Scope Considerations for Link Layer in Resilient Packet Rings

Pankaj K Jha
Cypress Semiconductor
pkj@cypress.com

RPR Scopes & Issues

- Need to develop a forward-looking RPR standard which can scale to meet carrier & subscriber requirements for many years
- Focus on Factors that help RPR Acceptance in other Optical Networking Development Groups
- Control and Data Plane Separation to provide a Unified Data Transport
- Ring and Mesh Networks

Scopes (contd.)

- Multiple RPRs connected by common node(s) and point-to-point links
- Propose no L3 changes, instead provide L2 hooks
- Leverage MPLS as Foundation for Traffic Engineering & QoS operations - avoid re-inventions
- Different rates at different spans in an RPR Network
- Considerations for a Packet Delineation Mechanism

Top-down Approach

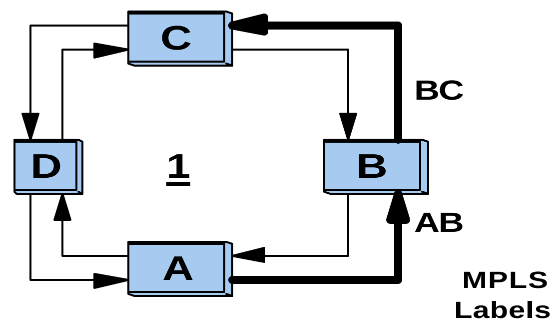
- Go top-down - important to first focus on scopes & requirements
- Then come up with a technology to achieve goals
- Suggest no bottoms-up approach - to force proprietary technologies to make them RPR
- A bottoms-up approach will limit our scope and functionality. High risk for RPR acceptance.

Traffic Engineering & QoS

- Traffic Engineering, Load balancing, and QoS - leverage existing and continuing work in MPLS
- Packets are classified into different Flows at the Ingress Node - using MPLS, for instance
- <u>No Need for Buffering Requirements in RPR Specification</u> (creating two-priority buffers, for example, are way insufficient)
- Buffering and 'Arithmetic' Traffic Balancing at Intermediate Nodes Balances all Types of Traffic - not desirable
- Let Nodes participate in Bandwidth Allocation - and Leave Buffer & Priority Management to the Nodes (Nodes may have hundreds of queues & buffers to direct flows - it is a local decision). No other Protocols mandate buffering.

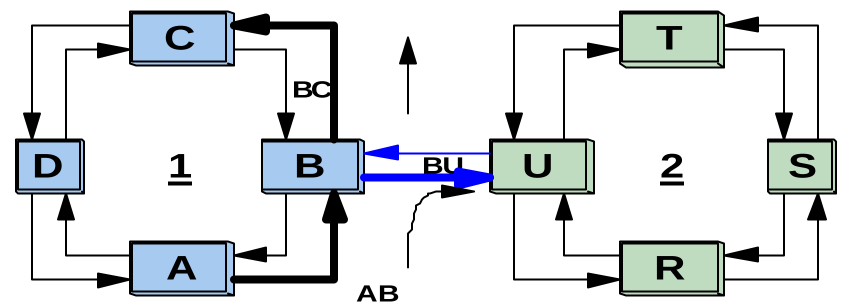
NBMA Nature of Optical Networks

SINGLE OPTICAL NETWORK



All Packets on A-B Link go to B

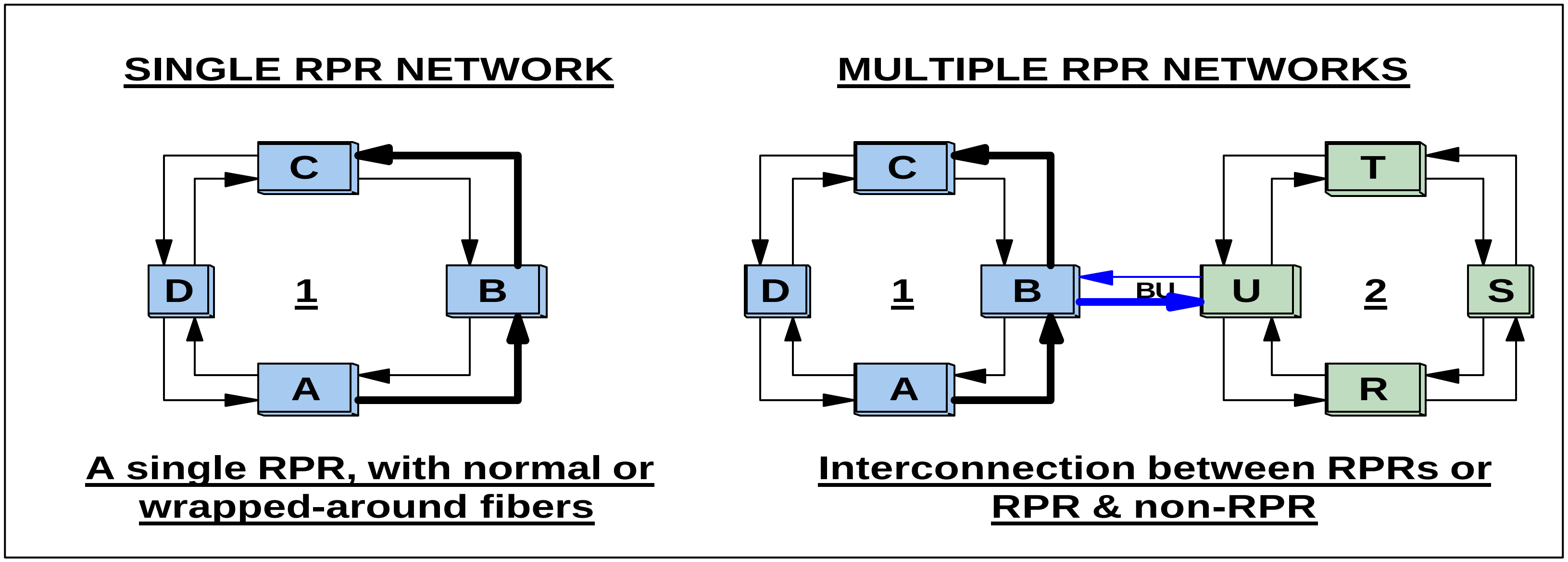
MULTIPLE OPTICAL NETWORKS



All Packets on A-B Link go to B.
B sends a packet to either U or C

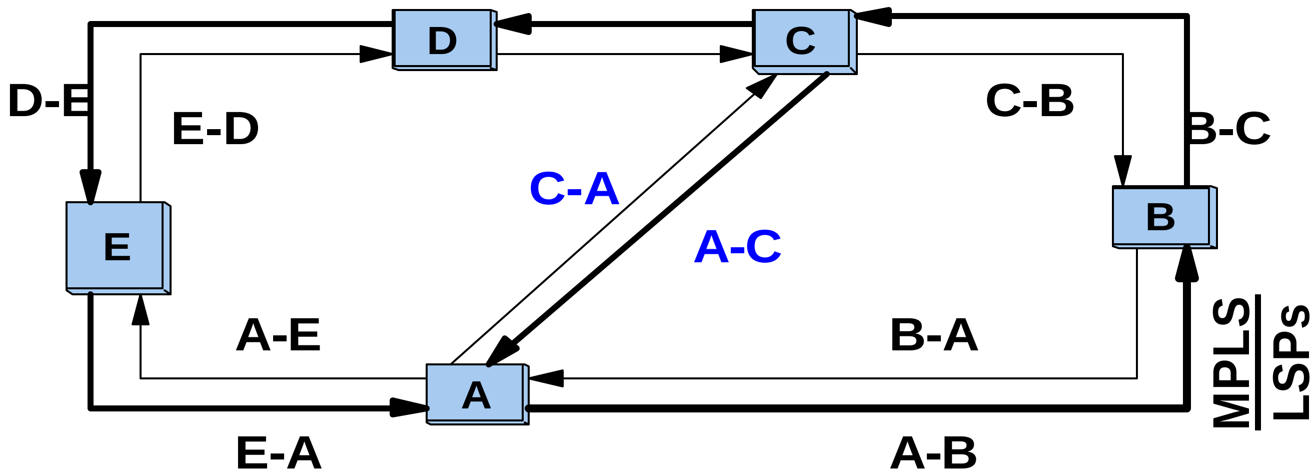
- **Every packet passes through one and only one node at a time at the input.**
- **Since packets are Rx then Tx, a node ('B') is always a node in A-B-C, A-B-U-T Paths**

RPR MAC Address Resolution



- Since Packets go to nodes down the ring one node at a time, RPR still behaves in an NBMA mode, never actually as a 'Broadcast LAN'
- 'A' cannot resolve 'C's MAC addressing through ARP (unlike LAN)
- A single RPR may be shared by multiple router nodes, each belonging to a different IP domain
- Hence we must develop method(s) to resolve Remote Destination MAC Address (e.g. NHRP)

RPR Links as (a set of) LSP(s)

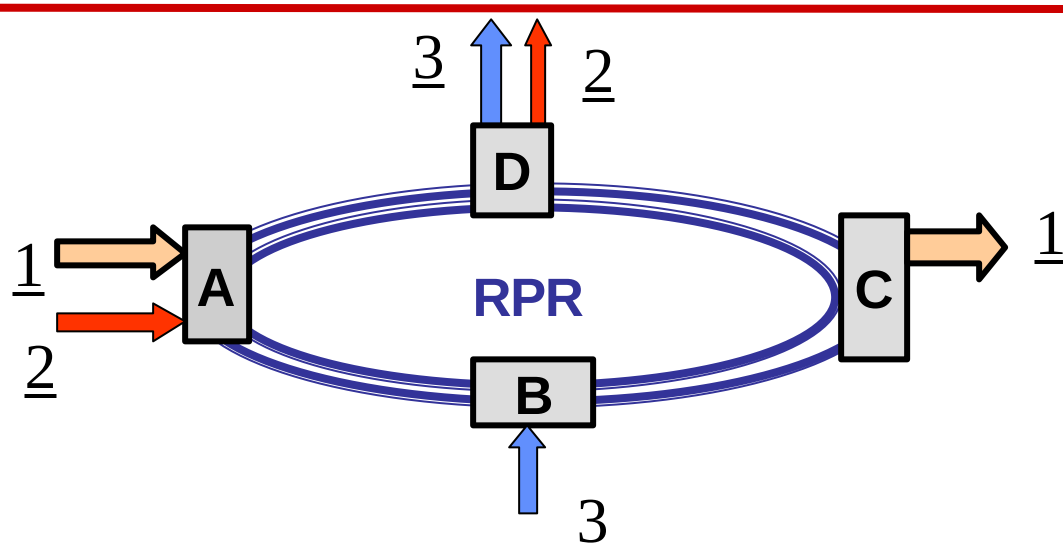


- An RPR Ring can be treated as one (or a collection of) LSPs.
- Any number of rings (not just 2) can be addressed using LSPs
- Mesh networks become just another type of LSP
- MPLS based fault protection and fault recovery MAY be used for RPR. Protection would easily work for 1:1 and 1:N, and any ring-mesh combination.
- Different traffic flows can easily be given different levels of protection

Control & Data Planes for RPR

- Control plane provides all of RPR features: fault-recovery, traffic engineering, QoS, load balancing, etc.
- Control and Data Planes for RPR should be de-linked to simplify designs for Resilience Packet Rings.
- Use MAC addresses for Control Plane
- A unified control plane will allow traffic pooling for different data types, and minimize links and keep alive messages
- Allow Native Packets to travel on RPR. Multiservice transport becomes easy in RPR.
- Native packets from non-RPR networks can pass through RPR network.

Control Plane in RPR

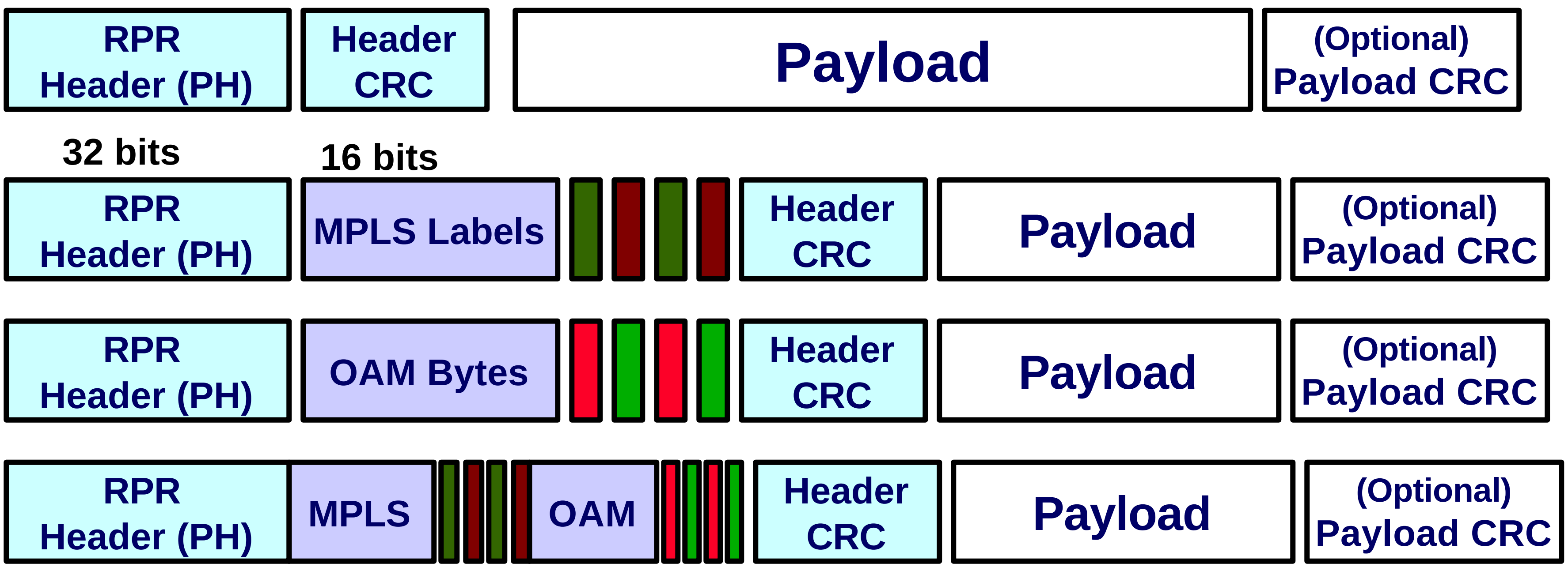


- Control plane may be used to set up MPLS path for a flow.
- Control Plane Packets may include TTL, etc. relevant for packet transport over rings
- Following header, normal Ethernet packet can be used
- MPLS LSPs are established for a flow.
- Link establishment, and fault detection

RPR Header

RPR Core Header is 32 bits wide, with a 16 bits Header CRC

TTL	Payload ID	Flags for MPLS, OAM, etc.



RPR Control Packet

Payload Header (PH)	Header CRC	Ethernet Frame with Special Payloads for RPR	(Optional) Payload CRC
32 bits	16 bits	IEEE802.3	

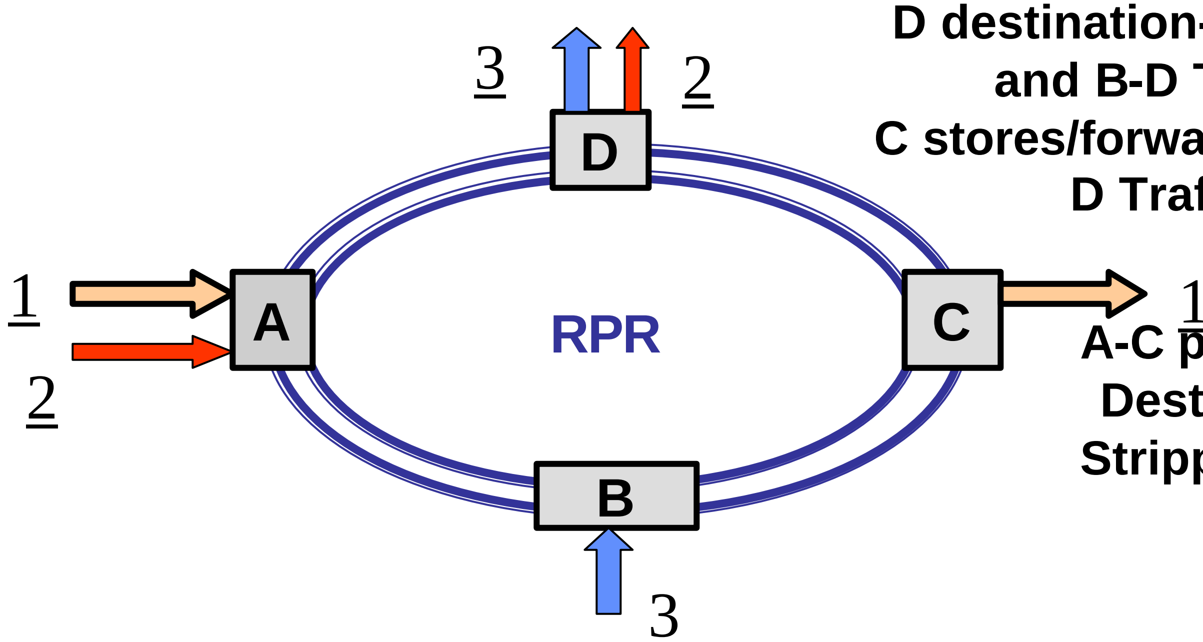
RPR control packets handle:

- **Topology Discovery**
- **Fault Discovery, Isolation, Recovery, and Restoration**

Data Plane: Addressing Methods

- Once Source and Destination Nodes are identified, traffic flows can occur either using MAC address or MPLS labels
- MPLS Labels can be used as Data Link Layer Address on RPR for Data Packets
- Use of MAC addresses requires high-speed gigabit CAM at all nodes
- Since MPLS Labels have Local Significance only, simple Logic at Nodes to look up labels and classify packets into different queues and schedulers. <u>No 48-bit lookups needed for high-speed data plane.</u>
- Fine-grained traffic engineering and QoS without any extra protocol changes.

Data Plane: MAC-addressable Nodes

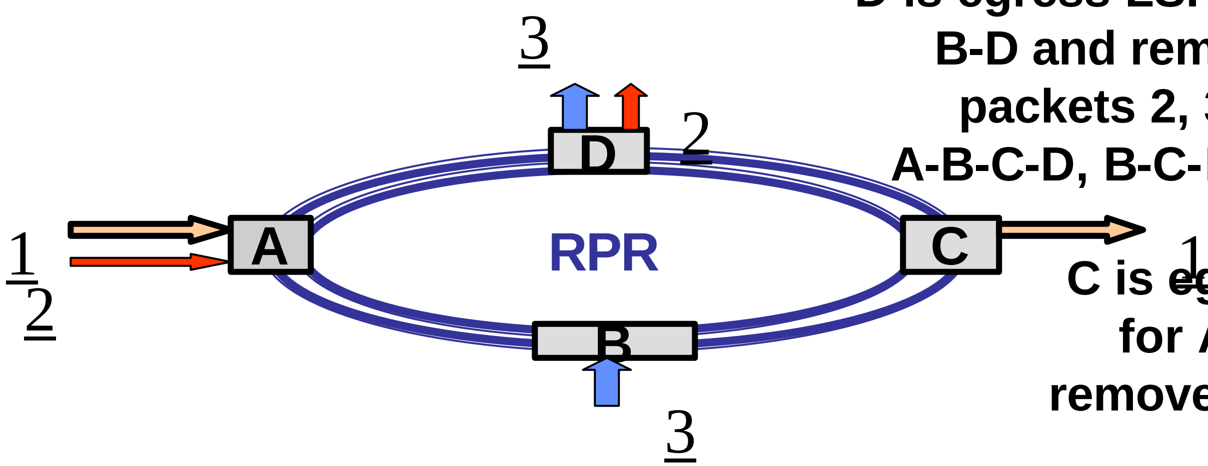


D destination-strips A-D and B-D Traffic
C stores/forwards A-D, B-D Traffic

A-C packet is Destination Stripped at C

- For MAC addressing, { Dest MAC, Src MAC } used for Packet Transfer
- Mega/Gigabit CAM logic required at <u>every</u> intermediate node, since each node must sift through hundreds of passing destination MAC values to determine stripping

Data Plane: RPR as MPLS LSP(s)

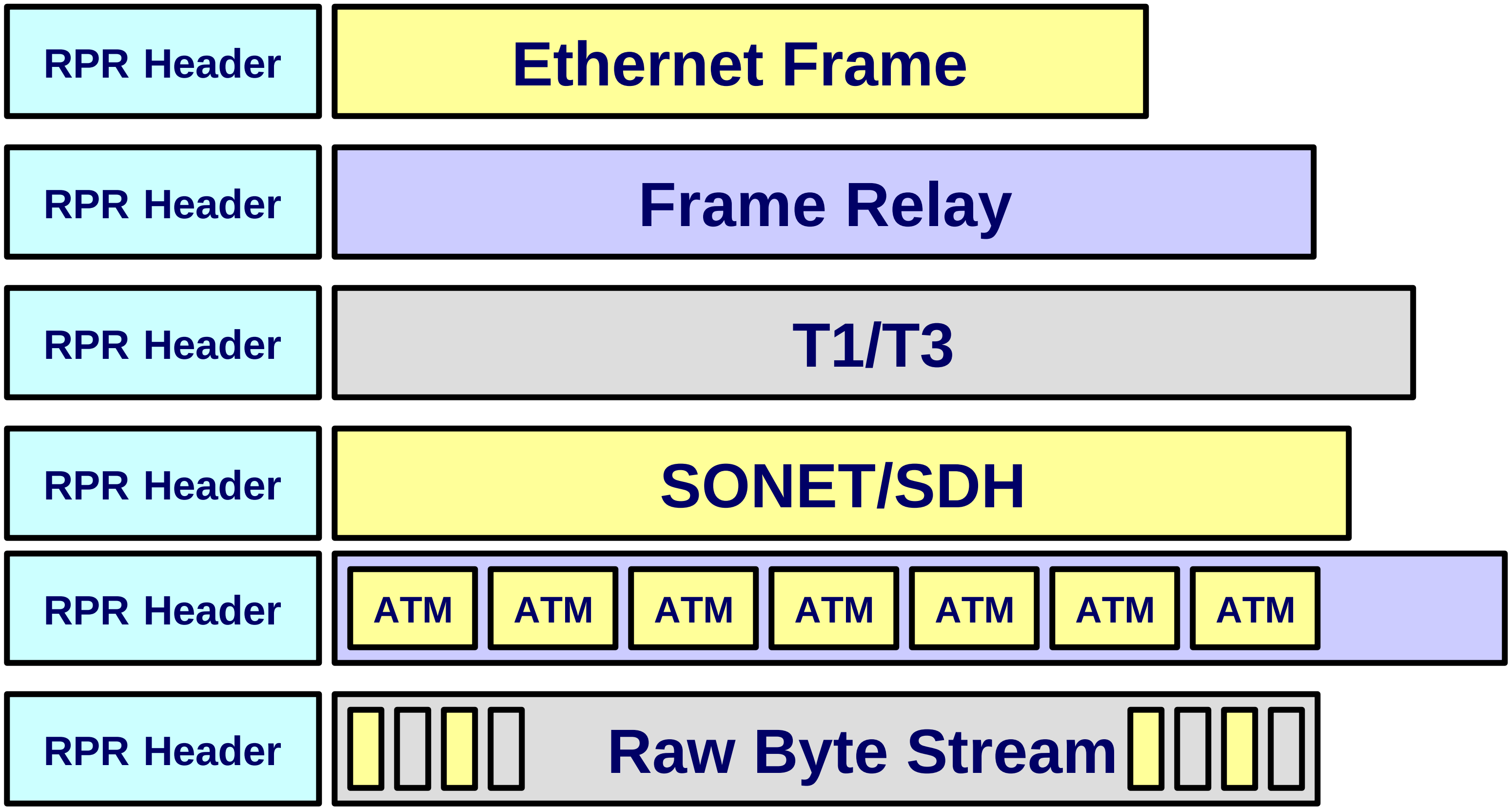


D is egress LSR for A-D, B-D and removes packets 2, 3 on A-B-C-D, B-C-D LSPs

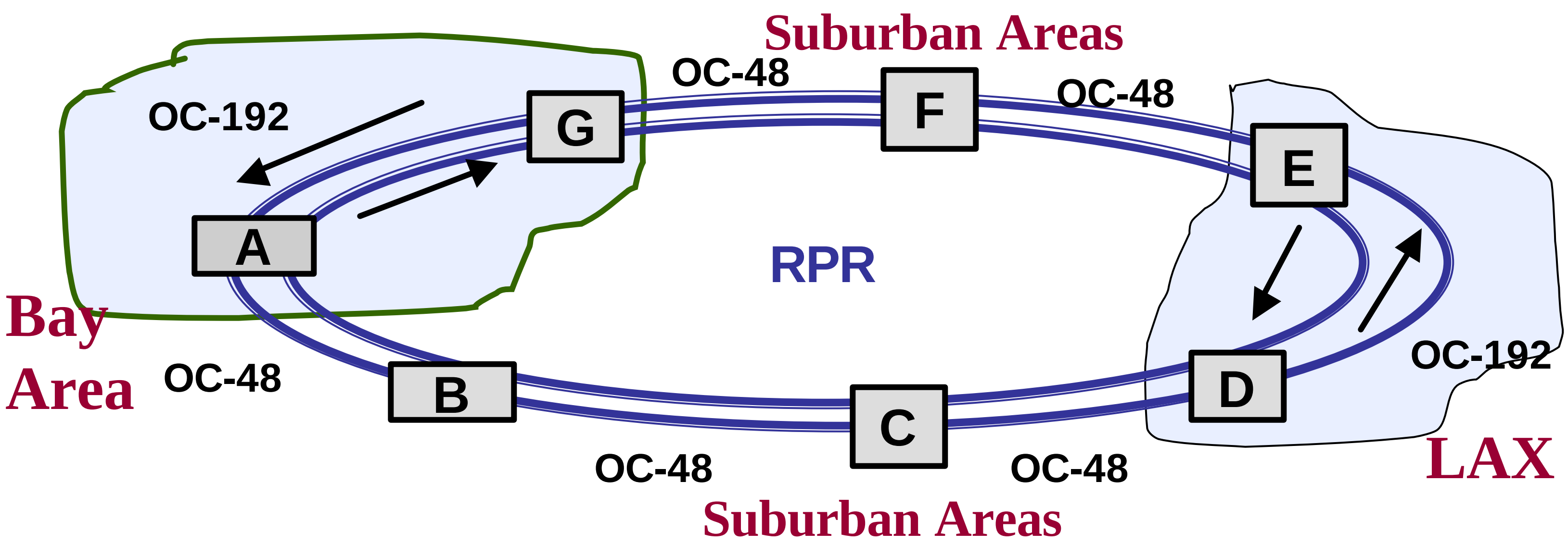
C is egress LSR for A-C and removes packet 1 on A-B-C LSP

- MPLS Labels carry packet from node to node, until packet reaches egress
- MAC addresses not needed for data transport due to NBMA nature of Optical Networks.
- Both directions of fiber can be used for data traffic, if needed.
- Unidirectional nature of LSPs fits well with fiber ring spans

Native Packet Transport



Different Rates on RPR Spans



- Traditional SONET/SDH TDM networks required ALL nodes to be upgraded to move to a higher speed. Expensive & Time Consuming
- With packet rings, it should be possible to increase rate of a section while keeping older nodes intact, with no impact to rest of the ring.
- When one fiber breaks, traffic flows are adjusted to fit the single fiber traffic rate.
- MPLS LSPs are unidirectional in nature, like optical fiber. Easy to send data packets on different LSPs with different speeds.

Frame Delineation Methods

- RPR must choose a robust frame delineation method
- Frame Delineation must provide a high recovery rate in high BER situations
- Avoid having to look at and de-stuff every byte - this is particularly problematic (high-speed hardware needed) at 10G+ speeds
- RPR packets can be sent on SONET/SDH as well as direct fiber.
- Two considerations compared here are - HDLC and SDL

Encapsulation with HDLC

- Byte-by-byte destuffing required
- Poor processing at higher speeds
- Loss of a single 0x7E can become a hopeless situation, since same value used both for SOP and EOP
- Hard to allocate queues/buffers - length of packet not known in advance.
- Poor delineation

Robust Transport with SDL

Simple Data Link (SDL) Protocol (rfc2823)

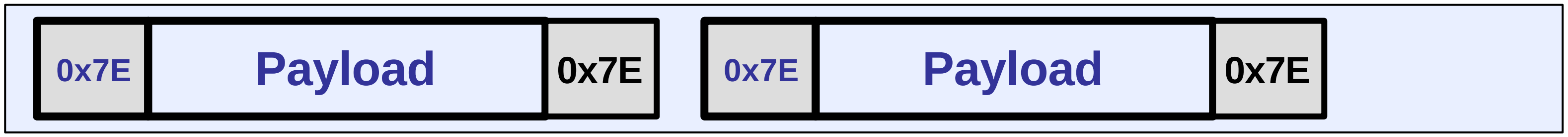
- Robust framing for any packet data over a point-to-point or a ring network.
- Packets are delimited using length/CRC construct instead of an HDLC-type framing with 0x7E at both ends
- Optional 16/32-bit CRC at the end of the packet
- No need to perform byte-by-byte de-stuffing at high optical data transport speeds.

Length	Length CRC	Payload (<= 64K)	(Optional) Payload CRC

Packet Delineation

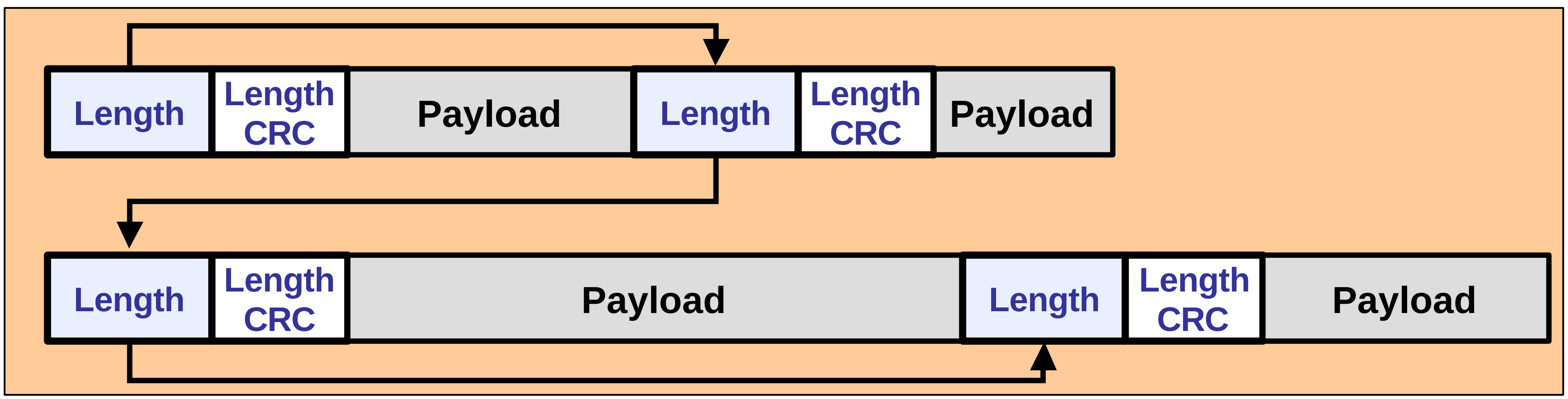
SDL vs. HDLC

HDLC



Byte-by-byte destuffing and checking for 0x7E and other control patterns
Loss of any 0x7E becomes a hopeless situation for recovery

SDL



Framers can jump Length bytes to get to next frame. No byte-by-byte lookup.

Conclusion

- RPR has a perfect opportunity to provide a unified mechanism for multiservice data transport over optical ring networks
- RPR would work for ring & mesh networks
- Native mapping for IEEE802.3 packets
- RPR can become the first protocol to formulate and allow use of MPLS as a L2 protocol for optical networks
- Use all existing protocols developed for Traffic Engineering with no modifications
- Native Data Transport for any Type of Data