IPoDWDM, 40G, 100G and beyond

Dirk Schroetter, Consulting Systems Engineer

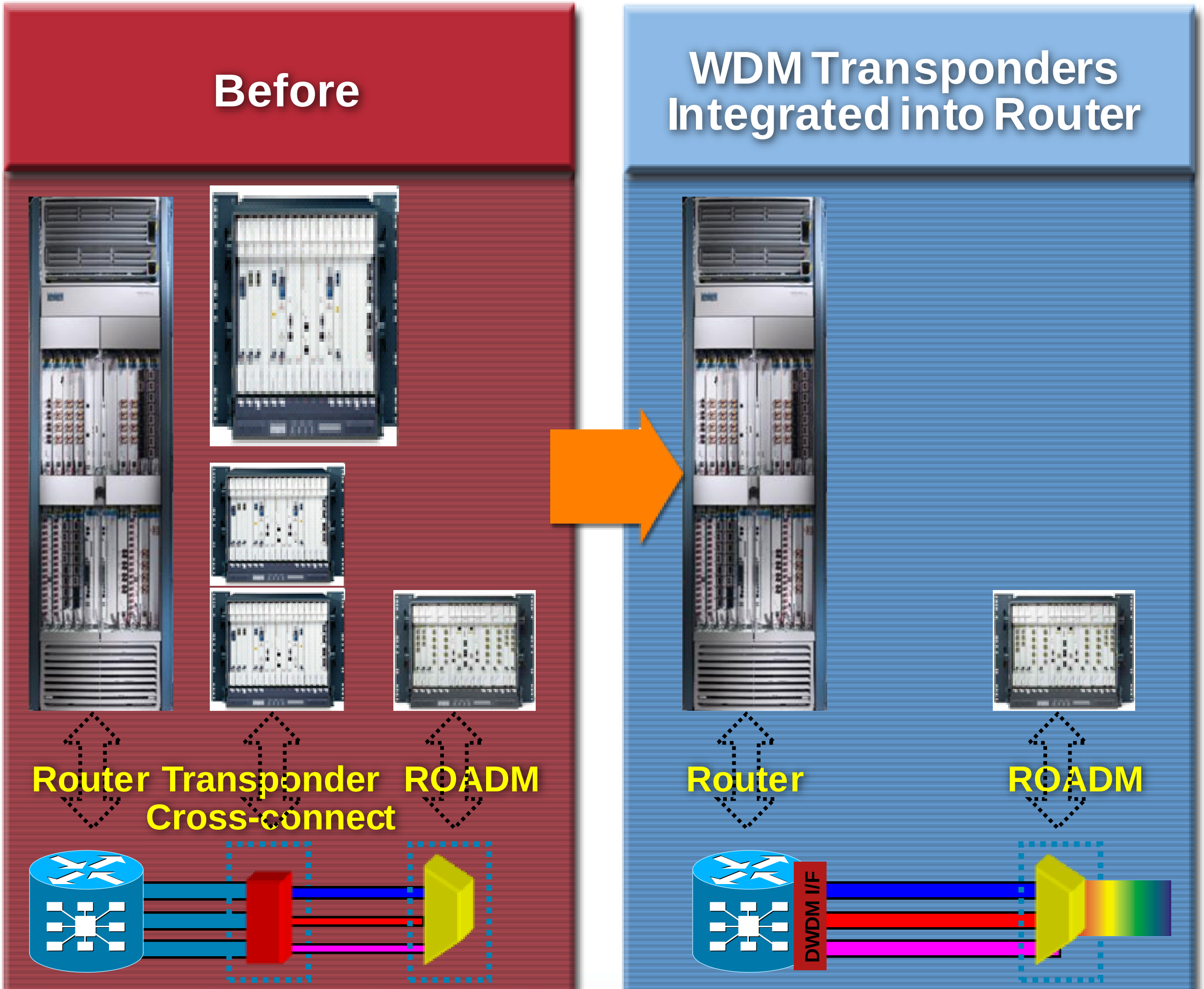
Welcome to the Human Network.

Agenda

- IPoDWDM overview
- Beyond the Basics – Proactive Protection
- Moving to 40G, 100G and beyond
- Q&A

IPoDWDM introduction

Basic idea of IPoDWDM on one slide



- Lower CapEx

  Elimination of OEOs

- Lower OpEx

  Space, power, management

- Enhanced resiliency

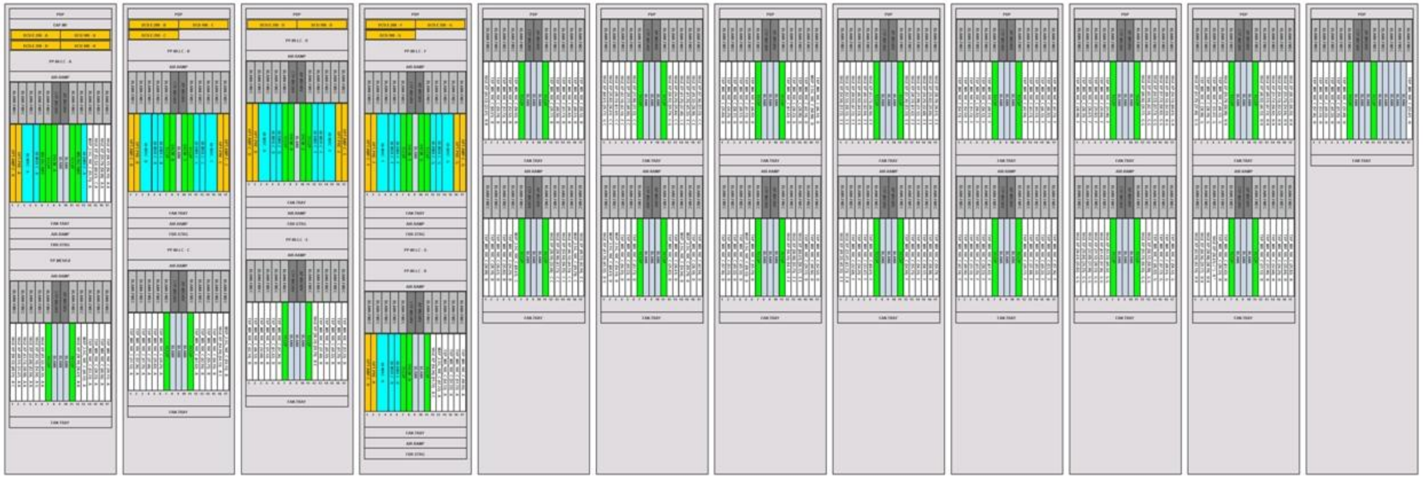
  Fewer active components

- Investment protection

  40G and beyond, interoperability over *existing* 10G systems

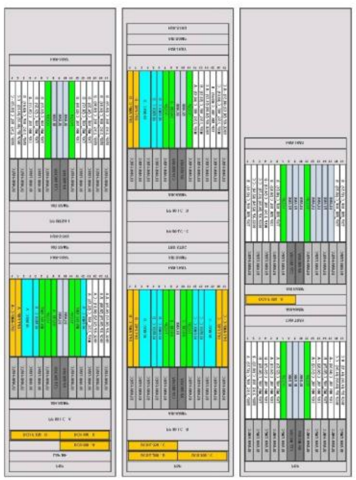
CapEx / OpEx / Green: Real-world example

Large EU SP



12 racks

DWDM model



3 racks

IPoDWDM model

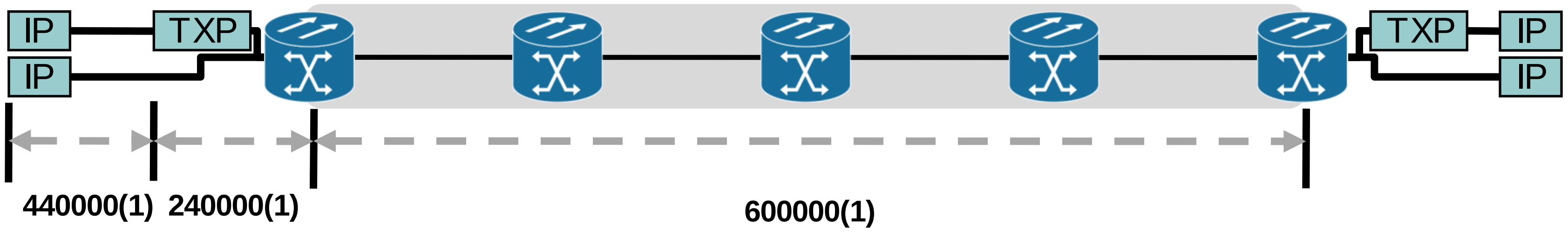
Power consumption traditional vs IPoDWDM

PoP1	*Max. Power (W)*	*Average Power (W)*
DWDM	18,084.1	11,913.73
IPoDWDM	4,250.2	2,813.86
Power Savings (W)	**13,833.9**	**9,099.87**
Energy Cost Savings (€/yr/PoP)	**15,391.32**	**9,894.42**

CO_2

Only Transmission equipment (DWDM+Transponders) shown. Router HW not shown. Except using IPoDWDM router line cards instead of traditional ones, there's no difference on the routers. Energy cost savings based on average price for industry of € 0.1255/kWh (04/2008)

Reliability comparison
Improving MTBF through IPoDWDM



- Transponder Case: yearly unavailability 7,7 minutes
- IPoDWDM case: yearly unavailability 3,3 minutes

Predicted MTBF values

IPoDWDM Solution Elements

IPoDWDM Solution

Interfaces

- Eliminates Transponders from DWDM Systems
- Increase Reliability
- Available on CRS-1, 12k (IOS only) and 7600

Virtual Transponders

- Allow for either integrated or separated management of IPoDWDM
- Solution element to overcome some pushback from transport department

Proactive Protection

- New feature allowing for zero to near-zero packet loss recovery
- Can lead to massive savings in bandwidth when used to rearchitect protection

CapEx/OpEx/Green

- Eliminates DWDM shelves, transponders
- Saves energy on power, cooling
- Saves space
- Lead-in to "Green message"

Foundation for "Intelligent Photonic Layer"

- Ties the IP Layer to DWDM system overhaul
- Can use "Optical Control Plane" to shift discussion on Optical from price to features

IPoDWDM is not just the interfaces

IPoDWDM Interfaces

Interface Availability

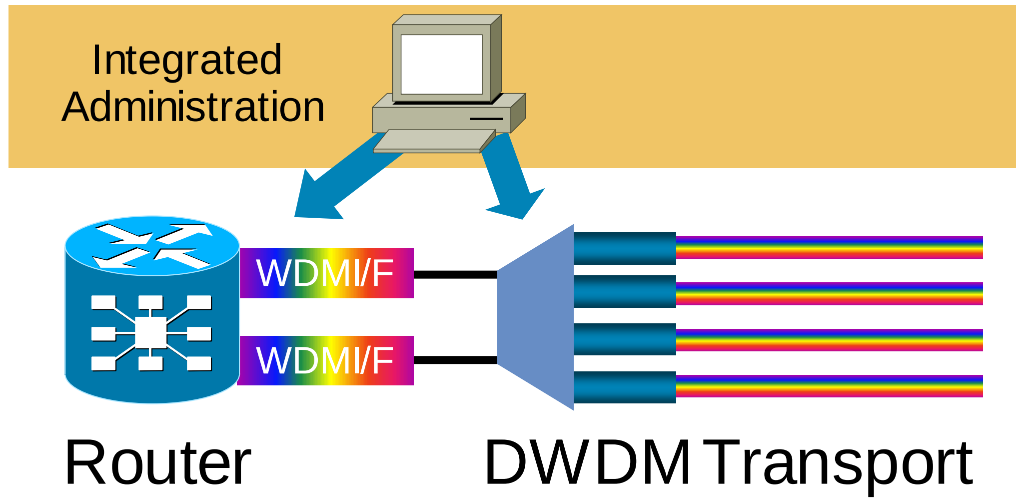
CRS-1	12000 Series	Cisco 7600

• 1-port 40G WDMPHY
  • Two versions available
  • ODB and DPSK+ modulation
• 4-port 10GE WDMPHY
	• 1-port 10GE WDMPHY
	• Currently only supported under IOS, not IOS-XR
		• 2-port 10GE WDMPHY
		• 4- port 10GE WDMPHY
		• Uses pluggable optics
		• Available Q1CY09

Virtual Transponders

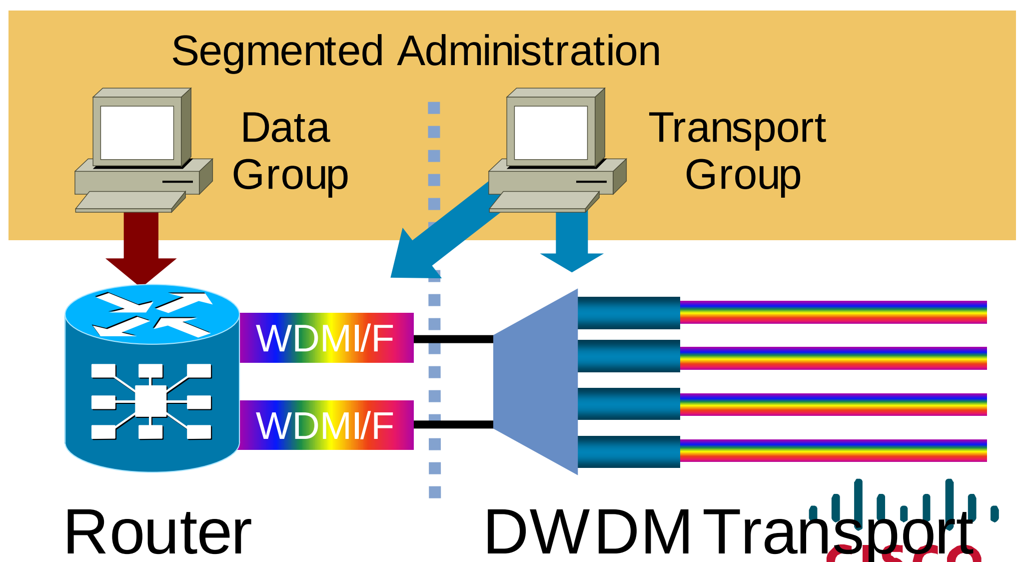
Integrated Administration

- End-to-end provisioning
- Better trouble shooting
- Reduced complexity

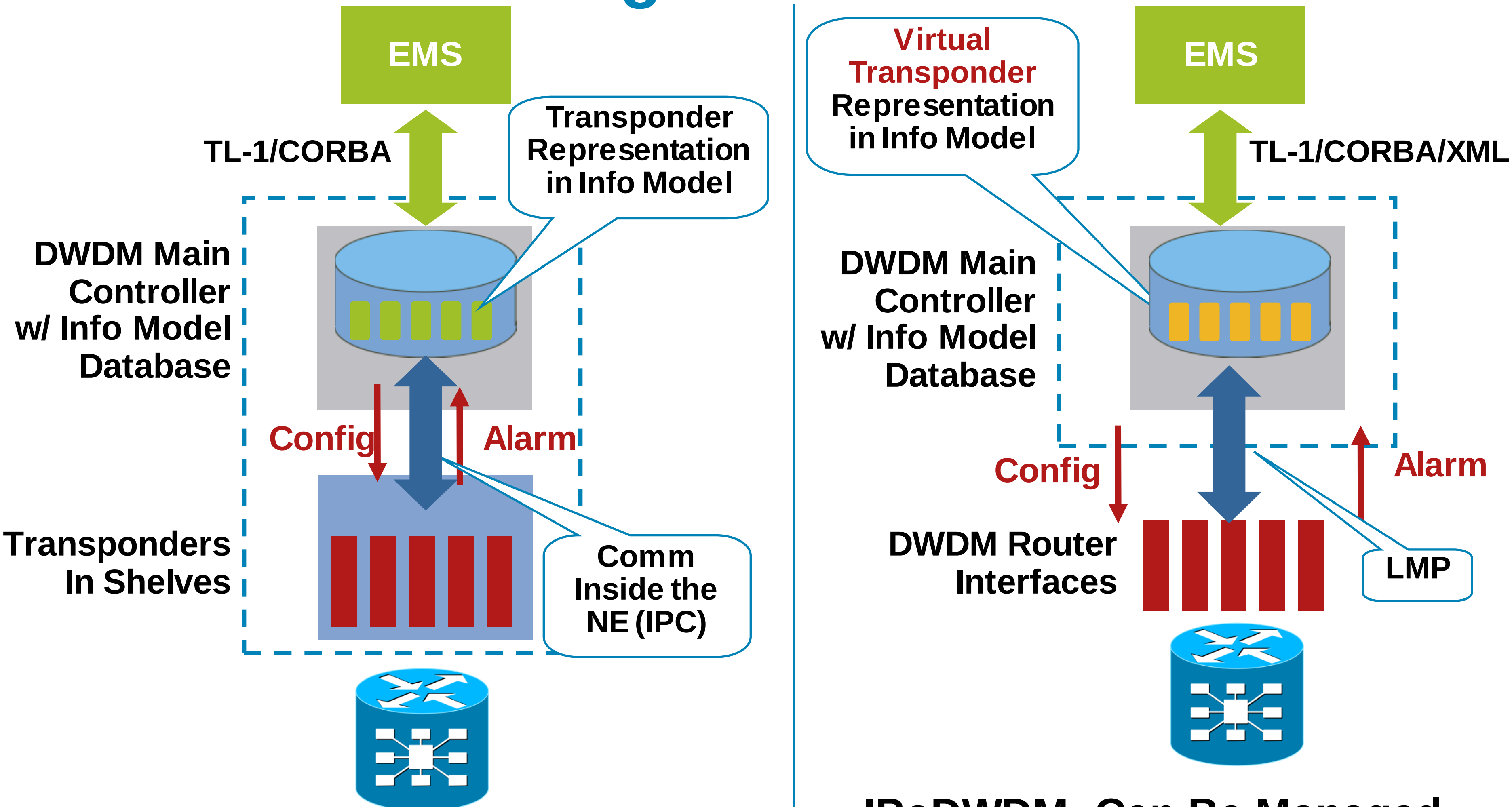


Segmented Administration

- Respect organization boundaries
- Data/transport group separation
- Restrict users through rule-based access control



VTs: Reintroducing a clear demarcation

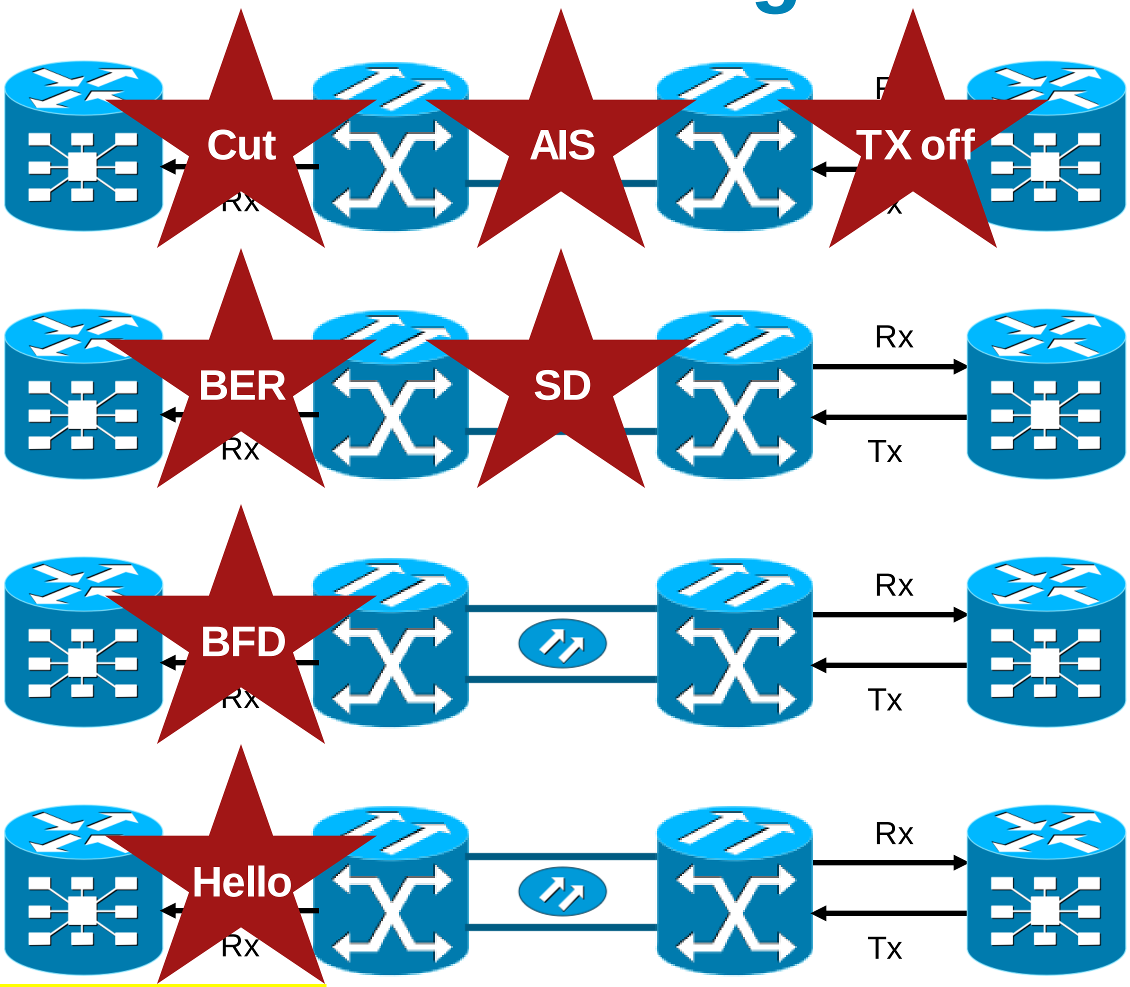


Traditional Network

IPoDWDM: Can Be Managed w/out Significant Changes

Beyond the basics – Proactive Protection

Factors influencing Failure Detection Time



Time to Detect ≈ 0 ms

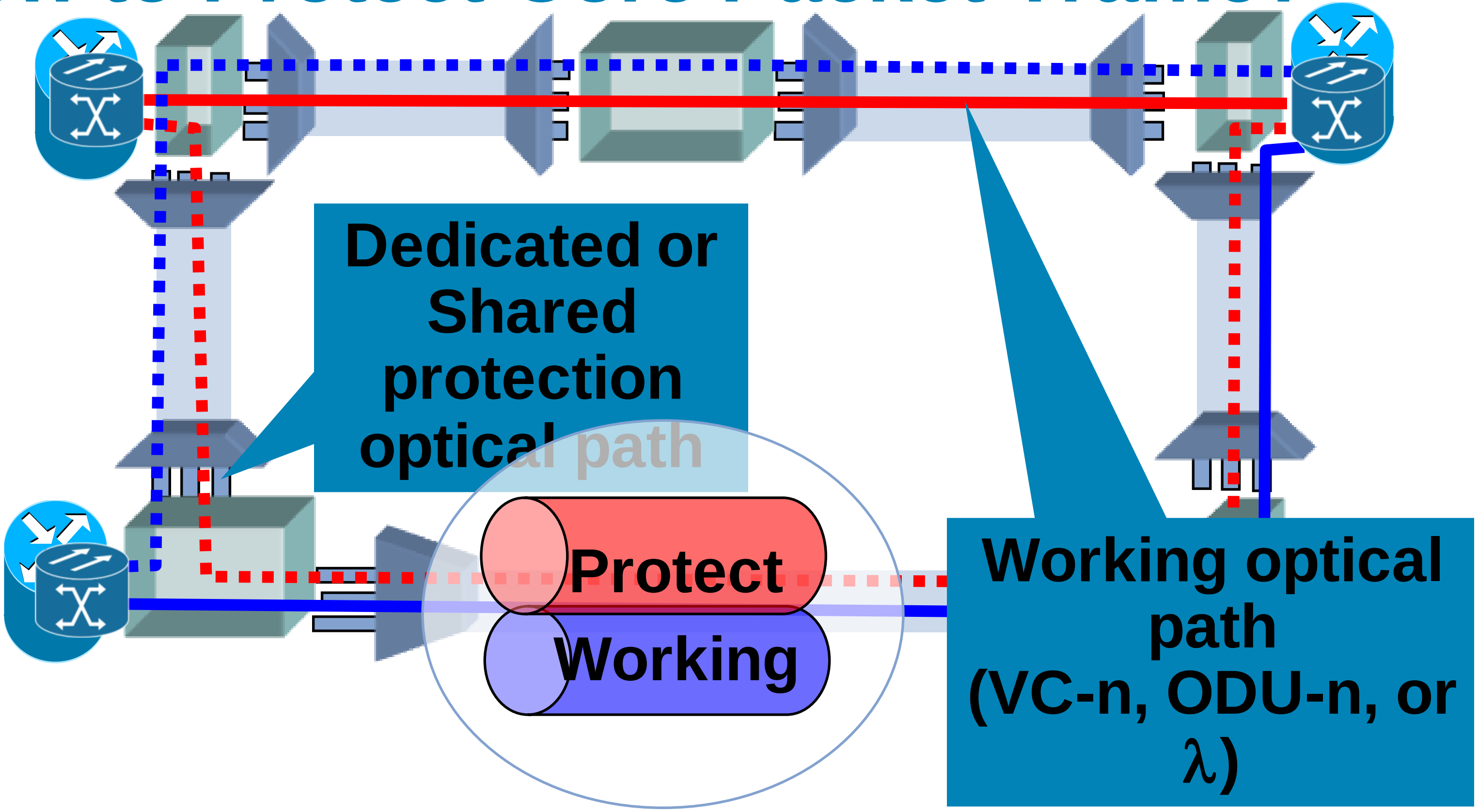
Time to Detect ??? ms

Time to Detect > 15 ms

Time to Detect > 1 s

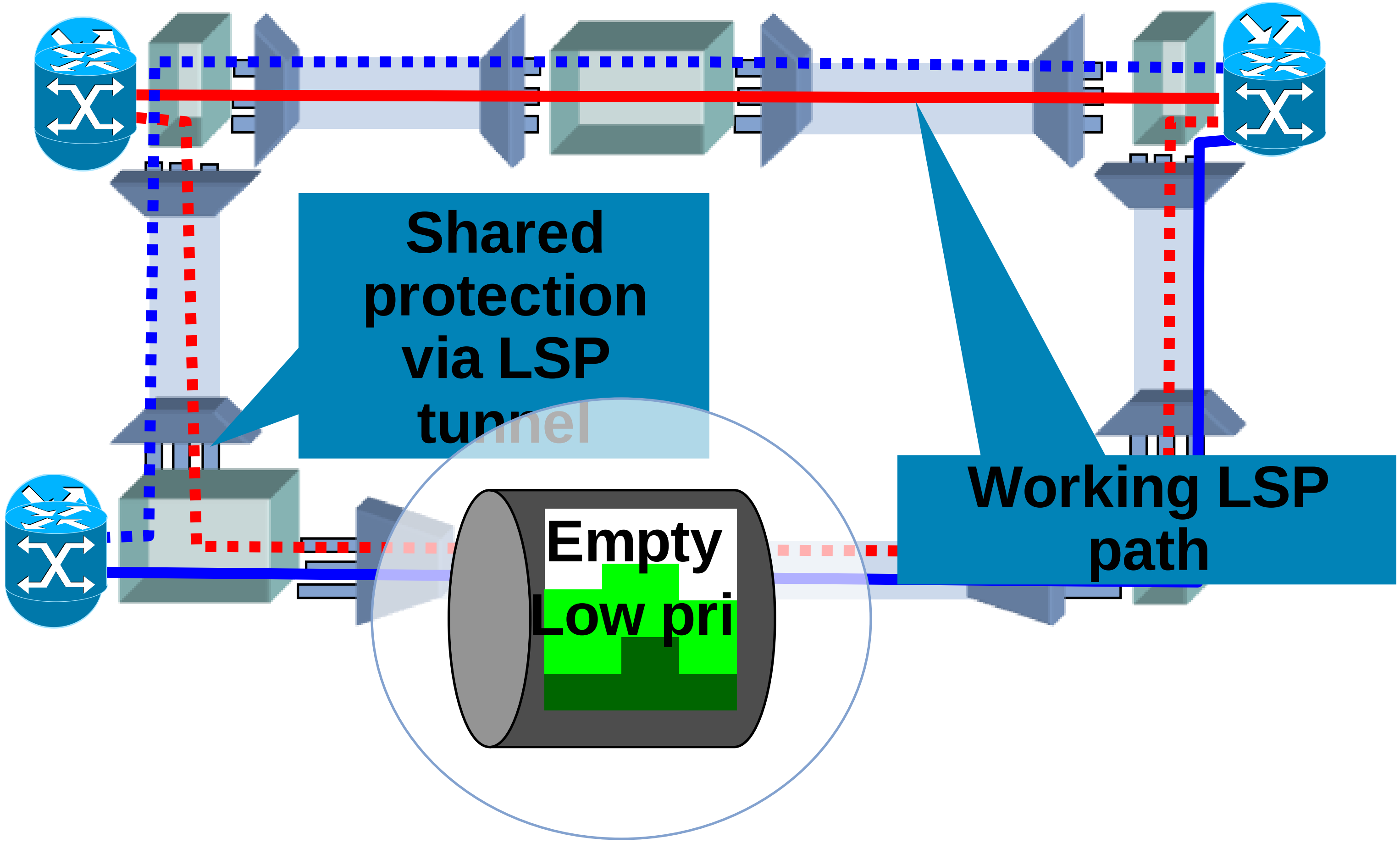
Conclusion: Hard errors are easy to detect, dribble conditions are hard

How to Protect Core Packet Traffic?



Conclusion: Optical layer protection requires additional bandwidth (30% typical in shared mesh architectures) and electrical switches

How to Protect Core Packet Traffic?



Conclusion: Protection at the packet level is as fast w FRR and supports a more cost effective architecture

Advanced Protection
Conclusions from Network Study

Model	10G Channels
IP Load Share	81
IP Shared HP Protection Bandwidth	90
IP Reserved HP Protection Bandwidth	97
TDM Shared Protection	106
TDM 1+1 Protection	111

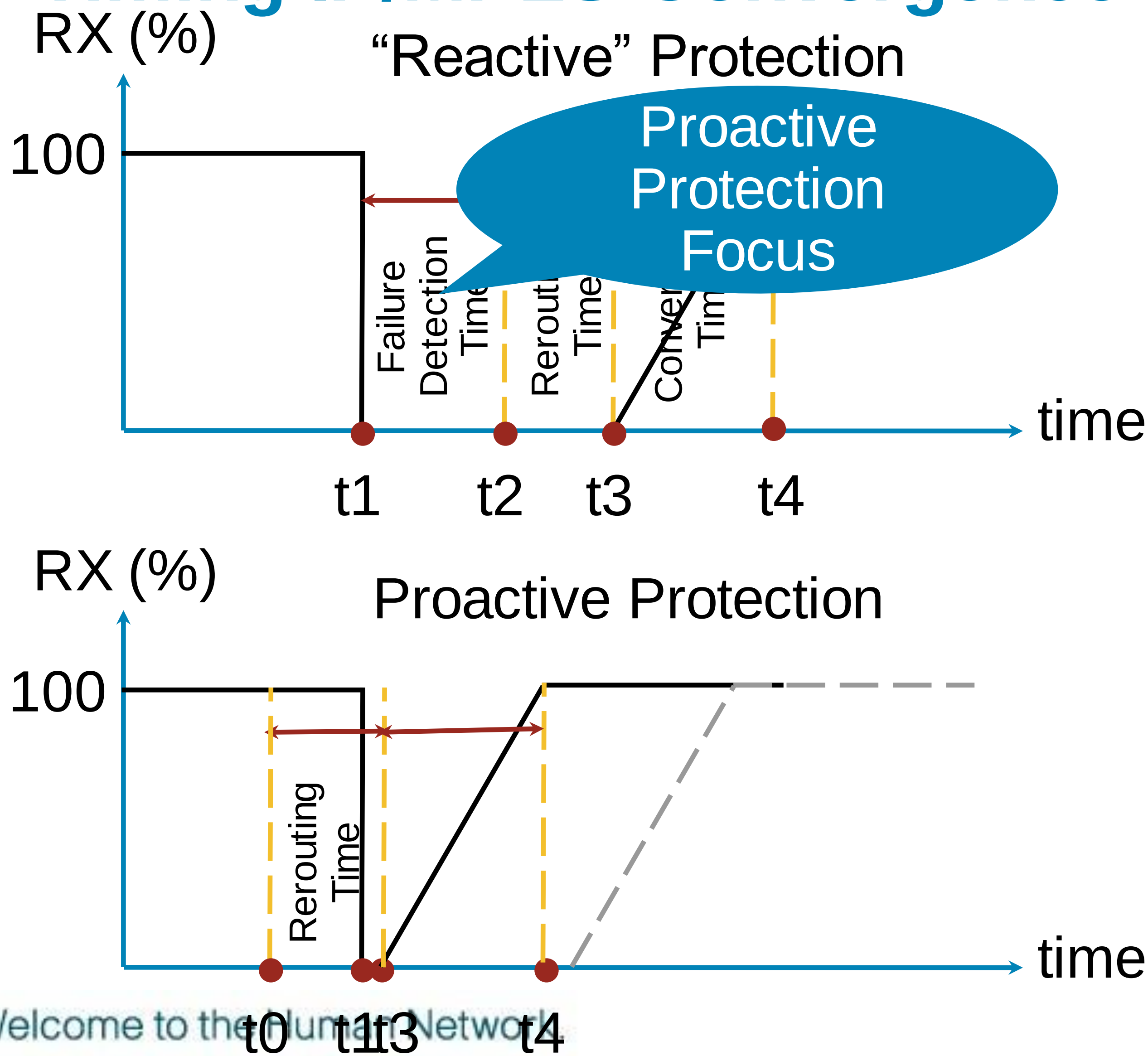
Conclusions:

1. IP/MPLS layer protection is much more bandwidth efficient if low priority traffic is preempted – see 1st row in above table

   Assumption: 30% traffic must be protected, the rest can be restored slowly

2. IP/MPLS protection is still very effective if low priority traffic is unprotected but not preempted – see 2nd and 3rd rows in above table
3. The CAPEX for IP/MPLS protection is within 10% of the most efficient TDM protection when WDMPOS interfaces are used
4. The CAPEX for IP/MPLS protection is 47% of the most efficient TDM shared protection when WDMPHY interfaces are used

Timing IP/MPLS Convergence



Failure Detection Time: Time Between Failure of Link Until Neighbor Is Declared "Down"

Rereouting Time: Time Between „"Neighbor Down" Event and Recalculation of Routes

Converging Time: Time Between Recalculation of Routes Until All Routers in Routing Domain Have the Same Routing Database

IP FRR Test Results

Protection Type	Fault Type	Convergence Time (ms)		
		Highest	Lowest	Average
Proactive	Optical Switch	11,50	11,18	11,37
Proactive	Noise Injection	0,02	0	0,00
Proactive	Fiber Pull (Tx)	25,48	0	12,39
Proactive	PMD-Injection	0,08	0	0,02
Standard	Optical Switch	11,54	11,37	11,43
Standard	Noise Injection	7404	1193	4305
Standard	Fiber Pull (Tx)	25,93	12,50	20,19
Standard	PMD-Injection	129,62	122,51	125,90

ISIS Fast Convergence Test Results

Protection Type	Fault Type	Convergence Time (ms)		
		C(500)	C(1000)	C(1)
Proactive	Optical Switch	170	220	163
Proactive	Slow Noise Injection	3	12	0
Proactive	Fast Noise Injection	3	9	0
Standard	Optical Switch	180	205	159
Standard	Slow Noise Injection	2990	3035	2880
Standard	Fast Noise Injection	596	620	540

MPLS FRR ISIS Test Results

Protection Type	Fault Type	Convergence Time (ms)		
		Highest	Lowest	Average
Proactive	Optical Switch	11,48	10,99	11,24
Proactive	Noise Injection	0,12	0	0,05
Proactive	Fiber Pull	14,95	0	4,95
Standard	Optical Switch	11,61	11,16	11,32
Standard	Noise Injection	2852	2602	2727
Standard	Fiber Pull	83,84	13,49	37,63

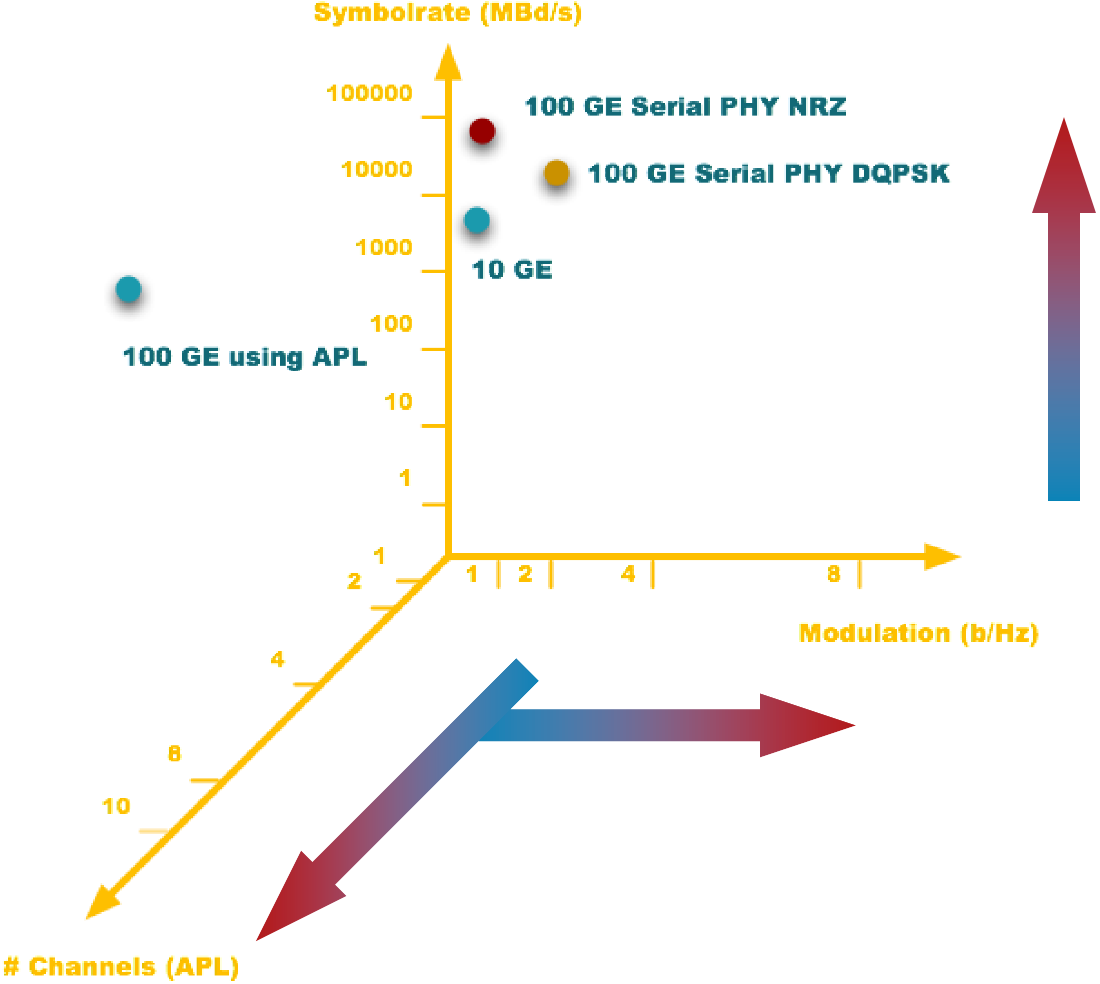
Moving to 40G, 100G and beyond

Welcome to the Human Network.

IEEE 802.3ba Objectives

- **Support full-duplex operation only**
- **Preserve the 802.3 / Ethernet frame format utilizing the 802.3 MAC**
- **Preserve minimum and maximum FrameSizeof current 802.3 standard**
- **Support a BER better than or equal to 10-12 at the MAC/PLS service interface**
- **Provide appropriate support for OTN**
- **Support a MAC data rate of 40 Gb/s**
  - **Provide Physical Layer specifications which support: 40 Gb/s**
    - –at least 100m on OM3 MMF
    - –at least 10m over a copper cable assembly
    - –at least 1m over a backplane
- **Support a MAC data rate of 100 Gb/s**
- **Provide Physical Layer specifications which support:**
    - –at least 40km on SMF
    - –at least 10km on SMF (likely to become 2-4km)
    - –at least 100m on OM3 MMF
    - –at least 10m over a copper cable assembly

40GbE / 100GbE solution space

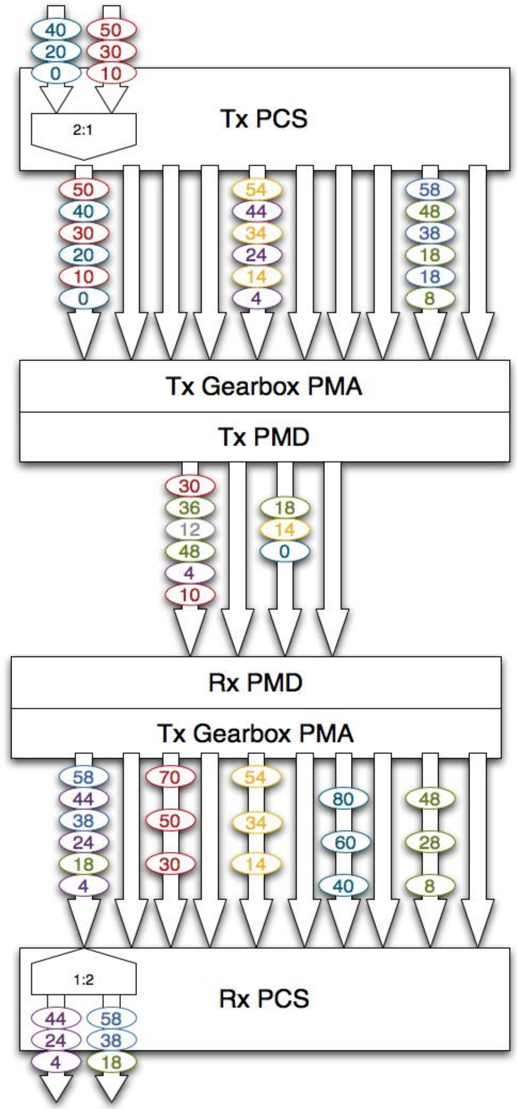


- Allows more implementable symbol rate
- 100 GBd not implementable in current CMOS
  - Well understood technology
  - Can reuse 10GE PMDs
    25 GBd (= 4 x 25 Gb) SiGe or
    Costly alternatives are under discussion for 4 x 2 x serial PHY
  - VCSEL (for MMF) arrays can overcome cost issue
- Optical impairments (PMD, CD, ...) square of symbol rate
- Not yet mature for optical transmission at large baud rate
    10G VCSELs just becoming available
    Reliability concerns for arrays
    In production for 850nm - 1310nm and 1550nm still very early
    May use different modulation type (e.g. DQPSK)
  - Poor spectral efficiency for WAN
- Advantageous for DWDM transmission wrt spectral efficiency
  - Requires some form of descew like EFM
- How to map that into SONET/SDH/OTN?
  - Has the strongest support in HSSG
    PCS mapping changes

Source: Based on IEEE HSSG contribution by Drew Perkins Sept. 2006

Key concept for 40/100GbE: Virtual Lanes

- Split 66B/64B stream into a number of Virtual Lanes
- The 40/100GB are then Bit Muxed over 1/4/10 physical lanes
- Electrical and optical scew will rearrange bits @ Rx but Virtual Lane Bit order is preserved
- MAC data rate preserved
- Alignment protocol reuses some BW from IPG

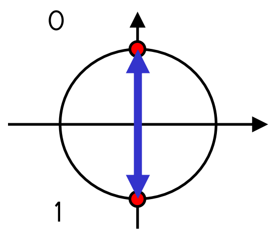


40G available technology

- DPSK

  Good Fiber – ~~Bad Fiber~~ – PMDC - Cost – Space and Power

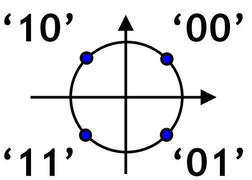
  Cost effective; requires external PMDC



- DQPSK

  Good Fiber – Bad Fiber – ~~PMDC~~ - Cost – Space and Power

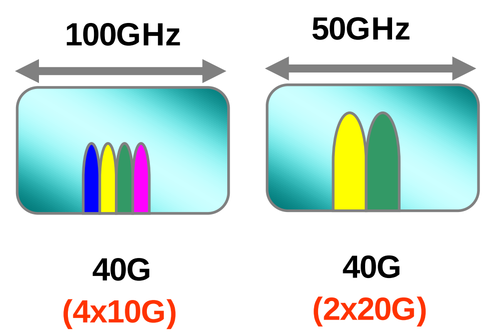
  Cost and Footprint penalty; improved PMD



- DMC (Dense MultiCarrier)

  Good Fiber – Bad Fiber – ~~PMDC~~ - Cost – Space and Power

  Good PMD performance and price/footprint

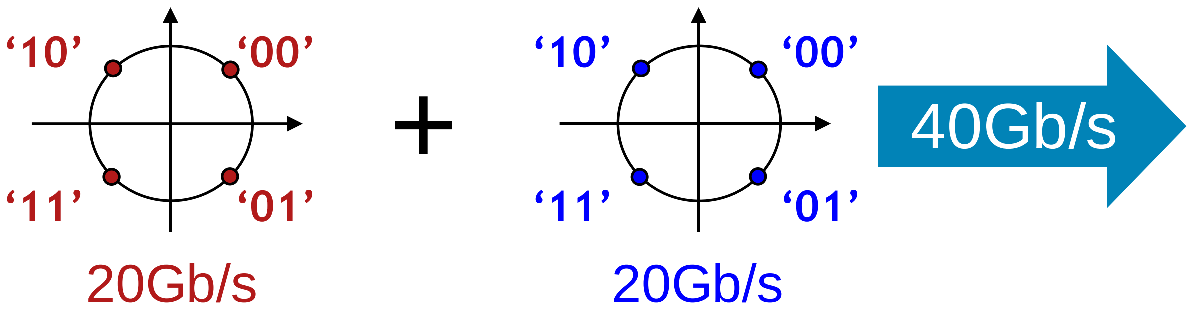


Legend: Black Neutral; Red Negative; Green Positive

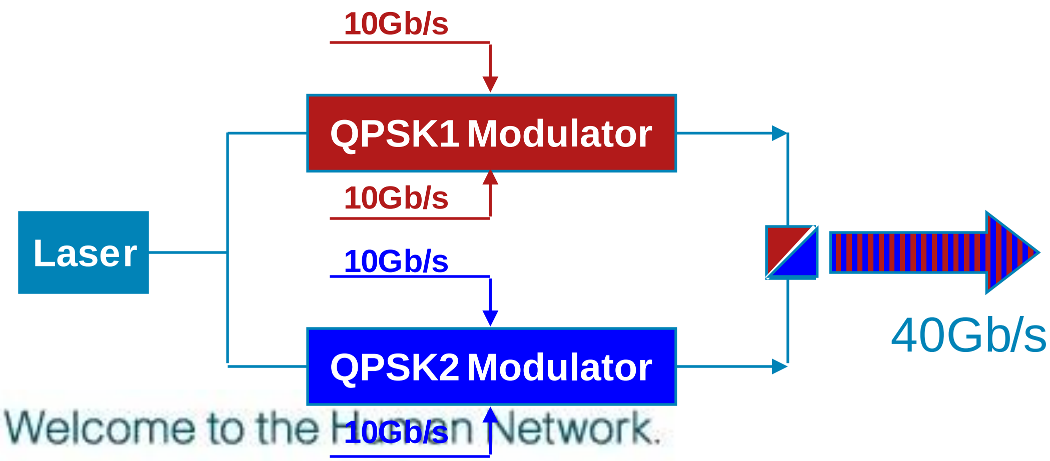
PolMux Solution

Dual Polarization QPSK with Coherent Detection

Transmitter: Two QPSK signals are muxed in polarization



Same Optical Bandwitdh as a 10G NRZ!!



40Gb/s = 10Gbaud

10Gbaud signal propagate into the fiber as a 10Gb/s signal

Technology Comparison

Parameter	10G (NRZ)	DPSK		DQPSK		Multi-Carrier		DP-QPSK
		w/o PMDC	w/PMDC	w/o PMDC	w/PMDC	4x10G (ODB)	2x20G (DQPSK)	
B-to-B OSNR (dB)	5	8.5	8.5	8.5	8.5	13	8.5	4
50GHz Compatibility	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
CD Robustness (ps/nm)		+/-650	+/-650	+/-650	+/-650			+/-4,000
- 1dB of OSNR margin	+/-500					+/-2,000	+/-500	
- 2dB of OSNR margin	+/-800					+/-3,000	+/-800	
PMD Robustness (ps)								> 25
- 1dB of OSNR margin	10	2.5		5		10	10	
- 2dB of OSNR margin	13	3.5	8	8	15	13	13	
Complexity		Low	Low	Medium	Medium	Low	Medium	High
Cost		100%	142%	141%	181%	102%	146%	118%

Welcome to the Human Network.

40GbE Variants

Approaches: Optimized for Area of Implementation

40GbE MAC

Backplane
Very Short Reach

- 4x10G Copper
  - 40GBASE-KR4

Data Center
Short Reach
100m

- 4x10G Copper
  - 40GBASE-CX4
- 4x10G Ribbon
  - 40GBASE-SR4

Campus
Intermediate Reach
10km

- 4x10G
  - 40GBASE-LR4

Long Haul

- ODU3

Current Focus of IEEE HSSG (Backplane, Data Center, Campus)

38.9GbE MAC with 64/66 coding
40.0GbE MAC with 64/66 coding

Expectation of 100Gig WDMPHY

- 100 Gig and above rates must meet minimum requirements:

  Target 10 Gig distances—1600–2000 Km reach

  Not simply a Greenfield technology, but plug and play over existing 10Gig networks

  Must be as open as possible, operate over third party DWDM networks

  Must operate over both 100GHz as well as 50GHz spacings

  Must be at a competitive cost point

  Power and footprint must be reasonable, can not redesign Router/ transport shelf due to blade

- To achieve must leverage/control:

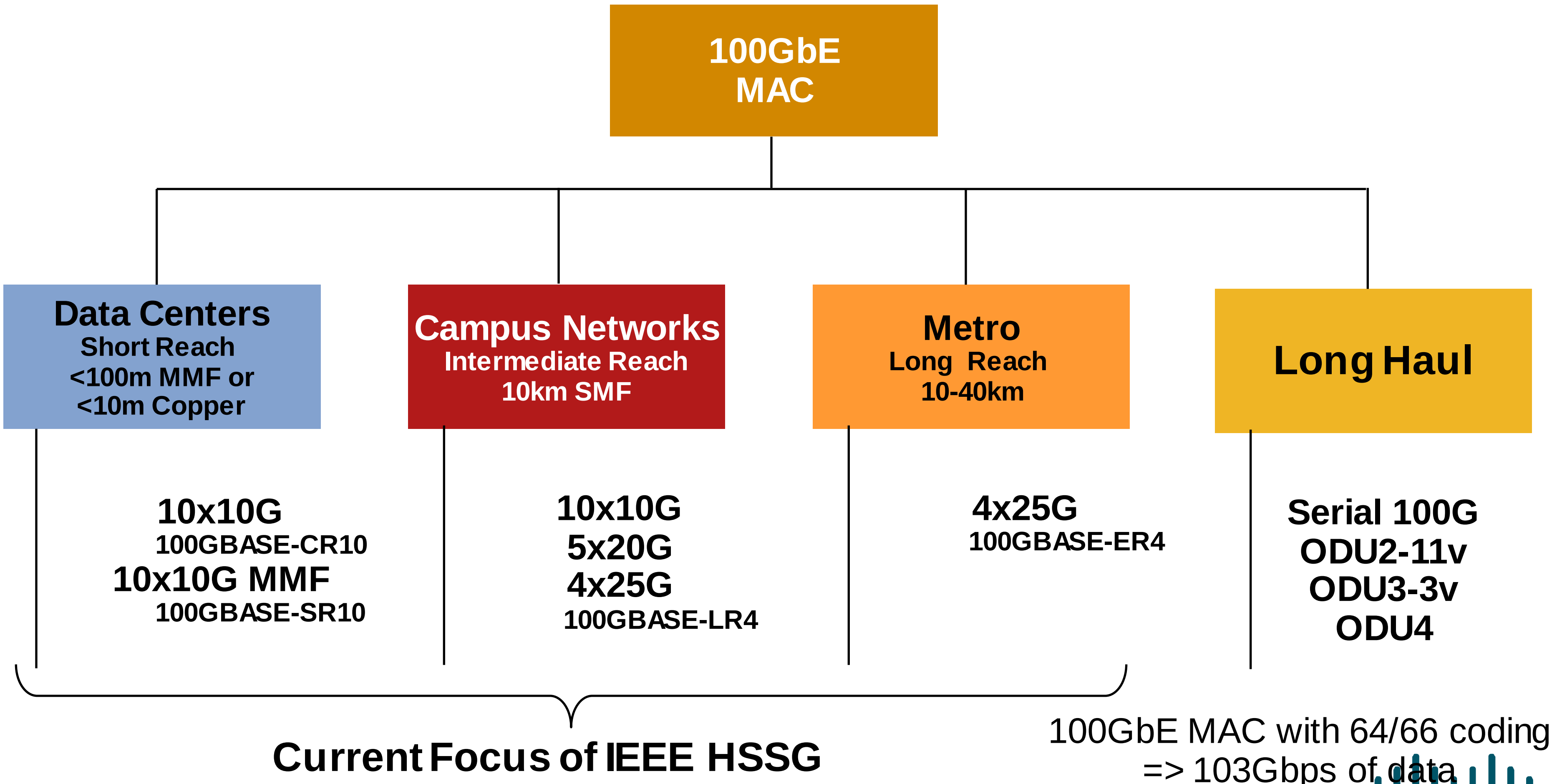
  1. Optical Impairments

  2. Modulations Schemes

Same as 40Gig Requirements!!

100GbE Variants

Approaches: Optimized for Area of Implementation



100GE and the ITU

- ITU agreed to optimize new OTU4 rate for 100GE transport
- Industry appears to have learnt its lesson from 10GE (i.e. LAN/WAN PHY)
  - Ideally want one mapping, one FEC and one modulation scheme !!!
  - some resistance from traditional transport vendors
- Transmission requirements are the same as 40G:
  - Must work over existing, installed DWDM common equipment;
  - Support 50GHz DWDM channel spacing;
  - Full band tunable lasers on 50GHz ITU grid
  - Optical reach of ≥ 1,500km;
  - Chromatic dispersion tolerance of >±800ps/nm;
  - Polarization Mode Dispersion tolerance of > 10ps (mean DGD);
  - Must be “plug and play” and installable by existing field technicians
- 40G was the first commercial application for advanced modulation:
  - 100G will build on similar approaches

100G Development Efforts

- Demonstrated first single-wavelength 100G over production network in June 2008
- Leading standardization effort within OIF (eco-system)
  - Strong supplier base and supply continuity (multiple vendor soln)
  - Compatibility among system vendors
  - Share development costs and reduce risks
- Structure/architecture of the solution
  - Separate cards for client and trunk
  - Regenerator uses 2 trunk cards
- Target Data rates / standards supported
  - OTU4 trunk as per ITU-T G.709 standard
  - Multiplexing of 10x10G (10GE, OC192, FC-1200, OTU2)
  - Multiplexing of 2x40G (OC768, 40GE, OTU3) plus 2x10G (10GE, OC192, FC-1200, OTU2)
  - 100GE Native

CISCO