Modern L2 VPNs –Implementing Network Convergence

By Scott Gross

NANOG 33

Jan.-Feb. 2005

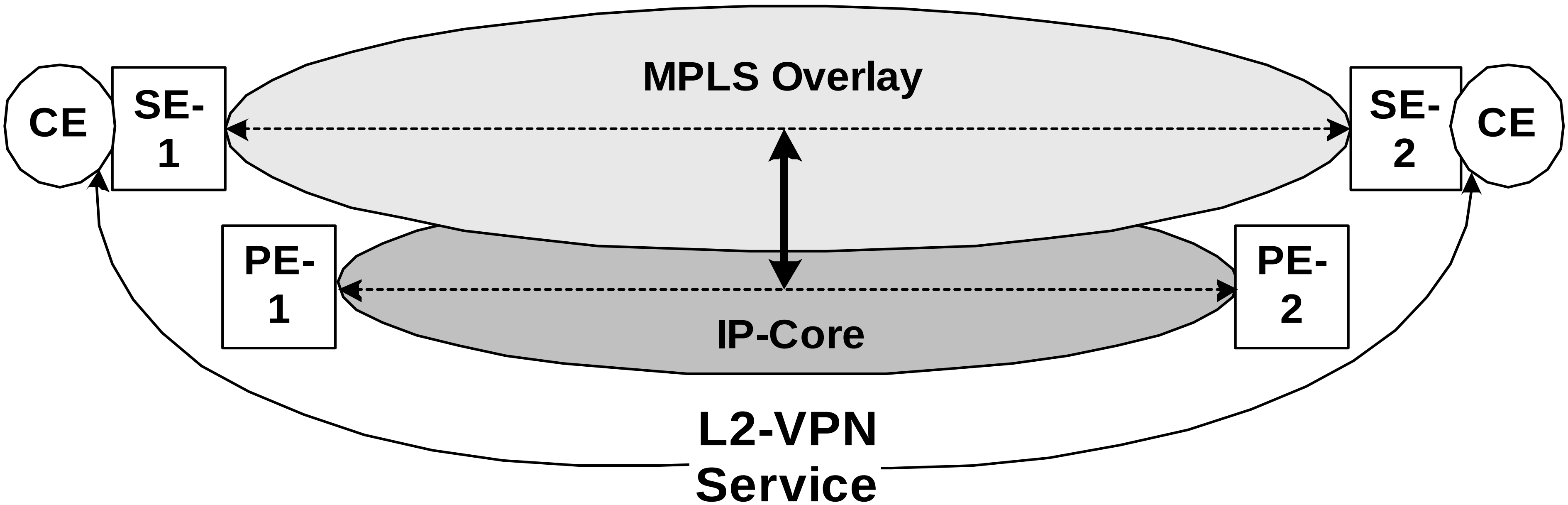
Network Convergence

- Our goal:
  - Interconnect Regional Networks to extend L2-VPN service with minimal cost
  - Point-to-Point
    - 802.1Q (tagged), Frame Encapsulation (FE), ATM
    - 802.1Q-FE; 802.1Q-ATM; and FR-ATM interworking
    - Implemented
  - Multipoint-to-Multipoint (broadcast/mutlicast)
    - Virtual Private LAN Service (VPLS)
    - Have yet to implement

Network Convergence

- Our solution:
  - Build an independent MPLS Overlay network using existing IP infrastructure
  - Customer Edge (CE) – Service Edge (SE) MPLS LSR- Provider Edge (PE) IP-Core Router

Convergence: General Model



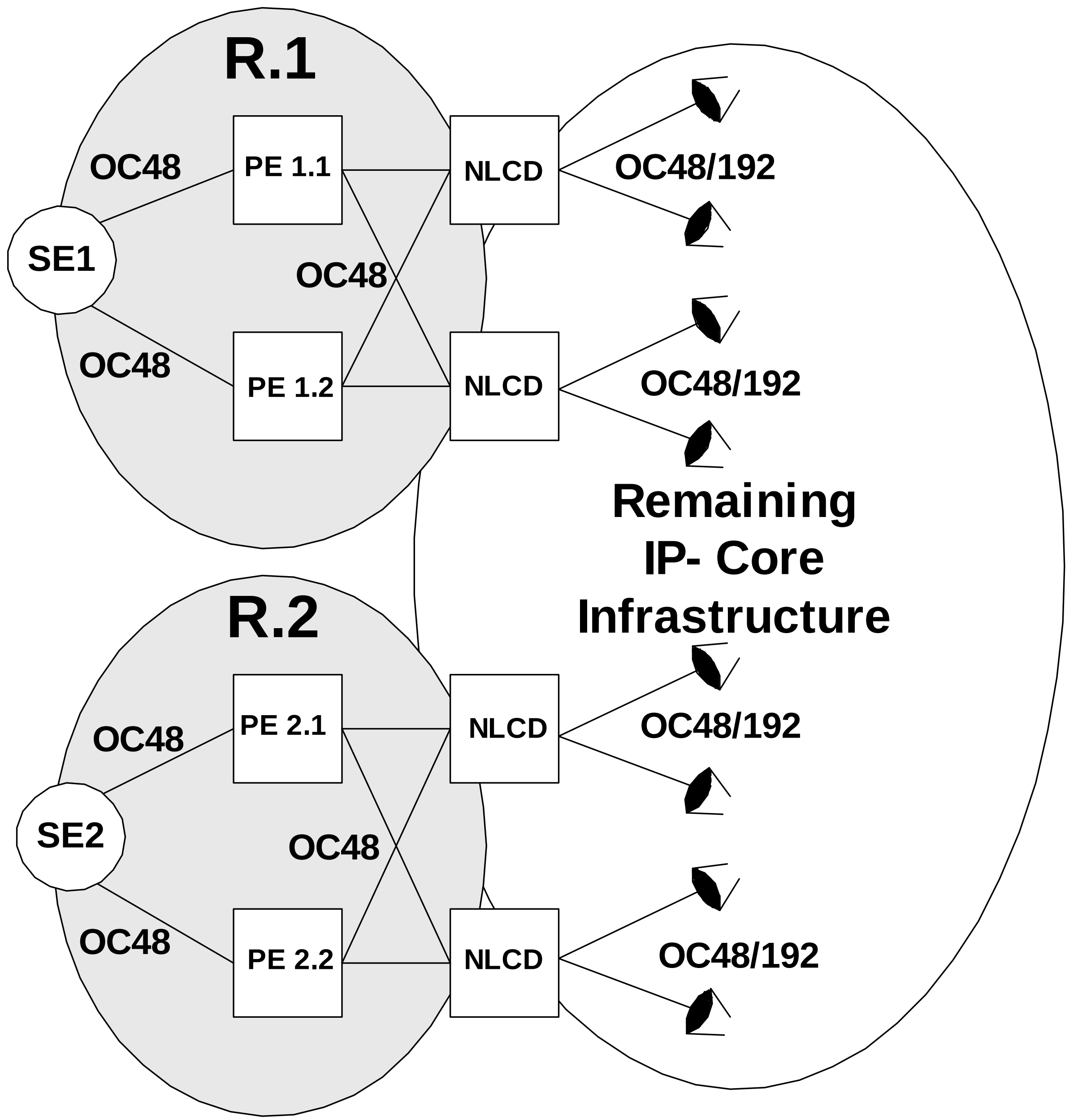
Network Convergence

- Requirements
  - Stable IP-core
  - Maintain Service separation among MPLS Overlay and IP-core
  - Independent routing protocol functionality among MPLS Overlay and IP-core

Network Convergence Requirements

- Stable IP-core
  - Must provide diversity
    - Multiple access nodes (PE) per region (R)
  - Must provide redundancy at SE-PE edge
  - Must possess internal redundancy
    - Multiple links PE to Next Level Core Device (NLCD)
  - Must possess sufficient bandwidth throughout
    - ≥ SE-PE attachment circuit

Regional Diversity/Redundancy



Network Convergence

- Requirements
  - Stable IP-core
  - Maintain service separation among MPLS Overlay and IP-core
  - Independent routing protocol functionality among MPLS Overlay and IP-core

Network Convergence Requirements

- Maintain Service Separation
  - Keep service expertise intact
    - Operations
    - Engineering
  - Viability depends on number of services offered
  - Operations complexity may increase
    - Organization of fault detection and isolation paths

Network Convergence

- Requirements
  - Stable IP core
  - Maintain service separation among MPLS Overlay and IP-core
  - Independent routing protocol functionality among MPLS Overlay and IP-core

Network Convergence Requirements

- Independent Routing Protocol Function
  - GRE tunnels chosen for simplicity and vendor availability
    - PPP chosen to terminate tunnel endpoints (SE-PE)
    - Static routes provide tunnel endpoint reach ability
  - Allows additional services to be added to IP-core without requiring changes to the MPLS Overlay
  - While routing protocols function separately, the MPLS Overlay IGP relies on IP-core IGP for transport of PDUs

GRE Tunnel Contents

GRE –RFC2784

32 bits

C	Reserved0	Ver	Protocol Type
Checksum (optional)			Reserved1 (optional)

C=Checksum Present Bit (bit 0): Indicates presence of optional Checksum and Reserved1 fields -Set to 0 indicating these fields are not present (per draft-ietf-mpls-in-ip-or-gre-08)

Reserved0 (bits1-12): Bits 1-5 are set to 0 or discarded -unless RFC1701 (N/A); bits 6-12 are set to zero (future use)

Ver=Version Number (bits 13-15): Set to 0 per RFC2784

Protocol Type (2 octets): Protocol type of Payload packet
In hex: IS-IS=00FE; IP=0800; MPLS=8847

Result: 4 byte header

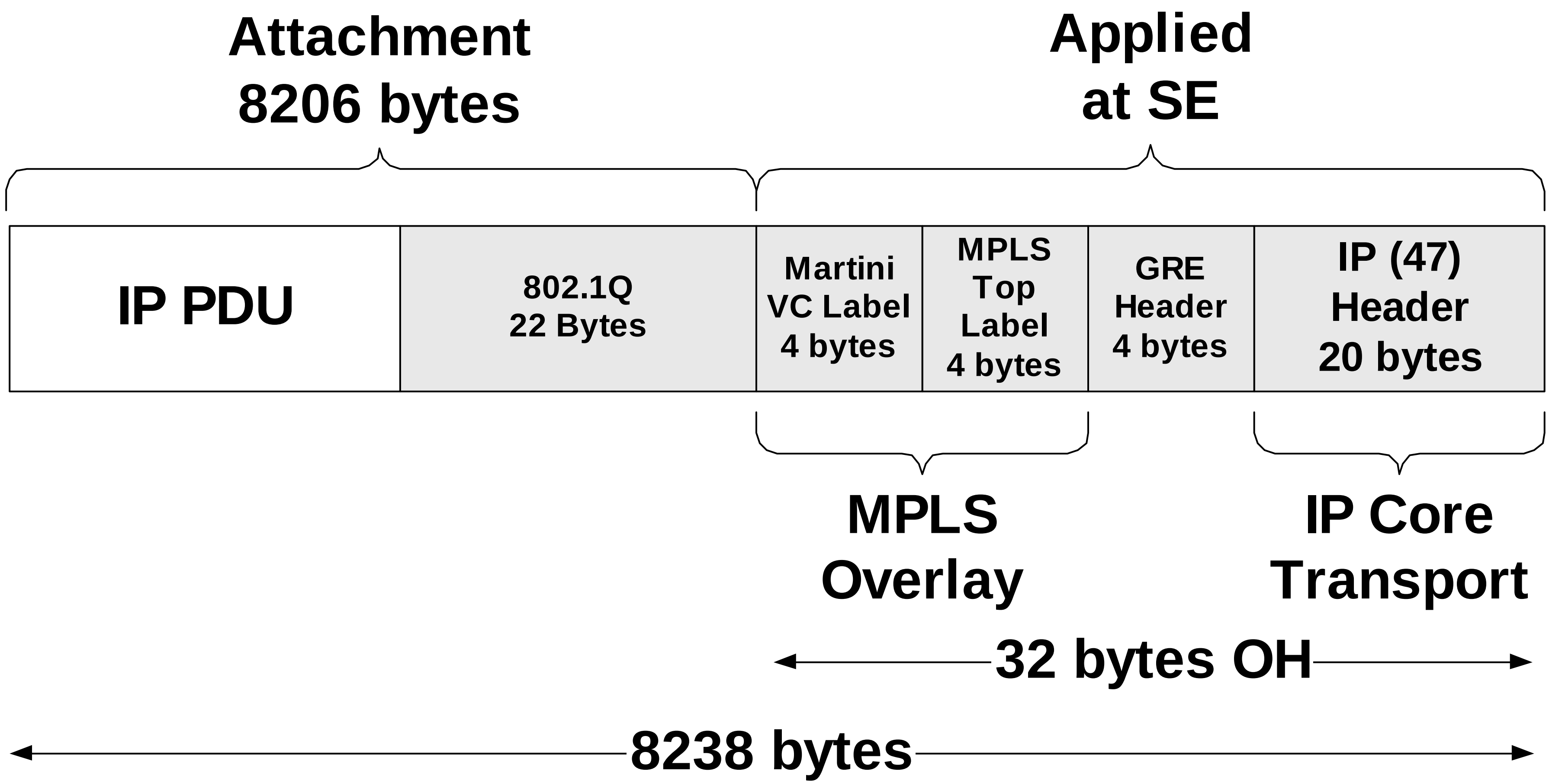
MTU Considerations

- Avoiding Fragmentation
  - Reassembly at GRE tunnel tail places an undesirable burden on resources (performance hit)
    - Path MTU Discovery (PMTU) is NOT performed (vendor implementation)
    - Don’t Fragment (DF) bit is set at tunnel head per draft-ietf-mpls-in-ip-or-gre-08 suggested default

MTU Considerations

– MTU settings in IP-core are set to accommodate largest possible packets from Service Edge
  - Using 802.1Q at 8206 bytes for attachment
  - Add MPLS: VC and Tunnel labels -8 bytes
  - Add GRE and IP header -24 bytes
  - Calculated encapsulation OH = 32 bytes

Encapsulation Overhead



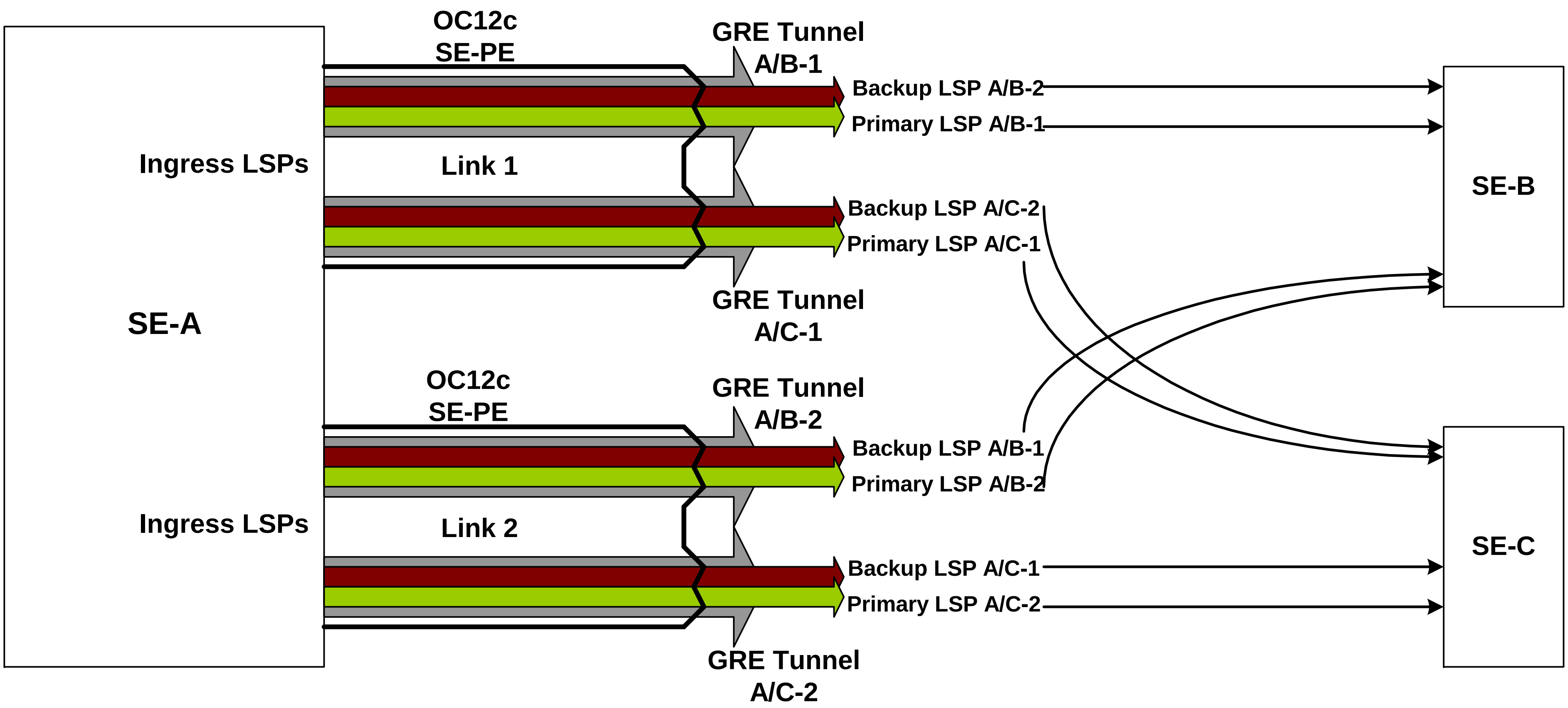
MTU Considerations

- IP-core IP MTU is set to 9100 with a media MTU of 9192
  - Accommodates largest packet from Service Edge
  - Allows L2 header growth up to 92 bytes (currently 4 bytes with PPP)

Design Details

- PPP→IP→ GRE→MPLS→L2 Services
  - Diversity and redundancy
    - SE-PE: PPP → IP
    - SE-SE: GRE → MPLS
  - Dual tunnels; fully meshed
  - Dual LSPs; fully meshed

Diversity and Redundancy Model



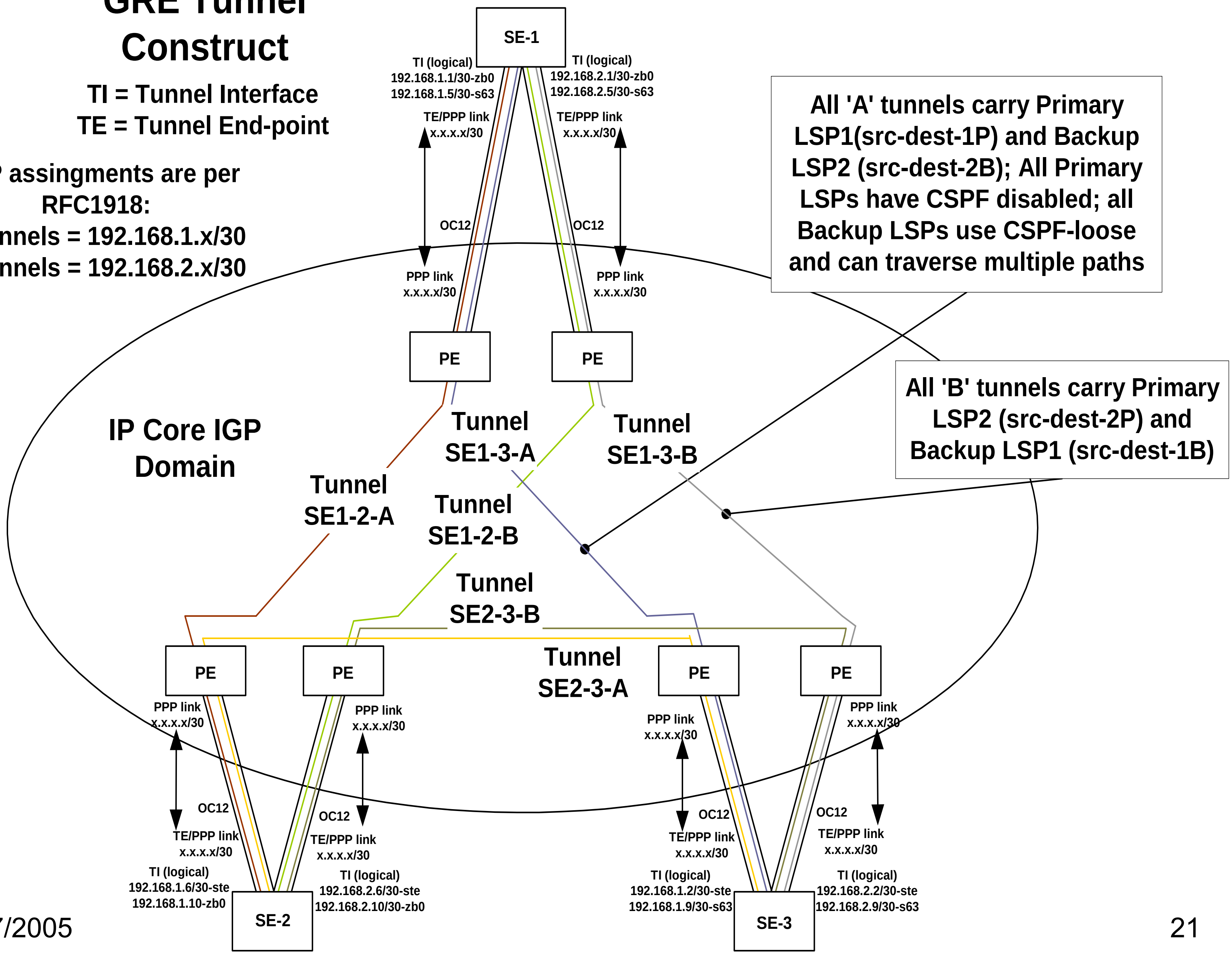
Design Details

- PPP→IP→GRE→MPLS→L2 Services
  - Diversity and redundancy
    - SE-PE: PPP → IP
    - SE-SE: GRE → MPLS
  - Dual tunnels; fully meshed
  - Dual LSPs; fully meshed

GRE Tunnel Construct

TI = Tunnel Interface
TE = Tunnel End-point

TI IP assingments are per RFC1918:
A Tunnels = 192.168.1.x/30
B Tunnels = 192.168.2.x/30

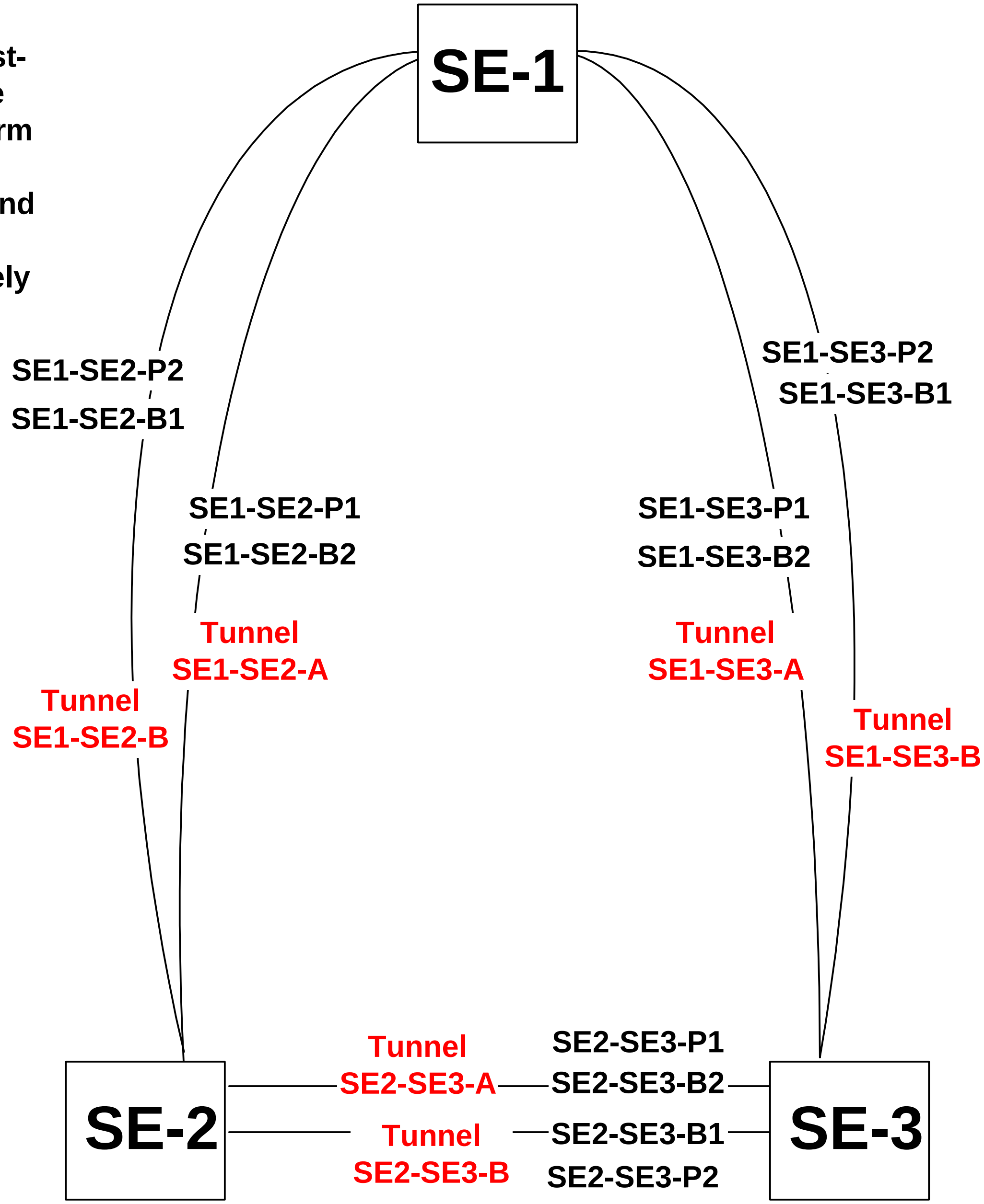


Design Details

- PPP→IP→GRE→MPLS→L2 Services
  - Diversity and redundancy
    - SE-PE: PPP → IP
    - SE-SE: GRE → MPLS
  - Dual tunnels; fully meshed
  - Dual LSPs; fully meshed

LSP Construct

LSPs defined as src-dest-P1 and src-dest-P2 are Primary LSPs that perform load-balancing; LSPs defined as src-dest-B1 and src-dest-B2 are active Backup LSPs respectively



Summary

- L2 VPN service was successfully extended coast-to-coast
  – Avg. RTD 77ms @ 4470 bytes
- Rapid service deployment
  – Less than a year from concept to reality
- Achieved at minimal cost
  – Pre-existing platforms used
- Operations & Engineering complexity increases

References

- IETF draft-martini-l2circuit-encap-mpls-08
- IETF draft-ietf-mpls-in-ip-or-gre-08
- IETF RFC2784 (GRE Specification)
- IEEE 802.1Q Virtual LANs

End

Scott@mae.net