LIMELIGHT NETWORKS

Delivering the Digital Lifestyle™

MPLS In Operation

NANOG 39 Toronto February 7 2007

Overview

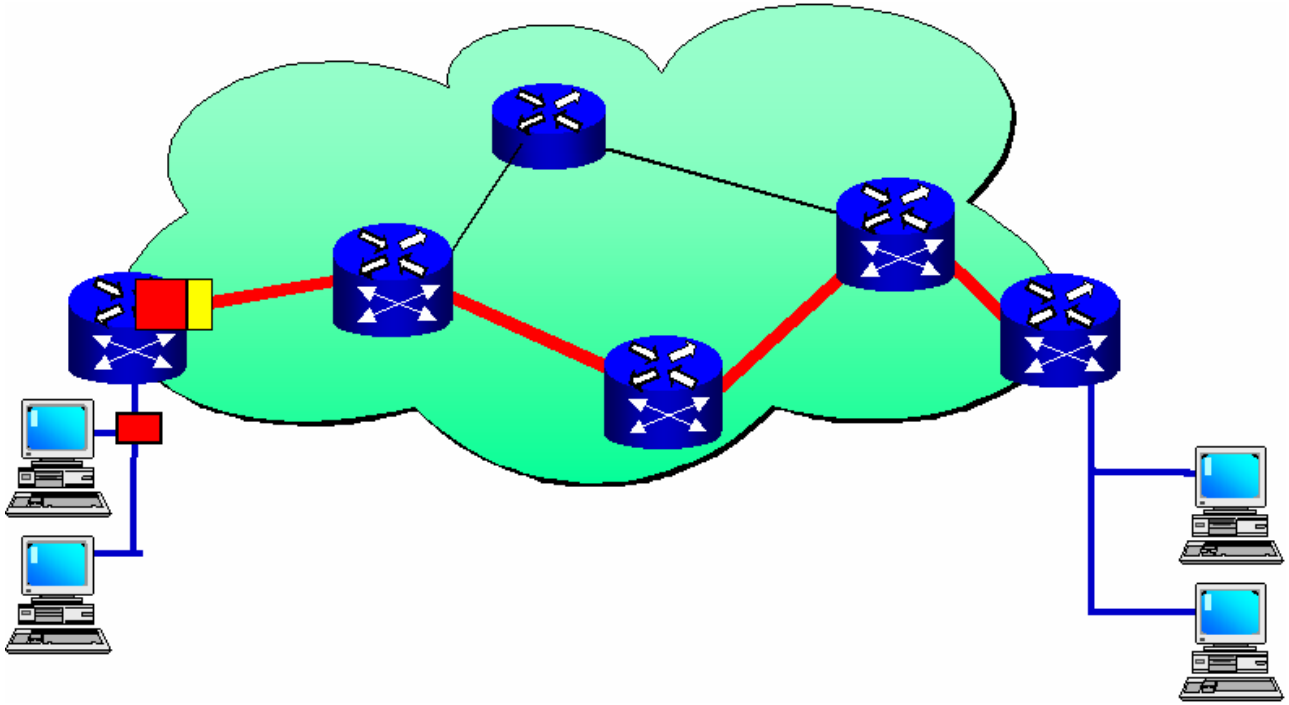
- What is MPLS?
- Why use Tunnels?
- Why use MPLS?
- Why use RSVP?
- Head-End Placement
- Online vs Offline Calculation
- Summary

MPLS Defined

- Multiple Protocol Label Switching
  - Packets are switched through a network through the cunning use of labels
  - All MPLS enabled routers have forwarding databases that contain a list of all the LSPs (Label Switched Paths (tunnels)) passing through the router, what label and interface a packet should come in on, and what label and interface it should send the packet out on

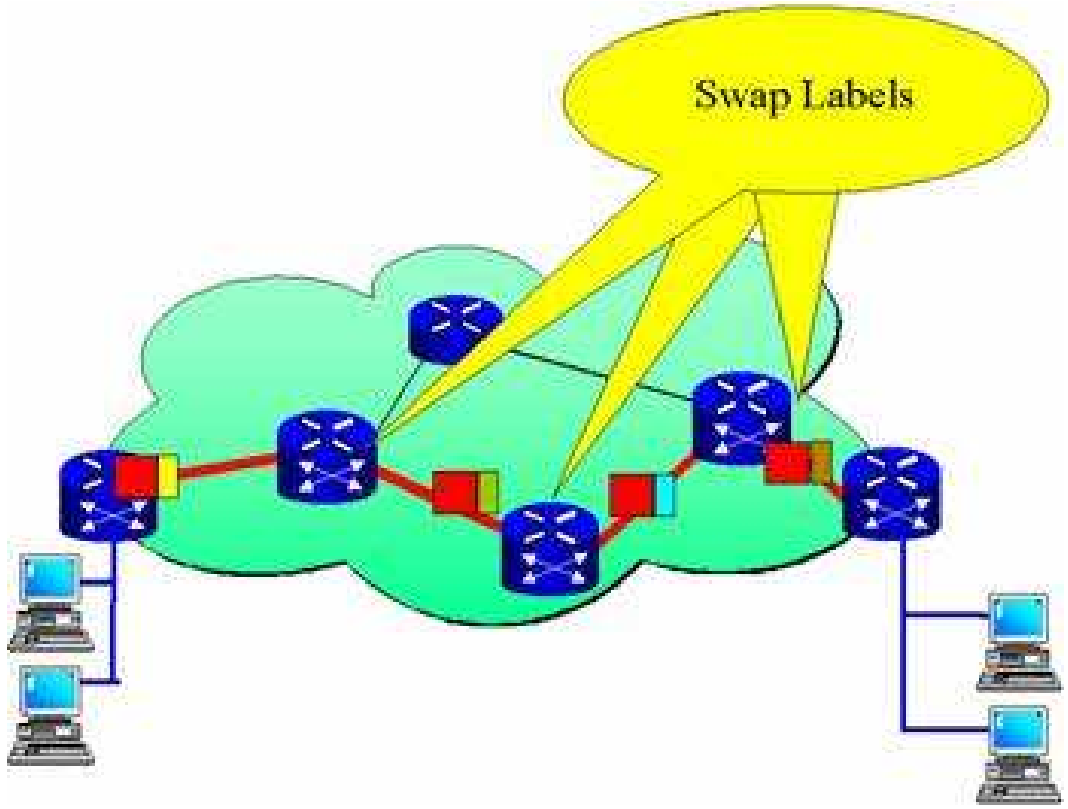
MPLS Forwarding

- An LER (Label Edge Router) determines that the next hop for a prefix is a tunnel and looks up the appropriate label to insert into the header



MPLS Forwarding

- An LSR (Label Switch Router) swaps the label and forwards the packet out the predetermined interface



Benefits of Tunnel Usage

- Accounting
  - How much traffic does your network have between Hong Kong and London?
- Traffic Awareness
  - Helpful when tracking down where bursts of traffic are sourced from or destined to (either during failure or DOS)

Benefits of Tunnel Usage

- Selective Traffic Engineering
  - Ability to move just a subset of traffic instead of all traffic to/from a particular path
- Loose vs Strict Explicit Hops
  - Ability to route traffic to its destination through a specific router or interface or along a specific path
    - Useful for SRLG aware tools

Benefits of Tunnel Usage

- Capacity Planning
  - Poor man's netflow
    - e.g. Peer loses a remote backbone circuit and both sides realize how much traffic is being exchanged in a suboptimal location. New peering established that saves resources and improves performace for both networks.
- Other Protocol Transport
  - VoIP
  - VPN
  - IPV6
    - No need to ipv6 enable the actual core

Drawbacks of Tunnel Usage

- More Complication
  - Routers need to run more protocols, have more moving parts (KISS this!)
- More Overhead
  - Routers need beefier CPUs, more RAM
    - RP upgrades
    - Staff Education
- Possible Disconnect Between Data vs Control Plane

TTL Propagation?

- no-propagate-ttl / no tag-switching ip propagate-ttl
  - Allow your operations to troubleshoot network issues instead of other companies' operations staffs
  - Hide core network events
- MPLS ping
- TCP Benefit of Fewer Hops
- Marketing advantage of fewer hops

Benefits of MPLS

- Handle increased size of edge capacity vs backbone capacity
  - Backbone used to have more capacity than what was being sold on the edge, but it's very common to have 10G edge ports alongside 10G backbone circuits
- Defer Upgrade Costs
  - Breathing room to implement upgrades while network is running hot
  - Use excess capacity in non-direct paths
- Allow Different QoS With Labels

Drawbacks of MPLS

- Asymmetrical Data Plane
  - Return path could be wildly different from forward path
    - IGP metrics help prevent most wackiness
  - More of something to be aware of than an actual drawback.
    - Most traffic is asymmetrical anyway.
- Blackholing of Traffic
  - All data in tunnel is lost when it breaks

Benefits of RSVP-TE

- Control Traffic Levels
  - Networks often have smaller backbone circuits in certain areas of their footprint.
    - e.g. NA company expanding in EU
- Usage of Uneven Parallel Circuits
  - Allows for quick bandwidth addition in a pinch
- Link Coloring
  - Only allow certain traffic to use particular circuits

Benefits of RSVP-TE

- Prioritization Of Traffic
  - Allow VoIP or VPN traffic priority on available network resources
- More resilient error detection (polling as well as message flooding)

Drawbacks of RSVP-TE

- More State
  - All routers need to be aware of available bandwidth of all interfaces of all other routers
  - Beefier hardware is needed (multiplier of drawback of tunneling in general)
- Traffic Crunch
  - When a tunnel fails, the routing control that is being counted on is not there

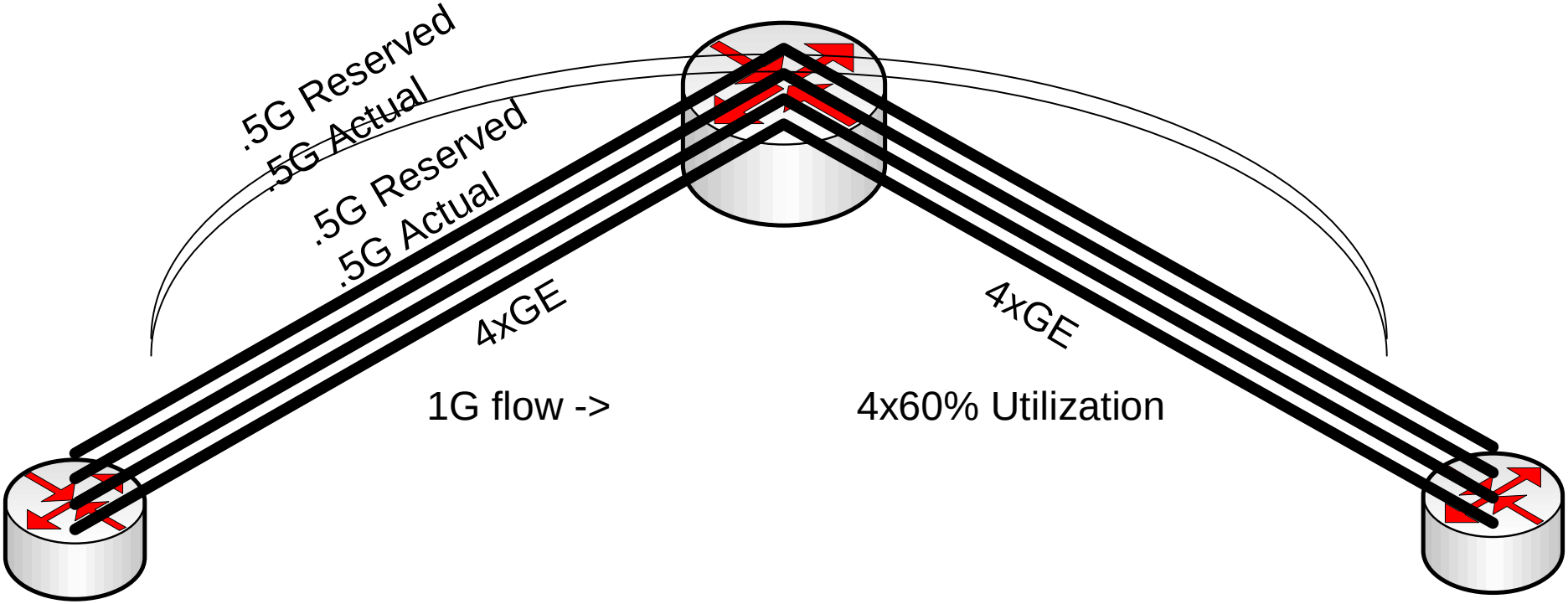
Vendor RSVP-TE Quirks

- Cisco
  - 15% until 90%, then 95, 96, 97, 98, 99
- Juniper
  - 10% at a time all the way to 100
    - Creates a problem after a circuit passes 90% utilization because the router won't signal to the rest of the network that it's nearing exhaustion
    - Configurable, but only by percentage, no granularity as reservations near maximum

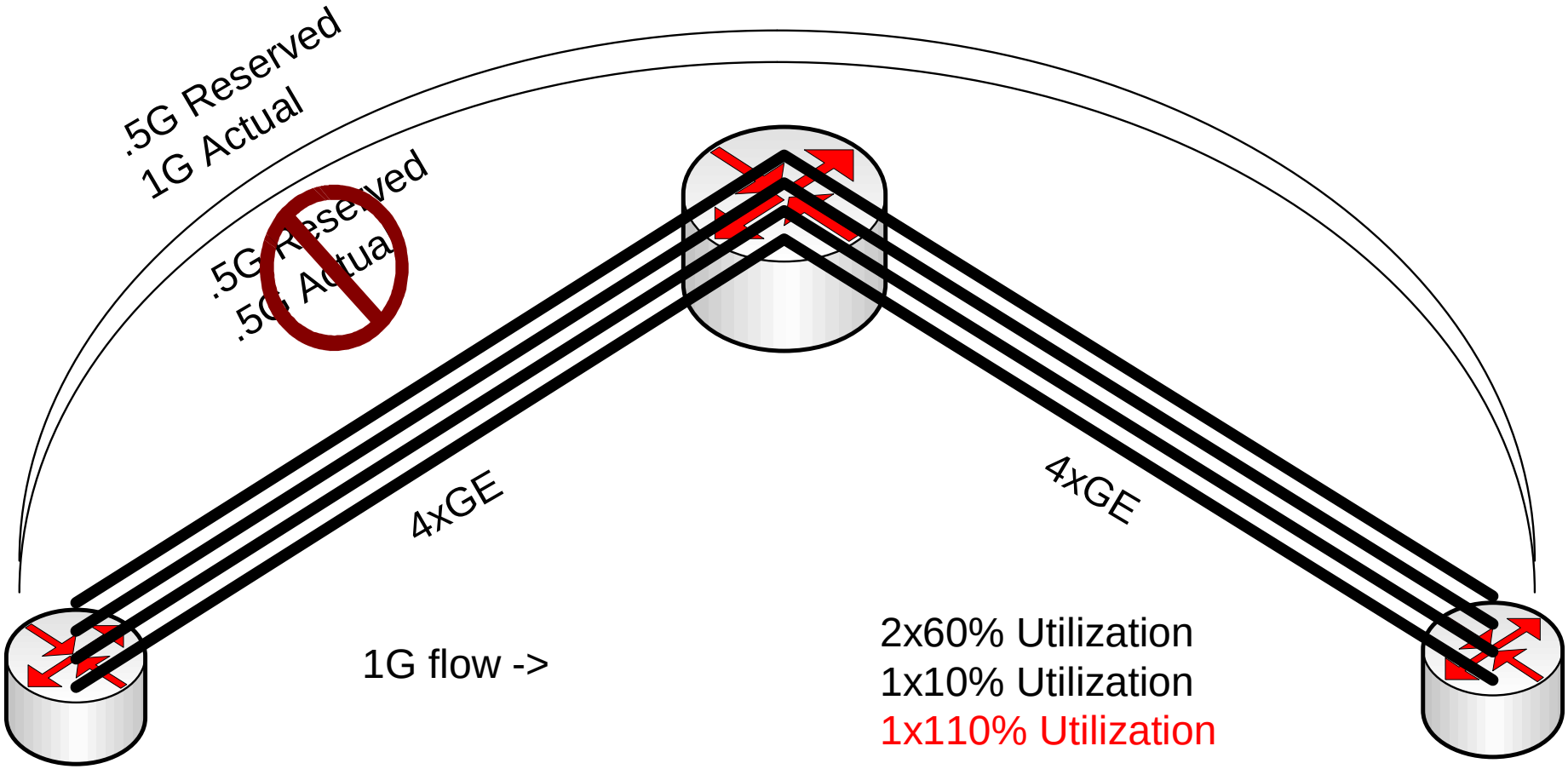
Handling Larger Flows

- What happens when your demands increase dramatically, but your budget hasn’t?
  - Allow your tunnels to keep growing unchecked
  - Convince your budget owner that you need millions of dollars of hardware upgrades immediately
  - Split demands into parallel tunnels

Parallel Tunnel Complications



Parallel Tunnel Complications



Tunnel Head-End Placement

- Benefits On Core
  - Fewer Tunnels
    - 100 devices require 9,900 tunnels according to N*(N-1) rule
  - Dumber Hardware For Edge
    - Cheaper aggregation devices that don’t need to run as many protocols

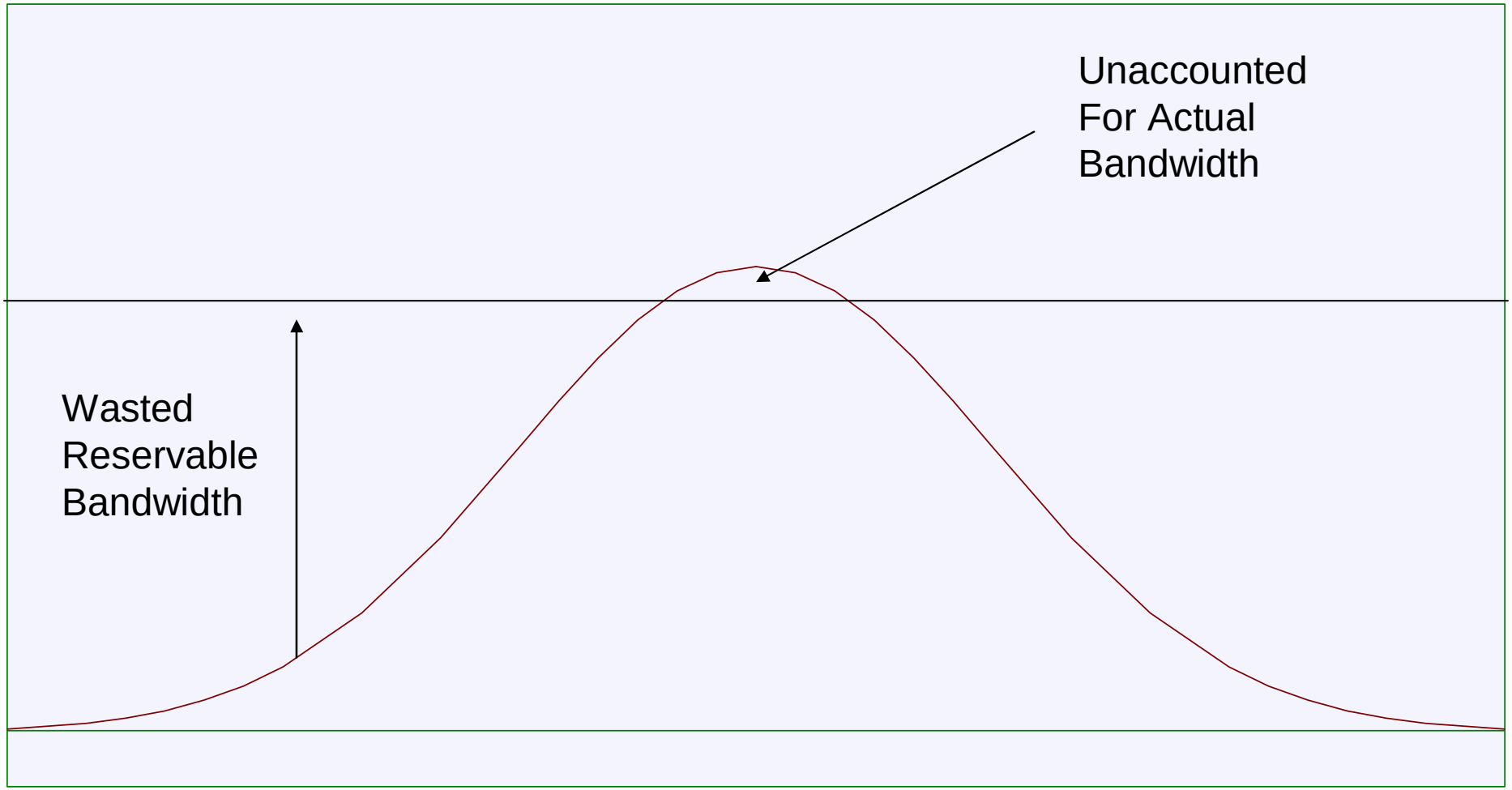
Tunnel Head-End Placement

- Benefits On Edge
  - Better Accounting/Control
    - Traffic can be viewed from router to router instead of pop to pop
    - Full utilization of backbone uplinks
  - Less “Slosh”
    - Traffic won’t bounce from uplink to uplink and tunnel to tunnel

Offline Calculation

- Benefits
  - Little/No Recalculation on Head-Ends
  - Network knows what it needs
  - Pathing is fairly constant
- Drawbacks
  - Inefficient Usage of Bandwidth During Off Peak Hours
  - Intensive Programming Effort
  - Highly Dependant on Stats Collection

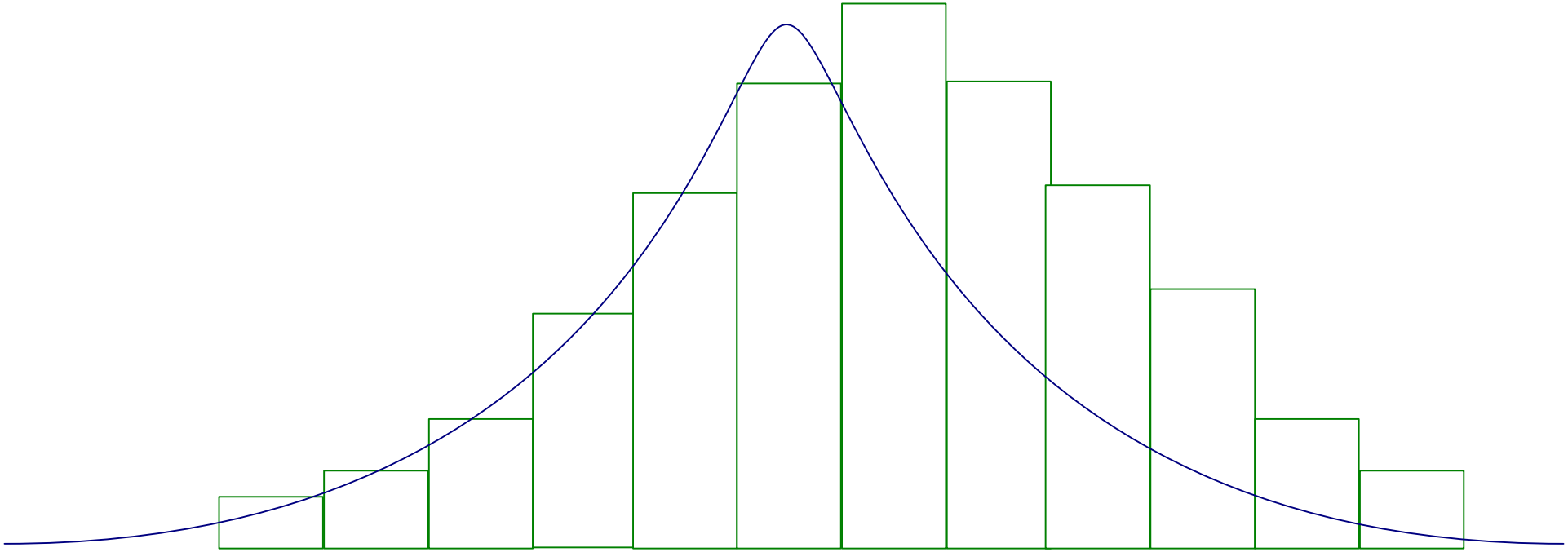
Static Reservations



Online Calculation

- Benefits
  - Better Reflection of Actual Traffic
  - Less Manual Intervention
  - Handles Sudden Surges of Traffic Better
- Drawbacks
  - Traffic Leads/Trails Reservations
  - Less predictable
  - Must use vendor’s heuristics

RSVP Trailing Actual Traffic



Causes of Tunnel Failures

- Circuit/IGP Failure
- Lack of Reservable Bandwidth
  - Configured Interface Bandwidth
  - Preemption
- Miscalculation of Necessary Bandwidth
- Protocol bugs

Soft-Preemption

- Make Before Break (like)
  - Tunnel signals and establishes while forwarding continues
- Vendor Support
  - Juniper : Yes
  - Cisco : Planned

Considerations

- Is it better to have a constant path (even if suboptimal) all the time, or have traffic move to optimal paths as bandwidth becomes available but move back off if the network congests?
- Is it better to have more smaller tunnels, creating more state and possibly running into ECMP issues, of smaller bandwidth or fewer larger tunnels?

Considerations

- Are tunnels better off on core or edge routers? (dumb core or dumb edge)
- Should tunnel sizes be calculated by the router or by the engineering staff?
- How do you plan for traffic that normally hairpins out of a router/city during maintenance/failure/dos?

Bottom Line

- Using tunnels allows for more control of traffic flowing through a network
- RSVP/MPLS TE can defer costs, however:
  - More moving parts
  - More engineering resources
- It takes money to make money

Questions?