MPLS Traffic Engineering in a Tier 1 Network

Chris Gibbings
Level 3 Communications

Level(3) COMMUNICATIONS

Overview | CommsUpdate | Sign Up

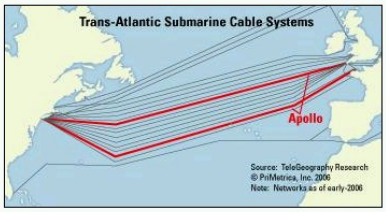
TeleGeography's CommsUpdate — Part of the GlobalComms Database

Today's Headlines | Article | Search | RSS | About | Advertise | Sign Up

World

Wednesday, 15 February 2006

600 billion bits across the Atlantic

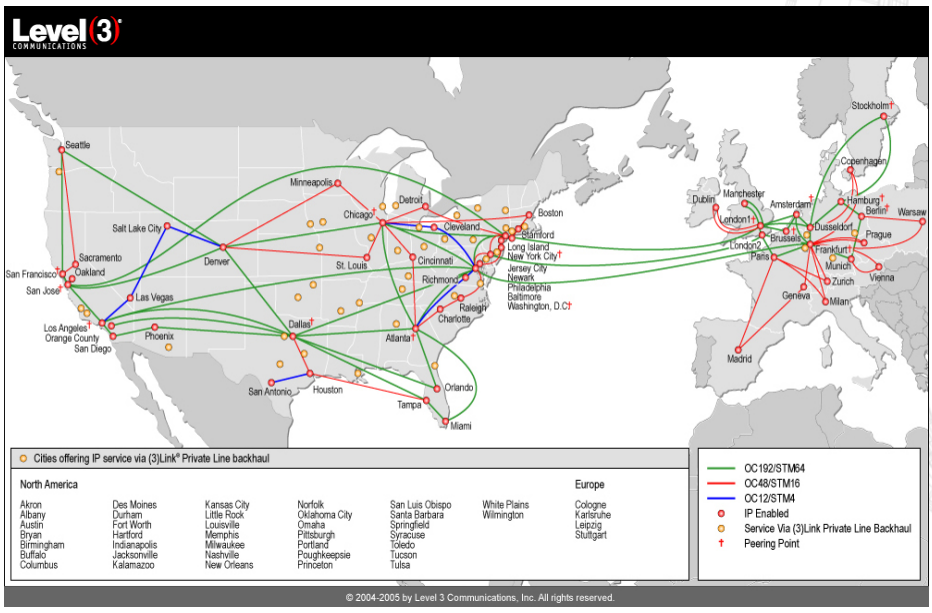


Level 3 made waves in the subsea capacity market last week when the company announced the purchase of up to 600 Gbps of capacity on the Apollo trans-Atlantic submarine cable system. With the immediate purchase of 300 Gbps and an option to acquire an additional 300 Gbps, this sale represents the single largest purchase of lit subsea capacity ever by an individual carrier. Overall the purchase of 300 Gbps represents 6% of the 5.2 Tbps of lit trans-Atlantic capacity.

2

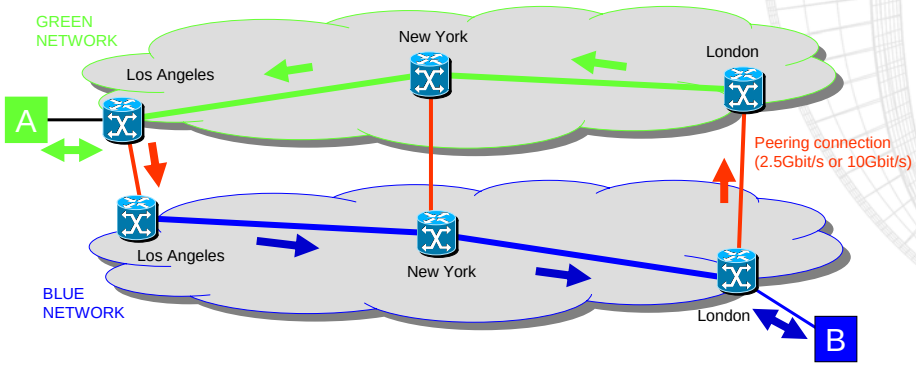
Agenda

- MPLS in a Tier 1 network
  - Tier 1 networks and transatlantic traffic
  - Advantages of MPLS
  - Design choices
- MPLS Traffic Engineering
  - Label switched path mesh
  - Link color
  - IP routing over LSPs
- Traffic balance on N x 10Gbit/s bundles
  - Internet traffic
  - Layer 2 flows



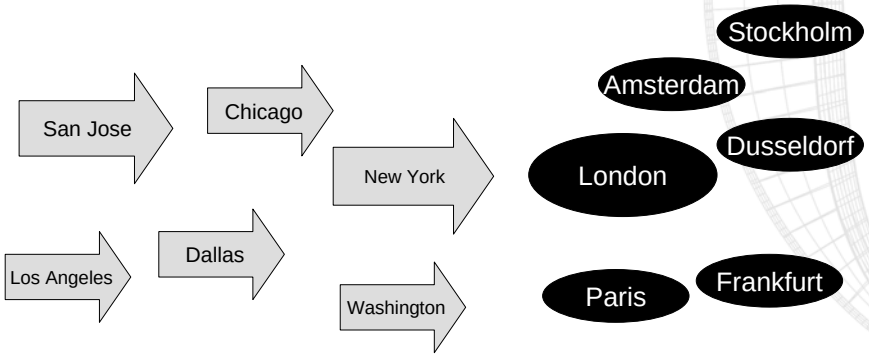
Tier 1 networks

- ‘A’ is a customer of the green network
- ‘B’ is a customer of the blue network



Traffic sources/sinks

- US to EU is the dominant direction for traffic
  - Imbalance reducing as EU to US is growing fast

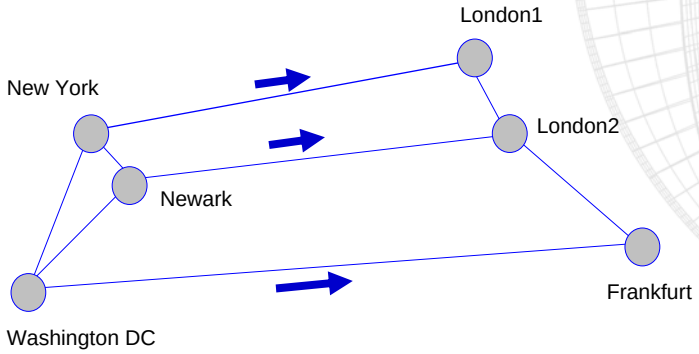


US major traffic sources

European major traffic sinks

Transatlantic connections

- Level 3 uses three cable systems
- MPLS traffic engineering used for balance
  - Steady state
  - Failure of one link

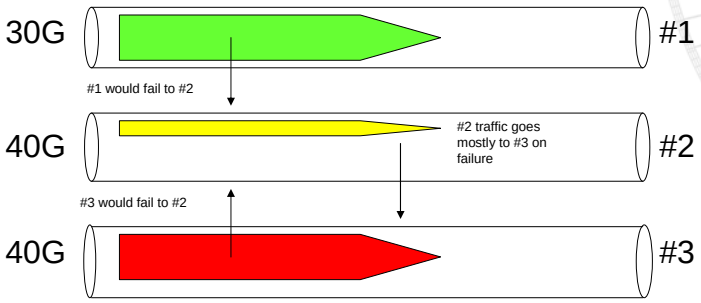


Some construction work in New York…

The morning after…

Transatlantic traffic

- How do we cope with any one of the three cable systems going down?
  - No perfect set of OSPF metrics
  - Need MPLS traffic engineering



MPLS advantages

- Traffic engineering
  - Share traffic across multiple 10Gbit/s links
  - Restoration paths when a link goes down
  - Traffic statistics
- Fast Re-route
  - Rapid switchover to alternative path when link goes down

Improve utilisation and restoration

- PWE3 (Martini) tunnels – (3)Flex
  - Ethernet
  - Frame Relay
  - ATM
  - Circuit emulation
- MPLS-VPN
  - RFC2547 virtual private network
- 6PE

New service opportunities

Network convergence

How to deploy MPLS

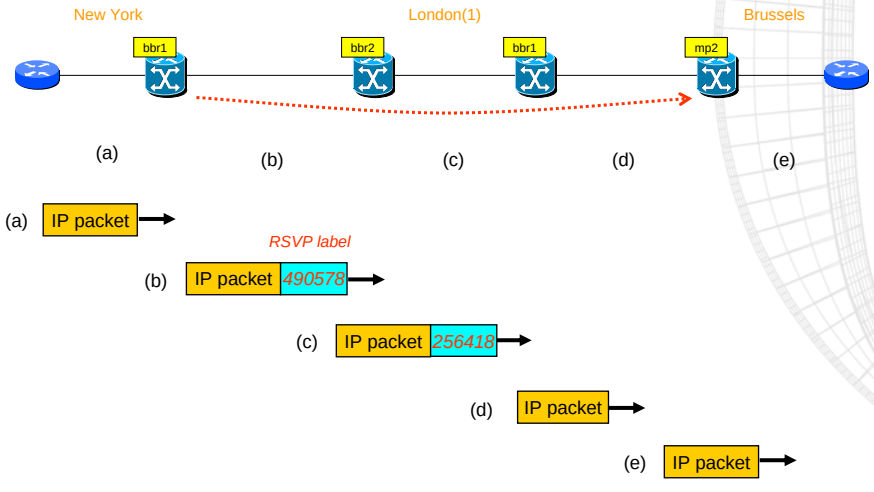
- RSVP – Resource Reservation Protocol
  - N-squared mesh to edge devices
  - Full traffic engineering
- LDP - Label Distribution Protocol
  - Easy to set up many edge devices
  - No traffic engineering

How to deploy MPLS

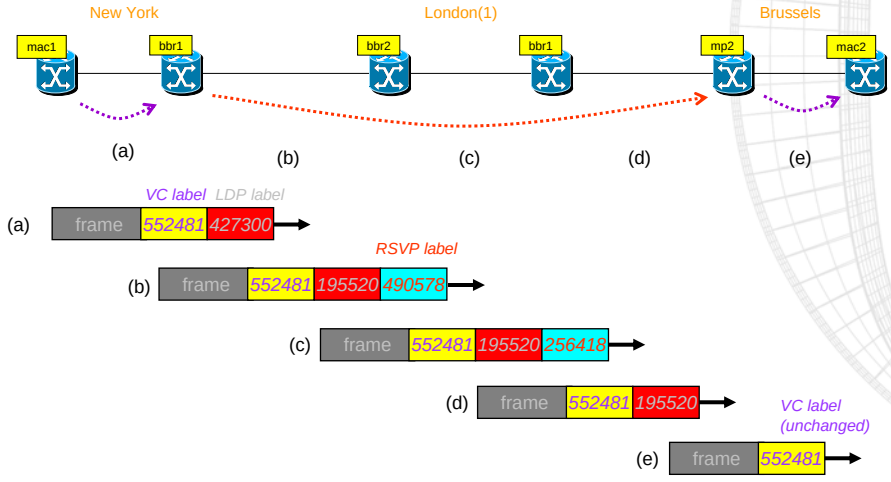
- RSVP – Resource Reservation Protocol
  - N-squared mesh to edge devices
  - Full traffic engineering
- LDP - Label Distribution Protocol
  - Easy to set up many edge devices
  - No traffic engineering

- LDP tunnelled over RSVP core mesh
  - Traffic engineering in core
  - Easy to add devices at edge

RSVP label switched path

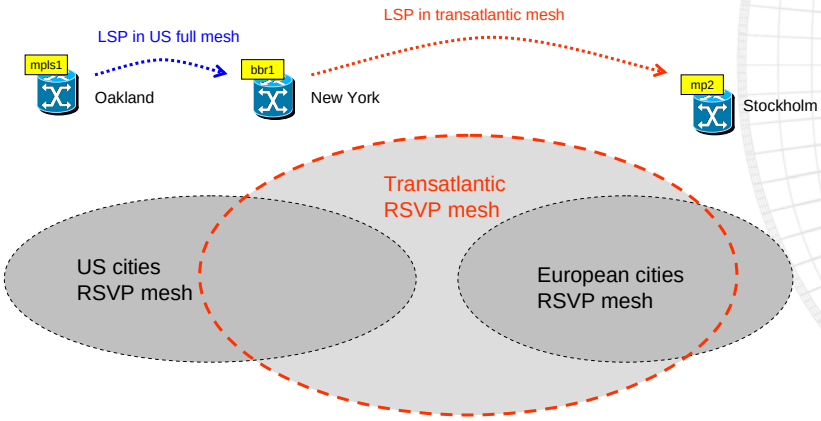


LDP tunneled over RSVP



15

Mesh of RSVP label switched paths



Mesh together cities with high traffic and/or Layer 2 transatlantic services

16

Link color

- Can add extra constraints

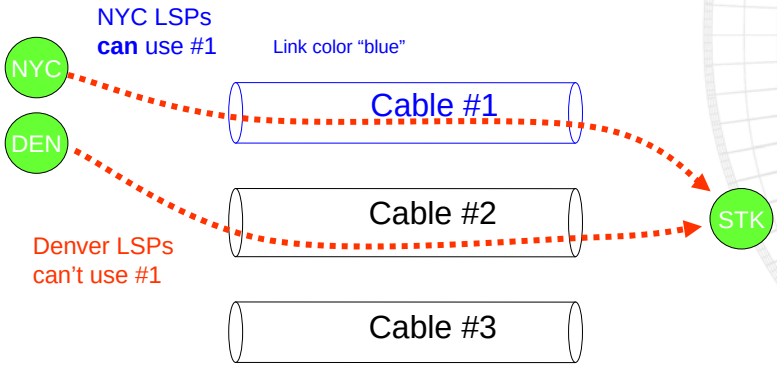
```
LSP name: Den-Ams  Destination: 1.1.1.1  Exclude: blue
LSP name: Den-Lon  Destination: 1.1.1.2  Exclude: blue
LSP name: Den-Par  Destination: 1.1.1.3  Exclude: blue
...
LSP name: NYC-Ams  Destination: 1.1.1.1  Exclude:
LSP name: NYC-Lon  Destination: 1.1.1.2  Exclude:
LSP name: NYC-Par  Destination: 1.1.1.3  Exclude:
...
```

Denver LSPs can't use 'blue' links

NYC LSPs **can** use 'blue' links

- Link color configured on router
  - OSPF-TE propagates link colour information through opaque LSAs
  - 32 colors available

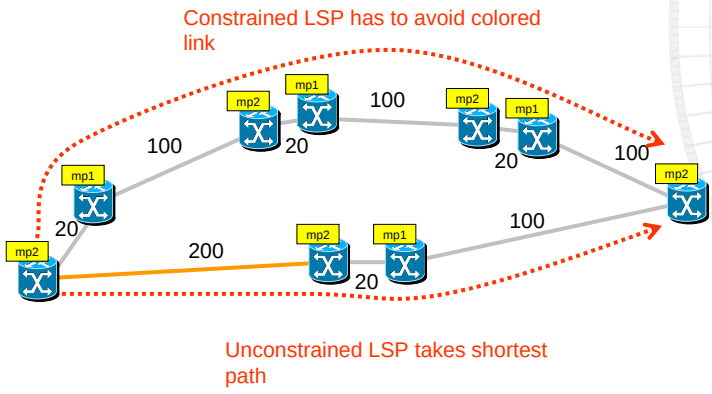
Link color



Modeling

- Maximize capacity by modeling traffic flows
- Model should include:
  - Link metrics
  - Link colors
  - LSP metrics and constraints
  - LSP traffic, city-to-city traffic, or AS-to-AS flows
  - Shared risk
    - logical links sharing same physical duct/fiber
  - LSP routing decision
- Steps in calculating routing
  - Calculate paths that LSPs take across network
  - Calculate IP routing decisions
    - LSPs used
    - Behavior under failure

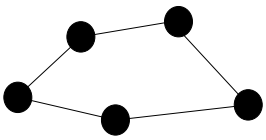
LSP modeling



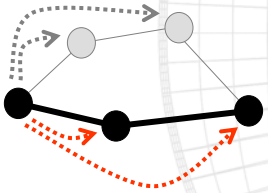
How IP uses LSPs – 3 options

A) Shortcuts

Step 1:
OSPF only

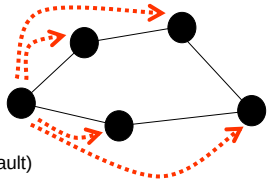


Step 2: Look for LSP shortcuts to routers on OSPF shortest path



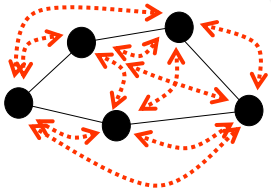
(Juniper default)

B) Locally-originated LSPs in SPF



(Cisco default)

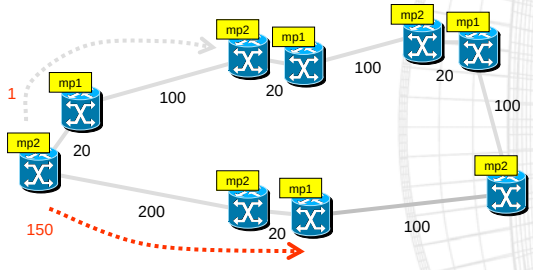
C) LSPs as links in OSPF area 0



IP routing examples

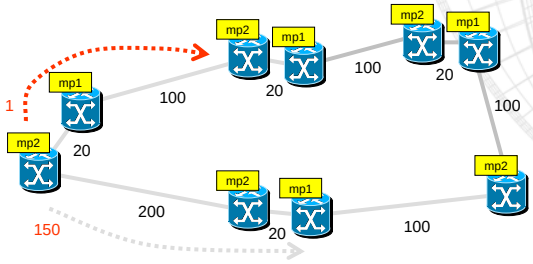
Shortcuts

1) Southern route selected by OSPF
2) Only LSPs to southern route hops considered as shortcuts



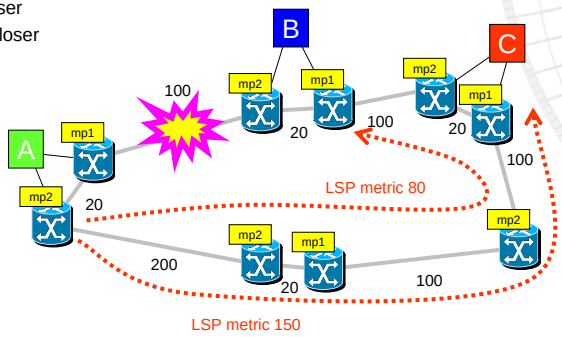
SPF includes locally originated LSPs

1) Northern route selected due to low metric LSP



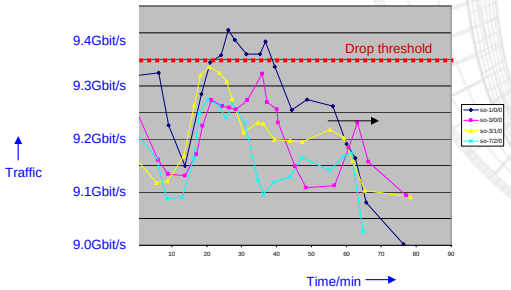
BGP next hop

- Which BGP next hop is closer to A after link failure – is it B or C?
  - C is the “IP” choice – closer with the new topology
  - B is the “ATM” choice – maintain existing traffic balance
- Technical choice – LSP or IP metrics
  - MPLS router Type 3 redistribution
    - IP so C is closer
    - LSP so B is closer



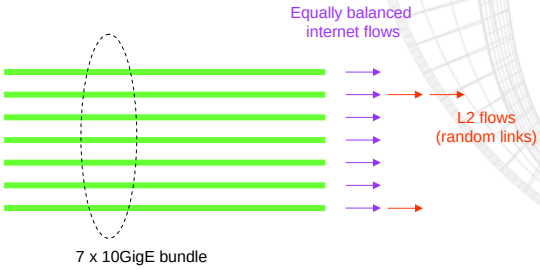
MPLS load balancing

- Internet IP traffic is balanced very well
  - Graph shows ~2% spread on 4 links
- Layer 2 streams are a problem
  - Tend to go on one link in bundle
  - N*10G bundle wastes (N-1)*flow



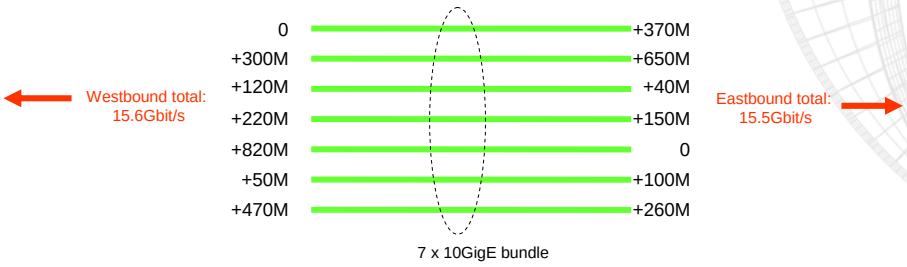
MPLS load balancing

- IP traffic carried on L2 tunnel
  - Should be OK to balance across multiple links based on IP addresses
  - Need to work with vendors
- Other large flows (IPSec, video…)
  - More challenging…



MPLS load balancing

- 30G (3x load balancing)
  - Clear correlation between 782M imbalance and large L2 flow (840M peak)
- Check on 70G link in both directions
  - Other factors may be involved, but unbalanced L2 flow is a concern



Conclusions

- MPLS is useful on large IP networks
  - Traffic engineering
  - Link Protection
  - New services
- Mesh RSVP label switched paths in the network backbone
  - Traffic engineering to avoid overload in failure
  - Link protection for rapid response to link failures
- Modelling traffic paths is important
  - IP routing over LSPs affects path taken
  - Layer 2 flows can unbalance N x 10G link bundles