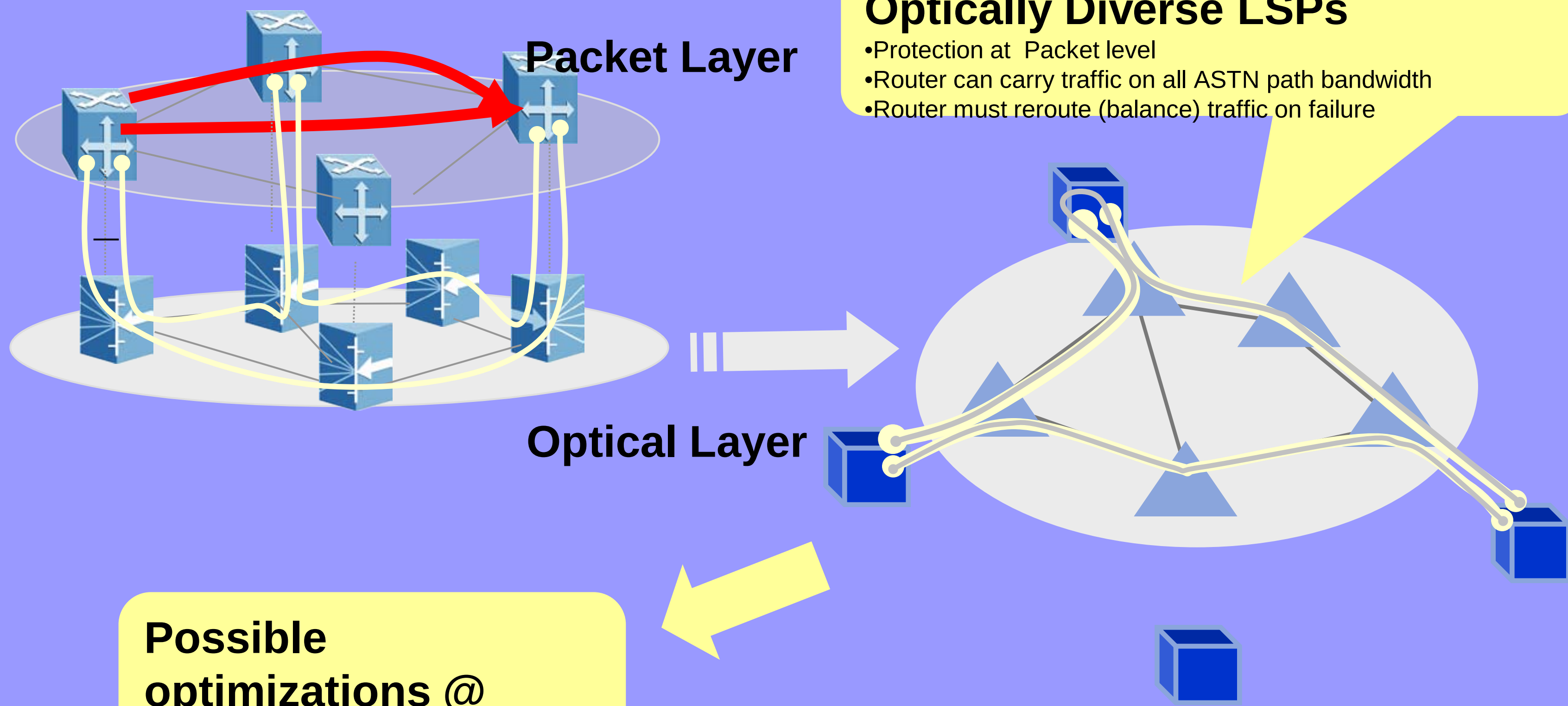
Eng MPLS:
- Static full mesh
- 2 LSPs for each (src, dest)

ONC: Net Controller
- Same as Eng. MPLS to start
- Nodes make decision
- Monitor, Model, Control

An Example of Strategic Offline Approach for Protection - Optically diverse LSPs

Optically Diverse LSPs

- Protection at Packet level
- Router can carry traffic on all ASTN path bandwidth
- Router must reroute (balance) traffic on failure



Possible optimizations @ optical level

- No optimization possible since ASTN cannot share protection bandwidth between ASTN paths