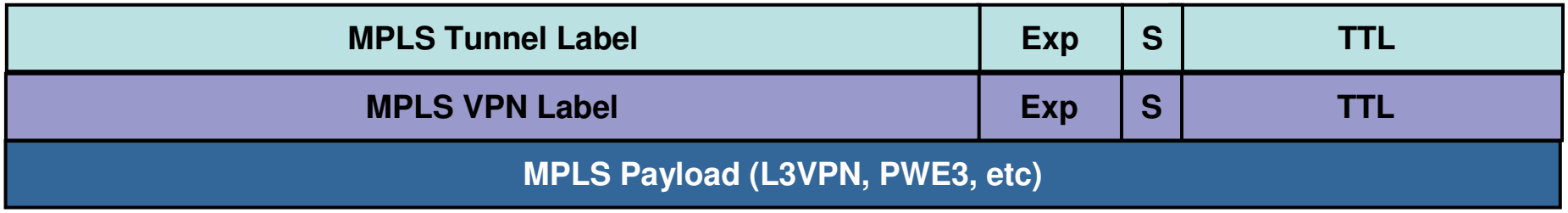
MPLS over IP-Tunnels

Mark Townsley
Distinguished Engineer

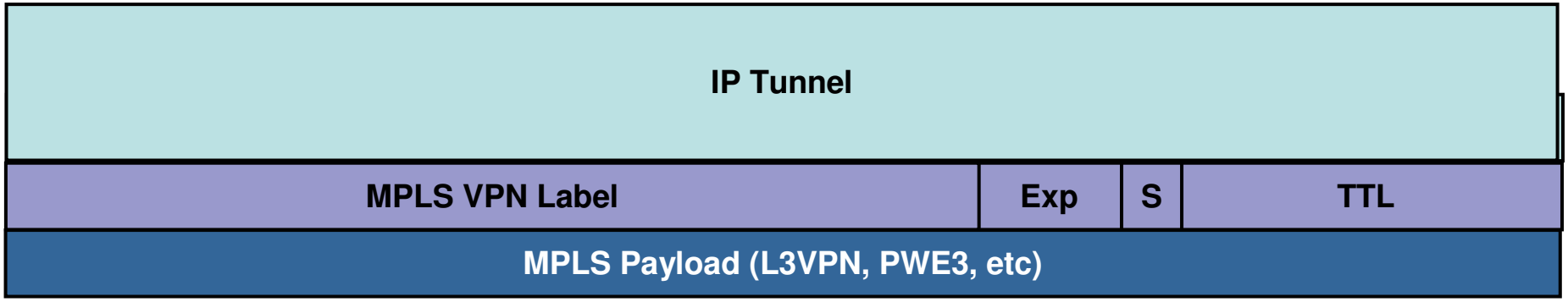
21 February 2005

MPLS over IP – The Basic Idea



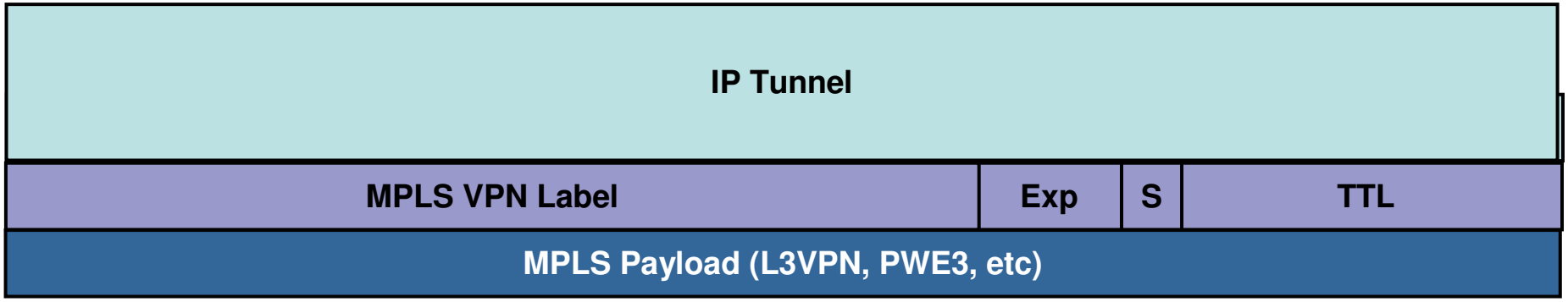
- MPLS Tunnel Label transports MPLS-labeled VPN packets between PEs. It is swapped along the LSP from one PE to another.
- MPLS VPN Label remains the same between PEs. It is exchanged via targeted LDP, MP-BGP, etc. and refers to a VRF, VPLS VSI, or PWE3 VC.

MPLS over IP – The Basic Idea



- MPLS Tunnel Label transports MPLS-labeled VPN packets between PEs. It is swapped along the LSP from one PE to another.
- MPLS VPN Label remains the same between PEs. It is exchanged via targeted LDP, MP-BGP, etc. and refers to a VRF, VPLS VSI, or PWE3 VC.

MPLS over IP – The Basic Idea

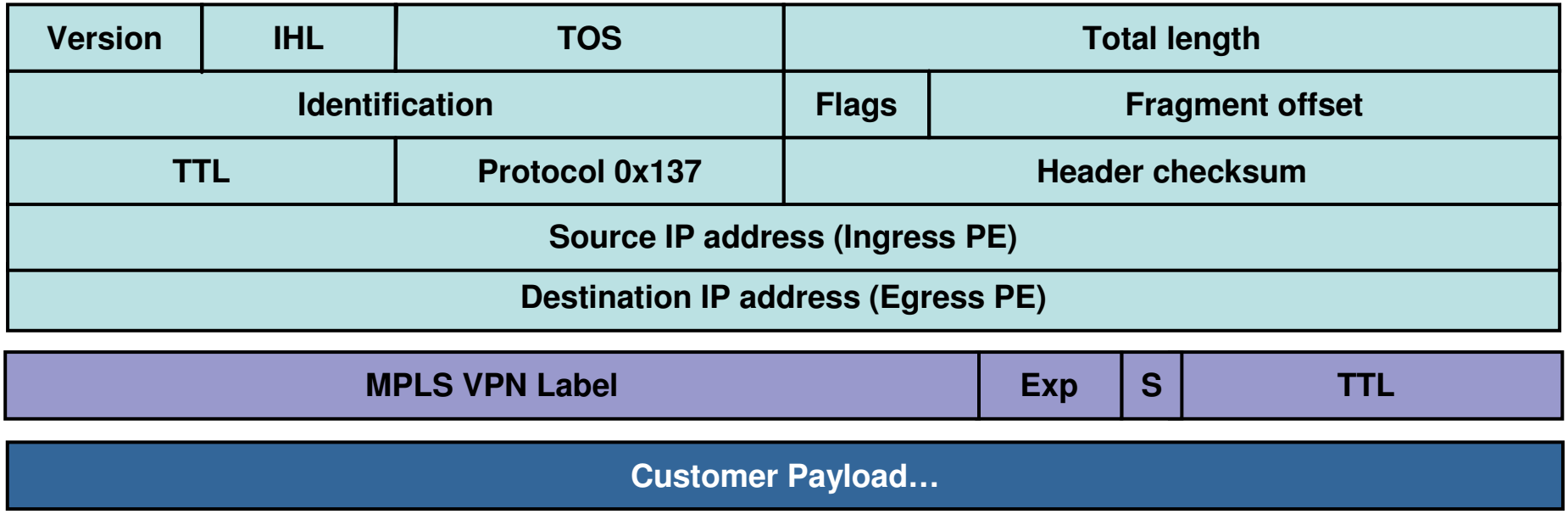


- MPLS Tunnel Label is replaced with an IP Tunnel, which performs the same function of getting the MPLS VPN label and payload between PEs
- Unfortunately, we have a few IP tunnels to choose from – each with different pros and cons

A Long Evolution Leading to Many Optoins…

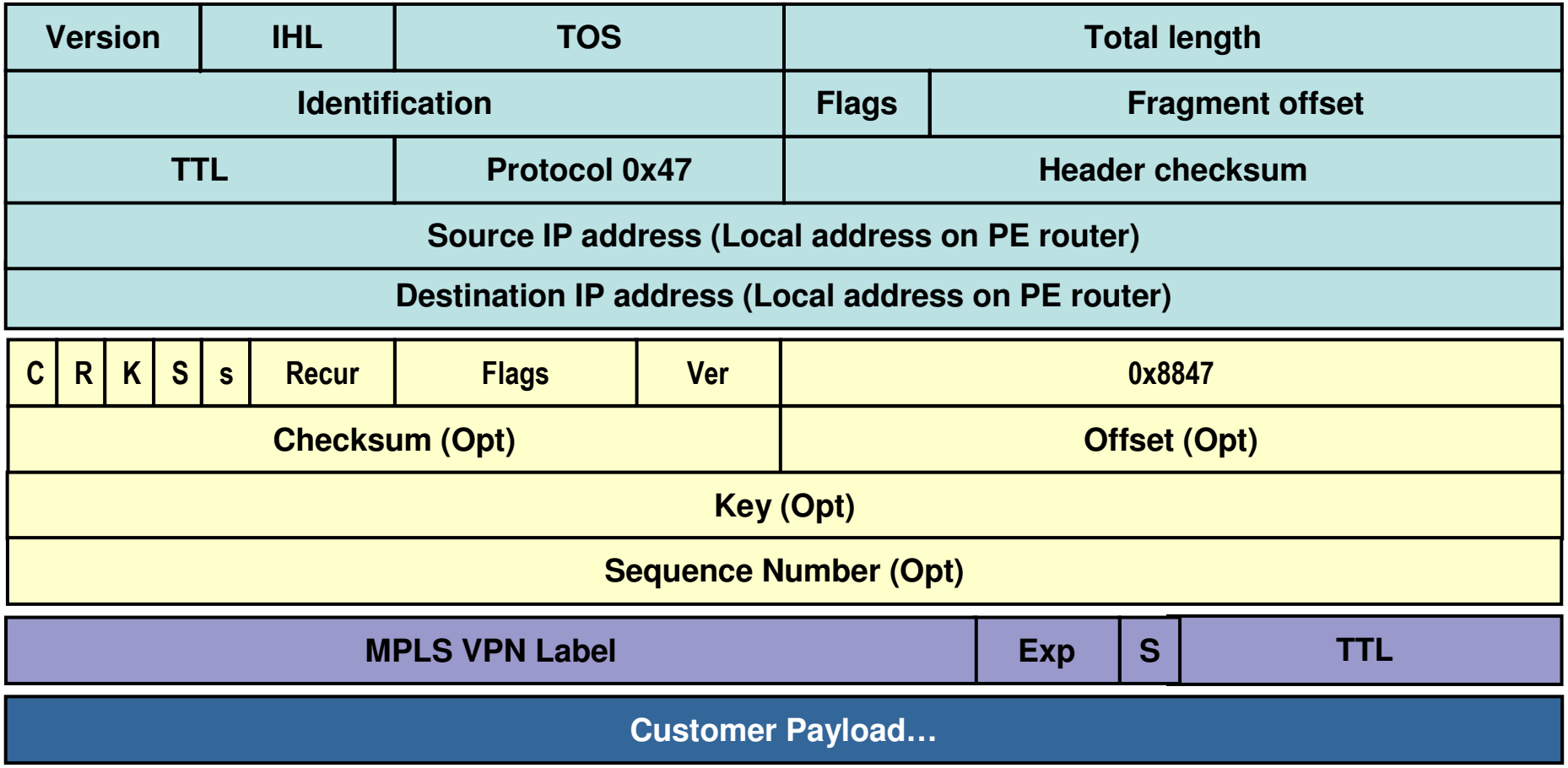
- Unfortunately, there are a lot of choices to wade through when it comes to MPLS over IP
  - MPLS directly over IP
  - MPLS over “Full” GRE/IP
  - MPLS over “Simple” GRE/IP
  - MPLS over L2TPv3 w/BGP Tunnel SAFI
  - Each of the above with IPsec
  - Point-2-Point vs. Point-2-Multipoint…
- This presentation will walk through the evolution of each of these methods of carrying MPLS over IP, leading us to where we are today

MPLS over IP Tunneling Technologies
MPLS over IP



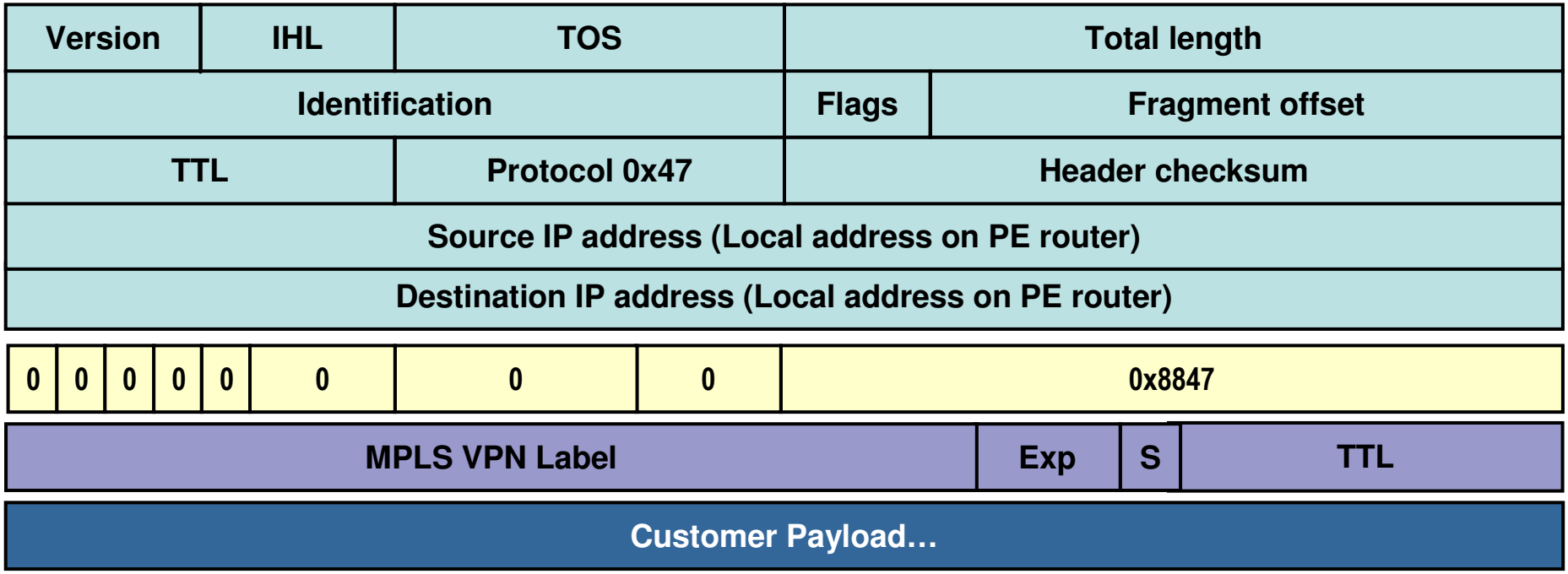
- Defined in draft-ietf-mpls-over-ip-or-gre-08.txt
- Smallest and simplest of MPLS over IP encapsulations (just +16 bytes)
- Not widely supported today

Tunneling Technologies
MPLS over "Full" GRE Header



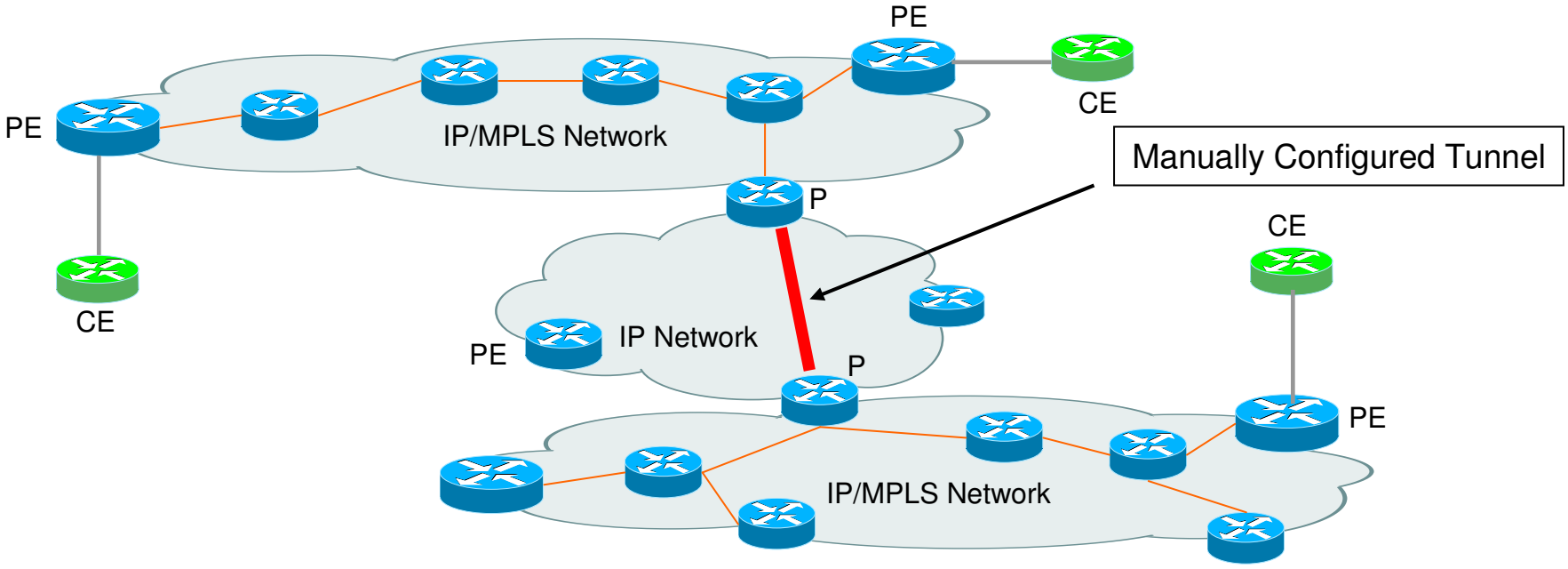
- Defined in draft-ietf-mpls-over-ip-or-gre-08.txt
- Also not widely supported today

Tunneling Technologies
MPLS over "Simplified" GRE Header



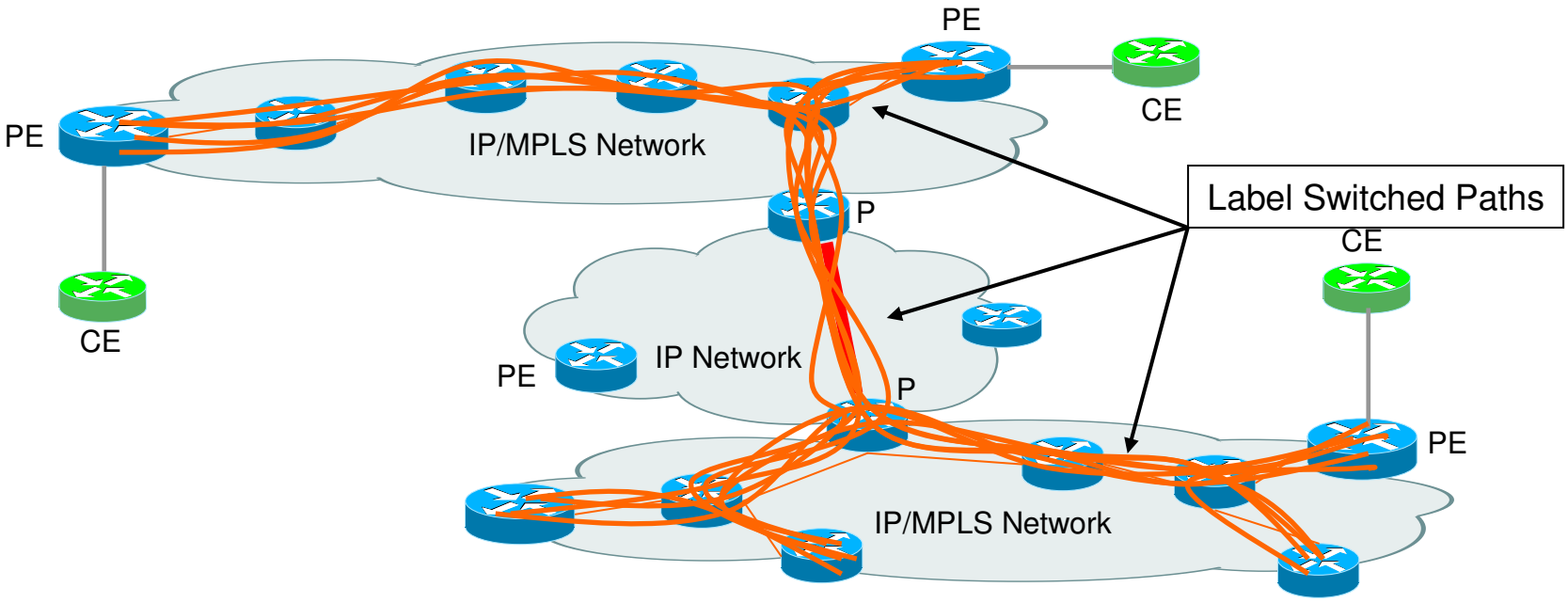
- **Most widely supported, particularly for manually configured, point to point tunnels**
- **Larger encapsulation than MPLS over IP, but with no tangible advantage as the GRE Header is simply reduced to a constant set of bits in each packet**

Manually Configured Overlay (GRE)



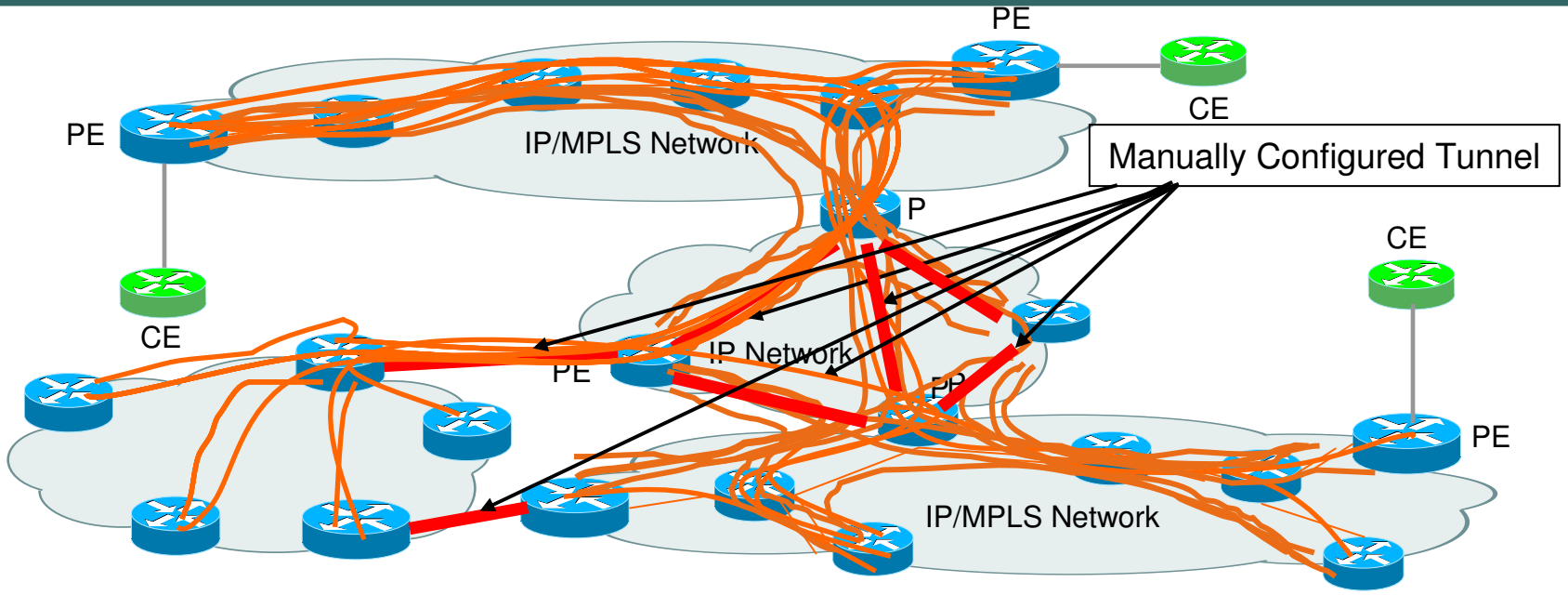
- **Manual Point-to-Point GRE Tunnel**
- **Connects disparate MPLS networks.**
- **Separate MPLS networks act as one, so all services enabled by MPLS are available across both clouds**
- **This was, and still sometimes is, a good thing… But…**

Manually Configured Overlay (GRE)



- **Number of LSPs are multiplied, setup between all nodes on BOTH networks**
- **IP-only PE Nodes Still Isolated**
- **Traffic may not traverse optimal path between PEs**

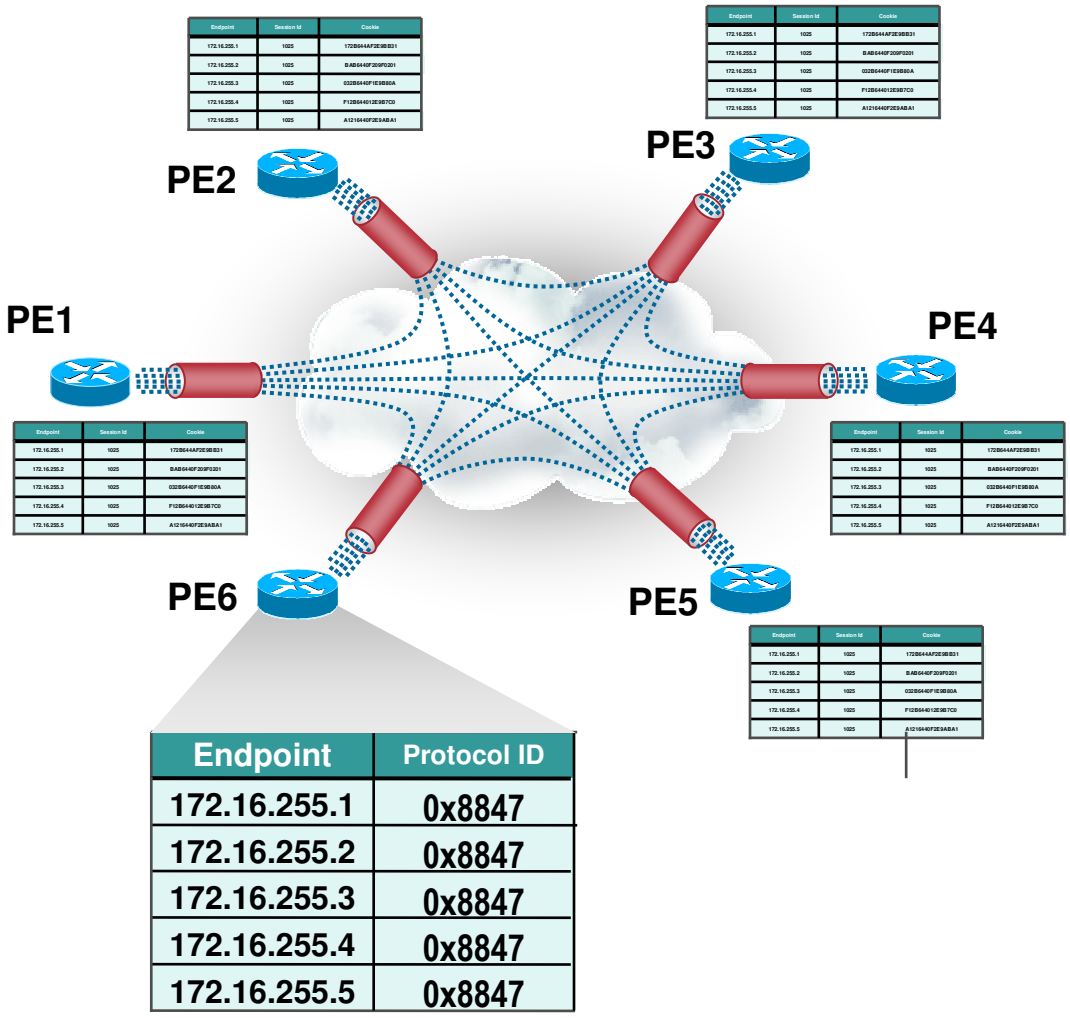
Manually Configured Overlay (GRE)



- **Each tunnel enlarges the single, flat, MPLS network. MPLS sees no hierarchy or partitioning.**
- **At high scale, manual GRE overlay network becomes cumbersome to manage and burdensome on number of LSPs and IGPs carrying /32 routes for all PEs**
- **Needed: Dynamic point-2-multipoint tunnels between PEs**

MPLS over Dynamic Multipoint GRE

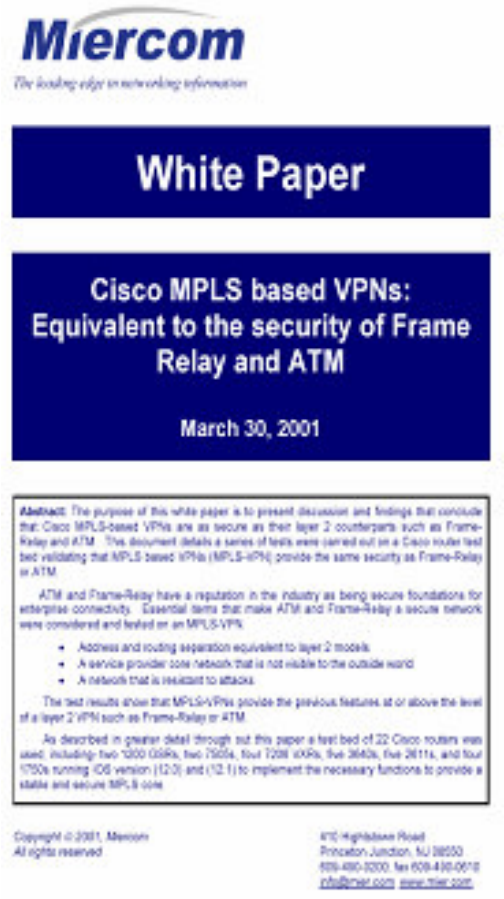
- One Multipoint GRE Tunnel is dynamically created on each PE for receiving traffic from other PEs
- But.. Mixed tunneling environments are not easily supported – if other PEs cannot decapsulate MPLS over GRE then VPN traffic could be blackholed
- Still Needed: A method for PEs to advertise if they are able to receive MPLS over IP traffic, and with what type of encapsulation



BGP Tunnel SAFI to The Rescue!

- draft-nalawade-kapoor-tunnel-safi-02.txt
- Defines a SAFI which binds a tunnel endpoint (PE IP address) to a set of tunnel capabilities:
  - Type 1 : L2TPv3 Tunnel information (Session, Cookie)
  - Type 2 : mGRE Tunnel information (Header Type, Key, etc)
  - Type 3 : IPSec Tunnel information (Security Association)
  - Type 4 : MPLS Tunnel information (Native MPLS)
- With this information being advertised along with the BGP Next Hop, PEs will only receive data for which they are able to properly decapsulate
- Policies may be defined – e.g., encrypt some tunnels, not others

What about Security?
Quick Review: MPLS VPN Security


Miercom

The leading edge in networking information

White Paper

Cisco MPLS based VPNs: Equivalent to the security of Frame Relay and ATM

March 30, 2001

Abstract: The purpose of this white paper is to present discussion and findings that conclude that Cisco MPLS-based VPNs are as secure as their layer 2 counterparts such as Frame-Relay and ATM. This document details a series of tests were carried out on a Cisco router test bed validating that MPLS based VPNs (MPLS-VPN) provide the same security as Frame-Relay or ATM.

ATM and Frame-Relay have a reputation in the industry as being secure foundations for enterprise connectivity. Essential items that make ATM and Frame-Relay a secure network were considered and tested on an MPLS-VPN.

- Address and routing separation equivalent to layer 2 models
- A service provider core network that is not visible to the outside world
- A network that is resistant to attacks

The test results show that MPLS-VPNs provide the previous features at or above the level of a layer 2 VPN such as Frame-Relay or ATM.

As described in greater detail through out this paper a test bed of 22 Cisco routers was used, including two 12000 GSRs, two 7500s, four 7206 VXRs, five 3640s, five 2611s, and four 1750s running IOS version (12.0) and (12.1) to implement the necessary functions to provide a stable and secure MPLS core.

Copyright © 2001, Miercom
All rights reserved

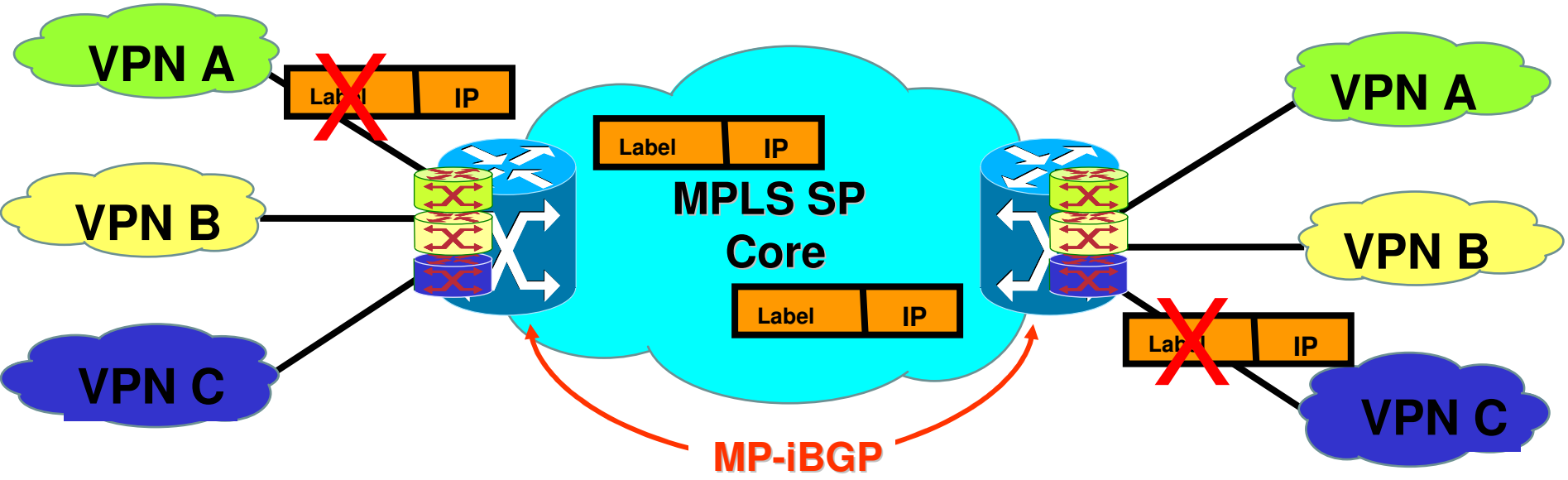
410 Highbdown Road
Princeton Junction, NJ 08550
609-490-0200 fax 609-490-0610
info@mier.com www.mier.com

Meircom performed testing that proved that MPLS-VPNs have met or exceeded all of the security characteristics of a comparable layer two based VPN such as Frame-Relay or ATM.

http://www.mier.com/reports/cisco/MPLS-VPNs.pdf

In addition : Cisco Security White Paper on MPLS http://www.cisco.com/warp/public/732/Tech/mpls/docs/0701_mpls_security_pu.fm.pdf

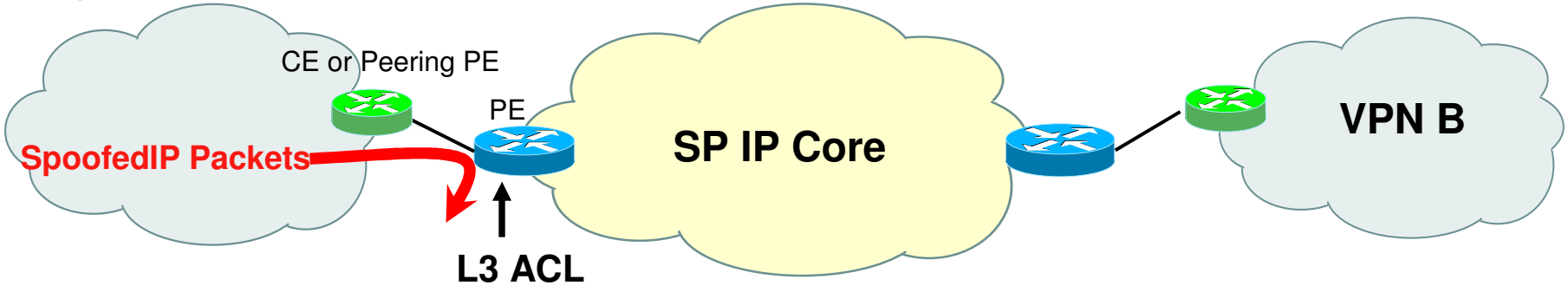
Quick Review: MPLS VPN Security



- Working assumption for MPLS VPN Security: The core network (PE+P) is secure
- MPLS-labeled packets will always be dropped on core boundaries.

MPLS over GRE
Security in an IP network

Single line of defense:



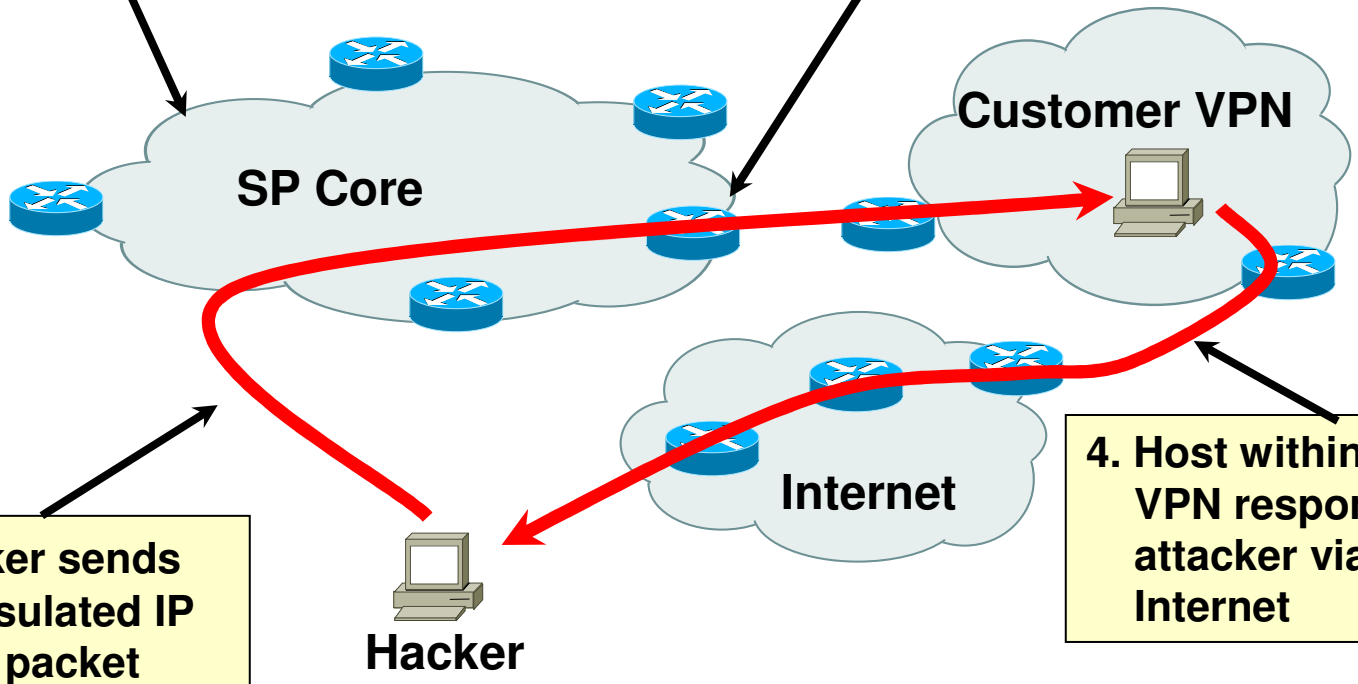
- **MPLS over GRE alone relies 100% on L3ACLs to protect VPN from spoofed data**
- **ACLs throughout the network can be operationally cumbersome (SA and DA address lists at each PE and border routers), could affect performance, subject to misconfiguration, etc.**
- **All it takes is *one* correctly spoofed MPLS label to infiltrate a customer VPN…**

VPN Services over IP Tunnels

Blind Insertion attack for VPN access

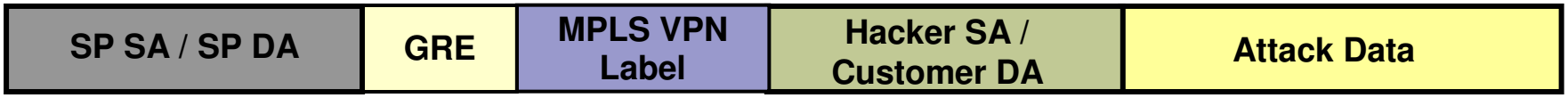
2. SP Core accepts encapsulated attack with valid IP source and destination of PEs

3. Egress PE receives an MPLS VPN over GRE packet and -- If the MPLS label is correctly chosen – routes the packet into the customer VPN



1. Attacker sends encapsulated IP attack packet

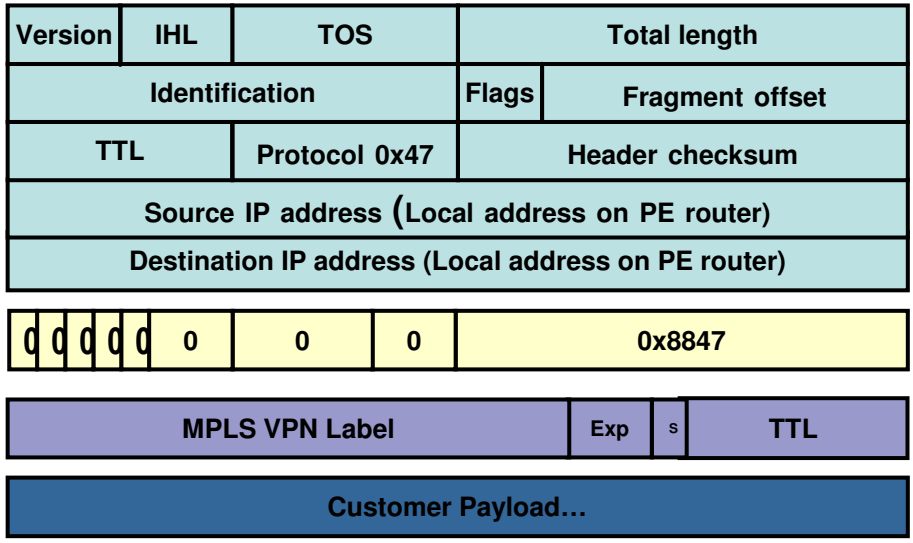
4. Host within customer VPN responds to the attacker via the Internet



Spoofing MPLS over GRE

- Service-Provider IP addresses can be discovered or easily guessed
- GRE Header contains constant, well-known values
- MPLS Label is 20-bits of variant data that must be guessed by hacker

- How quickly can a hacker guess a correct 20-bit MPLS label?



- 100 pps attack rate
- 100 active VPN labels (routes) on a PE

Answer:

1 minute, 45 seconds

Tunneling Technologies
Can we use IPSec?

- Of course! But you have to pay for it.
- IPsec is a very heavyweight solution, it requires p2p IKE key exchange, crypto acceleration hardware, etc.
- A number of MPLS over IPsec proposals were made in the IETF, in the end MPLS over IPsec is really MPLS over IP, GRE, or L2TPv3 used with IPsec in Transport Mode – IPsec is not “tunneling” it is just providing security for another type of tunnel
- IPsec can always be “bolted on” in places it is needed, particularly with the ability to advertise tunnel capabilities between PEs

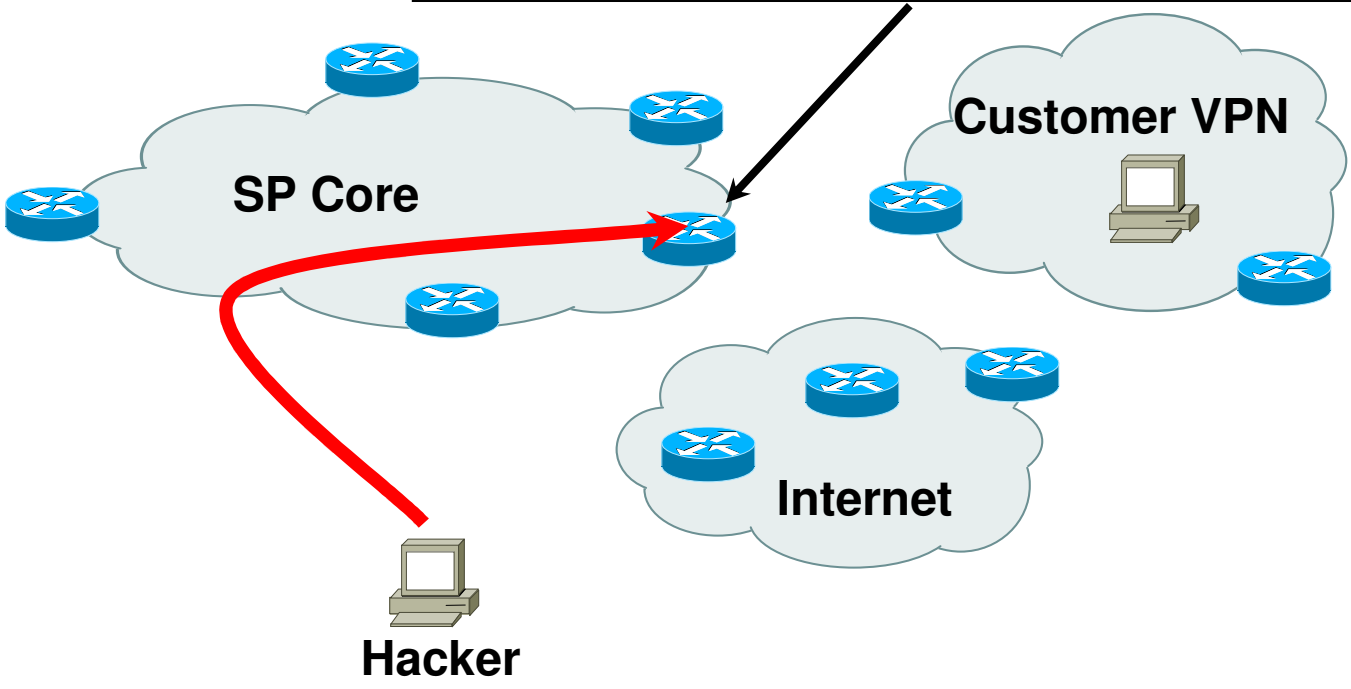
MPLS VPN over GRE Network Security

- Bottom Line: In order to avoid becoming a transit point for packets inserted into a customer VPN, IP ACLs alone are not a robust solution.
- IPsec may be used with any MPLS over IP tunnel type, but is expensive to both opex and capex
- Still Needed: An additional layer of protection to make spoofing far more difficult than it is today with GRE, but without the overhead of IPsec

VPN Services over IP Tunnels

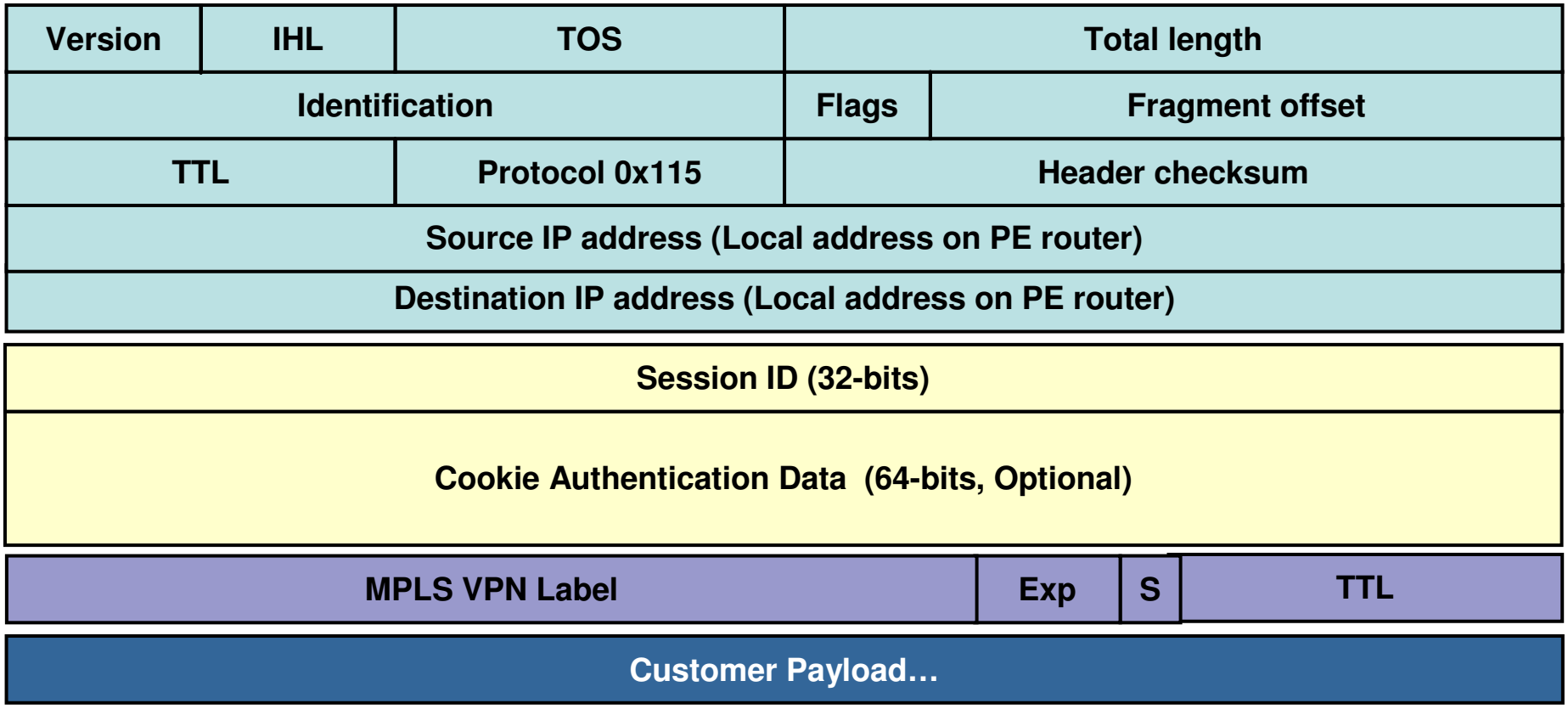
Where to apply additional layer of security

L2TPv3 provides a simple and efficient method to make simple packet spoofing attacks impossible. Protection occurs at the most important point, right before entering the Customer VPN



SP SA / SP DA	L2TPv3	MPLS VPN Label	Hacker SA / Customer DA	Attack Data

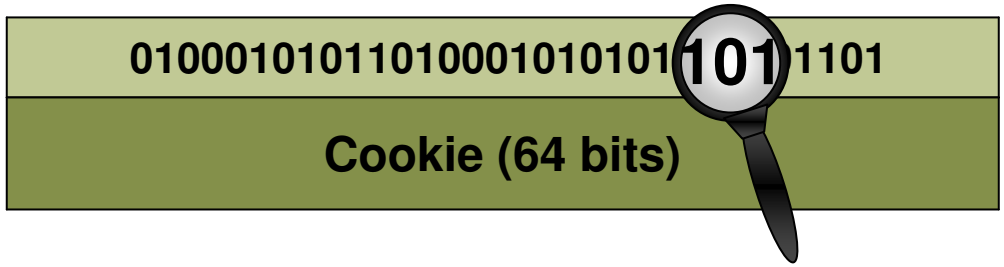
Tunneling Technologies
MPLS over L2TPv3 w/BGP Tunnel SAFI



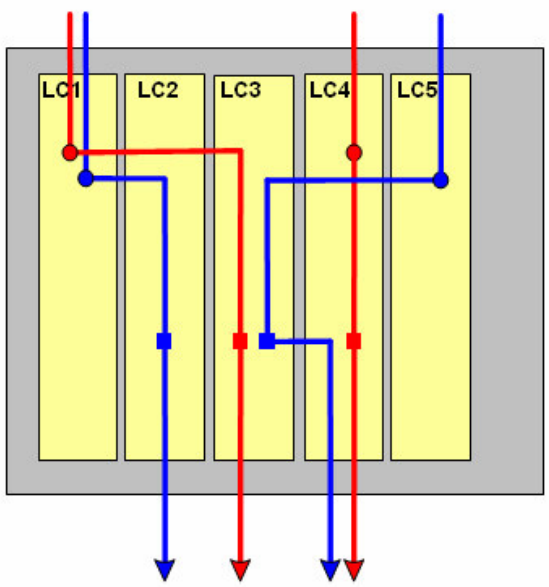
- draft-ietf-mpls-over-l2tpv3-00.txt & RFC3931
- Large scale deployments already exist today

MPLS over L2TPv3

L2TPv3 Distributed Session Processing

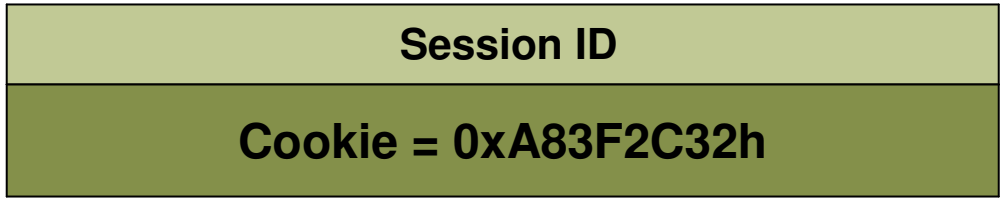


- On a distributed system the context for multiple services or multiple service instances can be balanced across resources
- Structure imposed on the the Session-ID bits can quickly vector the L2TPv3 packet to the resource servicing that context
- Processing of tunneled payload is based on the associated context – switch to an interface, route in a VRF, Bridge in a VSI…



MPLS over L2TPv3

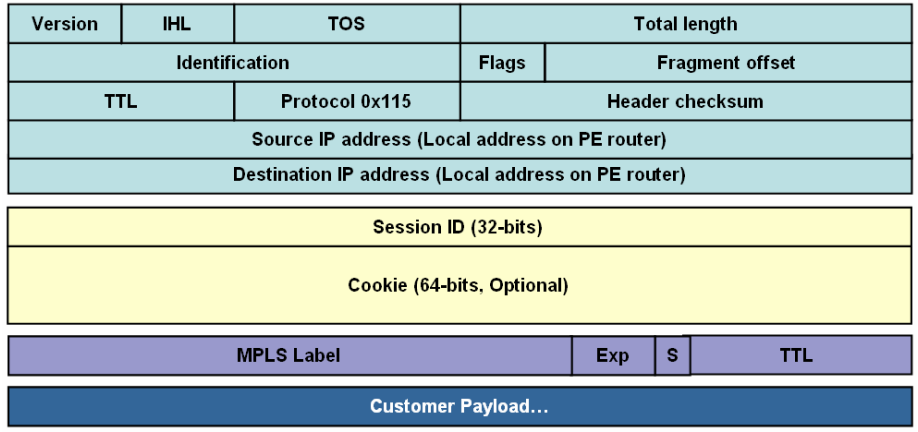
L2TPv3 Packet Authentication Check w/Cookie



- 64-bit value must match for each packet
- Not a 64-bit lookup! Just a very fast compare based on the Session ID lookup
- No encryption hardware needed
- Rather than checking an IP SA or DA, L2TPv3 “seeds” each packet with an unguessable value selected at random by each PE, and advertised to other PEs in the VPN via the BGP Tunnel SAFI
- Somewhat like an ACL, but simple to manage and virtually impossible for a hacker to guess

Spoofing VPNs over L2TPv3 Tunnels

- We assume that the L2TPv3 Session-ID may be known, as it could be predictable or even hard-coded to a constant for some services in order to optimize forwarding



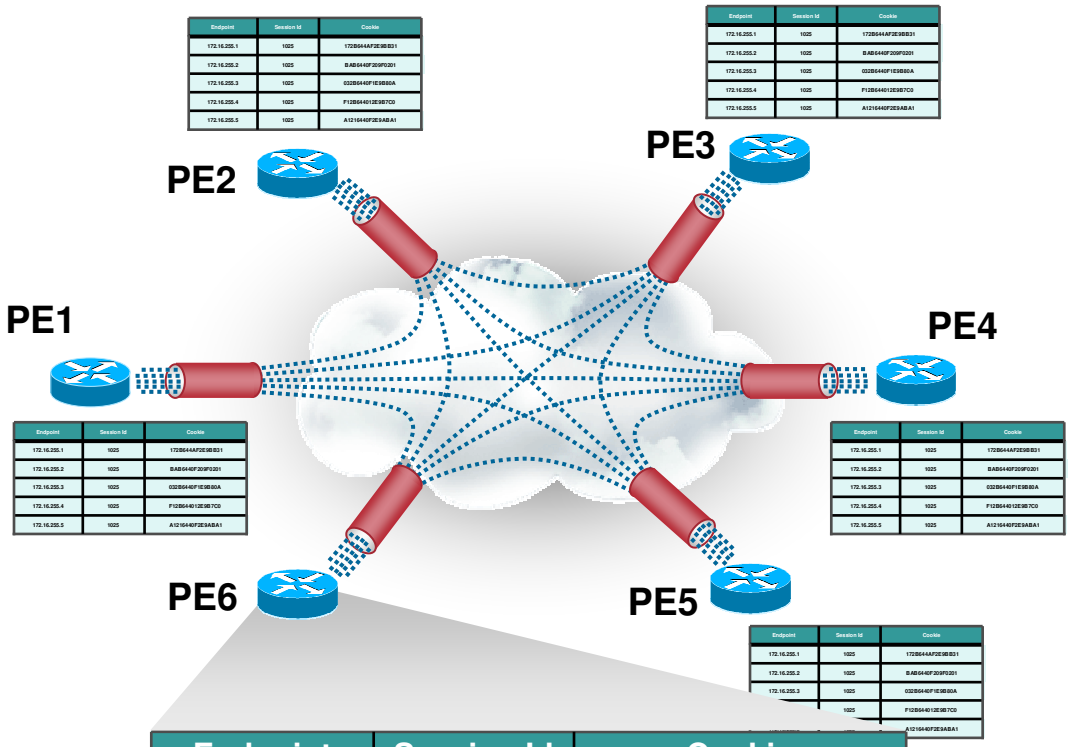
- How quickly can a hacker guess a correct 64-bit L2TPv3 cookie?

  - 10 Mpps attack rate
  - ANY VPN labels is considered valid

→ **Answer: 60 000 Years!**

Tunneling Technologies
L2TPv3 w/BGP Tunnel SAFI

- One L2TPv3 Multipoint session is dynamically created on each PE for receiving traffic from other PE's (point to point L2TPv3 signaling is not used)
- BGP advertises tunnel capabilities via Tunnel SAFI - MPLS over L2TPv3 traffic only sent to PEs which know how to handle it
- Tunnel SAFI also includes per-PE Session ID and Cookie pair

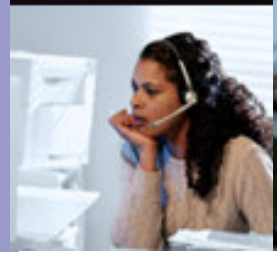


Endpoint	Session Id	Cookie
172.16.255.1	1025	172B644AF2E9BB31
172.16.255.2	1025	172B644AF2E9BB31
172.16.255.3	1025	172B644AF2E9BB31
172.16.255.4	1025	F12B644012E9B7C0
172.16.255.5	1025	A1216440F2E9ABA1

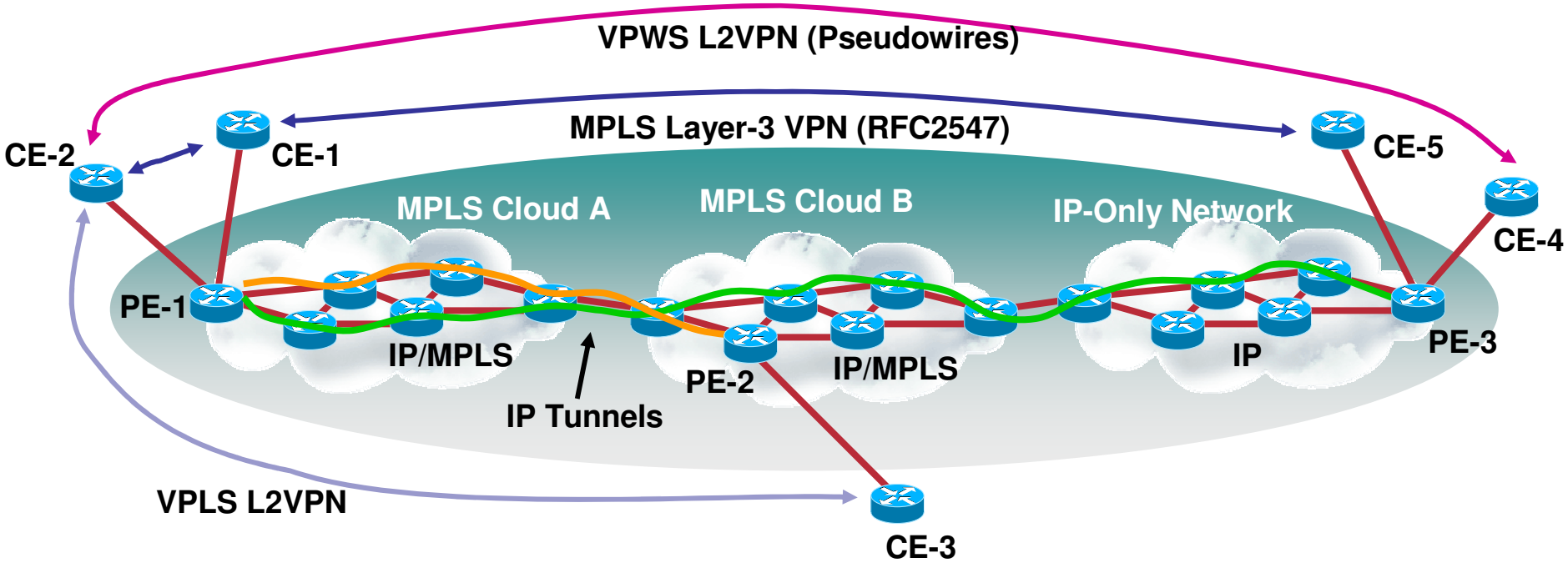
VPN Services over IP Tunnels
Review of capabilities

	Static IP	Static GRE overlay	Dynamic Multi-point GRE	L2TPv3 w/SAFI
Encapsulates MPLS over IP	Yes	Yes	Yes	Yes
Tested in a large active deployment	?	Yes	?	Yes
Avoids full mesh via scalable, dynamic, p2mp tunnels	No	No	Yes	Yes
Avoids blackholes by advertising tunnel capabilities	No	No	No	Yes
Encapsulation facilitates highspeed lookup and distributed processing assist	No	No	No	Yes
Simple, scalable, anti-spoofing protection built-in	No	No	No	Yes

MPLS over IP Tunneling Solutions

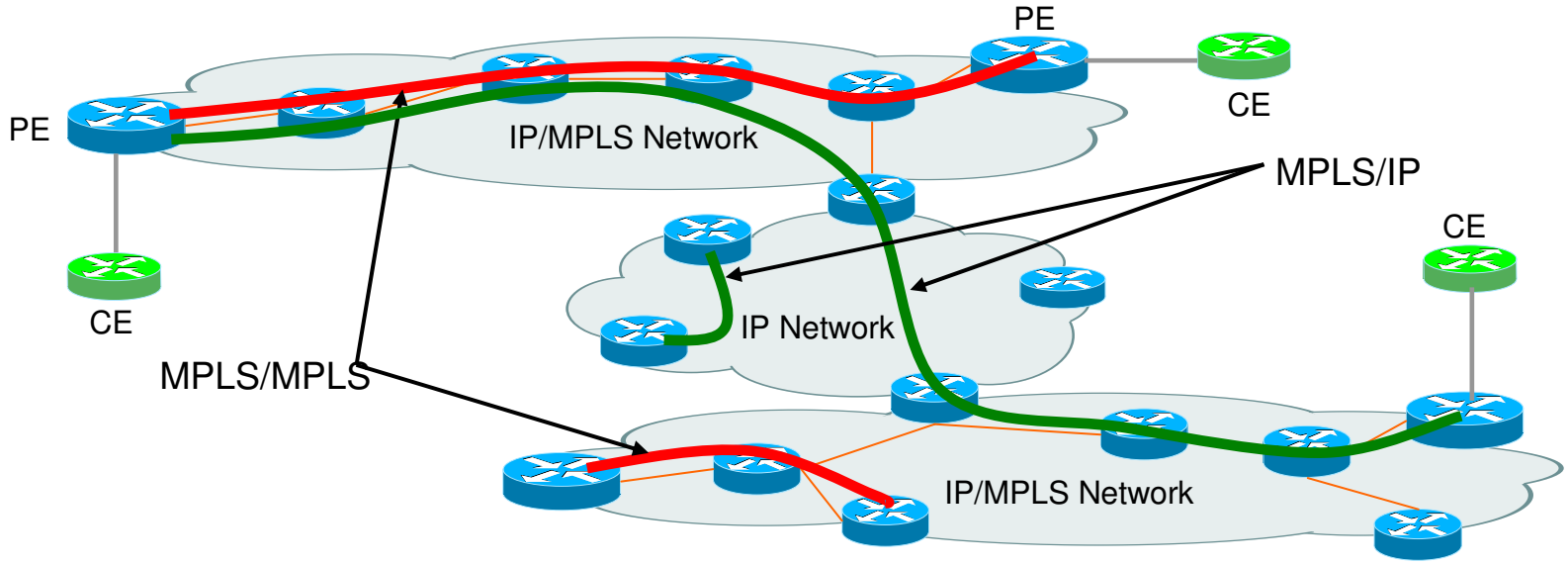


Extending the Reach of MPLS
MPLS over IP Tunnels



- **Multiple MPLS or IP networks, seamless global MPLS service presented to customers**

Step-by-Step Migration to MPLS

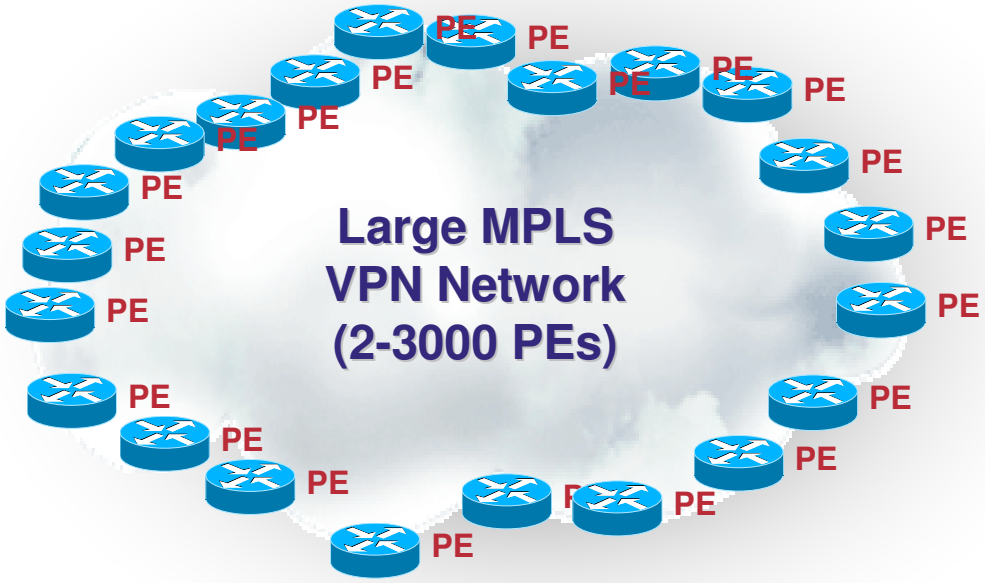


- **Native MPLS/MPLS is used when possible, MPLS/L2TPv3, MPLS/GRE or MPLS over IPsec where necessary, etc.**
- **Requires BGP Tunnel SAFI to advertise PE capabilities**

Operational Flexibility

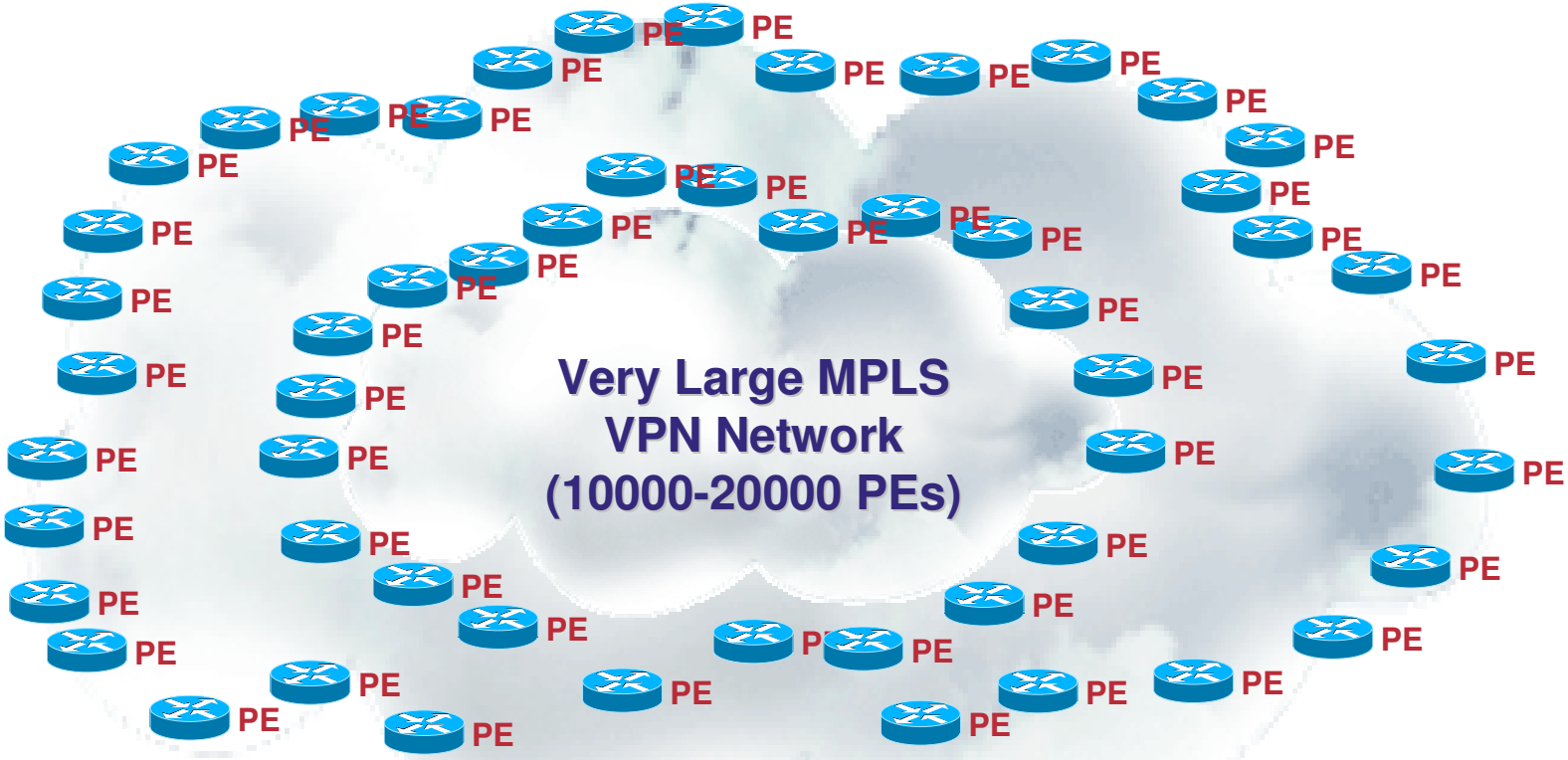
- There are many benefits to an MPLS core network, including Traffic Engineering, Fast Re-route, etc.
- However, IP networks without MPLS end-to-end can still be engineered well enough to deploy “Edge” MPLS-based services such as L2VPN and L3VPN
- Deploying Edge MPLS services may be decoupled from deploying MPLS “Core” features, allowing separate operational teams to migrate at their own pace

Scaling MPLS VPNS
Native MPLS VPNs (w/o IP Tunnels)



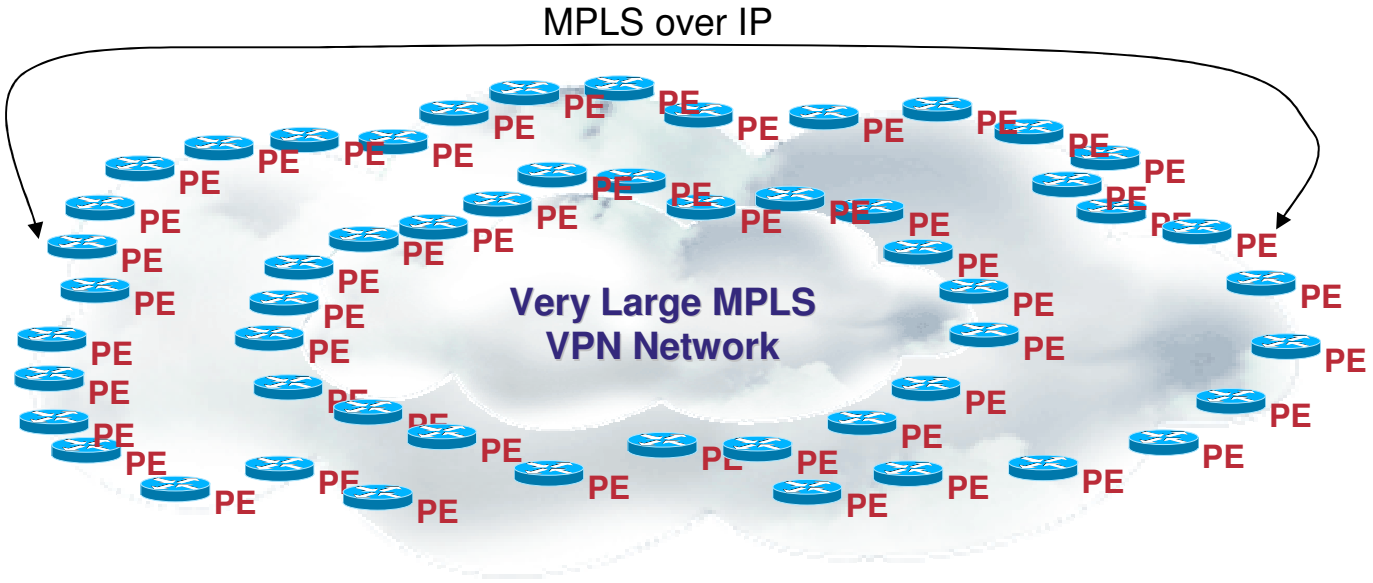
- Each PE must signal its own Tunnel LSP (i.e., LDP) and carry a /32 route within the IGP.
- To support 3000 PEs in one VPN, IGP must support 3000 /32 PE routes
- There are examples of this size Native MPLS VPN today

Scaling MPLS VPNS
Native MPLS VPNs (w/o IP Tunnels)



- 10000-20000 PEs means 10000-20000 /32 routes in an IGP.
- Seriously stresses existing state of the art in IGP protocols, may test absolute maximum (maximum in OSPF is 56000 routes)
- Inhibitor to pushing MPLS VPNs beyond the core and into access networks

Scaling
MPLS VPNs over IP Tunnels



- PEs are reachable via IP CIDR blocks, so no need for /32 LSPs between all PEs.
- "If you can ping the router, you can use it as a PE in the VPN"
- Not necessary to advertise /32 routes to PEs, so no need to carry 10000-20000 routes in the network IGP.
- Use Core MPLS features where you want, based on where you need FRR, TE, etc.
- Requires dynamic multipoint IP tunnels and built-in anti-spoofing security

MPLS over IP Tunnels

Summary of what you can do with this technology

- **Extending the Reach of MPLS**
  - **MPLS services (such as RFC 2547 VPNs) based on IP Tunnels can cross multiple providers (Inter-provider) or administrative domains (Inter-AS) to reach customers anywhere IP reaches**
- **Migration to MPLS**
  - **MPLS/MPLS where available, MPLS/IP where not**
- **Operational Flexibilty**
  - **Some service providers do not yet have (or do not yet want) MPLS in their core networks, but still want to offer their customers MPLS-based services**
- **Scaling MPLS VPN deployments**
  - **IP route aggregation allows for scaling MPLS VPNs across a very large number of PEs without increasing the number of PE-PE LSPs and associated /32 routes advertised in an IGP.**

MPLS over IP Tunnels
Summary of Available Tunneling Technologies

- Static MPLS over GRE may be used to connect a small number of isolated nodes or disparate MPLS networks, but is not recommended for high scale deployments
- Dynamic Multipoint Tunneling available with GRE or L2TPv3 solves the manual provisioning problem with static GRE tunnels, but still can allow blackholes
- The BGP Tunnel SAFI prevents blackholes to routers which cannot decapsulate a given type of IP tunnel, allowing staged migration to MPLS
- IPsec can provide strong security, but is expensive from an opex and capex perspective.
- L2TPv3 includes lightweight yet strong anti-spoofing protection, with zero additional opex complexity over mGRE, an no reliance on ACLs
- Conclusion: MPLS over L2TPv3 w/BGP Tunnel SAFI is the most feature rich and proven MPLS over IP Tunnel offering among the choices available