MPLS JAPAN 2004

The MPLS Meta-Revolution

Kireeti Kompella

Introduction

- We are in the midst of an MPLS revolution
  - Network core and edge architectures are dominated by the use and implications of MPLS
  - Network services take into account MPLS applications: VPNs, QoS/CoS, guaranteed services
  - “Convergence” and “Multi-Service Edge” (MSE) are the watchwords of new network designs
- Almost every new network build-out requires MPLS, either immediately or in the near future

Broader Implications of MPLS

- MPLS has gone beyond a data plane paradigm of label switching and label stacks
  - MPLS has had a major influence on the way we think about protocol design and development
  - MPLS underlies almost every new network service
  - MPLS is changing the way we think about network provisioning and configuration
  - MPLS is no longer confined to the Layer 3 network “core” and “edge” -- at one end, it has penetrated the Layer 2 access network and at the other, it is impacting the design of the new Layer 1 transport infrastructure, in the form of Generalized MPLS

Topics of this Presentation

- MPLS's impact on protocol design
- MPLS's impact on the provisioning model
- MPLS's impact on Layer 2 access networks, especially Metro-Ethernet

1. MPLS and Protocol Design

- Pragmatism
- Extensibility
- MPLS's suite of support protocols
- Generalized MPLS

Pragmatism and Protocol Design

Tenets of protocol design reinforced by MPLS:

- Reuse and extend existing protocols when they match the requirements reasonably closely
  - RSVP-TE, ISIS-TE, OSPF-TE, MP-BGP
  - LDP and LMP are good examples of exceptions
- "Rough consensus and running code"
  - Many aspects of MPLS are in production deployment long before standardization
  - RSVP-TE and LDP were deployed quite early
  - IP VPNs (RFC 2547bis), Layer 2 circuits ("martini") and VPLS all are revenue-generating services

Protocol Extensibility

- Every protocol associated with MPLS is designed to be easily extensible
  - Underlying philosophy is that *we do not know what new applications will be developed! Don't guess, but plan ahead for evolution and reuse*
  - When MPLS was first proposed, no one visualized the use of the MPLS paradigm in optical networks, or the application of MPLS to Ethernet switching (VPLS)
- When designing protocol mechanisms, make them as simple and general as possible

The MPLS Suite of Protocols

- MPLS does not consist of a single protocol, but rather a suite of helper protocols
  - These include: RSVP-TE, LDP, ISIS/OSPF-TE, MP-BGP, LMP, BFD …
  - Each protocol has different mechanisms and modes of operation as well as target utility
  - For a given function, one can choose among several protocols to get the required behavior
    - This improves the chances of intelligent reuse

Generalized MPLS

- In many ways, GMPLS is a triumph of the flexibility and forward-looking nature of MPLS
  - Only one new protocol was invented to enable GMPLS, namely the Link Management Protocol, LMP
  - All other protocols were reused and extended
    - RSVP-TE, ISIS-TE, and OSPF-TE
    - New requirements (ASON, UNI, protection and restoration, multi-domain) have all been met with the above protocols
- The amazing thing about this is that MPLS is viewed by many as a *data plane* paradigm!

2. MPLS and Provisioning

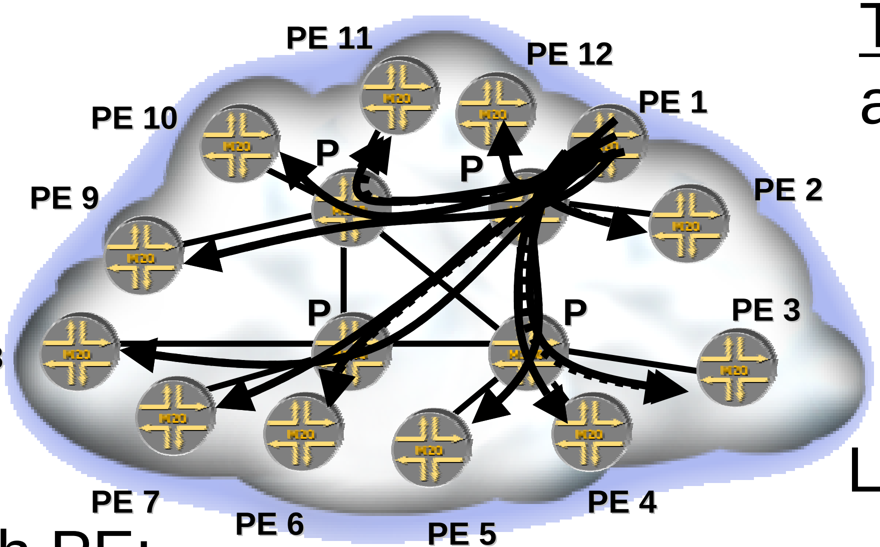
- Self-provisioning
- Auto-discovery
- Self-healing
- “Operational convergence”

Self-Provisioning

- MPLS for Traffic Engineering was initially seen as very operationally intensive
  - First, one had to produce a Traffic Matrix by measuring traffic between every pair of edge routers
  - Then, a full mesh had to be configured manually
  - This mesh of MPLS paths had to be monitored, both for liveness, and to verify the allocated bandwidth
  - Also, protection paths had to be manually configured
- There was no protocol help with any of this
  - This led to many carriers using LDP instead of RSVP

Self-Provisioning in Action

So much simpler! Total work is **O(N)** instead of **O(N²)**



The "old" way:
at each PE, do:

1. Measure b/w to all other PEs
2. Configure an LSP to every PE
3. Monitor all the (N-1) LSPs

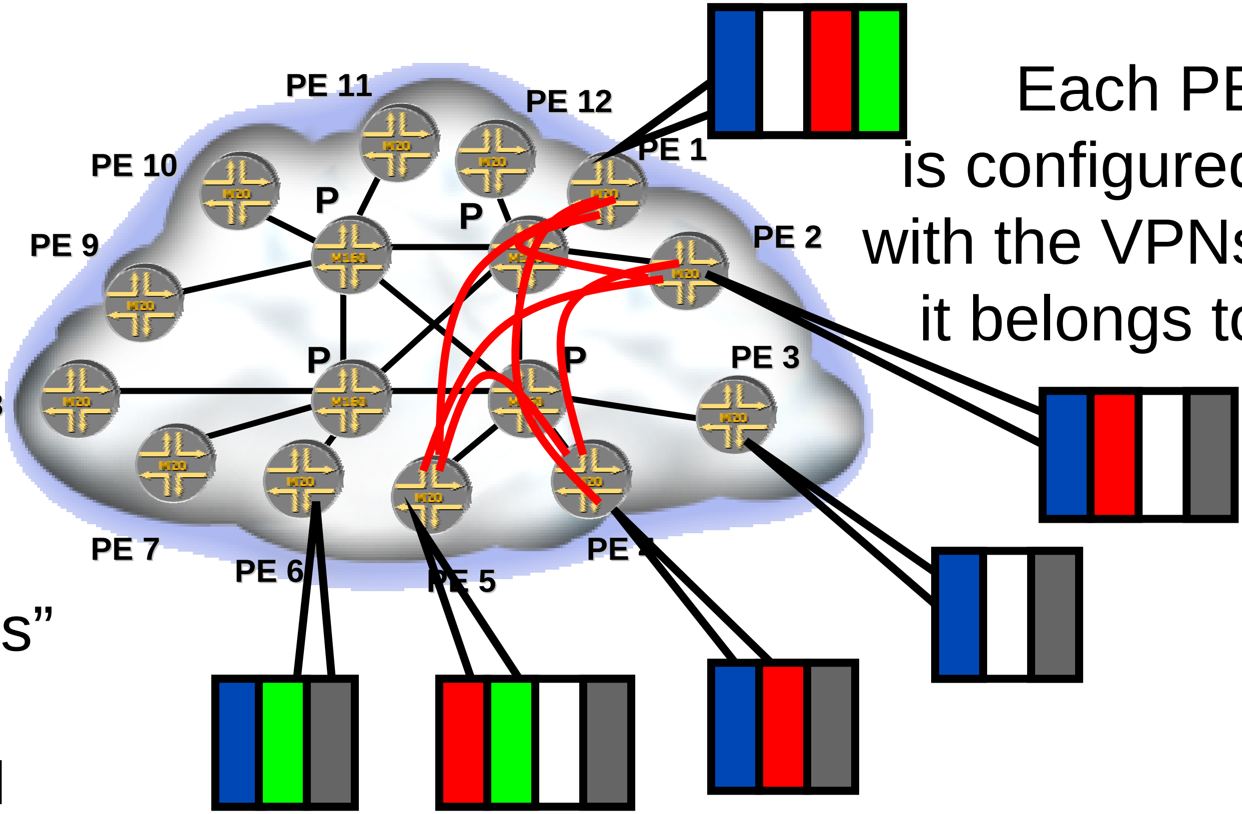
Instead, at each PE:

A. Configure "auto-mesh"
B. Configure "auto-bandwidth
C. Configure "fast reroute" (if desired)

Each PE will automatically create LSPs (with fast reroute) to all other PEs, monitor its bandwidth utilization and adjust reservatio

Auto-Discovery

Concept first introduced in BGP MPLS VPNs (RFC 2547)



Each PE is configured with the VPNs it belongs to

Each PE then "auto-discovers" other PEs in the same VPN

Again, an $O(N^2)$ task reduces to just O(N)

This same concept has been extended to Layer 2 VPNs, VPLS and even Layer 1 (Optical) VPNs

Self-Healing with MPLS

- MPLS enables self-healing
  - in the control plane using Graceful Restart techniques
  - as well as in the data plane, using fast reroute and/or standby paths
- Graceful Restart allows a node to gracefully recover its control plane state after a failure
- Fast Reroute allows for quick recovery from a link or node failure
- Standby Paths serve the same function, but slower, with less state and overhead

Operational Convergence

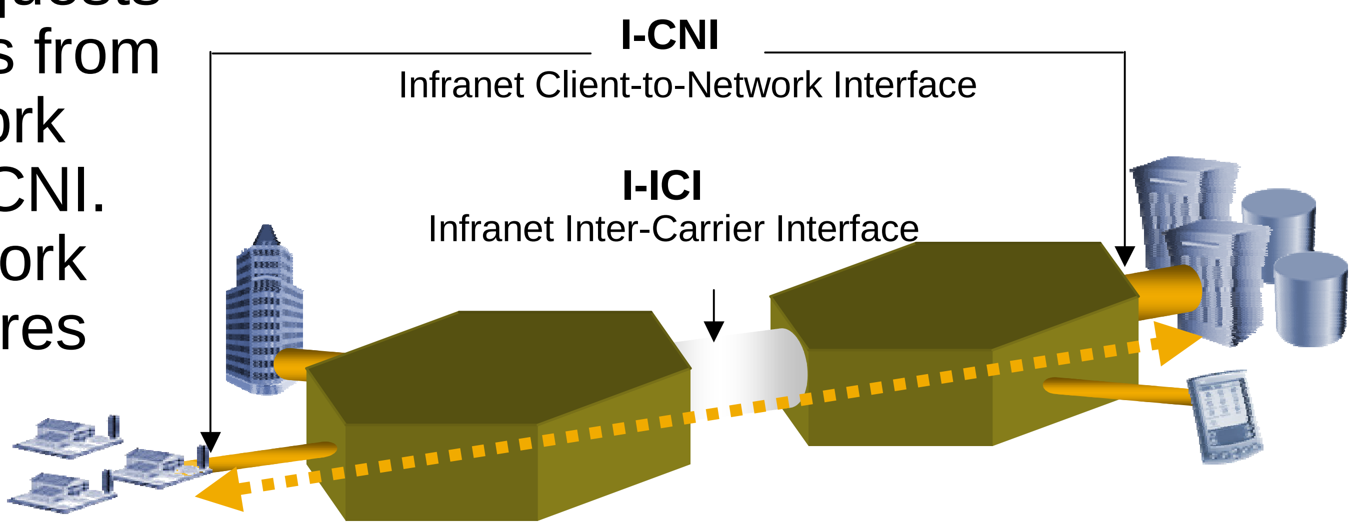
- There is much talk about "convergence" -- using a network (or a device) for multiple services
- Equally (or more) important is convergence of network **operations**
  - Using the same concepts for multiple services
- Examples in the context of MPLS abound
  - GMPLS uses MPLS concepts for transport networks
  - Auto-discovery works identically for many VPN types
  - New developments (point-to-multipoint LSPs, or inter-domain Traffic Engineering) reuse familiar concepts

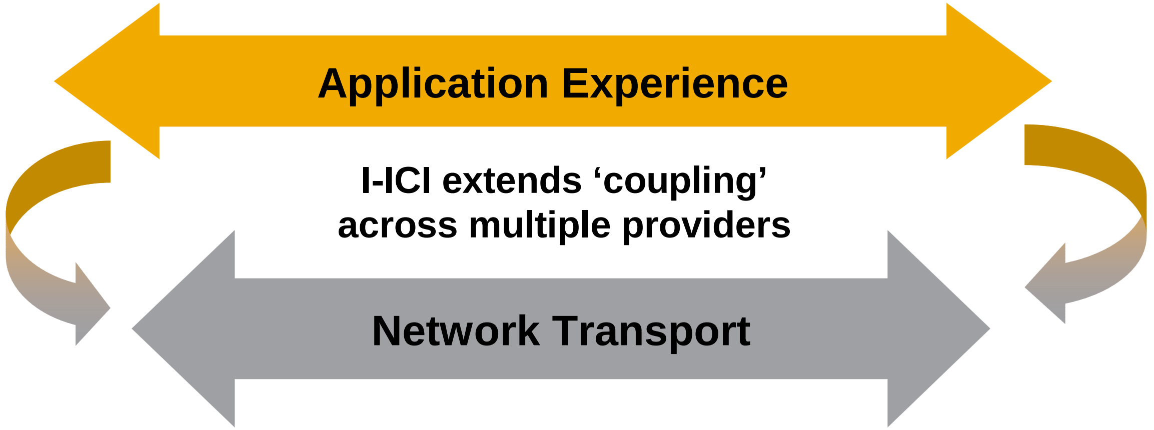
Self-Service Networking

- The next logical step is to off-load configuration management to the end customer
  - This both eases the provider's operational burden and offers more flexibility to the customer
- Doing this requires strong authentication and security (compare with ATM banking)
- Steps in this direction have been taken by the Infranet Initiative's Client-Network Interface
  - Similar efforts are underway at the DSL Forum

Infranet Client-to-Network Interface

Client requests resources from the network via the I-CNI. The network reconfigures itself in response





3. MPLS and Metro Ethernet

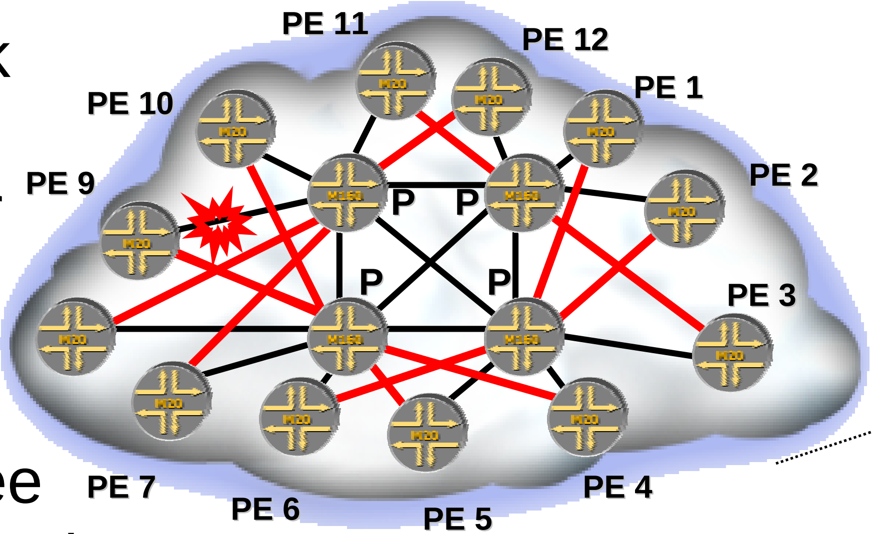
- Metro Ethernet is a very popular, fast-growing access technology, replacing Frame Relay/ATM
- However, Ethernet was not designed as a carrier-grade infrastructure for metro or WAN. Ethernet's deficiencies include:
  - Loops and broadcast storms are deadly
  - Spanning Tree doesn't allow load balancing
  - Detecting link failures is slow
  - Reacting to link failures is deliberately slow
  - Pure Ethernet switching does not scale
  - No means of offering services other than switching

Layer 2 Loops and Broadcast Storms

- If a loop occurs, packets loop for ever -- there is not “time-to-live” in Ethernet forwarding
  - This is exacerbated if packets are also replicated, for example, for broadcast traffic or flooding
- In many cases, the only way to recover is to physically turn off switches
  - Remote administration is often not possible
- This is completely unacceptable for carrier-class networks
  - It may be acceptable for a LAN within a building

Layer 2 Loops and Spanning Tree

Every carrier-class network must have extra links for redundancy



However, this will cause a devastating loop in an Ethernet environment

Spanning Tree Protocol is used to remove loops

But now, these extra links are wasted! They cannot be used for load balancing

If a link breaks, one of the extra links must be re-enabled to re-establish connectivity -- but this is done slowly and carefully

Scaling and Hierarchy (or Stacking)

- Scaling Metro Ethernet networks is very hard
  - VLANs have a limit of 4096, which is quite small
  - If customer MAC addresses are visible throughout the switched network, each switch will have to hold millions of MAC addresses
    - This is not scalable, secure or manageable
- Solutions to these issues include VLAN stacking and “MAC-in-MAC”
  - However, these solutions are very specific to Ethernet, and do not address manageability
  - Also, stacking is limited to one level

MPLS is the Right Approach

- Using MPLS in a Metro Ethernet network means
  - Ethernet switching is not used **at all**
  - Ethernet is just a transport (analogous to SDH)
  - Transient loops do not bring down the network
  - All links can be used; load balancing is possible
  - Fast detection and fast recovery are possible
- Furthermore, MPLS is needed in any case to offer <u>customer</u> services such as IP VPNs, VPWS and VPLS, as well as <u>infrastructure</u> services like fast reroute and Traffic Engineering

MPLS in the Future

- MPLS is at the perfect cusp of maturity and innovation
  - Five years ago, we were arguing about protocols, and still trying to understand where MPLS would be useful
  - Now, MPLS has proved itself, and is deployed in most networks, and in fact carries high-revenue services
  - Yet, MPLS innovations have not ceased, especially in the area of services and usability
    - MPLS Operations and Management (LSP ping)
    - Point-to-multipoint MPLS (multicast)
    - Virtual Private LAN Service (VPLS/MPLS LAN Emulation)

MPLS Dreams

- MPLS+GMPLS combine to offer self-managed, self-reconfiguring networks
  - Manual provisioning focuses on the service edge
- MPLS offers an effective access network
  - Legacy ATM-to-MPLS interworking
  - New metro Ethernet services
- MPLS continues to evolve and match new service needs and expectations

Thank You!