Metro technologies GE, SRP, RPR and more abbreviations

Lars Hanson, lha@cisco.com

Stefan Lindeberg, stefan@lindeberg.net

Agenda

- Introduction
- Gigabit Ethernet
- Cisco Spatial Reuse Protocol, SRP
- Building metro networks
  - SRP vs. Ethernet
- New technologies
  - Resilient Packet Ring
  - 10 GE
- Q+A

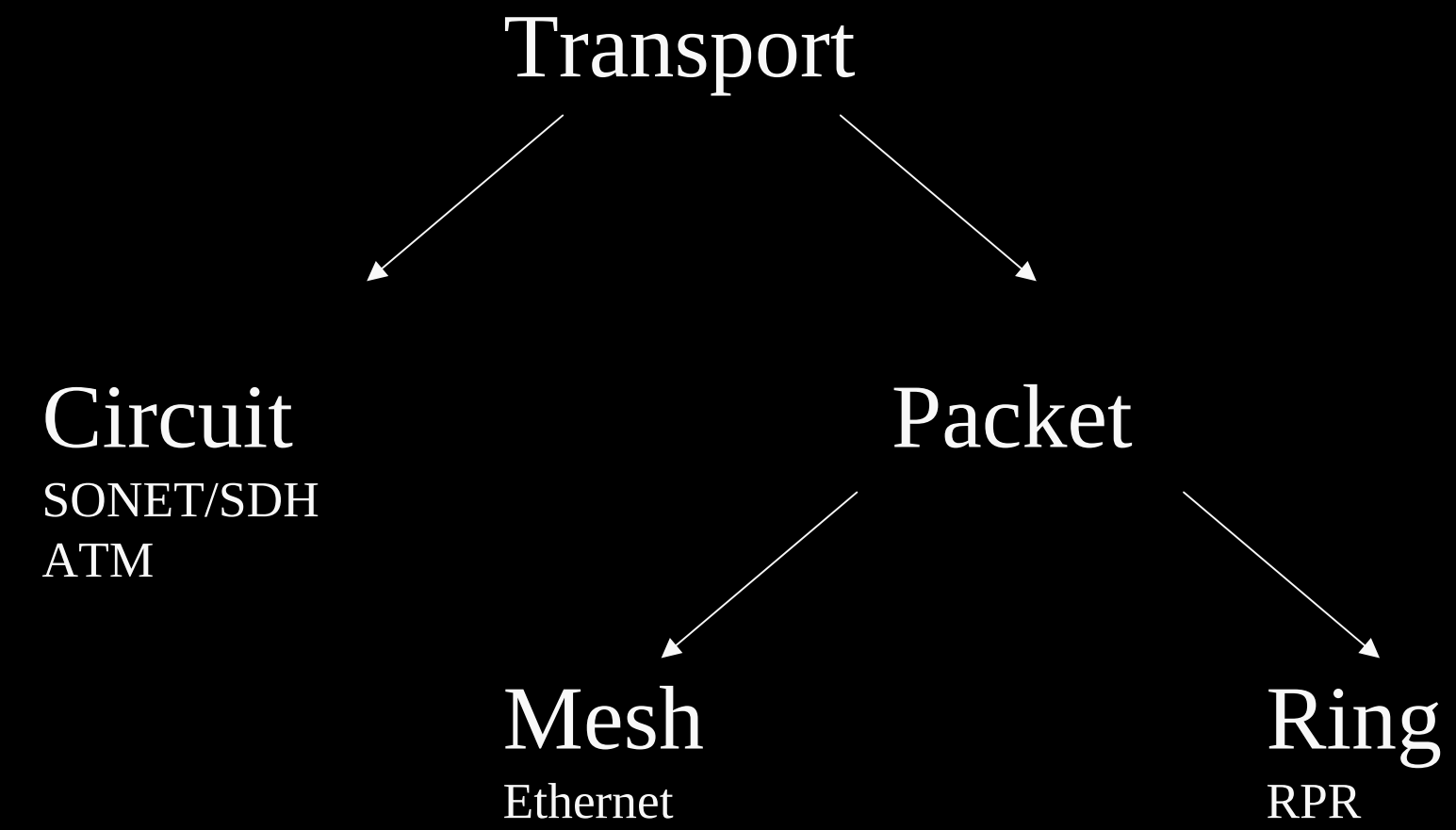
Stefan Lindeberg

- 16 years in datacom
  - IBM, NSC, Cisco
- Board member
  - Wavium, Xelerated, HyGlo, Comlase, Lumentis, PacketFront, Transmode
- Advisor to:
  - Dynarc, Startupfactory, Cygate

Lars Hansson

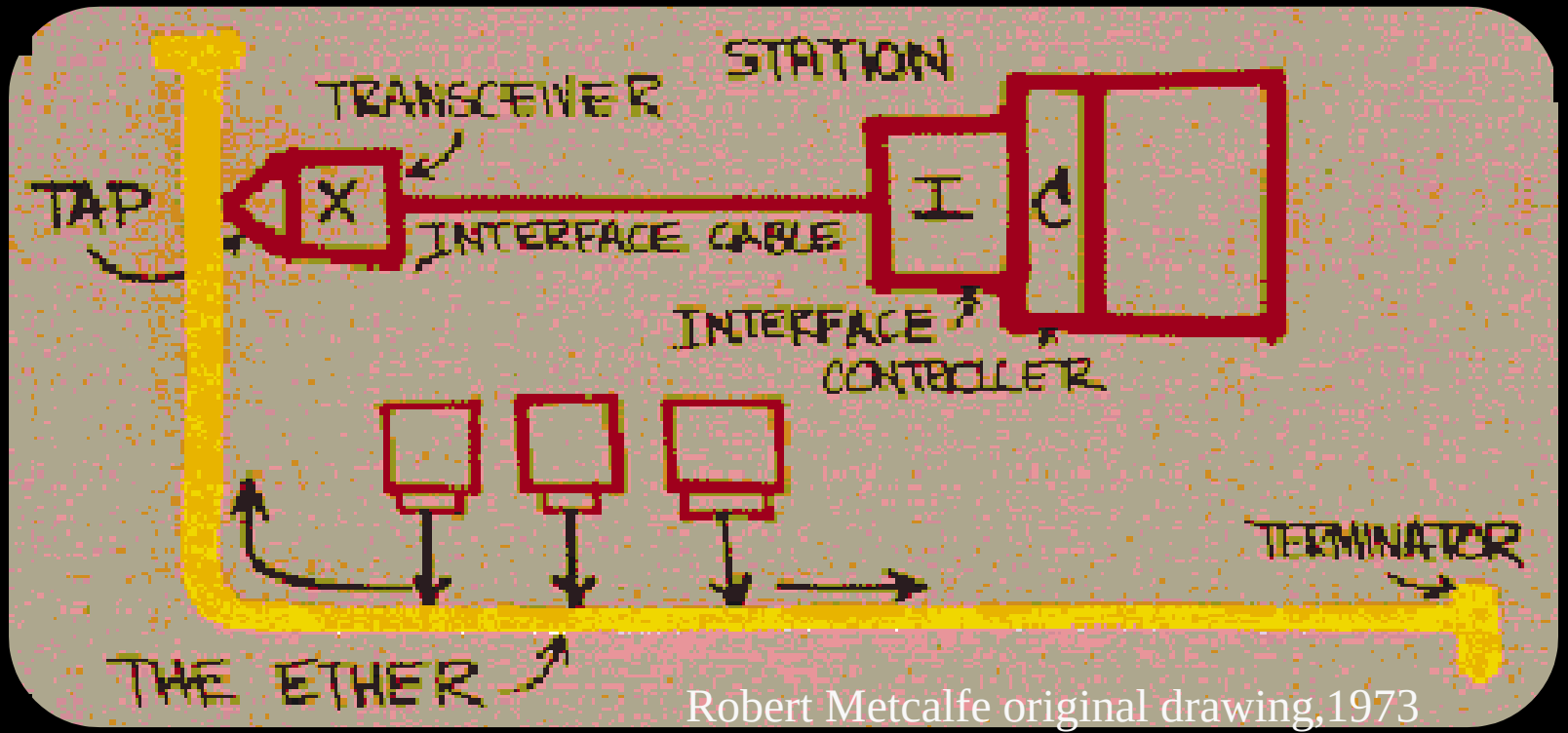
- 15 years in datacom
  - Upnet, Telia, Cisco
- Bla bla bla

Background Metro Technology



Ethernet

Ethernet



Robert Metcalfe original drawing,1973

”The value of the network expands exponentially as the number of users increases”

A reflection

- Remote Bridging
  - Remember Vitalink
  - Didn’t scale, needed layer 3 => router
- ATM LAN Emulation
  - Oops didn’t scale either
  - Lets add ½ a router => MPOA
  - Didn’t fly anyway
- MPLS L2 VPN
  - Here we go again!!!

SRP overview

SRP Overview

- SRP Protocol
- SRP Fairness
- SRP Protection

SRP Overview

- Cisco patent pending technology
- New MAC for LAN, MAN and WAN application
  - Spatial Reuse Protocol (SRP)
- Based on ring—dual counter rotating ring

SRP Overview

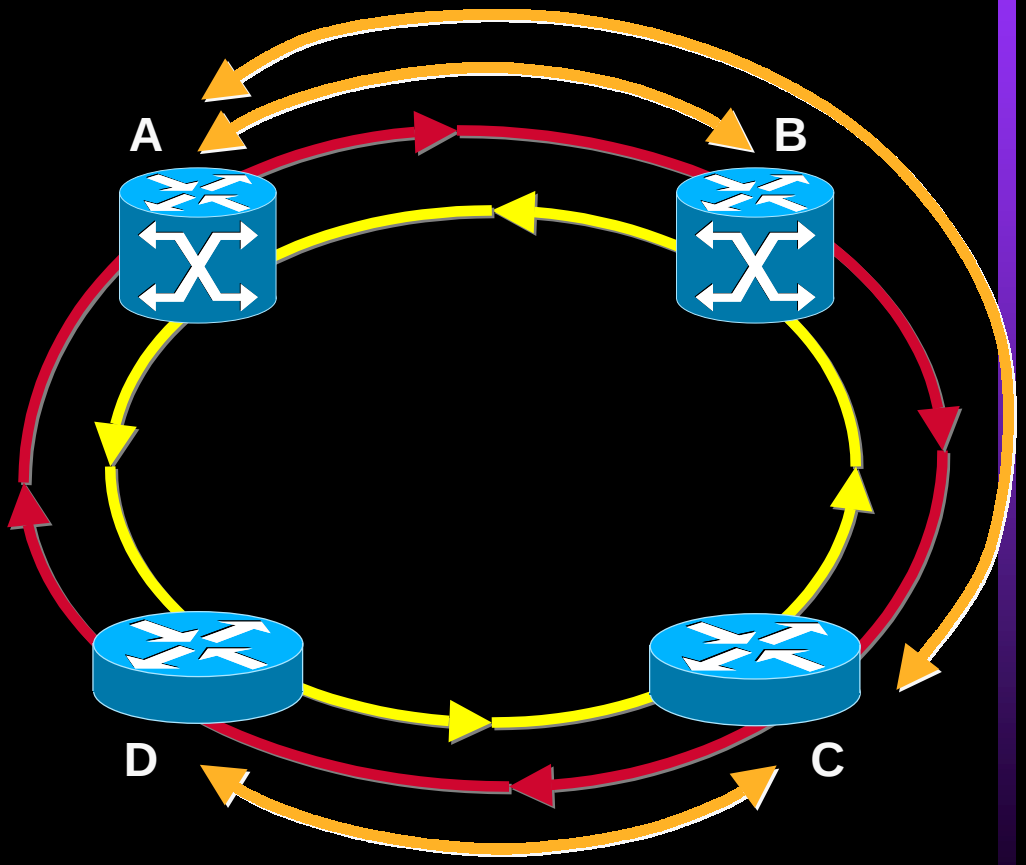
- Plug and play
- Multicast support
- Support traffic prioritization
- Uses the SRP fairness algorithm (SRP-fa) to control access to the ring and enforce fairness
  - No token—unlike Token Ring or FDDI
- Scalable to large number of nodes on the ring

SRP Overview

- Intelligent Protection Switching (IPS)
  - Survivability in the event of fiber facility or node failure, or signal degradation
- Media independent protocol
  - Initial implementation uses SONET/SDH framing

Spatial Reuse

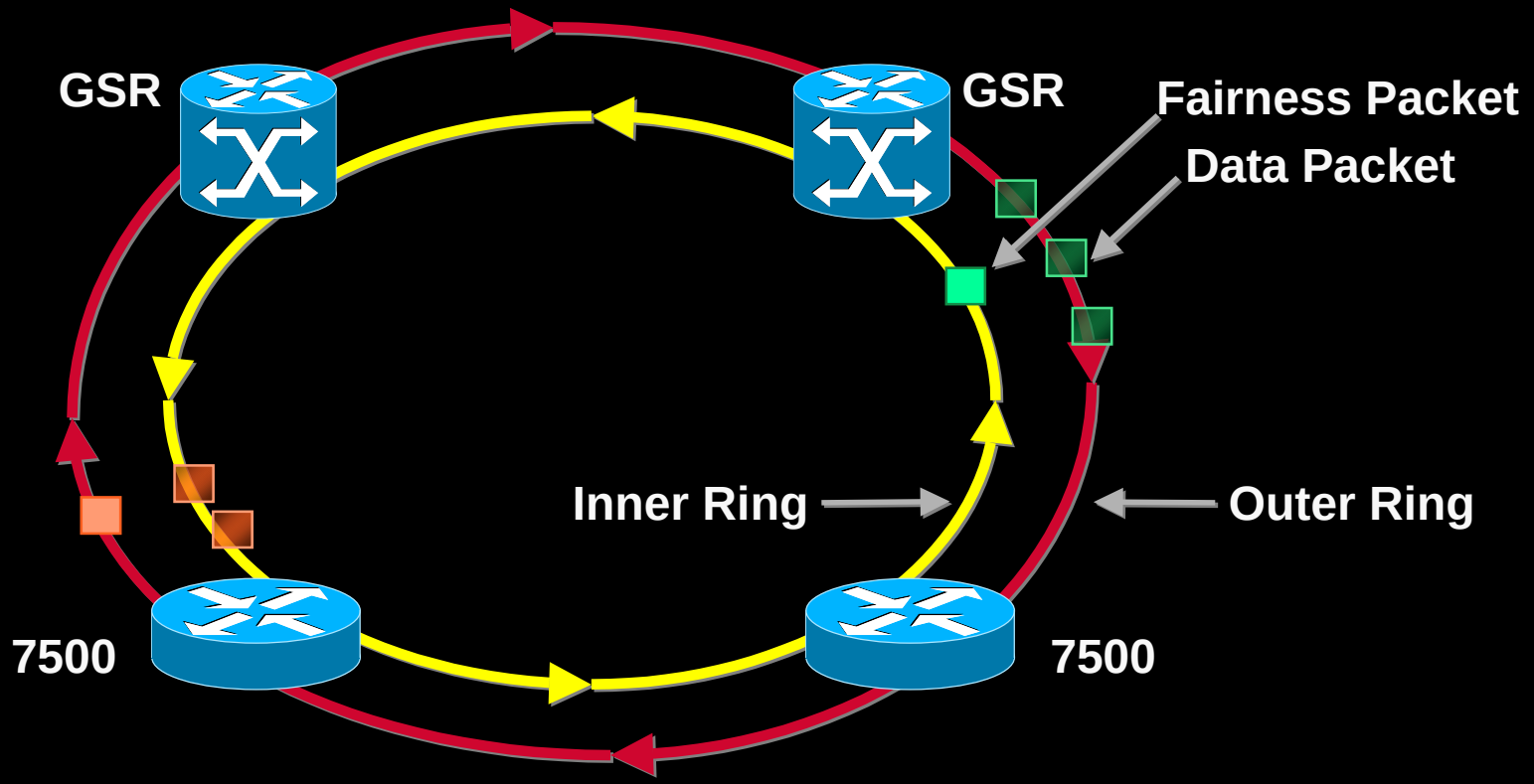
- The SRP protocol derives it's name from the spatial reuse capability
- Bandwidth consumed only on traversed segment
- Unicast packets travels along ring spans between the src and dest nodes only
  - –Destination stripping



SRP Ring

- An SRP ring is a bi-directional dual counter rotating ring
- The rings are referred to as Outer and Inner rings
- Both rings are used to transport data and control packets
  - Data packet is sent in one direction and the corresponding fairness packet is sent the opposite direction

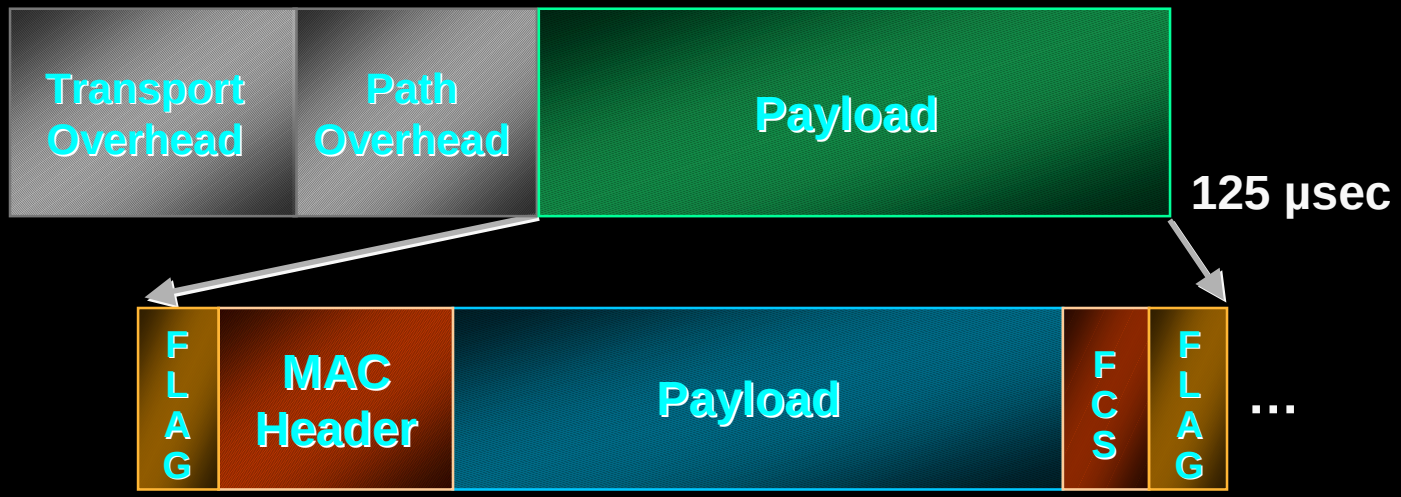
SRP Ring



Example of a 4 Node SRP Ring

SRP Packet

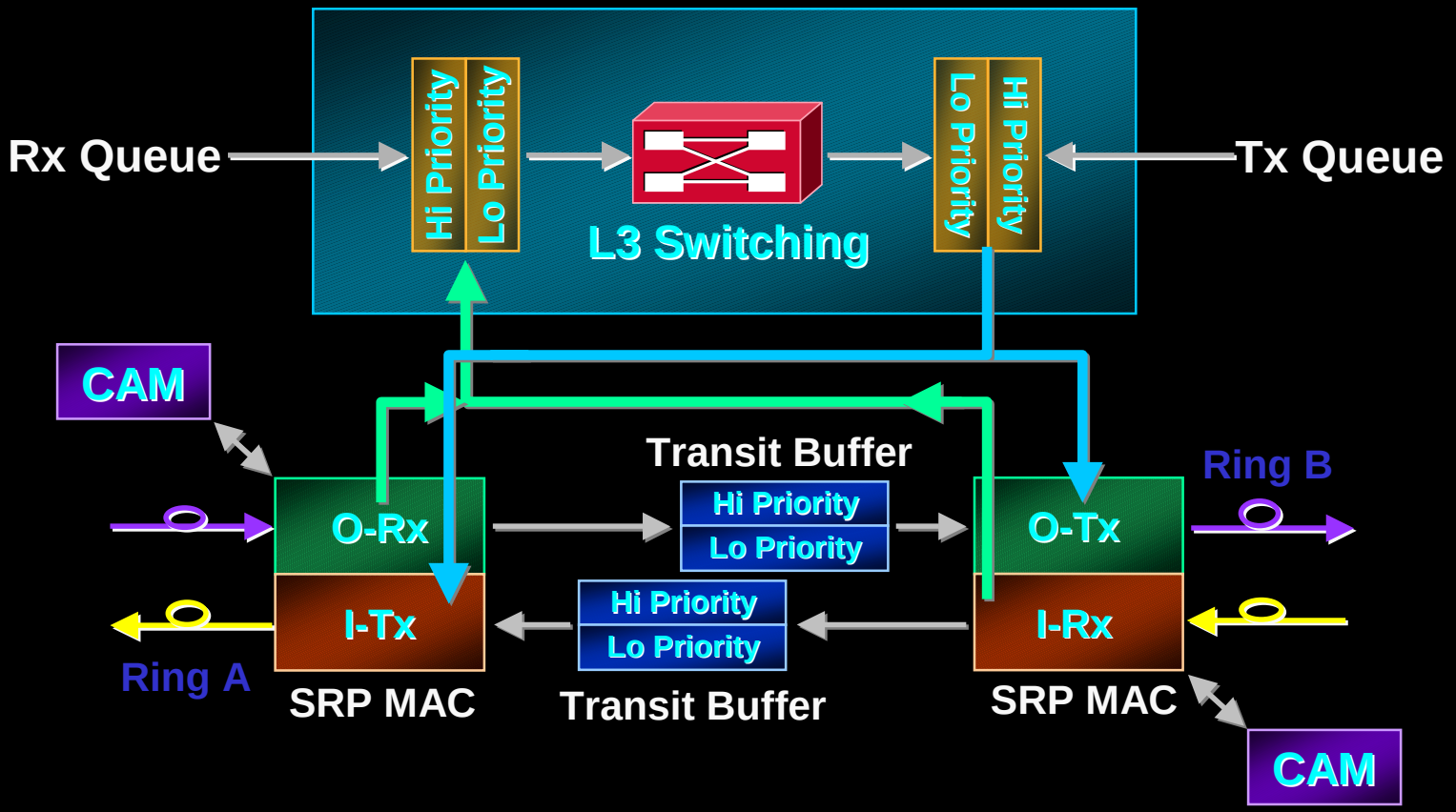
- SRP is a media independent MAC layer protocol
- The initial implementation utilizes SONET/SDH framing
- Concatenated payloads only



Single Subnet

- Both Outer and Inner rings are on the same IP subnet
- This enables rapid re-optimization of ring path selection and minimize route flaps in a ring wrap situation
- Ring wraps are handled by the lower layer and thus transparent to layer 3 routing protocols

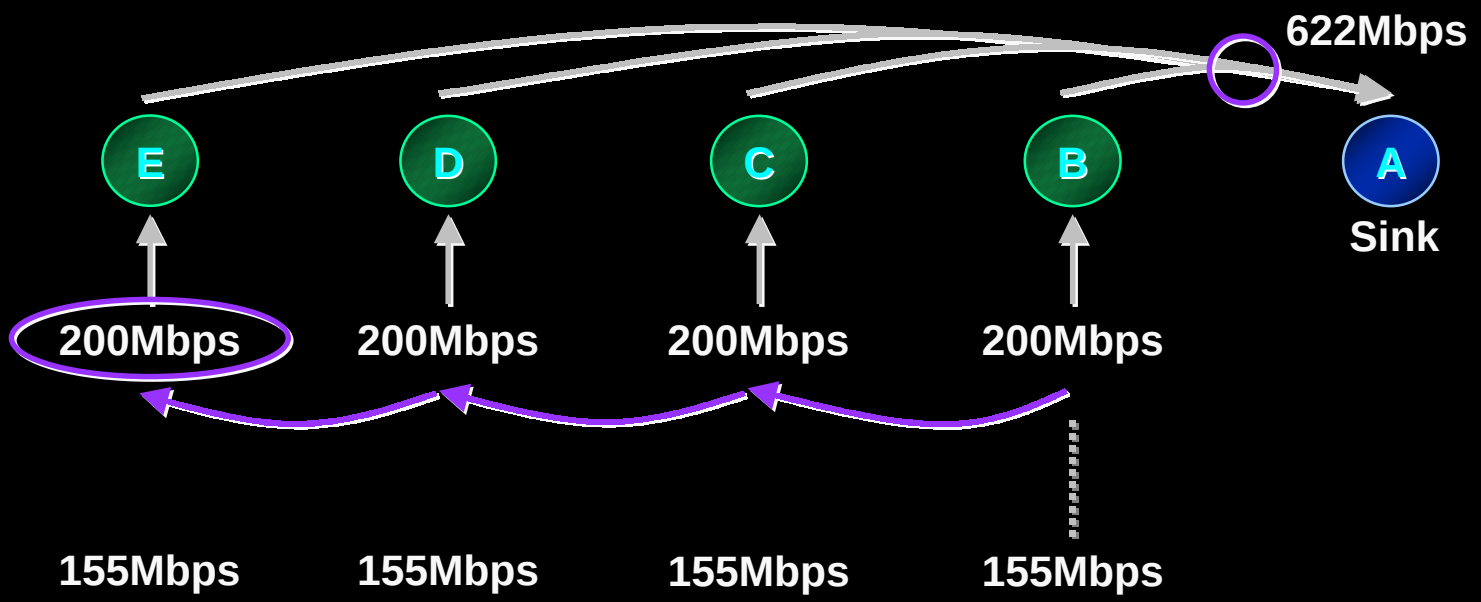
Node Packet Flow



SRP Fairness Algorithm

- SRP-fa is the mechanism that ensures
  - Global Fairness—each node gets a fair share of the ring bandwidth
  - Local Optimization—node maximally leverage the spatial reuse properties of the ring
  - Scalability—the ability to build large rings with many nodes that spans across large geographically distributed area

SRP-fa Operation Example



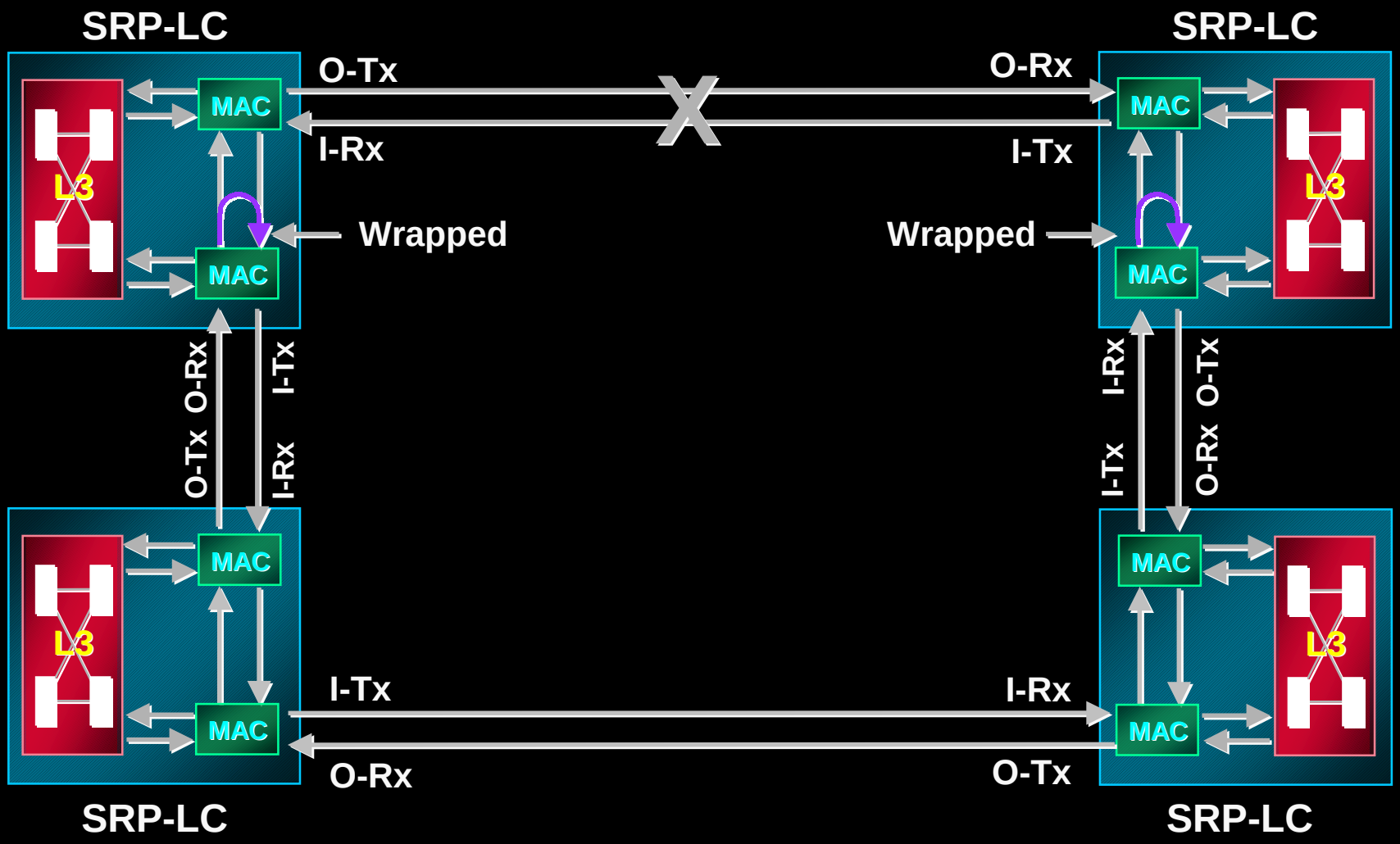
Intelligent Protection Switching

- IPS provides SRP with a powerful self healing feature which automatically recovers from fiber facility or node failure, or signal degradation
- IPS is analogous to the self healing properties of SONET/SDH rings
  - but without the need to allocate protection bandwidth

IPS

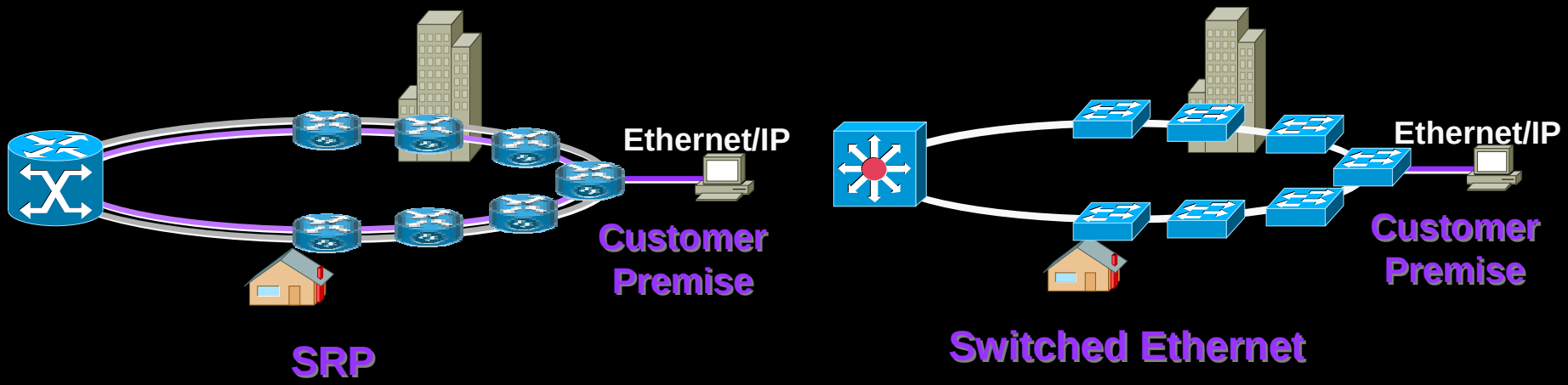
- Topology knowledge independence
- Ring wrapping to bypass failed fiber or node
  - Transparent to the layer 3 routing protocols
- Protection switching event hierarchy
- Ring restores in ≤ 50 msecs

Wrapped Packet Flow



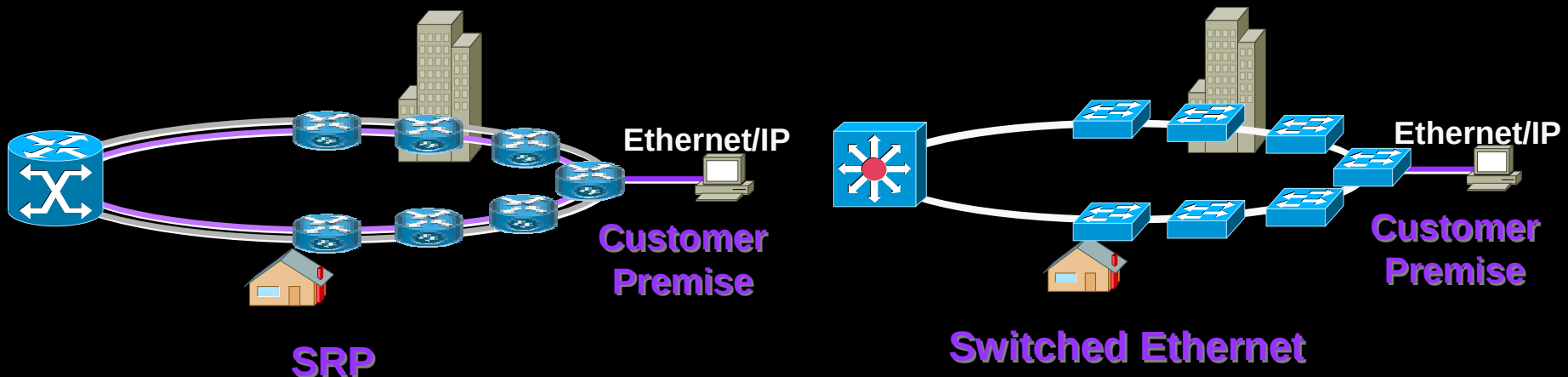
Cisco SRP vs Ethernet in a Metro Environment

SRP vs Ethernet - Scalability



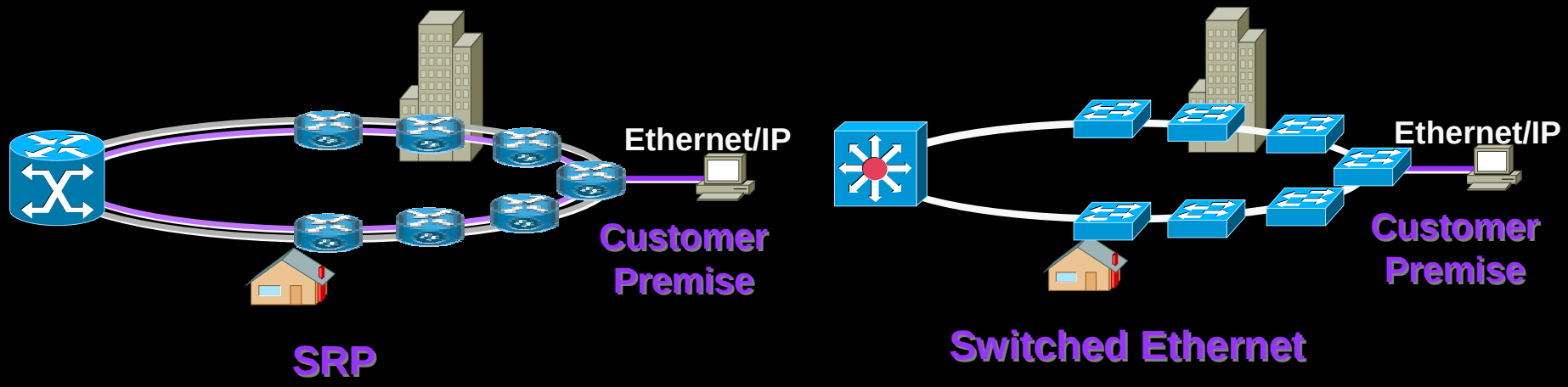
- 128 nodes per SRP ring vs. 7 nodes per Spanning Tree ring
  - Work in progress within IEEE to change, 802.1w

SRP vs Ethernet - Restoration



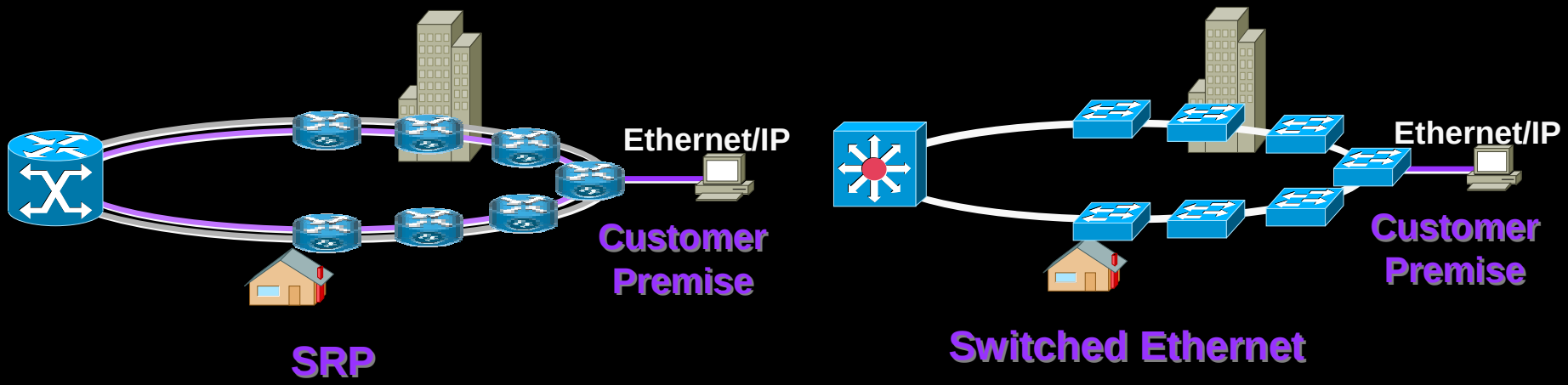
- 50ms SRP ring restoration vs. 5-50 seconds Layer 2 restoration
  - –Spanning tree work in progress IEEE 802.1w
- Layer 3 Ethernet and routing protocol, 2-90 seconds
- Head end must be layer 2 device to avoid partitioning
  - –Layer 3 can be separate device or integrated

SRP vs Ethernet



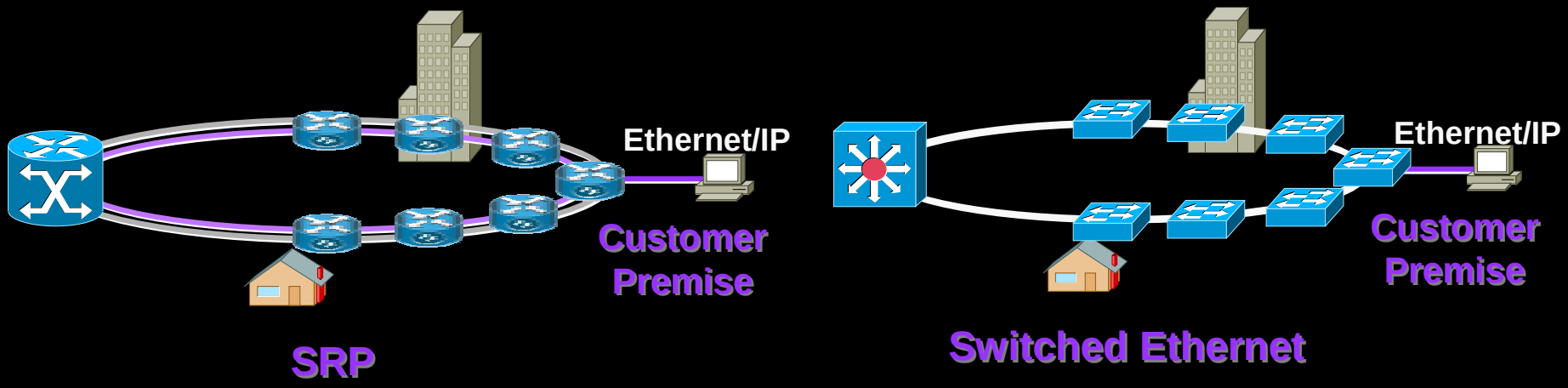
- Predictable latency and jitter with SRP
- No packet loss during ring transit
  - Packet on the ring will not be dropped

SRP vs Ethernet



- SRP - Extensive trouble-shooting capabilities due to SDH/SONET framing
- Ethernet – almost binary
- CAPEX vs. OPEX

SRP vs Ethernet



Other factors

- SRP more expensive
- Fewer vendors
- None standard technology

Metro Ethernet Forum

Looking at protocols and mechanism in standard ethernet that need to be improved for MAN applications, for example but not limited to:

- Spanning tree, STP
- VLAN
- QoS

802.17 Resilient Packet Ring RPR

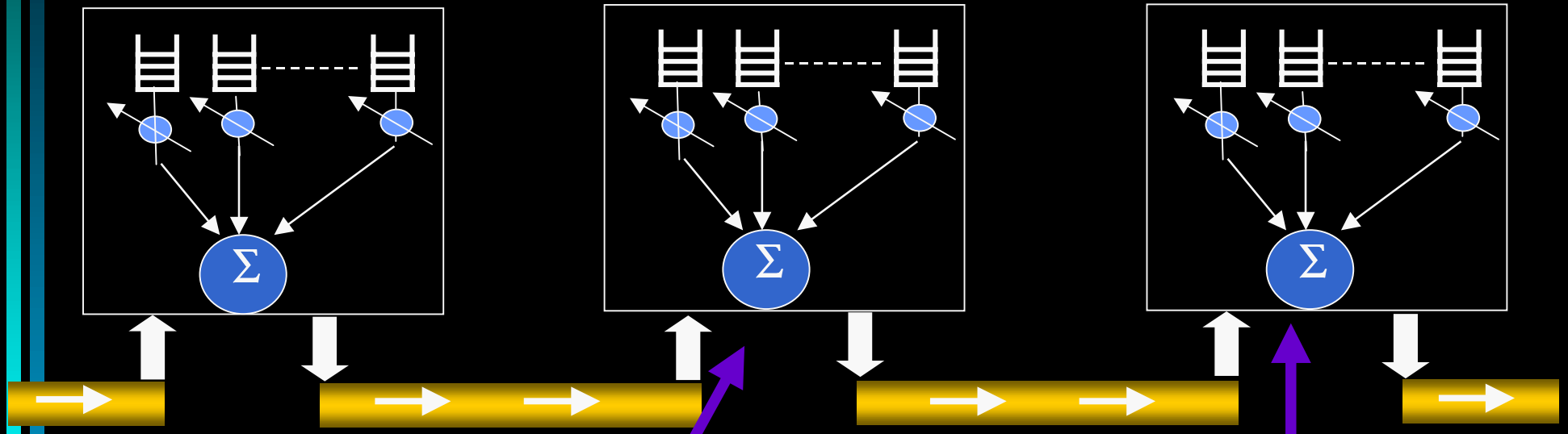
802.17 RPR

- IEEE 802 standardization effort for resilient packet rings
- Rules of the game, no existing technology can ”win”
  - Cisco SRP
  - Dynarc DTM
  - etc

802.17 Timeline

- Proposal cutoff November 2001
  - Currently two proposals
    1. Group 1 RPR proposal based on modified SRP
       - Cisco, Riverstone, Spirent Communications, AMCC, Mindspeed ……
    2. Group 2 RPR proposal
       - Alcatel, Dynarc, Nortel, Lantern, Luminous, NEC, ……
- First Draft January 2002
- Last addition May 2002
- WG Ballot July 2002
- Last change September 2002
- LMSC Ballot November 2002
- Standard March 2003

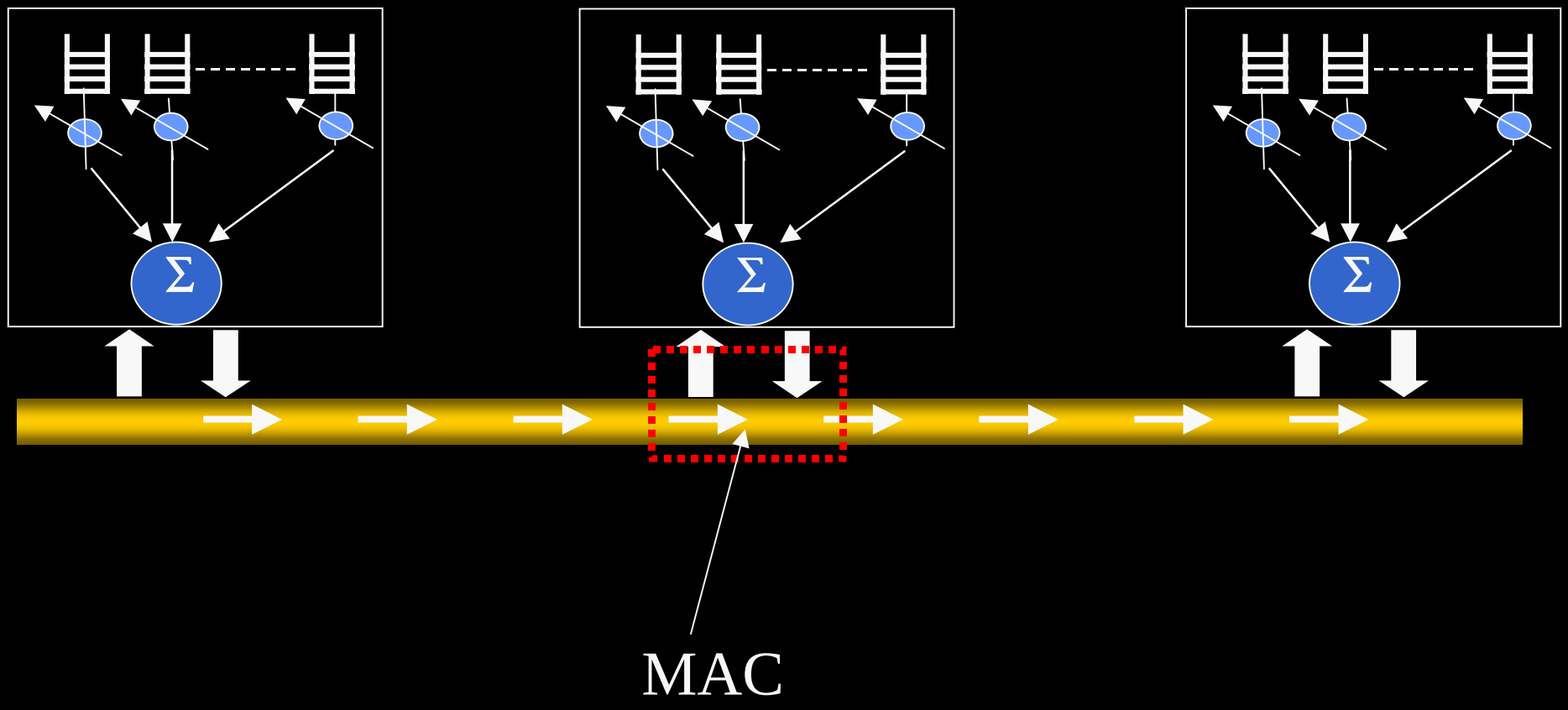
Packet switch architecture



Packet switches connected by point-to-point links

No shared medium

RPR MAC architecture



Group 1 RPR proposal

Group 1 RPR Proposal

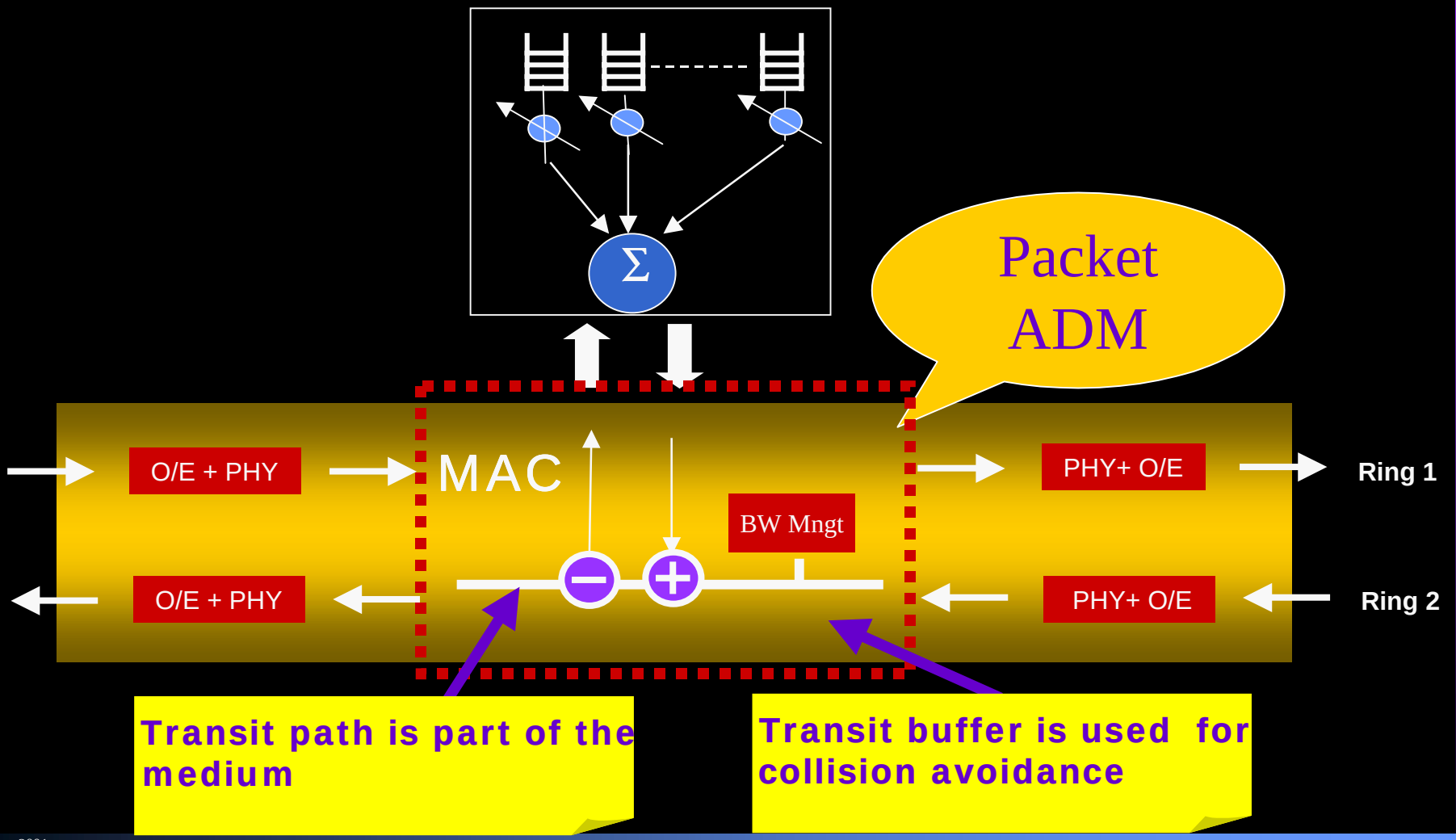
- Based on SRP
- Enhanced Fairness Algorithm
  - Knowledge of affected SPAN's
- Both Wrapping and Steering for protection
- New Frame Format – CRC on the packet header
- Physical Media – SONET/SDH, GFP and Ethernet

Group 2 RPR proposal

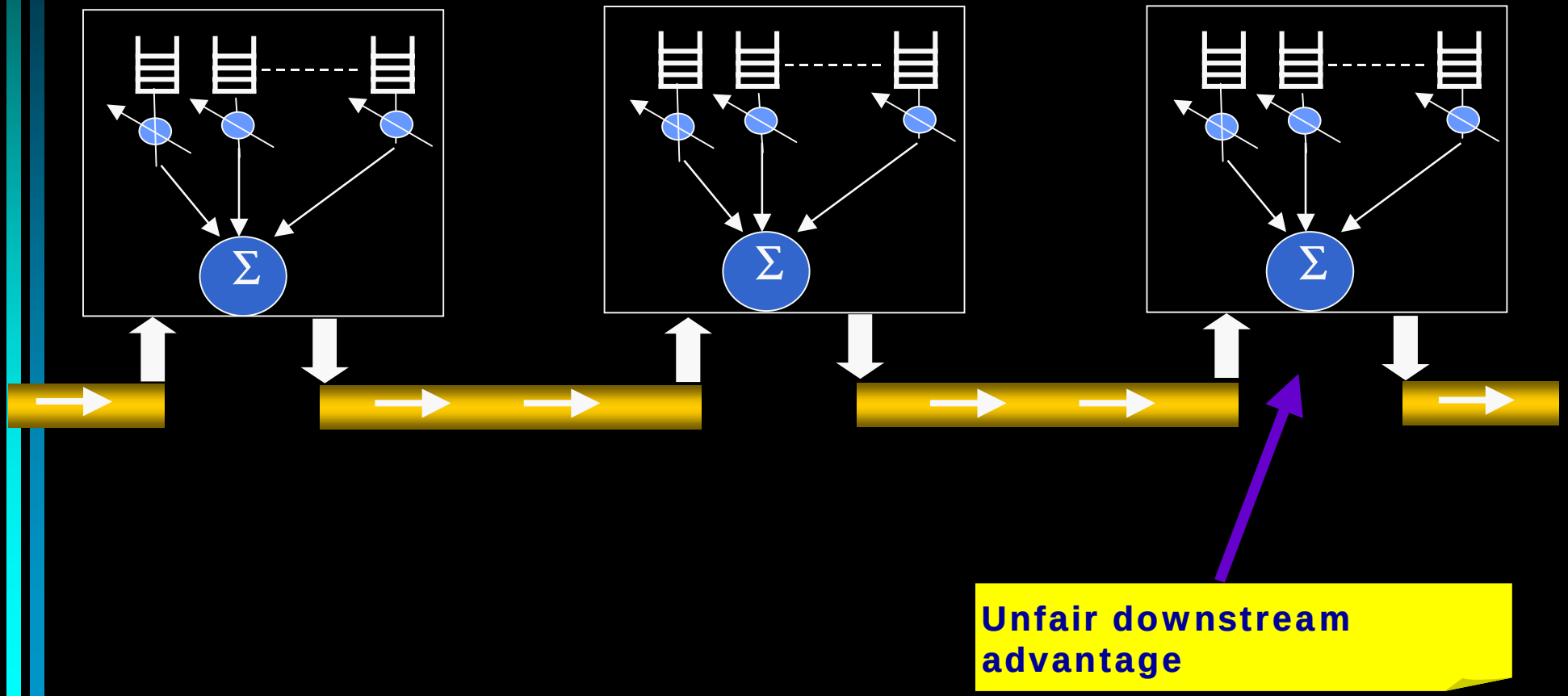
Fundamentals of Group 2 RPR proposal

- Shared media architecture
- Fair access
- Active dynamic bandwidth management
- Bandwidth-aware MAC
- Maximize throughput on all links
- Support for multiple rings
- Steering-based protection scheme

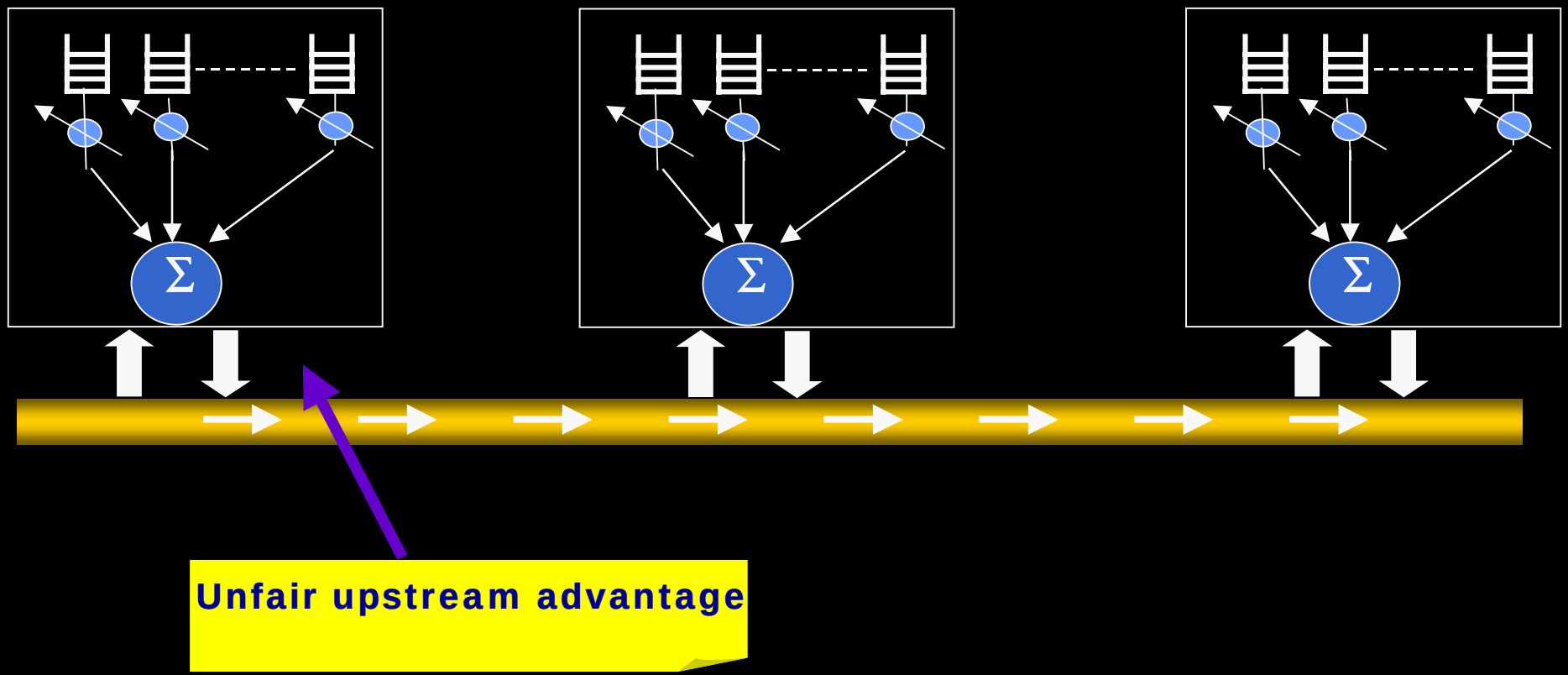
The RPR MAC


Σ
Packet
ADM
O/E + PHY
MAC
BW Mngt
PHY+ O/E
Ring 1
O/E + PHY
PHY+ O/E
Ring 2
Transit path is part of the medium
Transit buffer is used for collision avoidance

Fairness: Packet switches are not fair



Fairness: Packet ADMs can be unfair too



Fairness

Service received by a subscriber is independent of the subscriber's location on the ring

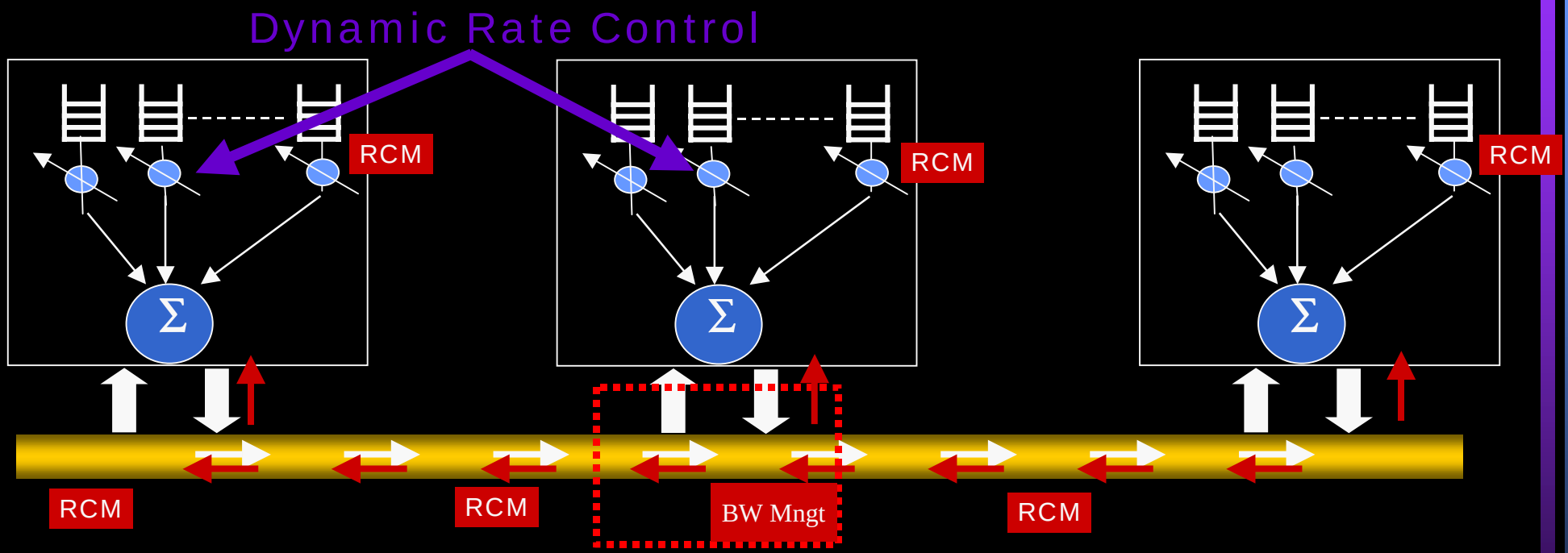
- Fairness is not equality
- Fairness = weighted allocation of resources

Active Bandwidth Management

Dynamic flow control is essential for maximum network utilization in shared network environments

- Closed loop flow control scheme
- Fairness algorithm is an integral component of bandwidth management scheme
- Network performance limited by robustness, responsiveness and precision of flow control

Bandwidth-aware MAC



RPR proposal BW management

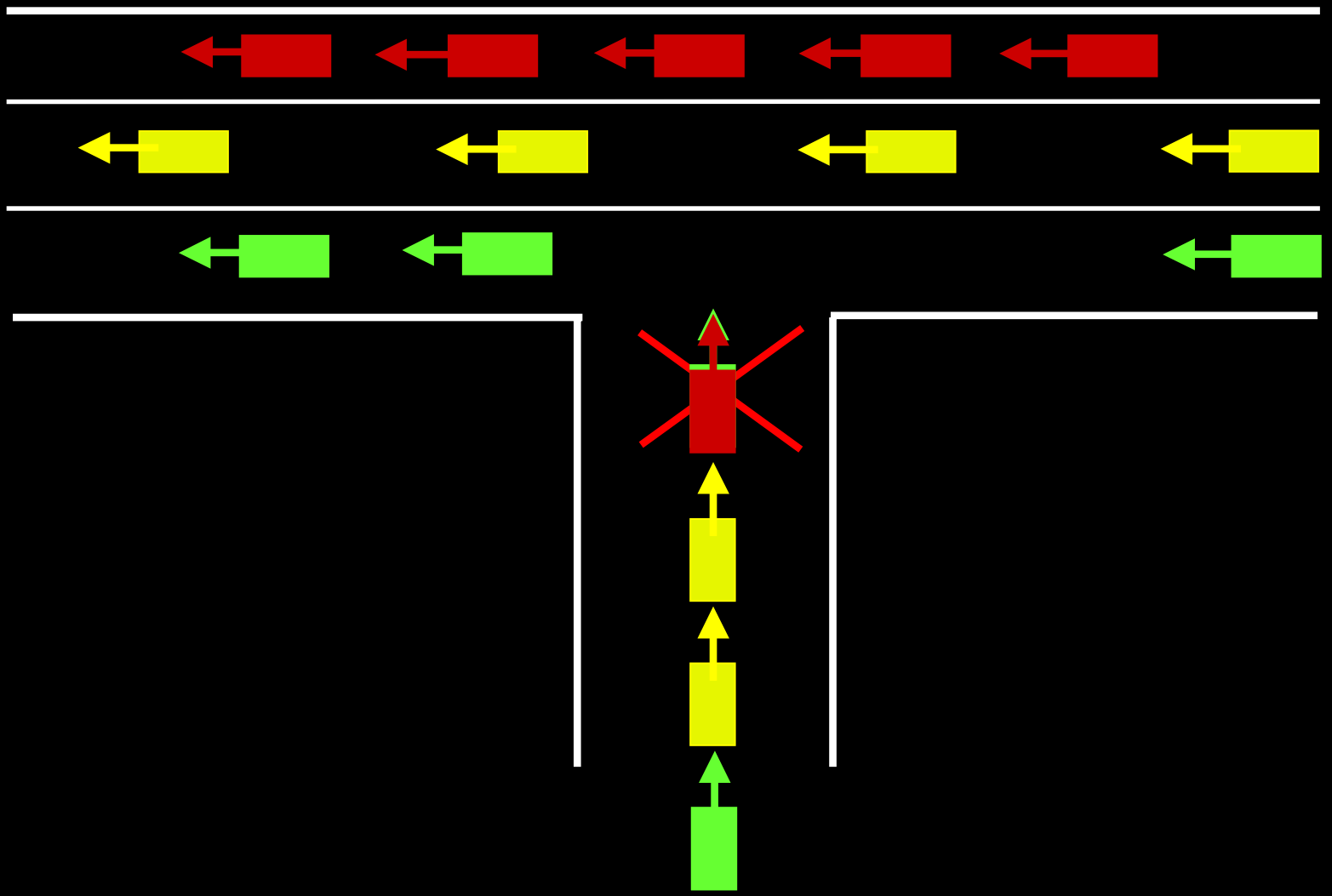
- Congestion avoidance
- Active BW management
- Fair access as traffic pattern changes
- Dynamic o Reallocates resources, high through put as traffic profile changes
- Support for VoQ and eliminate HoL blocking maximize throughput on all links
- Support for N+ 1 ringlets
- Support for weight fairness

Maximize throughput on all links

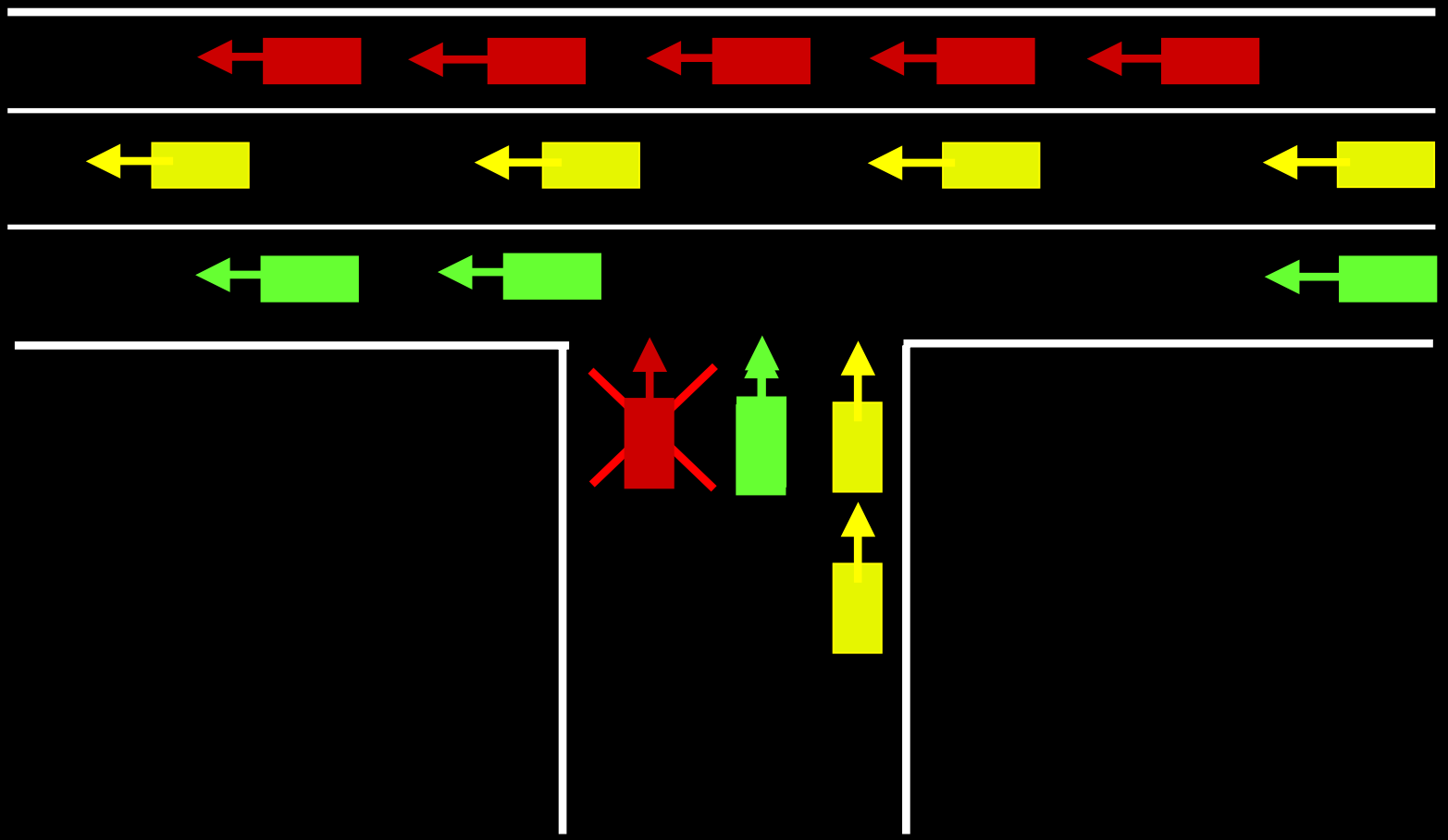
Traffic over un-congested links should not be throttled because of congestion happening on unrelated links

- Requires knowledge of destination

Head of line (HOL) blocking

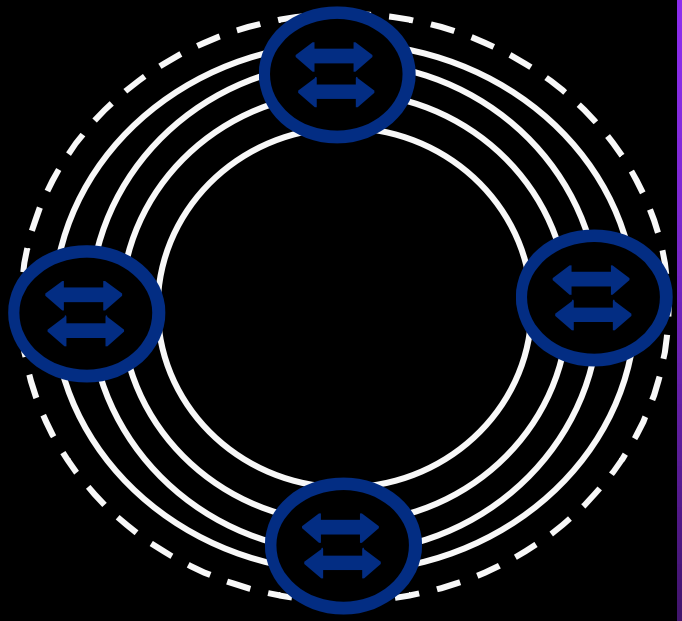


Throttle sources affected by congestion



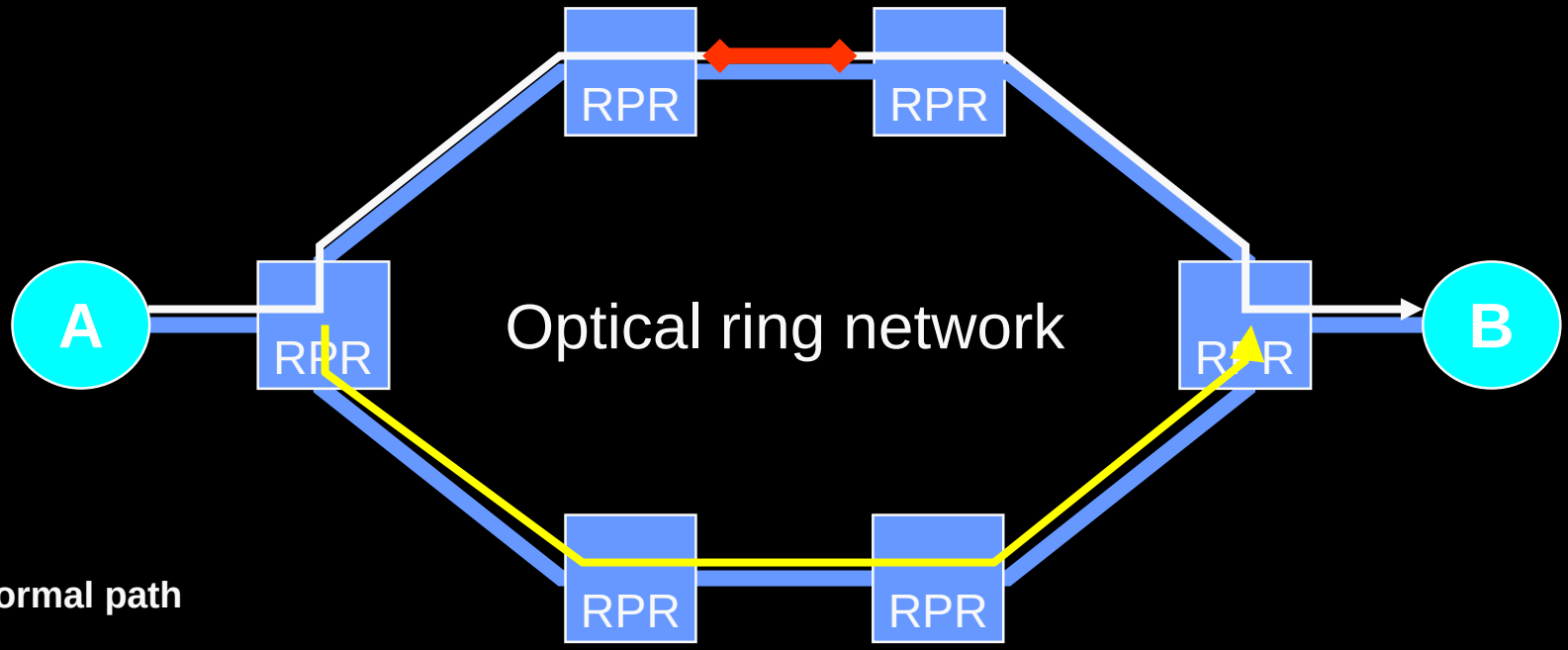
Multi-rings Scalability

- Problems:
  - Limited transmission capacity on fibers
  - high speed optics expensive
    - costly to scale by increasing link speed
  - linear increase
    - One ring at a time
- Conclusion:
  - scale by adding multiple rings
- Benefits
  - each additional ring increases the capacity
  - cheaper to add ring than to increase speed
  - individual rings can be operated at different speed
  - one logical Mac--several physical
    - the rings are managed as one aggregated link



Steering Ring Protection

- Optimal performance after failure
- Sub-50ms service restoration



Steering Ring Protection

- Each station knows of a ring segment failure and steers ring traffic away from the failure within 50ms of the failure
- Ring protection is initiated by all stations that become directly aware of a failure via local detection or through broadcast announcement
- Each station uses its knowledge of the topology of the ring to know how and when to steer ring traffic away from a failure

Comparison of the two proposals

IEEE 802.17 proposals comparison

	Group 1	Group 2
Protection	Wrapping or steering	Steering
Provisioning	Data centric	ADM like
Transit buffers	2 buffers priority levels	1 buffer collision avoidance
Fairness	Allow burst, manage when congested	Avoid congestion and allocate by request

Summary: RPR proposals

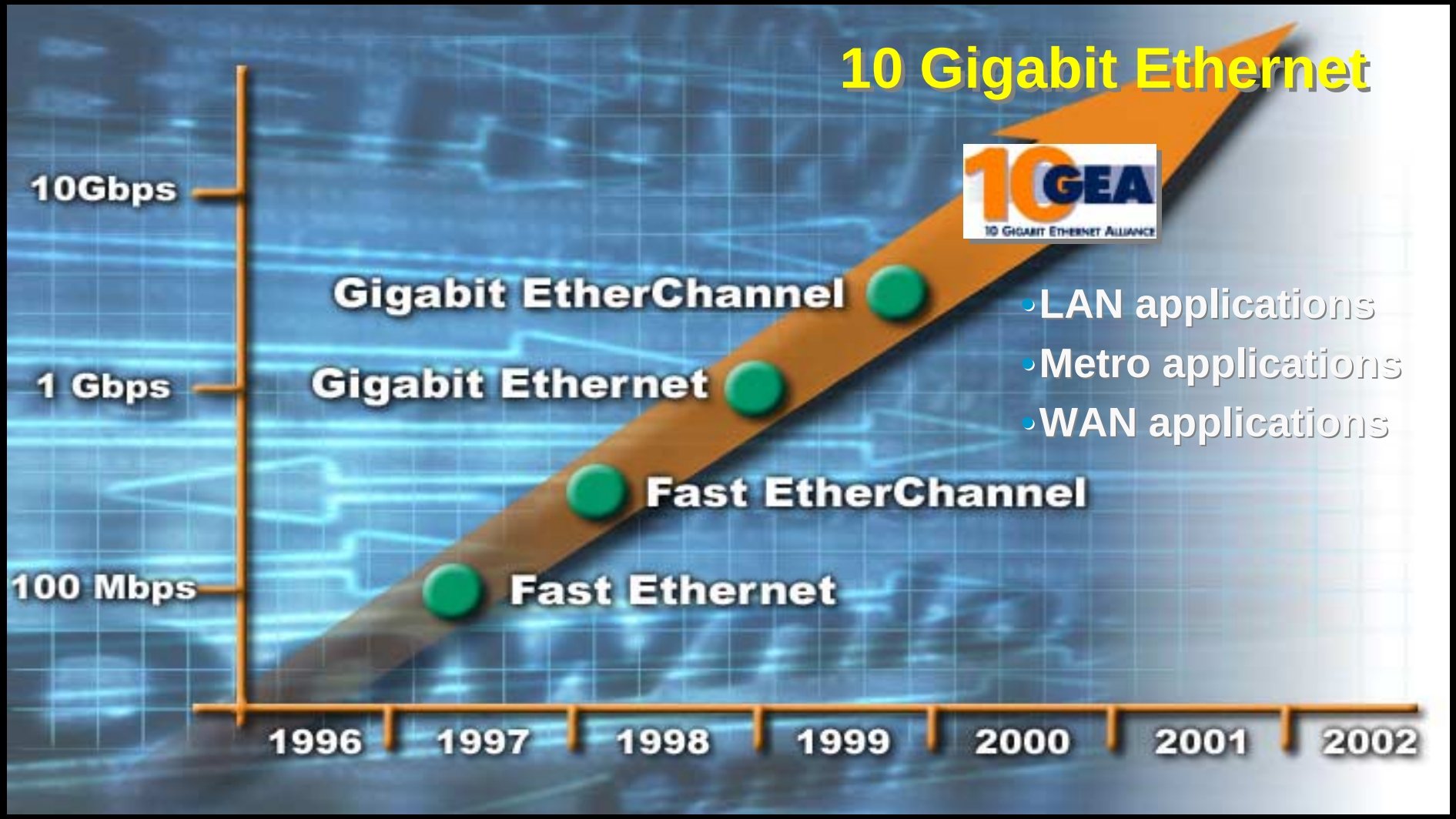
- 802.17 =|any current IP ring implementation
- Shared medium architecture
  - Transit path is part of the medium
  - Transit buffer is used for collision avoidance
- Active bandwidth management is required to provide fair access of ring capacity to stations
- Dynamic bandwidth management to avoid unused (wasted) capacity
- Bandwidth-aware MAC
  - Awareness of available capacity on links of the ring
- Maximize throughput on all links

10 Gigabit Ethernet
IEEE 802.3ae

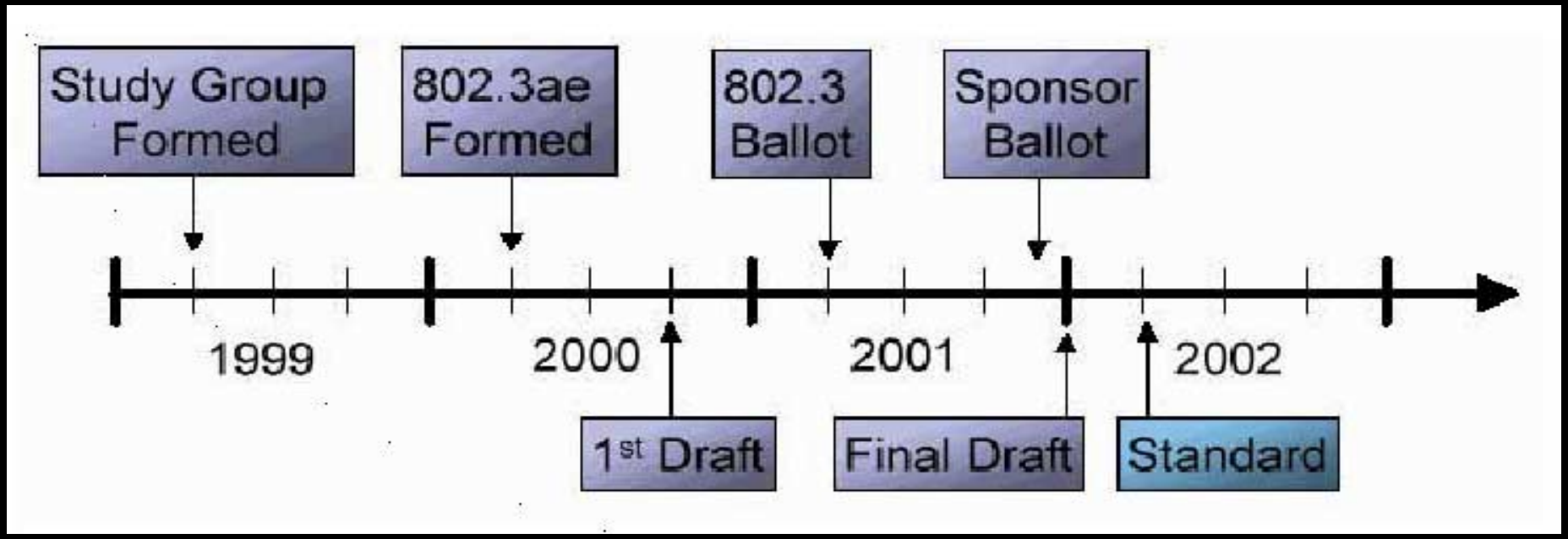
10 Gigabit Ethernet

- Timeline
- Overview
- PHY’s and PMD’s
- Status

Ethernet development



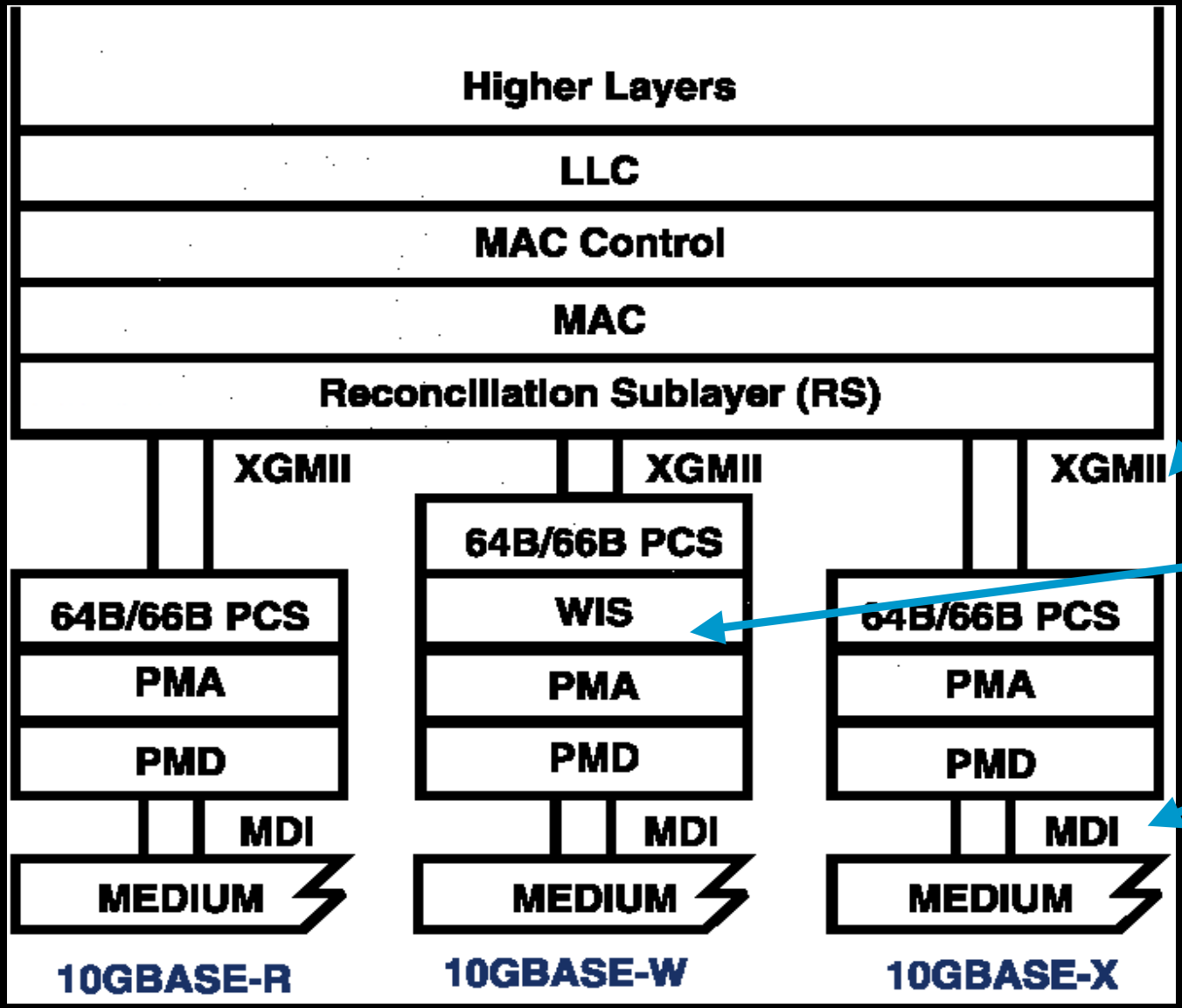
10GE


Study Group Formed
802.3ae Formed
802.3 Ballot
Sponsor Ballot
1999
2000
2001
2002
1st Draft
Final Draft
Standard

10GE Overview

- Uses 802.3:
  - MAC
  - Frame format
  - Min and max frame size
- Full duplex and fiber only
  - Doesn’t need CSMA/CD
- Physical Coding Sublayer
  - 64B/66B
- LAN PHY and WAN PHY

10 GE layers



Media Independent Interface

WAN interface sublayer

Media Dependent Interface

Putting it together

Stack	10 GE LAN PHY		10 GE WAN PHY
	Serial	Parallell	Serial
MAC	10Gb/s	10Gb/s	10Gb/s
PCS	64B66B	8B/10B	64B66B
PMD	1550nm 1310nm 850nm	1310nm	1550nm 1310nm 850nm
Line Rate	10.3Gb	4x3.125	9.953Gb/s

10 Gigabit Ethernet Optical Transceivers (PMDs)

PMD	Fiber Supported	Diameter (Microns)	Bandwidth (MHz*km)	Distance (Meters)
850 nm serial	multimode	50	500	65
1310 nm WWDM	multimode single mode	62.5 9.0	160 N.A.	300 10K
1310 nm serial	single mode	9.0	N.A.	10K
1550 nm serial	single mode	9.0	N.A.	40K

Questions?

Thank you!

Lars and Stefan