LAMBDA™ OpticalSystems

Presentation Version or Other Info
Thursday, May 25, 2006

Dr. Abdella Battou, CTO

The Lambda Grid – Mass Storage Systems over a Dynamic Optical Network

NASA/IEEE MSST2006

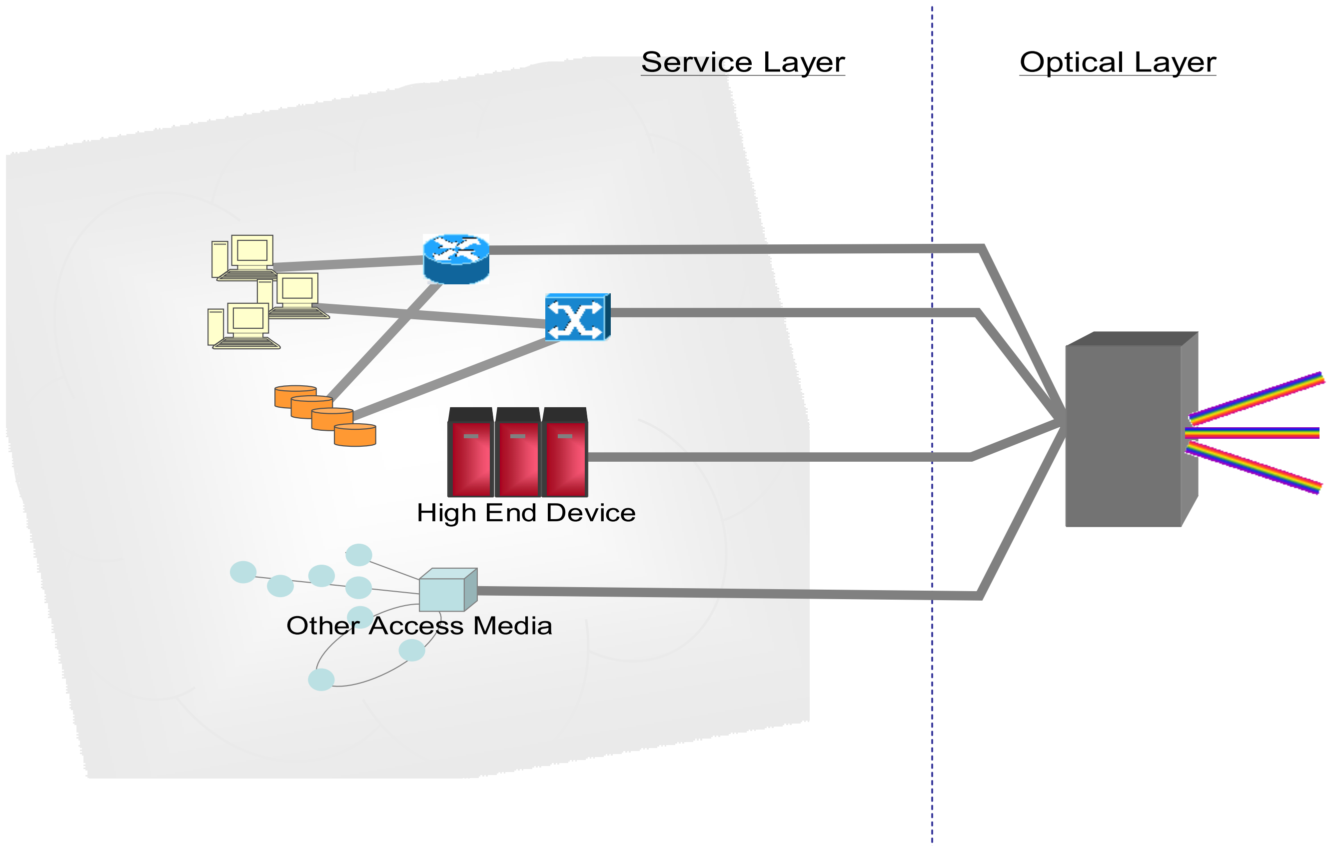
Why Storage over WAN?

- Consolidated backups and archives
- Disk mirroring, backups to disaster recovery sites
- High availability mission critical databases
- Distributed (logical) server clustering
- Disk virtualisation

Requirements - Performance

- What determines Storage performance ?
  - Available bandwidth
  - Latency
- What are the Wide Area Storage connection requirements ?
  - Guaranteed bandwidth
  - Guaranteed low delay
  - Guaranteed low error rate

2 Layer Network



Lambda Grid

Realize a metro network that can set up connections between any points, at any data rate, independent of format, temporarily or permanently.

Switched Transport Network
Dynamic Transport Network

Major attributes of this solution include

- Dynamic connectivity with high bandwidth and low latency using GMPLS to enable fast service discovery and allocation
- Ability to redistribute bandwidth statically or dynamically as new computing or storage element comes on-line or more bandwidth is required for an immediate large transfer.
- Low latency. There is no queuing in the path and minimum latency is guaranteed across the Lambda Grid .
- No congestion. The Lambda Grid uses either static dedicated wavelengths between servers and disk arrays or on-demand wavelengths to satisfy irregular large transfers.
- Wavelength services isolate traffic and provide immunity against congestions. Large clusters for example can grab wavelengths on-demand to satisfy huge transfers without affecting regular daily jobs.
- Last but not least, consolidation of all services over a single elegant, cost-effective, and scalable optical infrastructure.

What are the challenges ?

- Cost of the network – optical technology is available
- Today most of the transport is un-switched
- Try switching for a change - innovative switching architectures
  - Move around traffic at ease
  - Reconfigure network base on traffic demand
  - Turn up new bandwidth quickly
- Need Distributed Control Plane and a Service-Oriented Network Management System

DWDM is the technology of choice

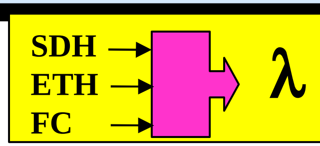
Performance

- Increase distance: EFEC, 2R, Dispersion, OA
- Increase bit rate : 40G and higher
- Increase number of wavelengths

High Capacity & Transparency

Integration

- **Switching + Transport**
- Multiservice, Multiprotocol (TDM + Packet)
- Tunables & Pluggables (XFP, SFP)



Automation

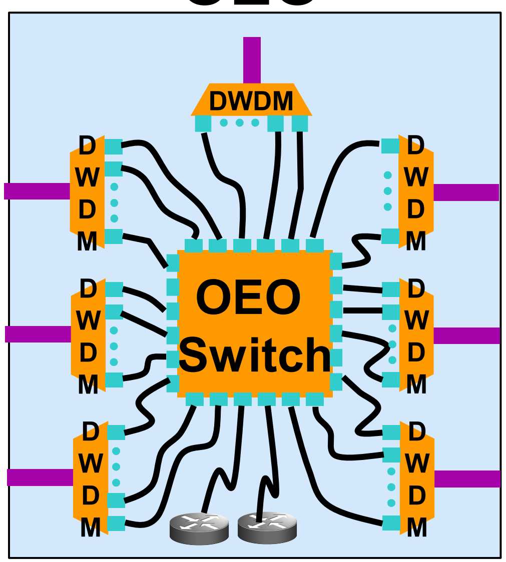
- Dynamic provisioning
- Fast installation
- Reduce operator errors

GMPLS & OIF U

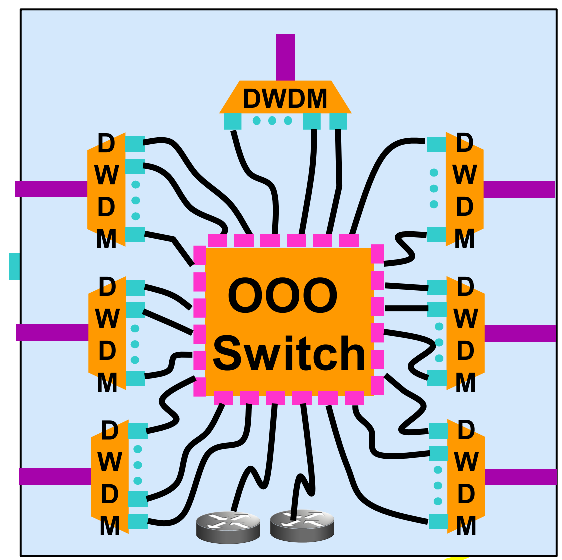
Integrated DWDM and Optical switching

An all-optical switch with integrated DWDM, under a GMPLS control plane, provides the highest levels of simplification and cost savings.

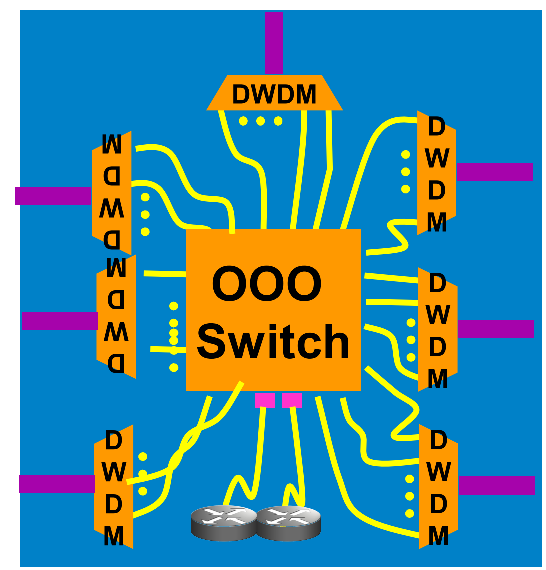
OEO



OXC + external DWDM



OXC+internal DWDM

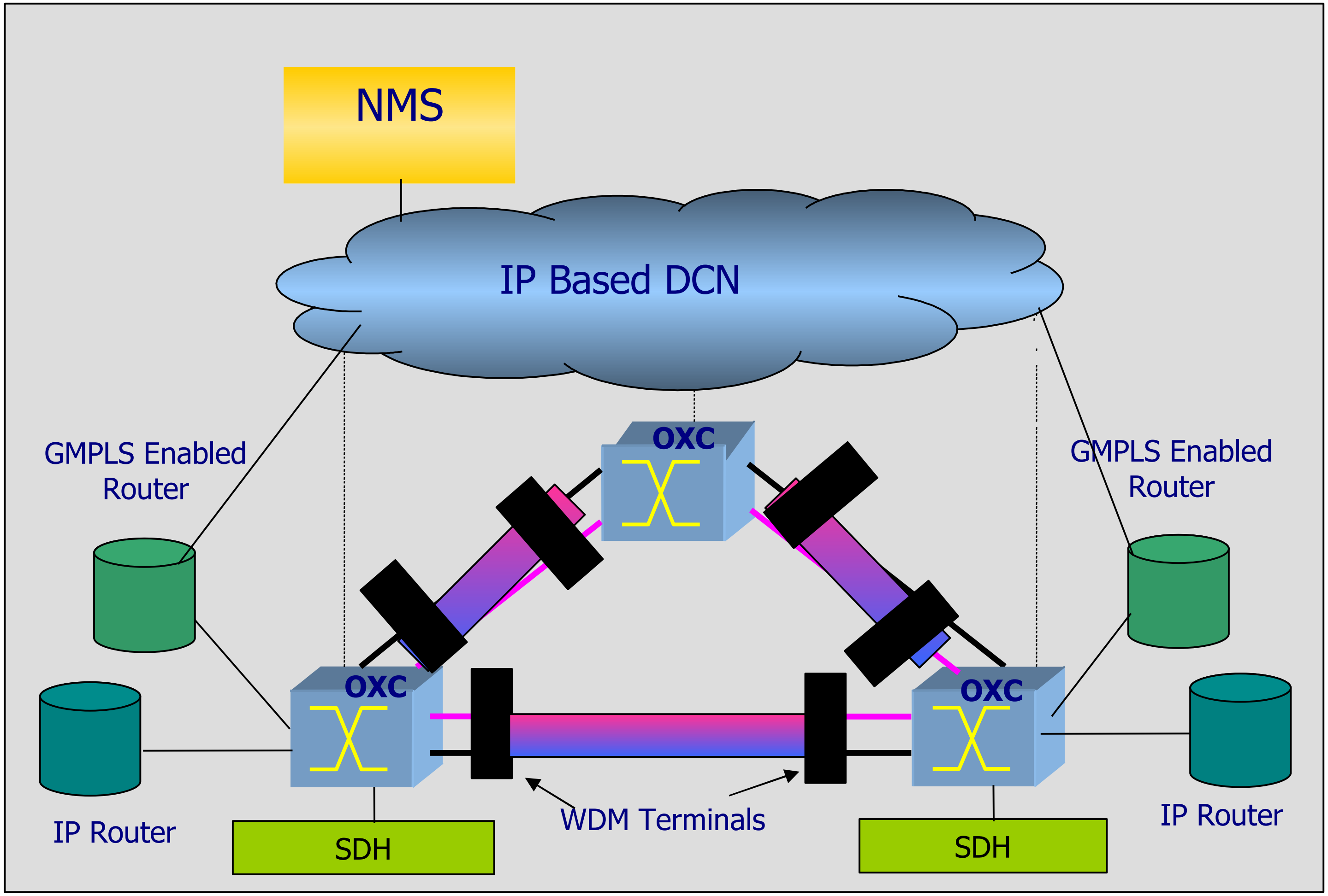


=$ 10G, 40G

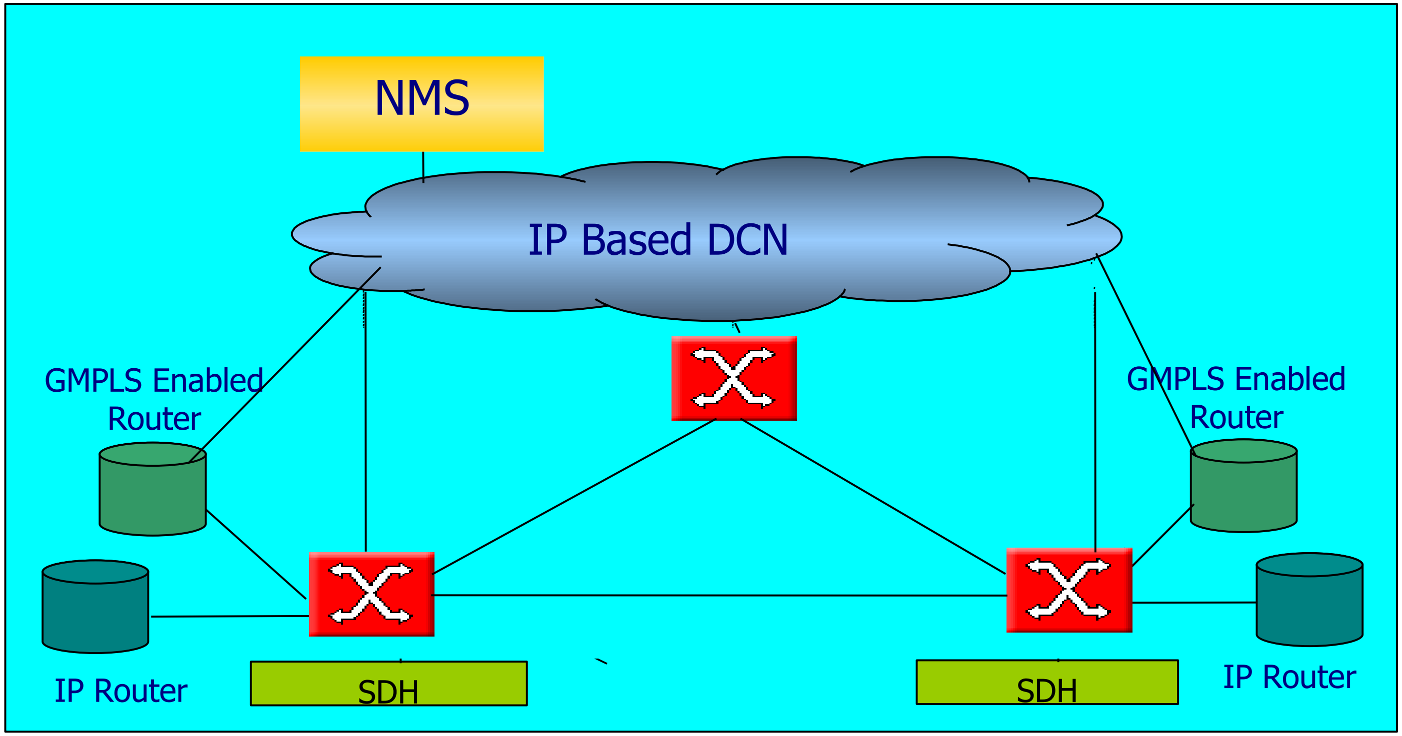
OEO tax.

$$$ OPEX.

OXC approach



OXC + DWDM approach



Wavelength Switching Scalability

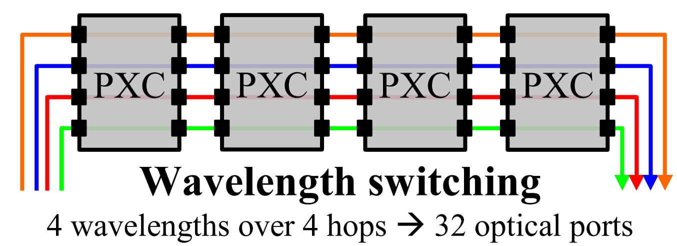
- Grid-scale applications will ultimately press even wavelength switching – Example:

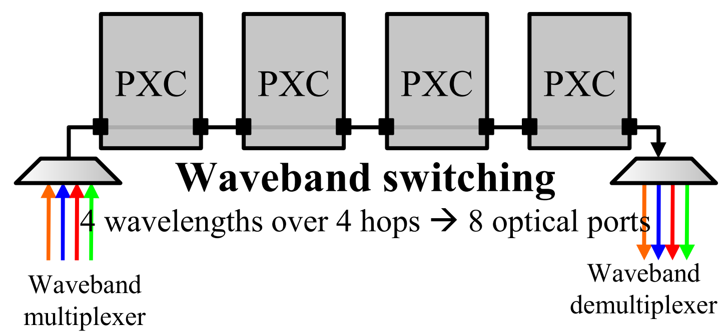
Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10	DWDM; 1 + 10 GigE Integration
2005	10	2-4 X 10	λ Switch; λ Provisioning
2007	2-4 X 10	~10 X 10; 40 Gbps	1st Gen. λ Grids
2009	~10 X 10 or 1-2 X 40	~5 X 40 or ~20-50 X 10	40 Gbps λ Switching
2011	~5 X 40 or ~20 X 10	~25 X 40 or ~100 X 10	2nd Gen λ Grids Terabit Networks
2013	~Terabit	~MultiTbps	~Fill One Fiber

Require too many optical ports to provide non-blocking connectivity! (2003–2011)

Source: Larry Smarr, "The Optiputer - Toward a Terabit LAN ," The On*VECTOR Terabit LAN Workshop Hosted by Calit2, University of California, San Diego - January 2005

- Similar to any other switching technology, aggregation is essential for scalability of wavelength switching
- Emergence of transparent multigranular (wavelength and waveband) switching architectures





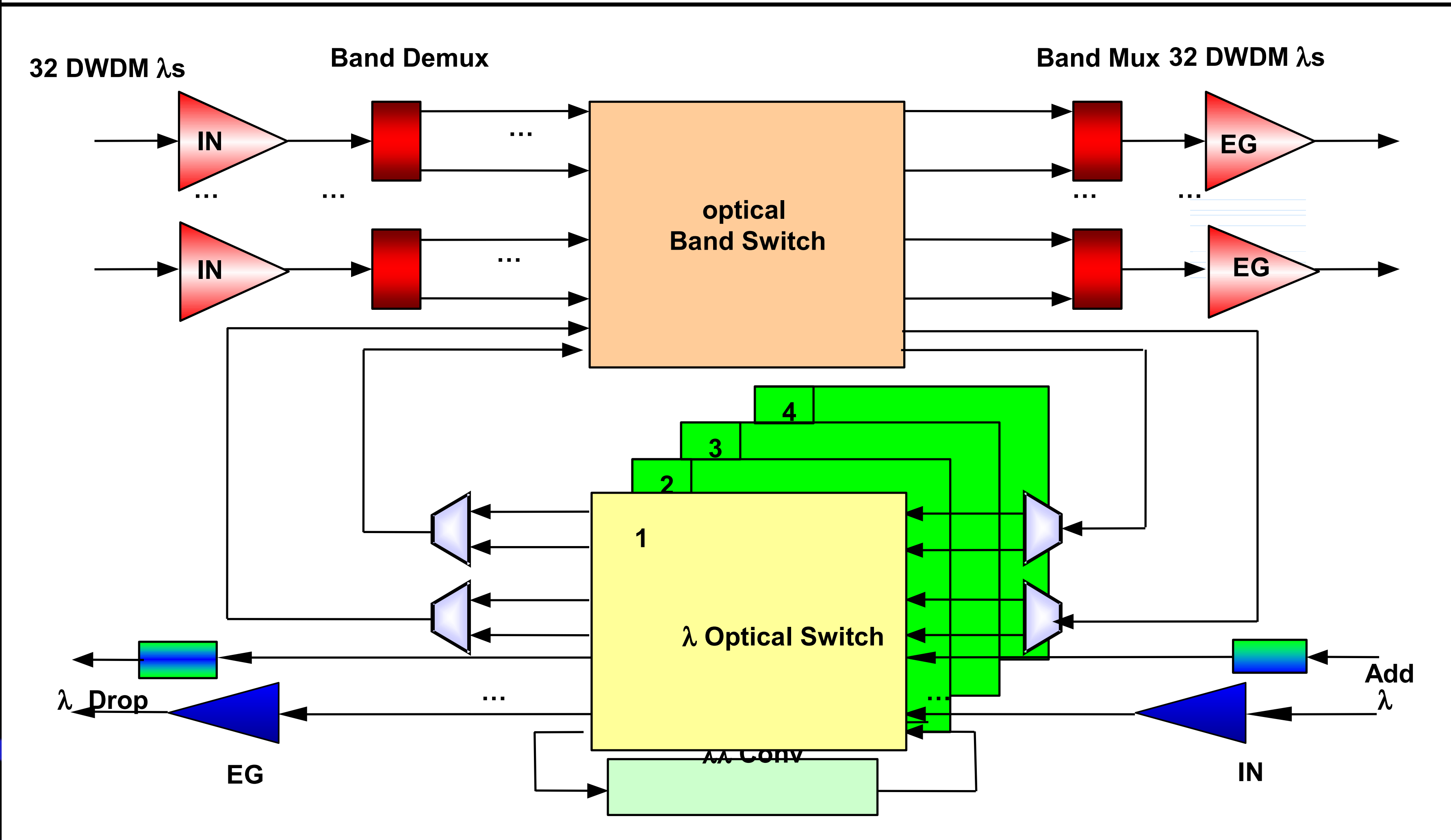
What happens when traffic grows ?

- Growth in DWDM traffic → large number of wavelengths → large photonic fabrics.
- Larger OXC →higher cost and complexity → unproven reliability → hinder deployment.

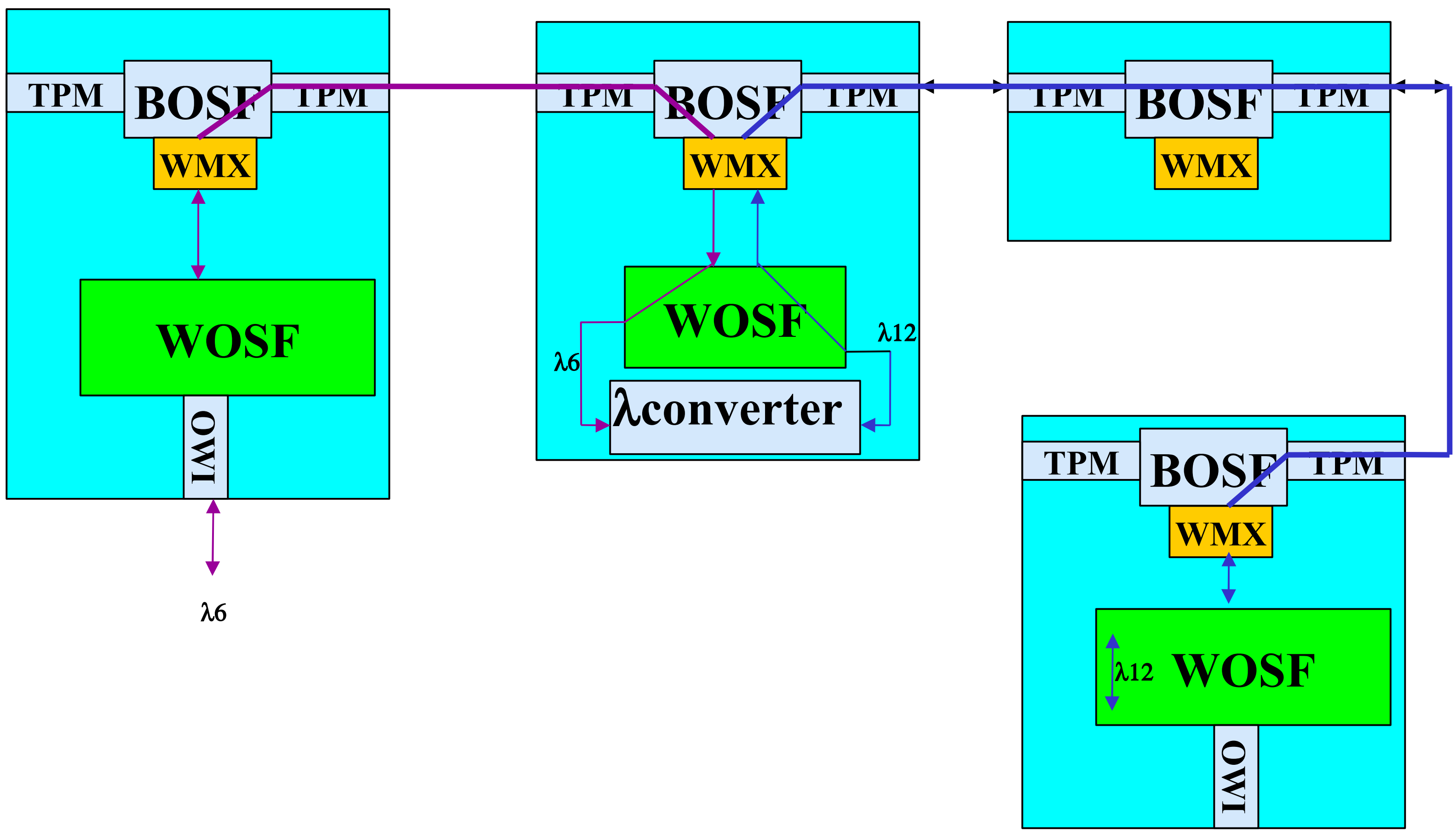
Use Wavebands

- Several wavelengths are switched as a band using a single port
- Port reduction → smaller fabrics → more scalable & less power consumption

Waveband and Wavelength switching

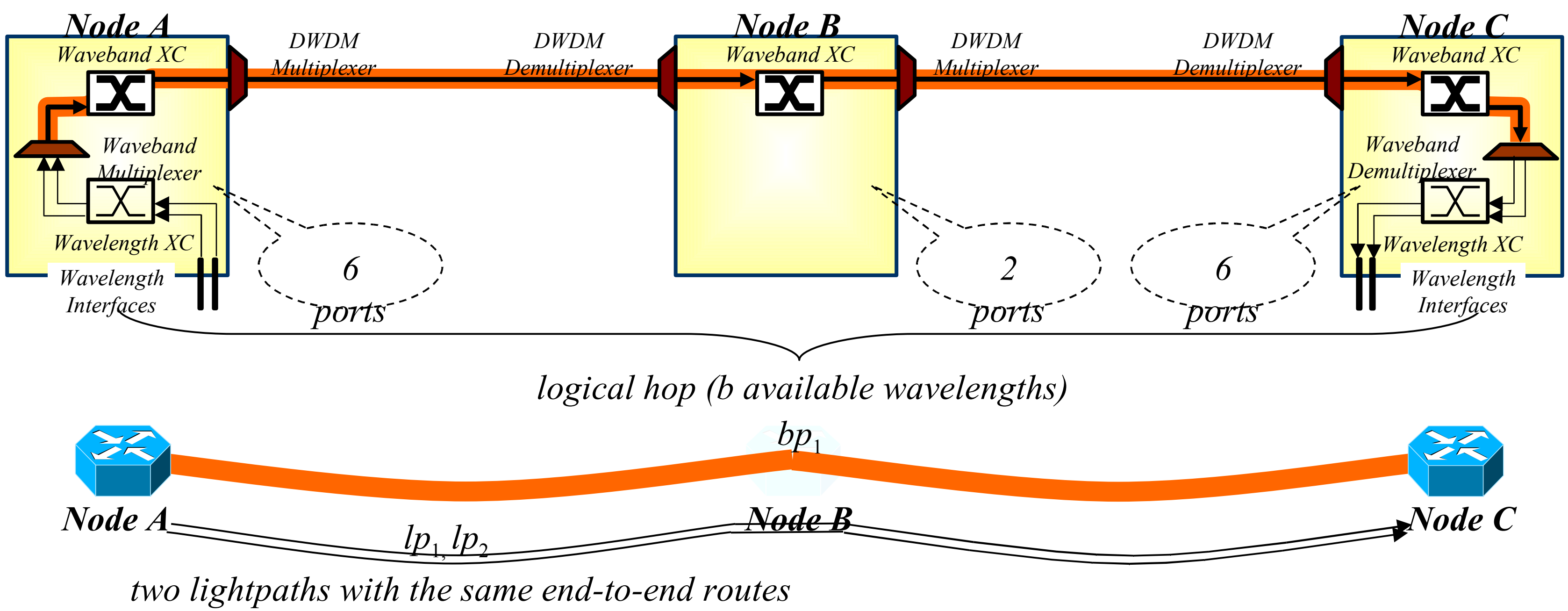


Circuit with λConversion



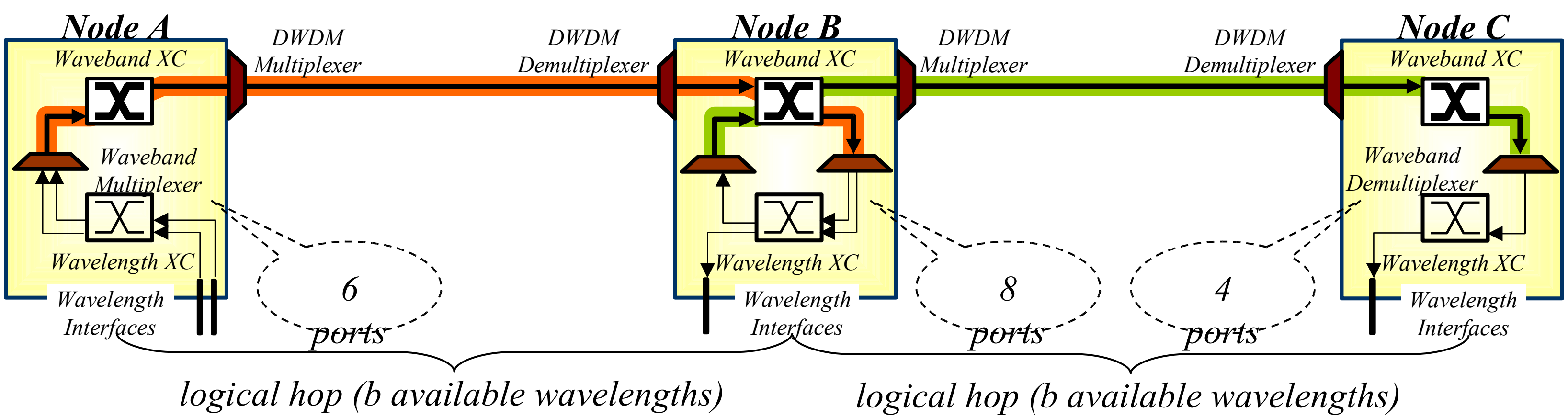
Wavebanding – Simple Case

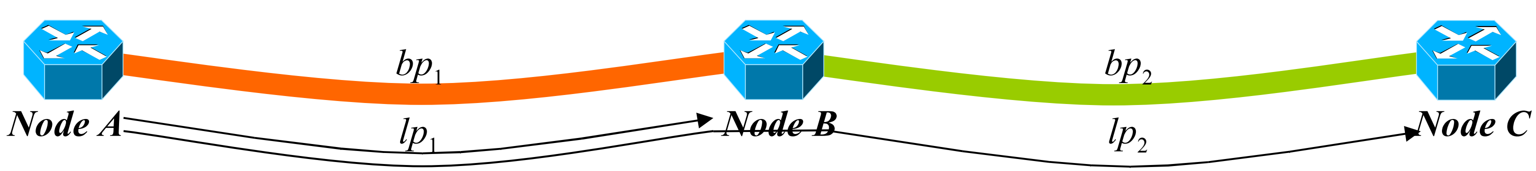
- In the simple case, wavelength circuits (lightpaths) with the same source and destination nodes are grouped together in a waveband
- Logically, these lightpaths can be thought of as being routed on a *logical link* made of one or more waveband circuits (bandpaths)
- Transit nodes switch the signal at waveband level and therefore take only two optical ports for each switched waveband
- End nodes have to terminate the waveband and therefore need more ports



Wavebanding – More Complex Case

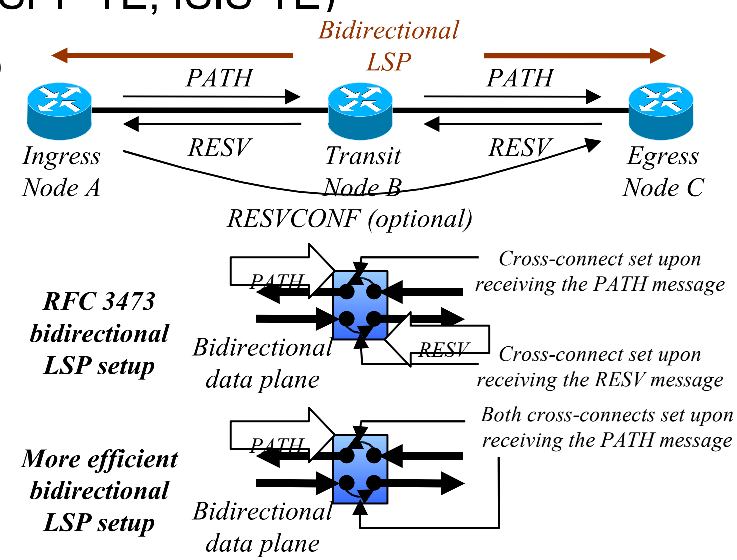
- In the simple case, wavelength circuits (<u>lightpaths</u>) with the same source and destination nodes are grouped together in a waveband
- Logically, these lightpaths can be thought of as being routed on a *logical link* made of one or more waveband circuits (<u>bandpaths</u>)
- Transit nodes switch the signal at waveband level and therefore take only two optical ports for each switched waveband
- End nodes have to terminate the waveband and therefore need more ports



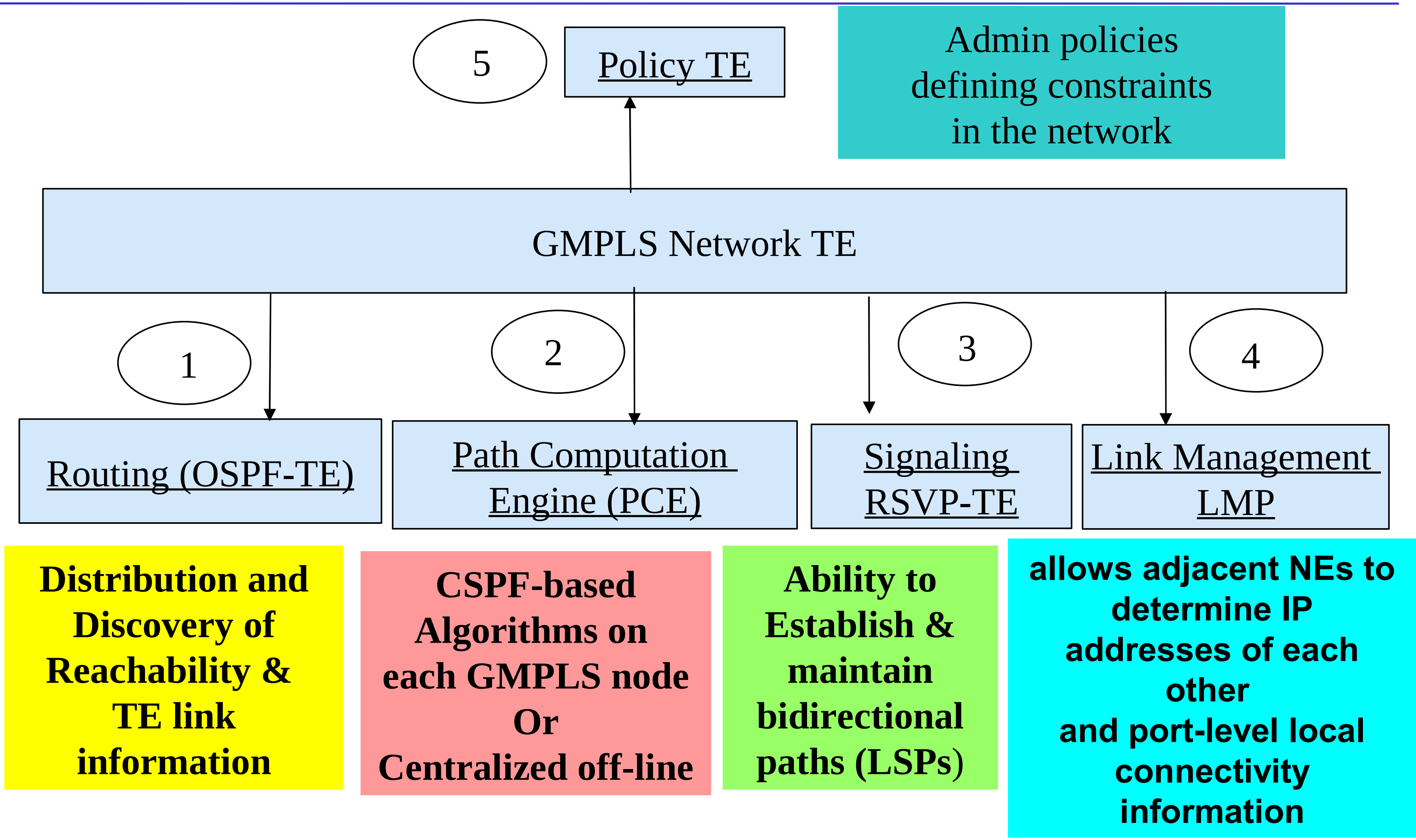


Generalized Multiprotocol Label Switching

- IP-based control plane paradigm to control packet, time slot (TDM), wavelength, waveband and space (fiber) switching across multiple switching layers, and across multiple domains.
- Developed by IETF – CCAMP workgroup with liaison work with OIF and ITU-T
- Mature standard now (RFC 3945) with various extensions for different switching technologies (Layer 2, wavelength/waveband, SONET/SDH,…)
- Basic functionalities/protocols
  - Neighbor discovery/link management (Link Management Protocol - LMP)
  - Routing with traffic engineering extensions (OSPF-TE, ISIS-TE)
  - Signaling (RSVP-TE with GMPLS extensions)
- Applications/solutions
  - Recovery (protection, restoration)
  - Make-before-break
  - Layer 1 VPN (L1VPN working group)



GMPLS functional components



GMPLS – RSVP-TE

- PATH request
  - what's a Label, Bandwidth for an Optical Switch ?
  - How to differentiate between a Fiber, Waveband, Wavelength connection request ?
  - How to differentiate between SONET or ETHER connection request ?

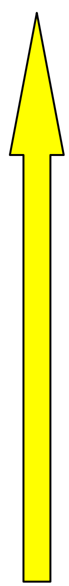
Generalized Label: Switching capability and encoding type

Exple: LSC, SONET-SDH, OC-48

- Alarm generation suppress and graceful teardowns ➔ ADMIN_STATUS
- Data plane not affected by control plane faults ➔ graceful restart

RSVP-TE Scalability

- Set regions based on Switching Capability
- LSP hierarchy – Interface switching Capability



- **Fiber Switch Capable (FSC)**
- **Band Switch Capable (BSC)**
- **Lambda Switch Capable (LSC)**
- **Time Division Multiplexing Capable (TDM)**
- **Packet Switch Capable (PSC)**

Optical Networks – GMPLS based QoR

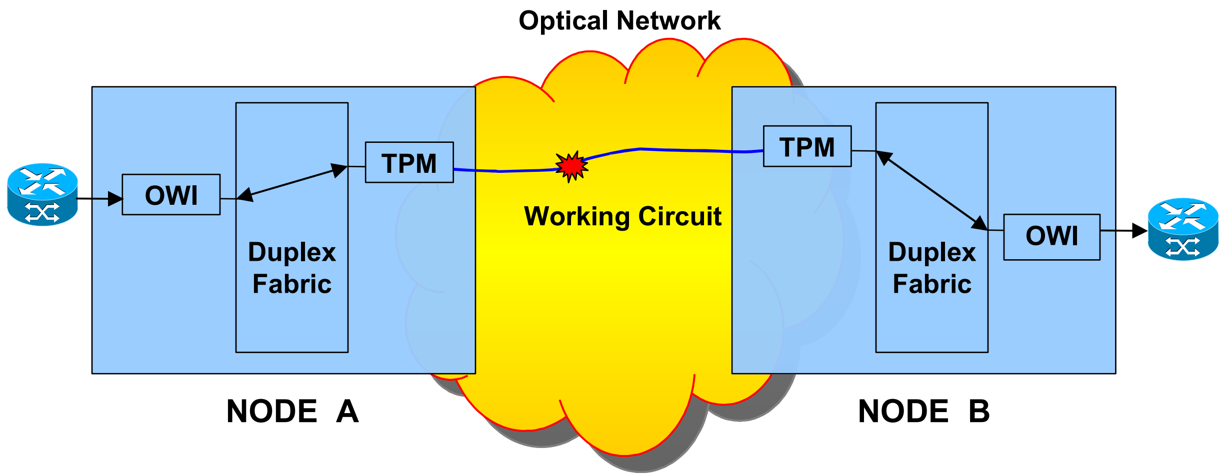
- **Low Priority** – no restoration, no protection, pre-emptable
- **Basic** – no restoration, no protection , not pre-emptable
- **Auto-Restore** – no protection , not pre-emptable
- **1:1** – protected. Protection path may be used for low priority traffic. Both protection and working paths have restoration.
- **1:N** – protected. Protection path is shared and may be used for low priority traffic. Upon failure of the working path, a switchover to the protection path occurs if and only if that path is not in use by another 1:N path, preempting any low priority traffic. Auto-restoration is also provided for both the working and protection paths.
- **1+1** – protected. Both working and protection paths carry data. Upon failure of the working path, a switchover to the protection path occurs. Auto-restoration is provided for both the working and protection paths.

Application intelligence (replication, migration)

Network intelligence (1+1 protection)

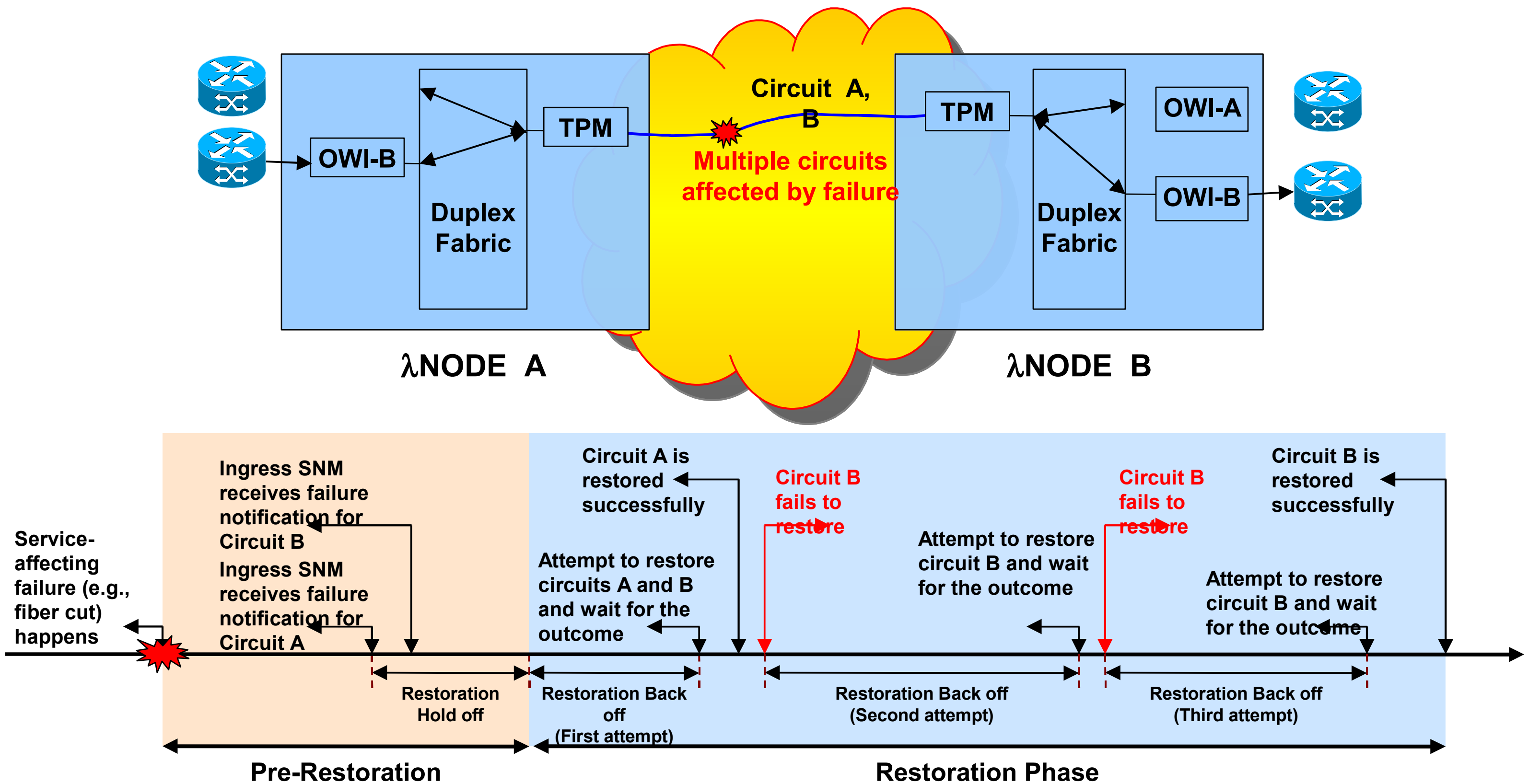
Basic Service Level

- No recovery, service is torn down if its circuit is not repaired before a certain time.



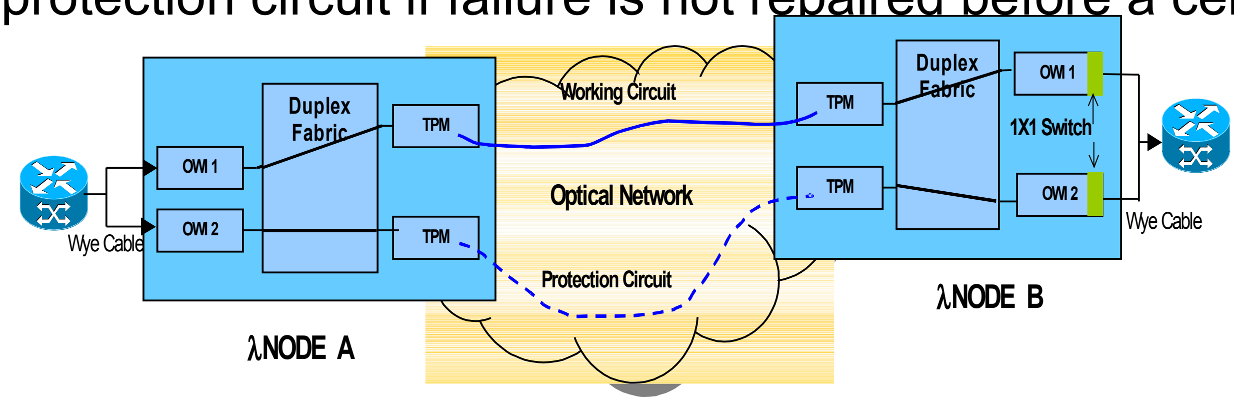
Auto-Restoration Service Level

- Recovery is in the form of restoration; service is restored on new circuit if the failed circuit is not repaired before a certain time.
- Multiple failed services with the same ingress node are restored at the same time, allowing more efficient use of resources
- A random back-off mechanism is provided to handle resource contention



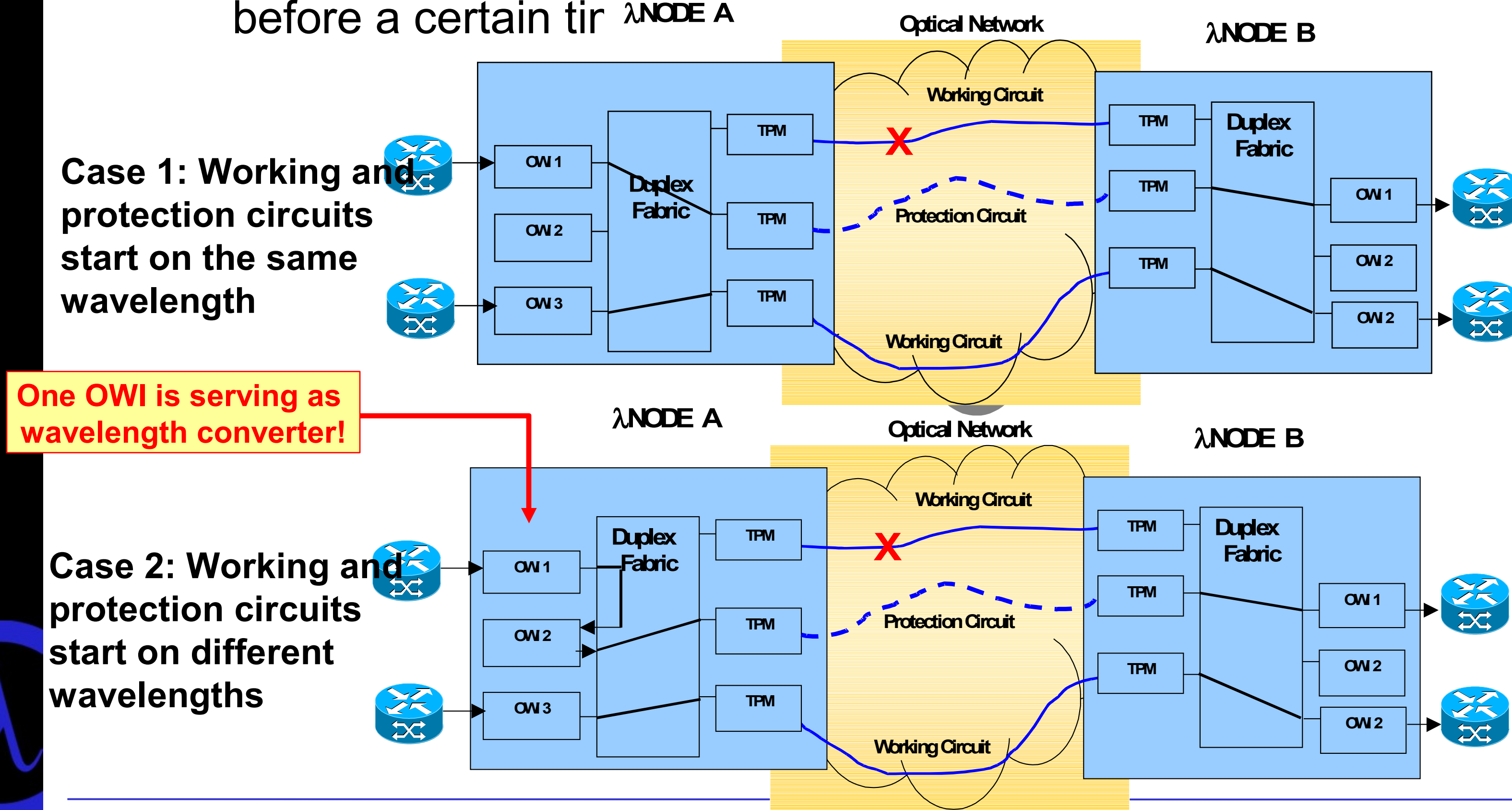
1+1 Path Protection Service Level

- Strongest recovery performance with dedicated protection
- Service is protected against transponder failure as well
- Switching in the failed direction done in less than 50 msec
- Switching is always bidirectional, but switching in the working direction may take more than 50 msec (no effect on service)
- Switchover to working or protection circuit can also be be done manually through the management plane
- 1+1 protected status is restored by establishing a new protection circuit if failure is not repaired before a certain time

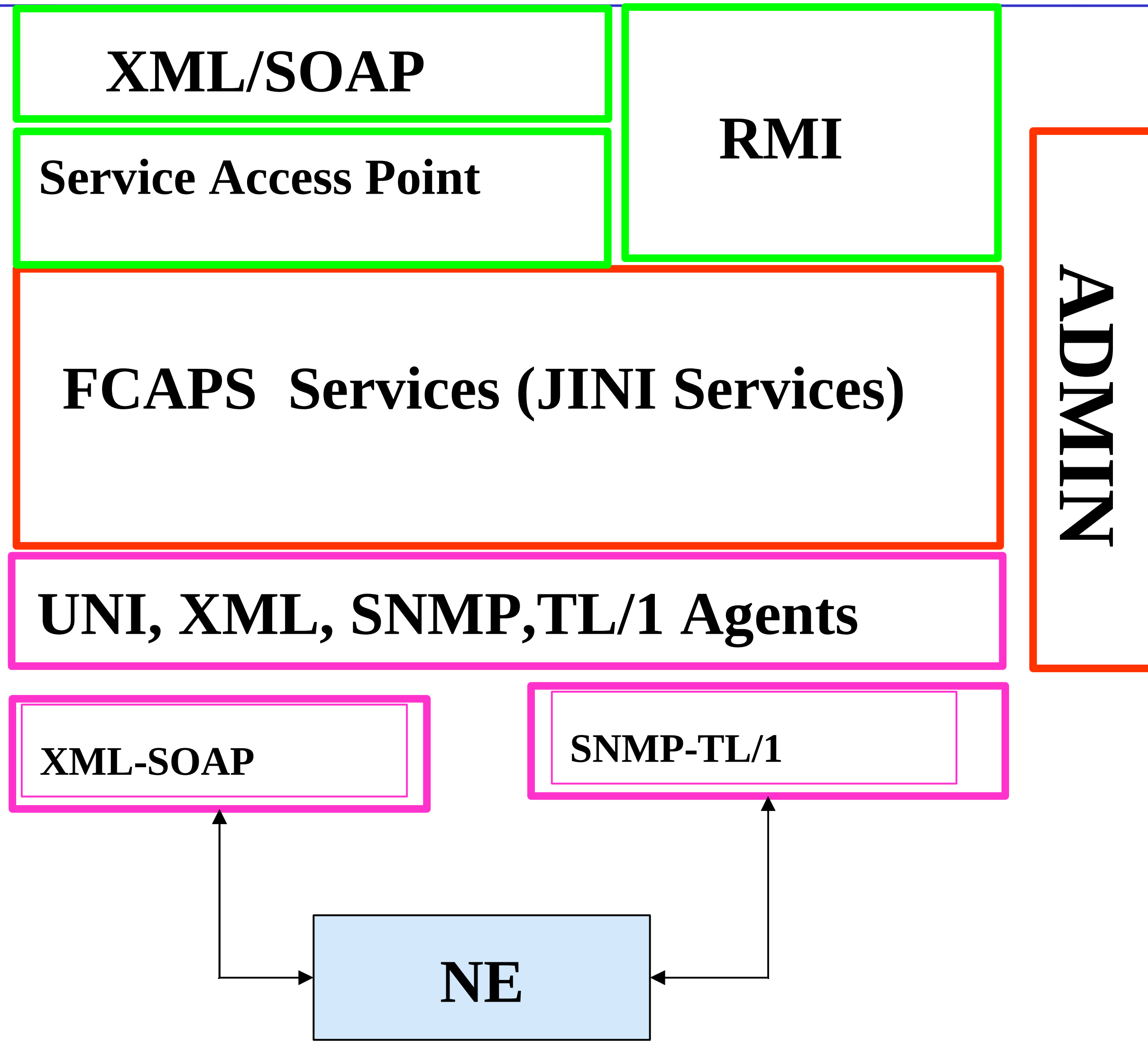


1:N Path Protection Service Level

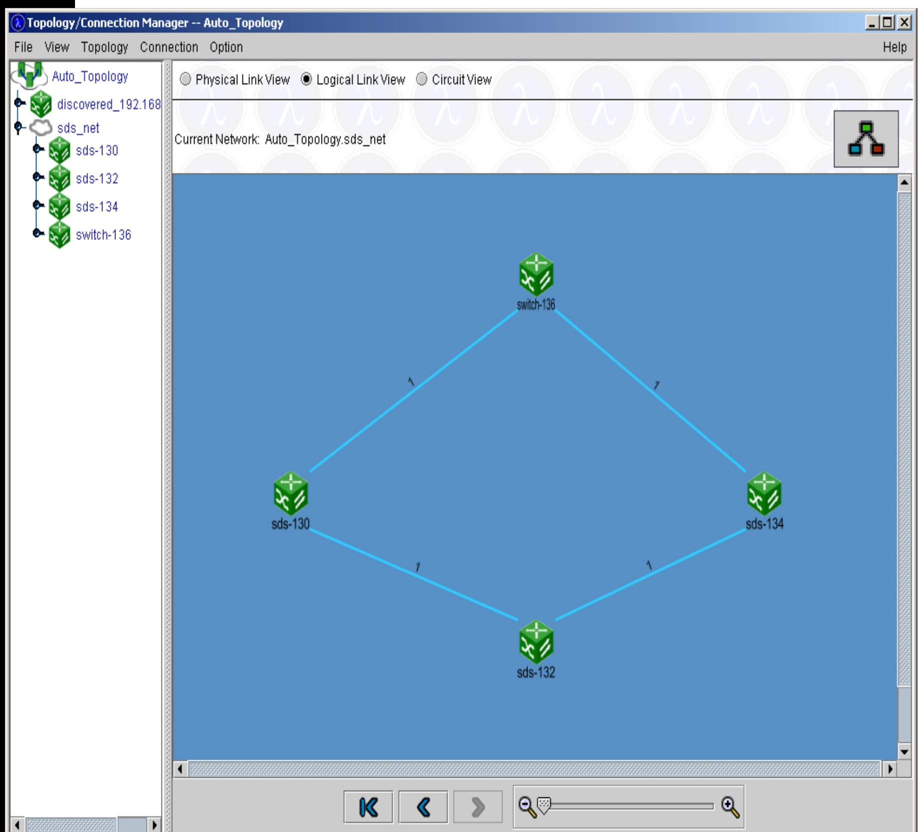
- Multiple services share one protection circuit
- To avoid inefficient use of transponder, service is revertive
- 1:N protected status is restored by moving back the working circuit to the repaired circuit, or another new circuit if failure is not repaired before a certain tir



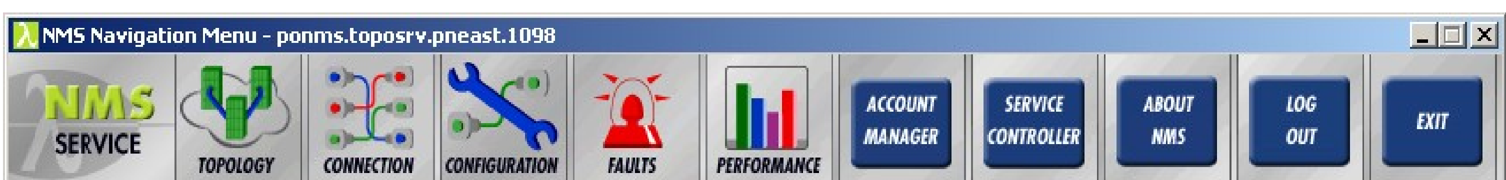
SDS – Service Delivery System



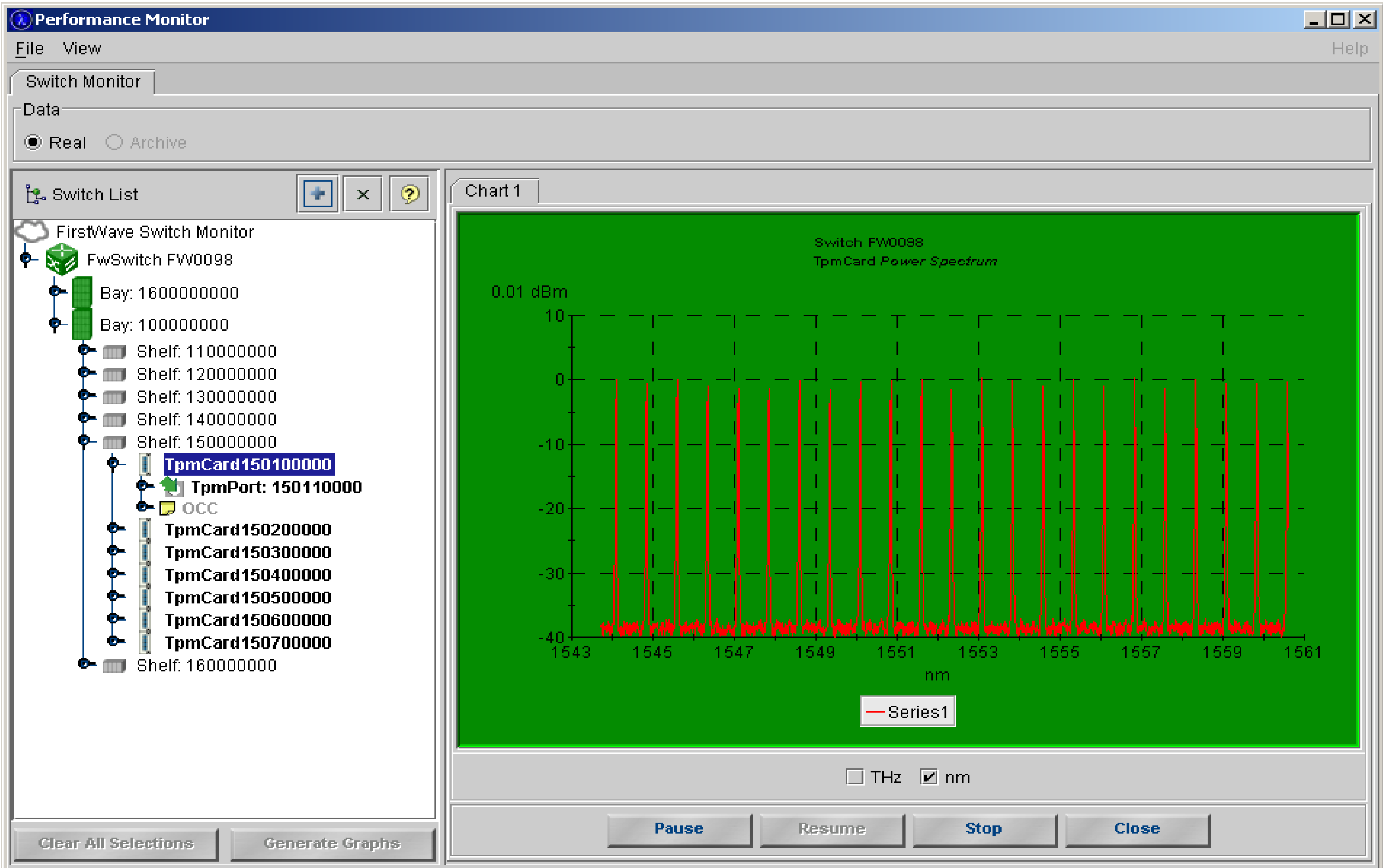
Network Management with JINI



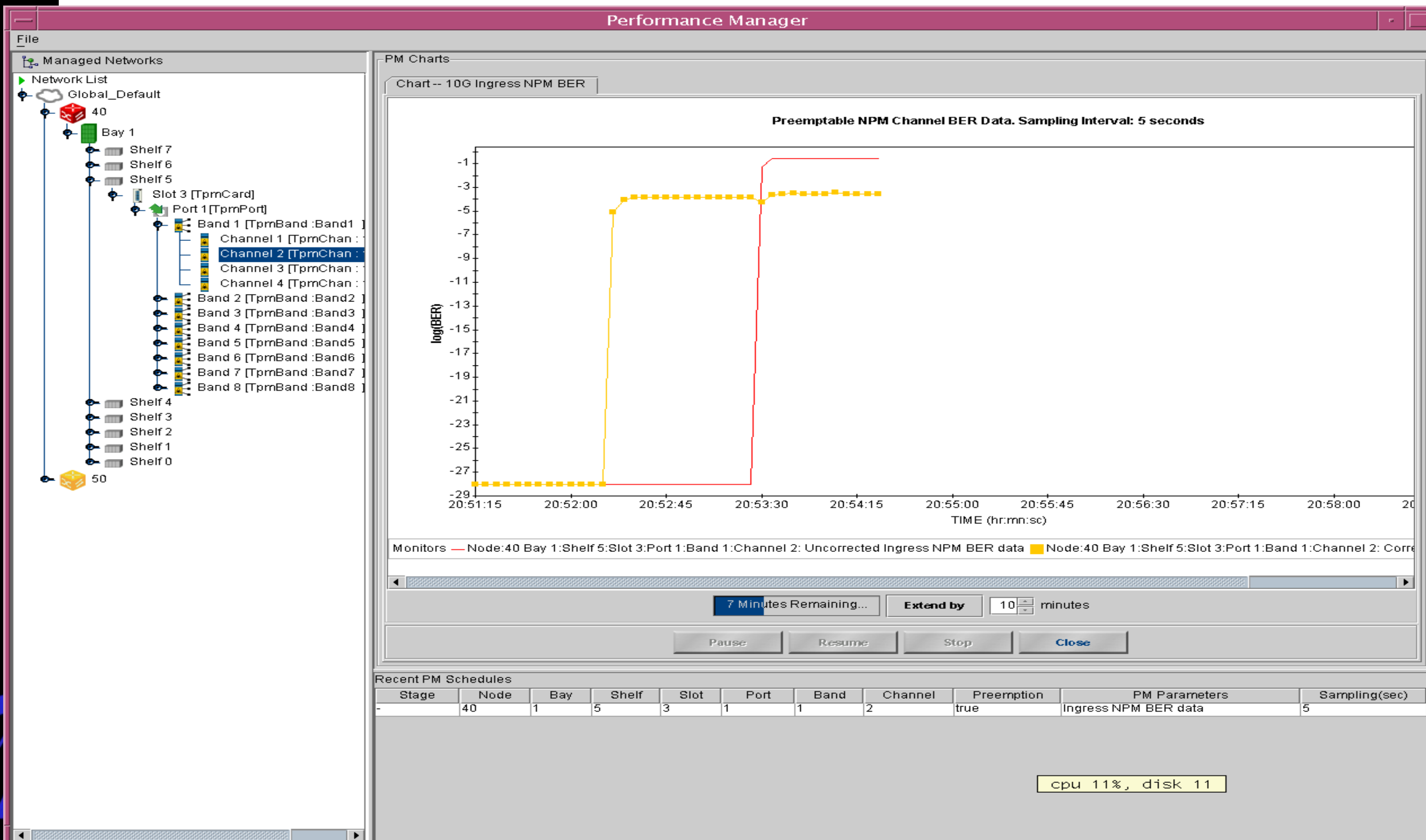
- Extensibility – easy & fast
  - **Deploy new services at run-time, find each other automatically**
- Exchange of services on-the-fly
  - **leasing to replace old services without power-down**
- Fault-tolerance
  - **services entries are leased**
  - **deploy multiple instances - redundancy**
- Scalability
  - **deploy multiple instances of same service**
  - **hierarchical federations**



Performance Management View

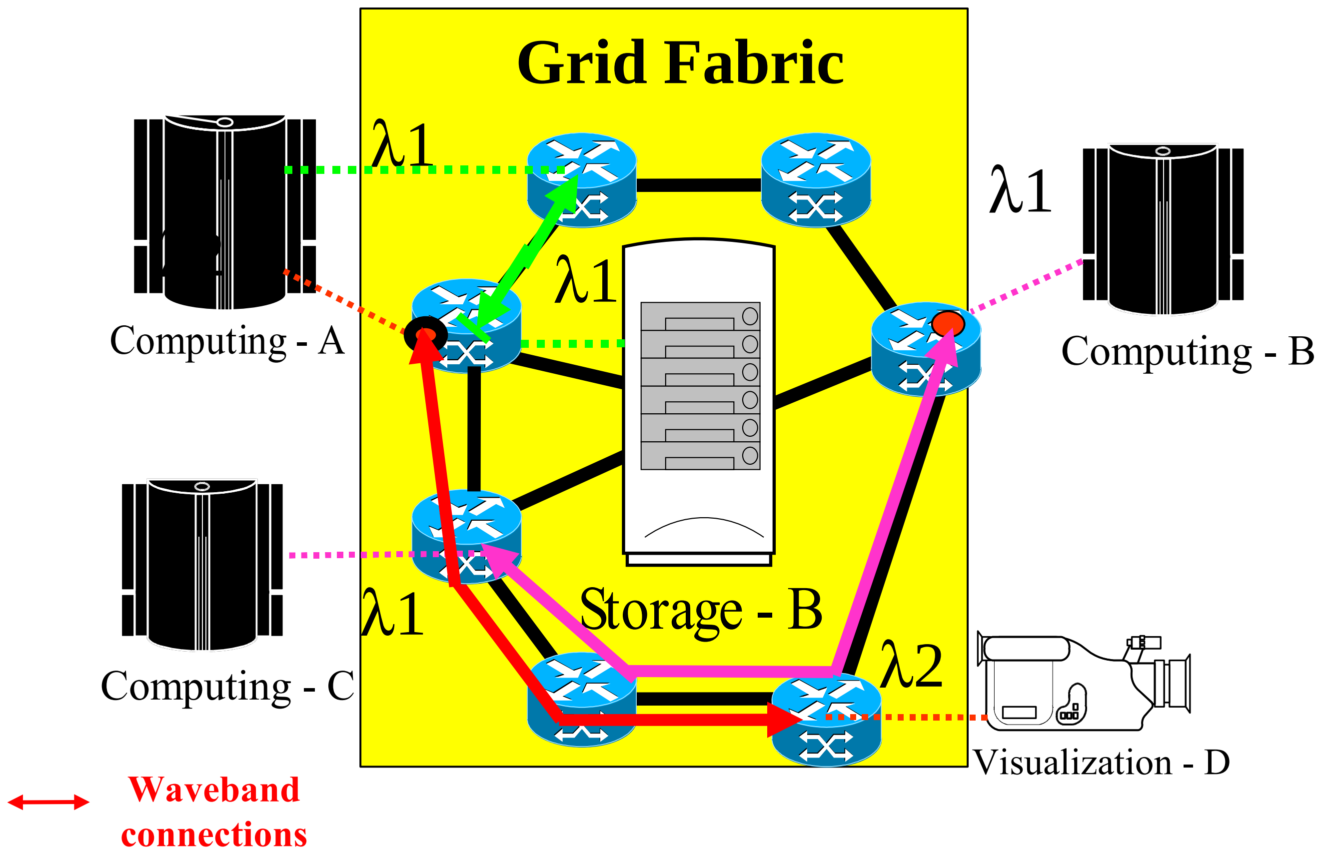


Digital performance



Dynamically direct BW to where it is needed

- Update dynamically the logical topology
- Dynamic wavelength connection



Lambda OpticalSystems Product Family

LambdaNode 2000

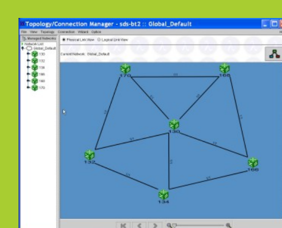
- Multi-degree intelligent all-optical switch for regional and metro core applications
- Integrated DWDM transport and optical amplifiers
- GMPLS Control Plane offers Opex savings and Mesh Protection
- Up to 256 wavelengths in one rack, 40Gbs ready

LambdaNode 3000

- Intelligent optical cross-connect
- Carrier grade solution
  - Redundant fabric, control, line cards
- Up to 128x128 port capacity
- GMPLS Control Plane

LambdaNode 200

- All-optical switch with GMPLS control plane
- 64x64 ports (bidirectional)
- Ports run at any optical speed
- Access, campus applications



LambdaCreate

- GUI-based Network Management System
- Full FCAPS: fault, configuration, accounting, provisioning, security
- SNMP, TL1 and TMF-814 Northbound interface