MPLS High Availability

IEEE CQR Workshop – London, June 7-9, 2006

Harmen van der Linde

Product Manager – MPLS

Cisco Systems, Inc.

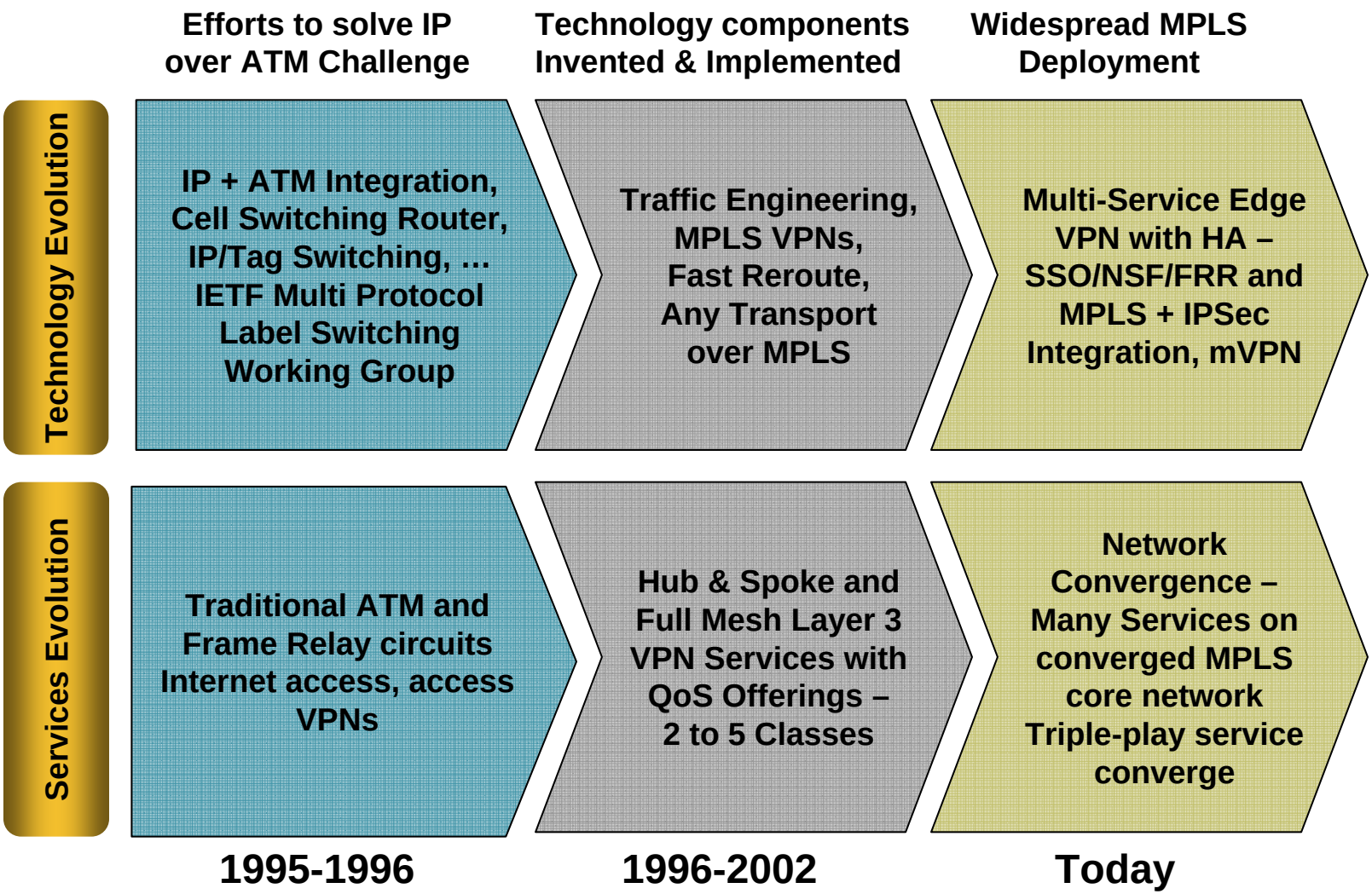
Internet Technologies Division

havander@cisco.com

Topics

- **MPLS overview**
- **High Availability (HA)**
- **HA mechanisms**
- **MPLS HA mechanisms**
- **Conclusions**

MPLS Network Evolution



Multi-Protocol Label Switching (MPLS)

- **MPLS has become key infrastructure technology for Service Provider networks and new emerging broadband networks**
- **MPLS architecture can be decomposed into two functional planes**
  - **Control plane**
  - **Forwarding plane**
- **MPLS control plane**
  - **Distributes labels and establishes label switched paths**
  - **Multiple control protocols; LDP, BGP, and RSVP-TE**
- **MPLS forwarding plane**
  - **Used for MPLS labeled data packet forwarding**
- **MPLS Applications**
  - **Layer-3 VPNs, Layer-2 VPNs, Traffic Engineering (TE)**

Forwarding Plane v.s. Control Plane

- **Traditional routing/switching platforms**
  - **Software-based control and forwarding plane**
- **Next-generation routing/switching platforms**
  - **Separate control and forwarding plane**
    - **Control plane resides on Route Processors (RPs)**
    - **Forwarding plane resides on Line Cards (LCs)**
- **Certain failures confined to control plane**
  - **Hardware failure on active RP**
  - **Software failure on active RP**

High Availability (HA)

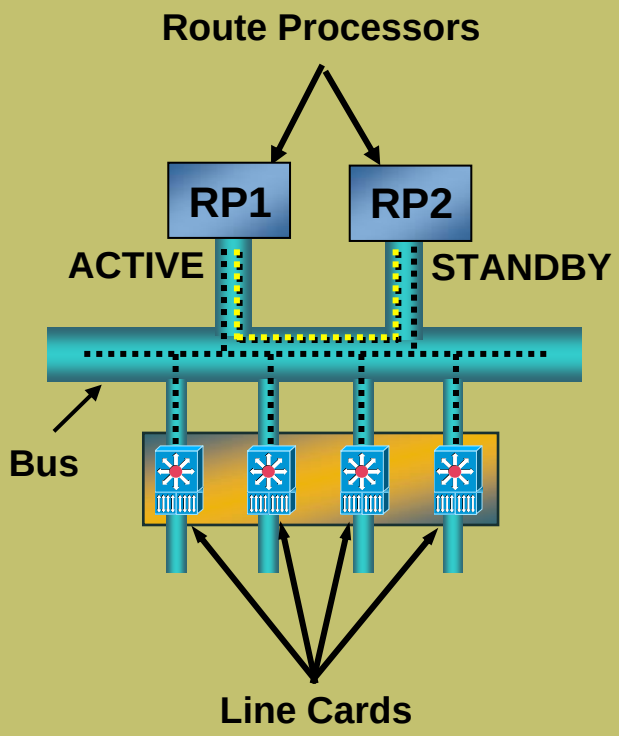
- **End-to-end protection against network failures**
- **High Availability focused on maximizing network uptime and minimize effects of planned and unplanned network outages**
  - **Preserve end-to-end network service connectivity**
- **For network high availability, disruption in forwarding plane must be kept to an absolute minimum**
  - **Isolate control plane failures from forwarding plane**
- **Separation of control and forwarding plane should allow forwarding to continue while control plane recovers (NSF)**

HA Mechanisms

- **Component and Device Level Resiliency**
  - **Hardware and software component resiliency**
    - **Distributed line cards, route processors, Modular operating software**
  - **Stateful Switch-Over between RPs (SSO)**
  - **Control/forwarding plane decoupling; Non-Stop Forwarding (NSF)**
- **Network Level Resiliency**
  - **Optimized convergence algorithms, speeding up network recovery**
  - **Intelligent protocol fabric with network-wide forwarding awareness**
- **Operations and management**
  - **Embedded event management for proactive maintenance**
  - **Embedded, lightweight measurements of availability metrics**
  - **In-service software upgrades (ISSU)**

NSF With SSO

- **Non-Stop Forwarding (NSF): minimal or no packet loss**
  - **Packet forwarding continues during reestablishment of peering relationships**
  - **No route flaps between participating neighbor routers**
- **Stateful Switch-Over (SSO): zero interruption to protocol sessions**
  - **Active RP synchronizes information with standby RP**
  - **Session state maintained for high availability-aware protocols on standby RP**
  - **Standby RP takes control when active RP is compromised**



MPLS HA – Component Resiliency

- **MPLS High Availability features extend NSF with SSO capabilities for:**
  - **MPLS Forwarding**
  - **Label distribution protocol (LDP)**
  - **Layer-3 Virtual Private Networks (MP-iBGP)**
  - **Traffic Engineering (TE)**
  - **AToM and L2VPNs**
- **Minimal disruption to MPLS forwarding plane due to route processor control plane failures**
  - **Includes MPLS control plane failures (LDP, BGP, RSVP)**

MPLS HA – Network Resiliency

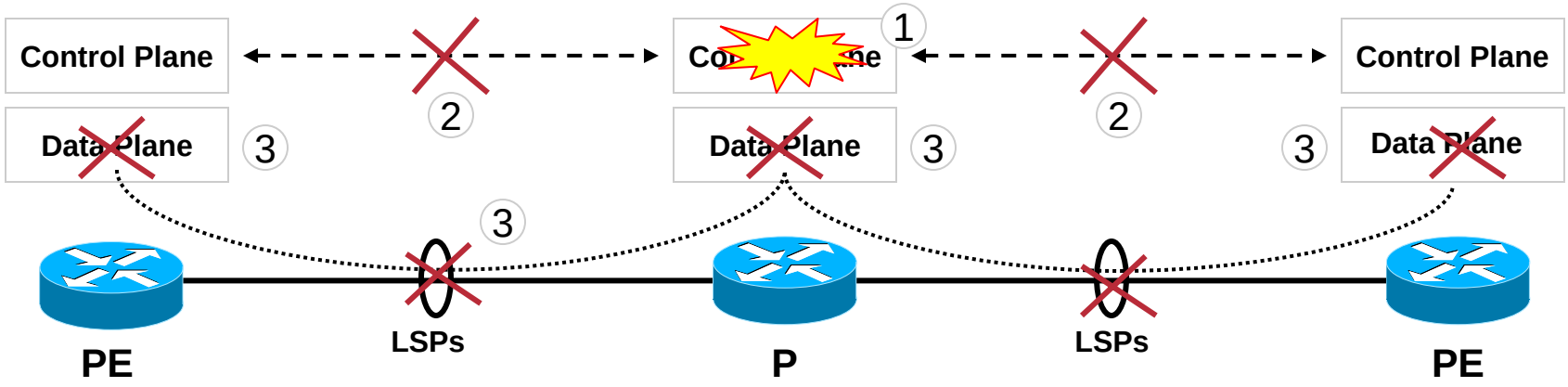
- **MPLS control plane protocol enhancements to improve failure detection time and network convergence**
- **Graceful Restart (GR)**
  - **LDP, MP-BGP**
- **Fast Convergence**
  - **LDP (IGP sync), MP-BGP**
- **TE FRR**
  - **Link protection**
  - **Node protection**

MPLS Graceful Restart (GR)

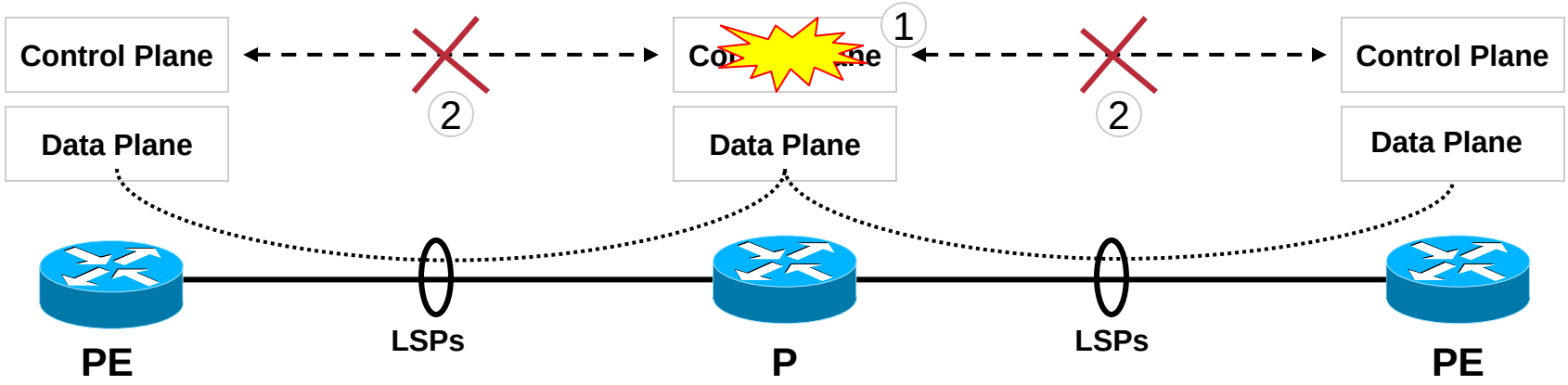
- LDP and BGP use TCP as a reliable transport mechanism for exchange of protocol messages
- TCP session between LDP/BGP peers may go down due to a RP switchover or HW/SW failures
- On detection of TCP session failure (if HW/SW failure), existing LDP and BGP control plane components would remove their forwarding state
- Graceful Restart mechanism enables continuous MPLS traffic forwarding during MPLS Control Plane failure and recovery
  - Temporary use of old MPLS forwarding information until refresh of forwarding entries

MPLS Graceful Restart + NSF/SSO

No MPLS HA support



MPLS HA Support: NSF/SSO + Graceful Restart

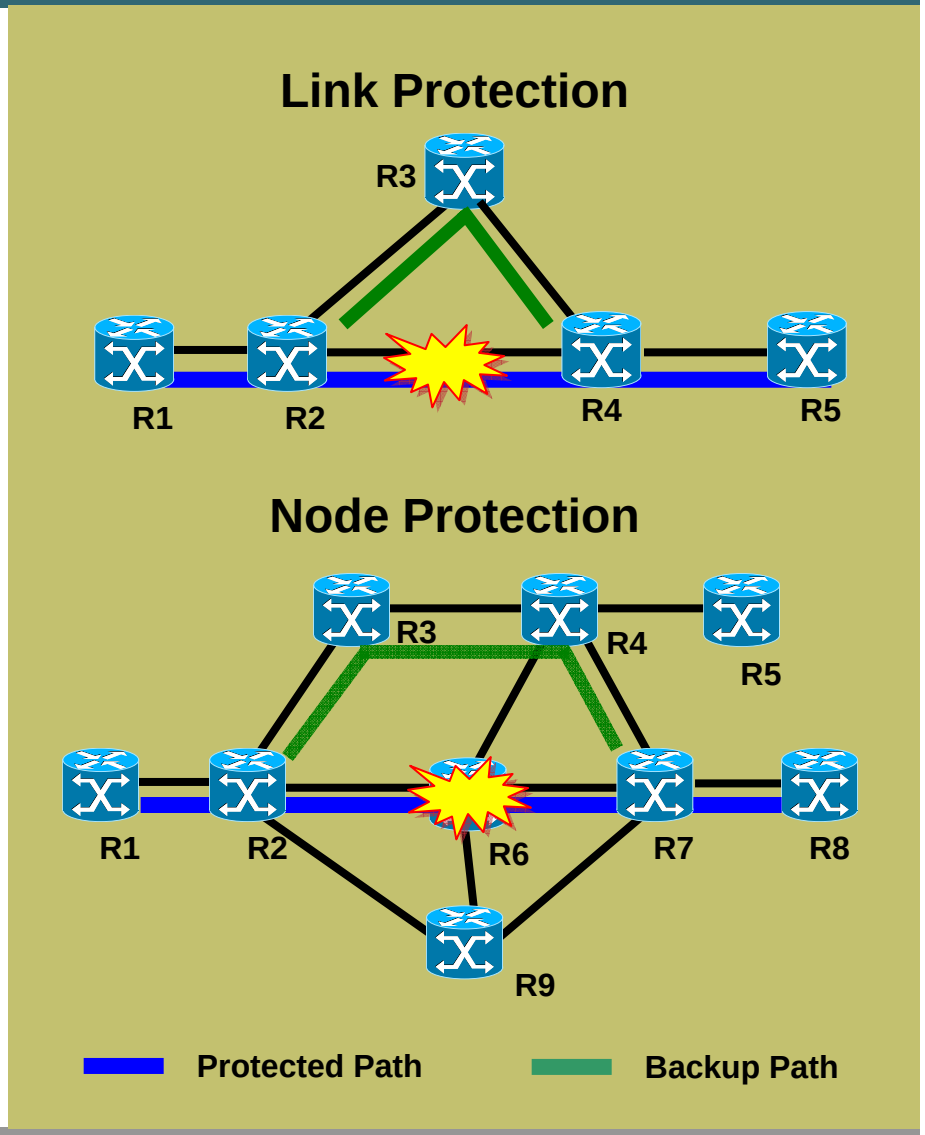


Fast Convergence: LDP-IGP Sync

- **LDP maintains MPLS label database (LIB), which is linked to the routing database maintained by IGP (e.g., OSPF)**
  - **LIB content used by LDP to write MPLS forwarding (LFIB)**
- **LDP and IGP operation loosely coupled**
  - **Separate independent synchronization mechanisms for exchange of IGP routes and LDP label information (between IGP and LDP peers, resp.)**
- **Mismatch between IGP routing database and LDP label/forwarding database can potentially lead to MPLS packet loss**
  - **E.g.; IGP on link re-converges before LDP completes label exchange and LFIB updates, resulting into missing label forwarding (LFIB) entries**
- **LDP-IGP synchronization aimed to minimize potential MPLS packet loss as a result of mismatch between IGP and LDP**
  - **IGP instructed to delay (via hold timer) bringing up IGP adjacency on primary link**
  - **IGP instructed to advertise max-metric for link (resulting in possible upstream re-routing)**

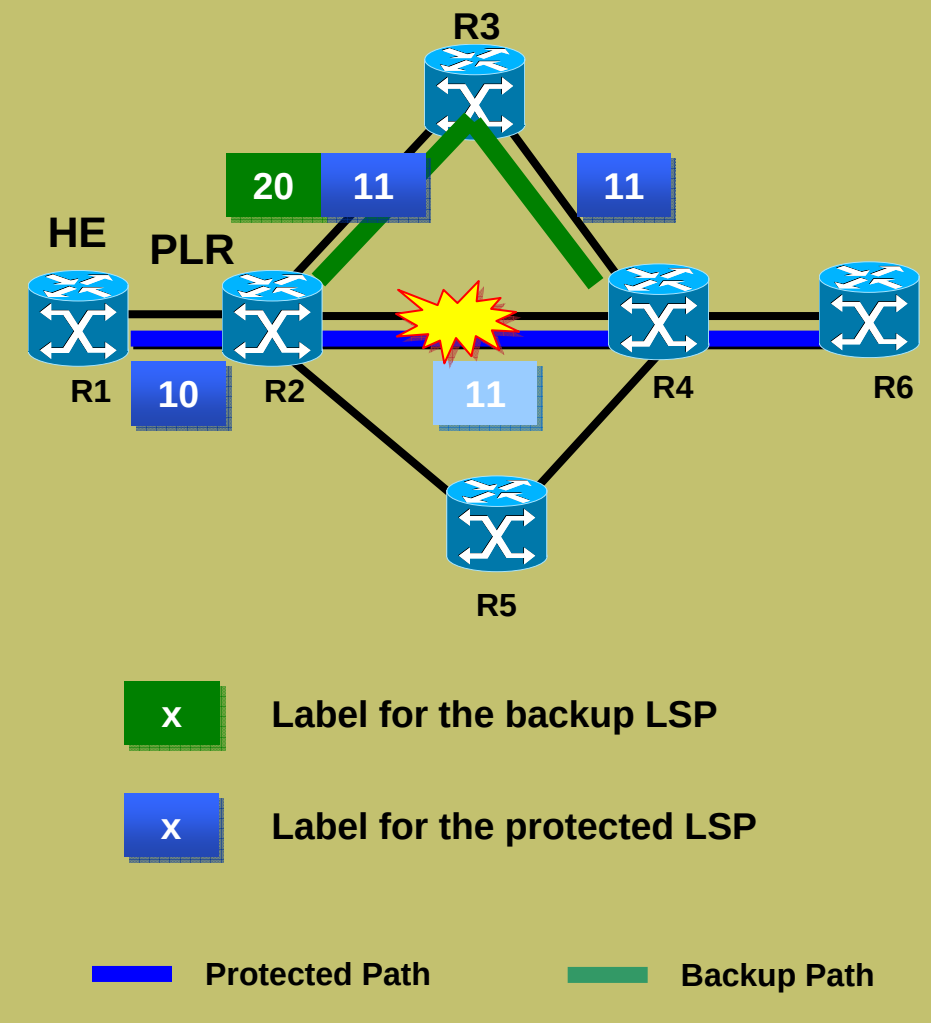
MPLS TE Fast Re-Route (FRR)

- IP routing protocols (e.g., OSPF, BGP) may be tuned to convergence within a few seconds
- Some traffic (e.g. voice) will require more aggressive convergence time
  - Typically 50 ms or less
- MPLS TE FRR offers protection against network failures
  - Link protection
  - Node protection



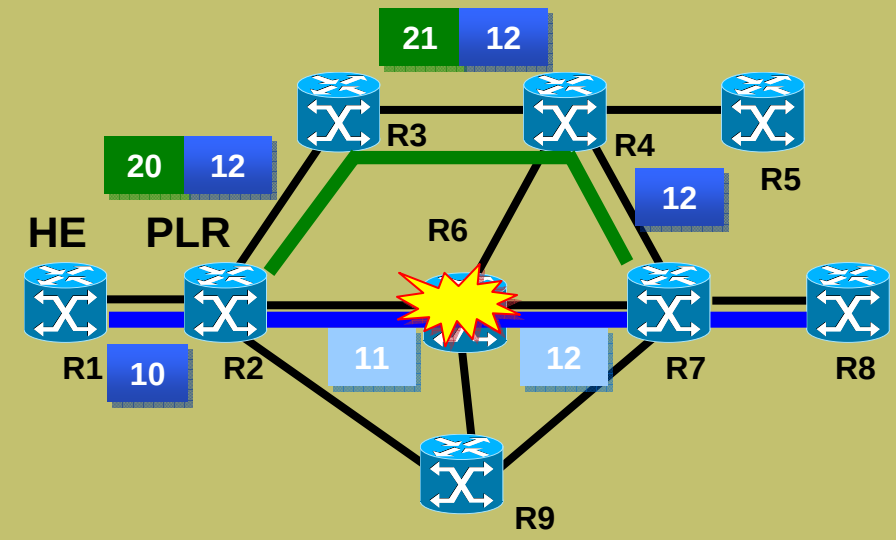
FRR Link Protection

- **Creation of Next-hop Backup Tunnel**
  - **Parallel path around protected link to next hop**
- **On link failure detection Point of Local Repair (PLR) swaps label and pushes backup label**
  - **Traffic sent over backup path**
- **PLR notifies TE Head End (HE), which triggers global TE path re-optimization**



FRR Node Protection

- Creation of Next-next-hop Backup Tunnel
  - Parallel path around protected node to next-next-hop
- Node failure triggers Point of Local Repair (PLR) to swap label and push backup label
  - Traffic sent over backup path around failed node
- PLR notifies TE Head End (HE), which triggers global TE path re-optimization



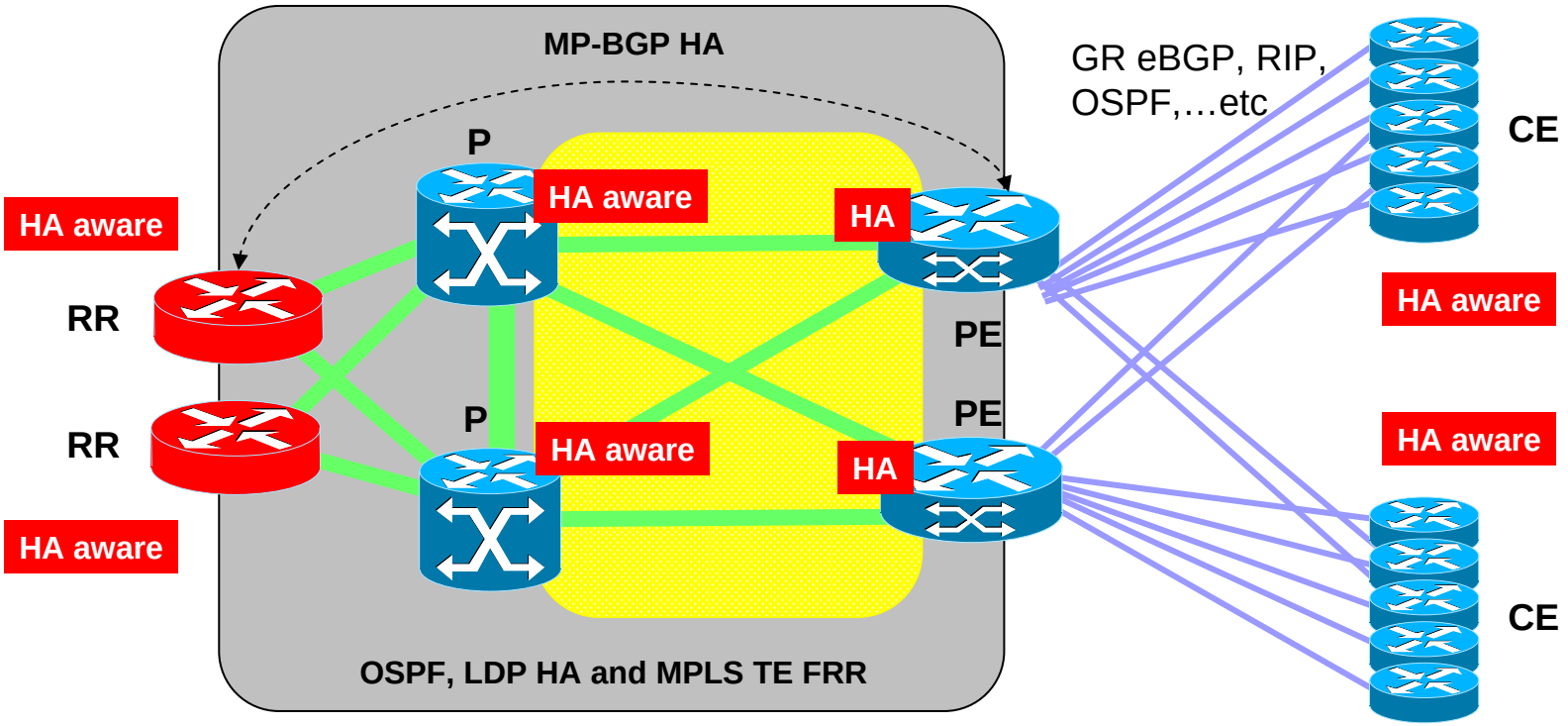
x Label for the backup LSP

x Label for the protected LSP

Protected Path Backup Path

HA in MPLS Network Infrastructure

HA Aware Devices	HA Capable Devices
Graceful Restart Functionality for all related control protocols on CE, P (core), and Route Reflectors (RRs)	Full NSF/SSO functionality on PE nodes (PE1 and PE2)



Conclusions

- **High Availability focused on maximizing network uptime and minimize effects of planned and unplanned network outages**
- **MPLS HA functional areas include Component and Device Level Resiliency, Network Level Resiliency, and Operations and management**
- **MPLS HA focused on separation of control and forwarding plane to allow forwarding to continue during control plane failures and recovery**
- **Network and layer-2/3 service convergence are driving need for High Availability in Service Provider MPLS networks**

Acronyms

ATM	Asynchronous Transfer Mode
AToM	Any Transport over MPLS
BGP	Border Gateway Protocol
FRR	Fast Re-Route
GR	Graceful Restart
HA	High Availability
HE	Head End (for TE path)
HW	Hardware
IGP	Interior Gateway Protocol
IP	Internet Protocol
ISSU	In-Service Software Upgrade
LC	Line Card
LDP	Label Distribution Protocol
LSP	Label Switched Path
LIB	Label Information Base
LFIB	Label Forwarding Information Base
MPLS	Multi-Protocol Label Switching
MP-iBGP	Multi-Protocol Interior BGP
NSF	Non-Stop Forwarding
PLR	Point of Local Repair
RP	Route Processor
RSVP	Resource Reservation Protocol
SW	Software
SSO	Stateful Switch-Over
TCP	Transmission Control Protocol
TE	Traffic Engineering
VPN	Virtual Private Network