IPoDWDM, 40G, 100G and beyond

Dirk Schroetter, Consulting Systems Engineer
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

**Power consumption traditional vs IPoDWDM**

<table>
<thead>
<tr>
<th>PoP1</th>
<th>Max. Power (W)</th>
<th>Average Power (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWDM</td>
<td>18,084.1</td>
<td>11,913.73</td>
</tr>
<tr>
<td>IPoDWDM</td>
<td>4,250.2</td>
<td>2,813.86</td>
</tr>
<tr>
<td>Power Savings (W)</td>
<td>13,833.9</td>
<td>9,099.87</td>
</tr>
<tr>
<td>Energy Cost Savings (€/yr/PoP)</td>
<td>15,391.32</td>
<td>9,894.42</td>
</tr>
</tbody>
</table>

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**

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Virtual Transponders</th>
<th>Proactive Protection</th>
<th>CapEx/OpEx/Green</th>
<th>Foundation for “Intelligent Photonic Layer”</th>
</tr>
</thead>
</table>
| • Eliminates Transponders from DWDM Systems  
• Increase Reliability  
• Available on CRS-1, 12k (IOS only) and 7600 | • Allow for either integrated or separated management of IPoDWDM  
• Solution element to overcome some pushback from transport department | • New feature allowing for zero to near-zero packet loss recovery  
• Can lead to massive savings in bandwidth when used to rearchitect protection | • Eliminates DWDM shelves, transponders  
• Saves energy on power, cooling  
• Saves space  
• Lead-in to “Green message” | • 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
## Interface Availability

<table>
<thead>
<tr>
<th>CRS-1</th>
<th>12000 Series</th>
<th>Cisco 7600</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 1-port 40G WDMPHY</td>
<td>• 1-port 10GE WDMPHY</td>
<td>• 2-port 10GE WDMPHY</td>
</tr>
<tr>
<td>• Two versions available</td>
<td>• Currently only supported under IOS, not IOS-XR</td>
<td>• 4-port 10GE WDMPHY</td>
</tr>
<tr>
<td>• ODB and DPSK+ modulation</td>
<td></td>
<td>• Uses pluggable optics</td>
</tr>
<tr>
<td>• 4-port 10GE WDMPHY</td>
<td></td>
<td>• Available Q1CY09</td>
</tr>
</tbody>
</table>

*Uses pluggable optics*
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

EMS

TL-1/CORBA

Transponder Representation in Info Model

EMS

TL-1/CORBA/XML

Virtual Transponder Representation in Info Model

DWDM Main Controller w/ Info Model Database

Config

Alarm

Comm Inside the NE (IPC)

DWDM Main Controller w/ Info Model Database

Config

Alarm

LMP

IPoDWDM: Can Be Managed w/out Significant Changes

Transponders In Shelves

Welcome to the Human Network.
Beyond the basics – Proactive Protection
Factors influencing Failure Detection Time

- **Cut**: Time to Detect ≈ 0 ms
- **AIS**: Time to Detect ??? ms
- **TX off**: Time to Detect > 15 ms
- **BER**: Time to Detect > 1 s
- **SD**: Time to Detect > 1 s
- **BFD**: Time to Detect > 15 ms
- **Hello**: Time to Detect > 1 s

Conclusion: Hard errors are easy to detect, dribble conditions are hard.
How to Protect Core Packet Traffic?

Dedicated or Shared protection optical path

Working optical path (VC-n, ODU-n, or λ)

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

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".

Rerouting 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.

“Reactive” Protection

Proactive Protection Focus

Proactive Protection
## IP FRR Test Results

<table>
<thead>
<tr>
<th>Protection Type</th>
<th>Fault Type</th>
<th>Convergence Time (ms)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Highest</td>
<td>Lowest</td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Optical Switch</td>
<td>11,50</td>
<td>11,18</td>
<td>11,37</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Noise Injection</td>
<td>0,02</td>
<td>0</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Fiber Pull (Tx)</td>
<td>25,48</td>
<td>0</td>
<td>12,39</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>PMD-Injection</td>
<td>0,08</td>
<td>0</td>
<td>0,02</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Optical Switch</td>
<td>11,54</td>
<td>11,37</td>
<td>11,43</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Noise Injection</td>
<td>7404</td>
<td>1193</td>
<td>4305</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Fiber Pull (Tx)</td>
<td>25,93</td>
<td>12,50</td>
<td>20,19</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>PMD-Injection</td>
<td>129,62</td>
<td>122,51</td>
<td>125,90</td>
<td></td>
</tr>
</tbody>
</table>
## ISIS Fast Convergence Test Results

<table>
<thead>
<tr>
<th>Protection Type</th>
<th>Fault Type</th>
<th>Convergence Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>C(500)</td>
</tr>
<tr>
<td>Proactive</td>
<td>Optical Switch</td>
<td>170</td>
</tr>
<tr>
<td>Proactive</td>
<td>Slow Noise Injection</td>
<td>3</td>
</tr>
<tr>
<td>Proactive</td>
<td>Fast Noise Injection</td>
<td>3</td>
</tr>
<tr>
<td>Standard</td>
<td>Optical Switch</td>
<td>180</td>
</tr>
<tr>
<td>Standard</td>
<td>Slow Noise Injection</td>
<td>2990</td>
</tr>
<tr>
<td>Standard</td>
<td>Fast Noise Injection</td>
<td>596</td>
</tr>
</tbody>
</table>
## MPLS FRR ISIS Test Results

<table>
<thead>
<tr>
<th>Protection Type</th>
<th>Fault Type</th>
<th>Convergence Time (ms)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Highest</td>
<td>Lowest</td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Optical Switch</td>
<td>11,48</td>
<td>10,99</td>
<td>11,24</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Noise Injection</td>
<td>0,12</td>
<td>0</td>
<td>0,05</td>
<td></td>
</tr>
<tr>
<td>Proactive</td>
<td>Fiber Pull</td>
<td>14,95</td>
<td>0</td>
<td>4,95</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Optical Switch</td>
<td>11,61</td>
<td>11,16</td>
<td>11,32</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Noise Injection</td>
<td>2852</td>
<td>2602</td>
<td>2727</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Fiber Pull</td>
<td>83,84</td>
<td>13,49</td>
<td>37,63</td>
<td></td>
</tr>
</tbody>
</table>
Moving to 40G, 100G and beyond
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 FrameSize of 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

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

- 100GE not implementable in current CMOS
- Can reuse 10GE PMDs
- Costly alternatives are SiGe or InP
- Optical impairments (PMD, CD, ...) ~ square of symbol rate
- May use different modulation type (e.g. DQPSK)
- Advantageous for DWDM transmission wrt spectral efficiency
- Has the strongest support in HSSG
- Poor spectral efficiency for WAN
- Requires some form of descwel like EFM
- How to map that into SONET/SDH/OTN?
- PCS mapping changes

© 2008 Cisco Systems, Inc. All rights reserved. Cisco Confidential
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

20Gb/s

\begin{align*}
'10' & \quad '00' \\
'11' & \quad '01'
\end{align*}

+ 

\begin{align*}
'10' & \quad '00' \\
'11' & \quad '01'
\end{align*}

40Gb/s

Same Optical Bandwidth as a 10G NRZ!!

10Gb/s signal propagate into the fiber as a 10Gb/s signal
## Technology Comparison

<table>
<thead>
<tr>
<th>Parameter</th>
<th>10G (NRZ)</th>
<th>DPSK</th>
<th>DQPSK</th>
<th>Multi-Carrier</th>
<th>DP-QPSK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o PMDC</td>
<td>w/PMDC</td>
<td>w/o PMDC</td>
<td>w/PMDC</td>
<td>4x10G (ODB)</td>
</tr>
<tr>
<td>B-to-B OSNR (dB)</td>
<td>5</td>
<td>8.5</td>
<td>8.5</td>
<td>8.5</td>
<td>13</td>
</tr>
<tr>
<td>50GHz Compatibility</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>CD Robustness (ps/nm)</td>
<td>+/-650</td>
<td>+/-650</td>
<td>+/-650</td>
<td>+/-650</td>
<td>+/-2,000</td>
</tr>
<tr>
<td>- 1dB of OSNR margin</td>
<td>+/-500</td>
<td>+/-800</td>
<td>+/-500</td>
<td>+/-800</td>
<td>+/-3,000</td>
</tr>
<tr>
<td>- 2dB of OSNR margin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMD Robustness (ps)</td>
<td>10</td>
<td>2.5</td>
<td>5</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>- 1dB of OSNR margin</td>
<td>13</td>
<td>3.5</td>
<td>8</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>- 2dB of OSNR margin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Cost</td>
<td>100%</td>
<td>142%</td>
<td>141%</td>
<td>181%</td>
<td>102%</td>
</tr>
</tbody>
</table>
40GbE Variants

Approaches: Optimized for Area of Implementation

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

Current Focus of IEEE HSSG
Expectation of 100Gig WDM PHY

- 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

100GbE MAC

Data Centers
- Short Reach: <100m MMF or <10m Copper
  - 10x10G 100GBASE-CR10
  - 10x10G MMF 100GBASE-SR10

Campus Networks
- Intermediate Reach: 10km SMF
  - 10x10G
  - 5x20G
  - 4x25G
  - 100GBASE-LR4

Metro
- Long Reach: 10-40km
  - 4x25G 100GBASE-ER4

Long Haul
- Serial 100G ODU2-11v ODU3-3v ODU4

Current Focus of IEEE HSSG
- 100GbE MAC with 64/66 coding => 103Gbps of data

Welcome to the Human Network.
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