Optical Networks for Future Internet Design

'Trends in Communications'
CN2S - A.A. 2010-2011

Prof. Carla Raffaelli
Prof. Walter Cerroni
DEIS – University of Bologna

Foreword

FP6 IST "Broadband for all"
Network of Excellence

e-Photon/ONe
"Optical Networks: Towards Bandwidth Manageability and Cost Efficiency"

1st phase: 2004-2006
2nd phase: 2006-2008

38 partner institutions:
- 32 academic institutions
- 4 telecom operators
- 2 manufacturers

with broad European coverage (from Portugal to Turkey)
~400 researchers actively involved in the NoE
Coordinator: Politecnico di Torino

What is a Network of Excellence?

From report on EC IST projects:

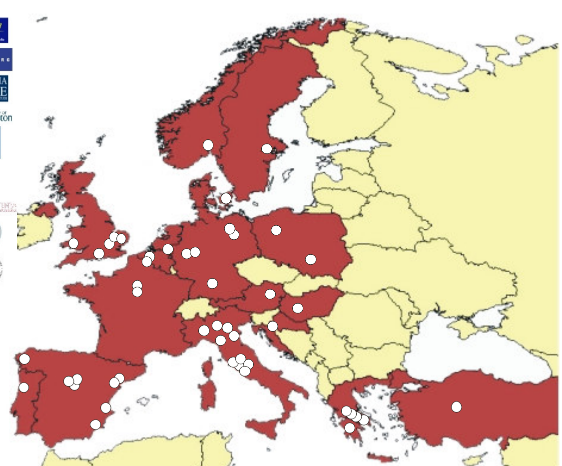
"Networks of Excellence should be designed as an instrument to cover different forms of collaboration and different sizes of partnerships"

e-Photon/ONe and BONE aimed at and succeded in "integrating and focusing the rich know-how available in Europe on optical communication and networks, both in universities and in research centres of major telecom manufacturers and operators" using the following structure:

- strong integration of a *core membership*
- active involvement of all partners in the NoE
- involvement of external institutions ("Collaborating Institutions") outside Europe

- sample outcomes:
  - Joint publications
  - Researcher mobility action
  - Master curriculum definition on Optical Networking
  - Rodamap on optical transport network technology public deliverable available at http://www.e-photon-one.org
  - ....

Follow up: NoE BONE Consortium (2008-2010)



49 partners
17 countries
523 researchers

See BONE portal: http://www.ict-bone.eu

Lecture outline

Part I:

- *Introduction to optical networks*
  - *Motivations*
  - *Basics*

Part II:

- *Optical switching architectures*
  - *Hardware and software building blocks of a network node*

Part III

- *Towards new network concepts: programmable router and experimental activities*

Discussion

Introduction to Optical Networks

Service drivers

- High bandwidth video-based services
  - HD TV, Video conferencing, cinema services
- Massive narrowband services diffusion like VOIP, messaging and email
- Increasing demand of bandwidth in access networks

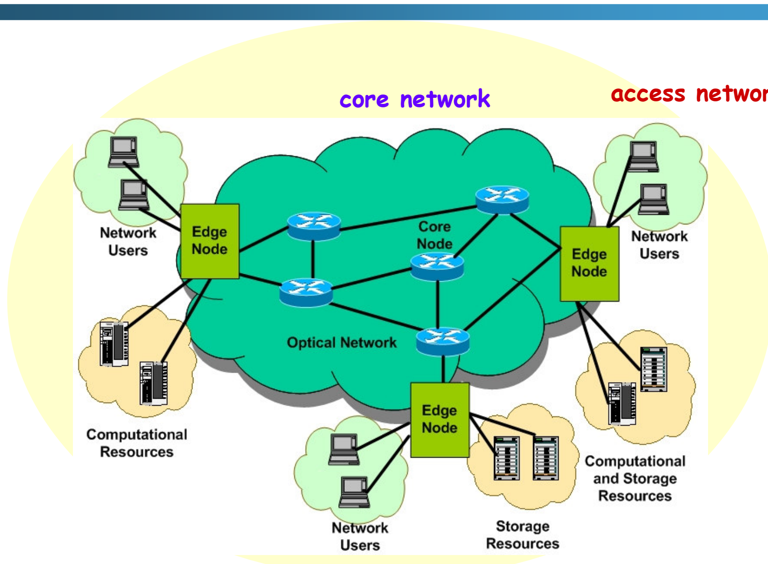
Network providers' drivers

- More efficient use of network resources
- Enhanced configuration and management capabilities
- Enhanced quality networks
- Enhanced automation in the service provisioning processes
  - Costumer control and re-configuration capability
- Enhanced monitoring capability and resiliency

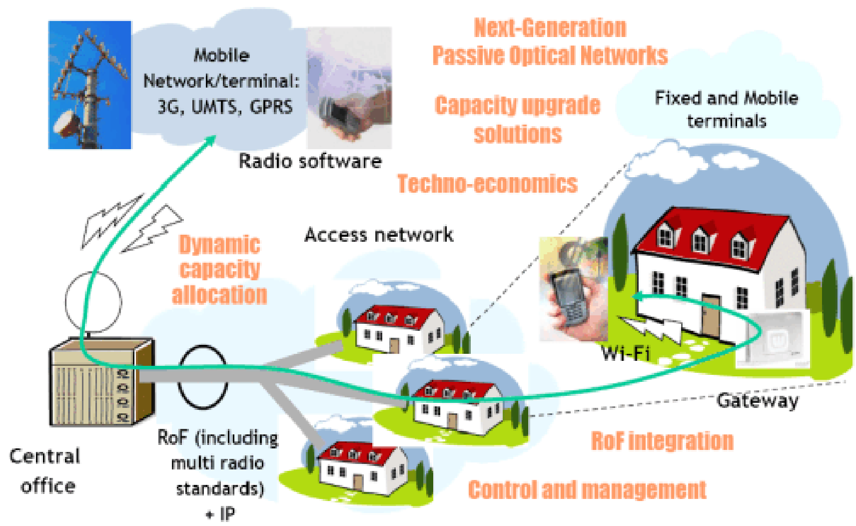
Networks of the future

- New technology approaches are being developed for access and core networks
- Access networks
  - advanced passive optical network solutions: bandwidth delivery over long distance
  - Interworking with radio access: to reach people on the move
- Core networks
  - SDH and NG-SDH based solutions
  - Enhanced Ethernet to support packet services in the metro area
- Need for higher capacity systems increases
  - Flexible control of DWDM capacity
  - Need for switching in the optical layer of the network

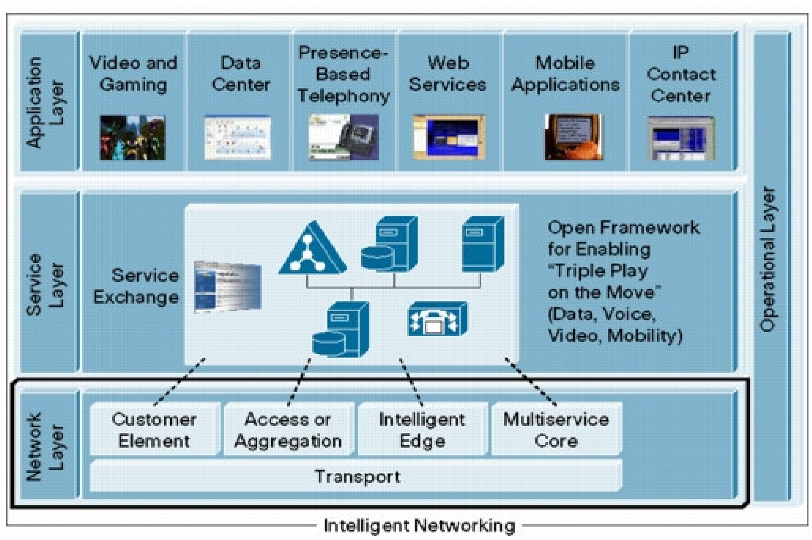
A picture of optical network



Focus on network access



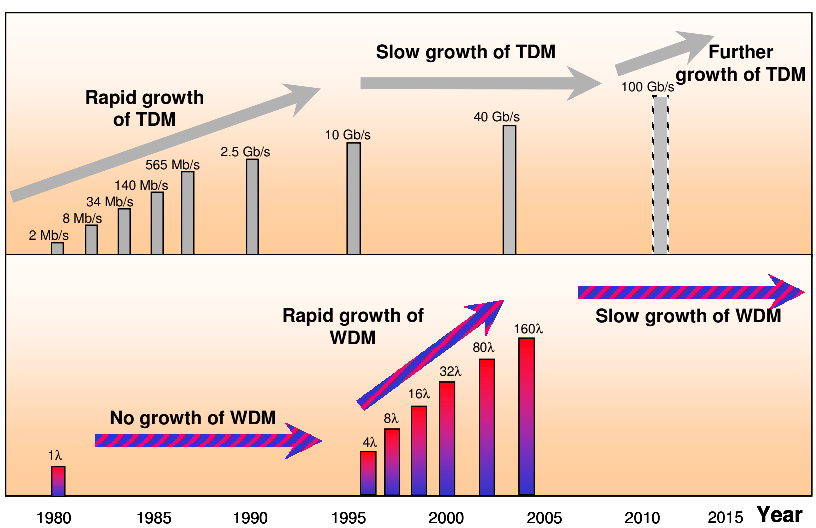
A layered view of optical network



High speed transmission

- Current networks employs amplified DWDM (Dense Wavelength Division Multiplexed) systems
  - Channel bit rates up to 40 Gb/s
  - Connect main switching centres
- Line rates and router interface rates follow a predictable increase driven by
  - Capacity demand
  - Technology economics
- Further increase driven by
  - Research on feasible technology
  - Specific network goals

Capacity evolution

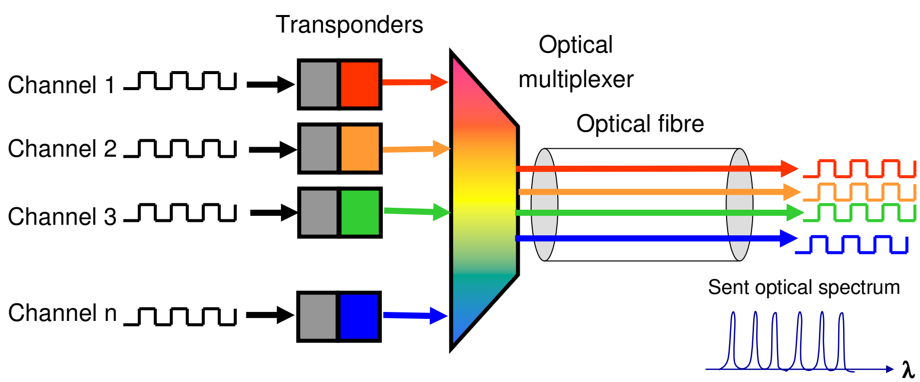


Advances in WDM Networking

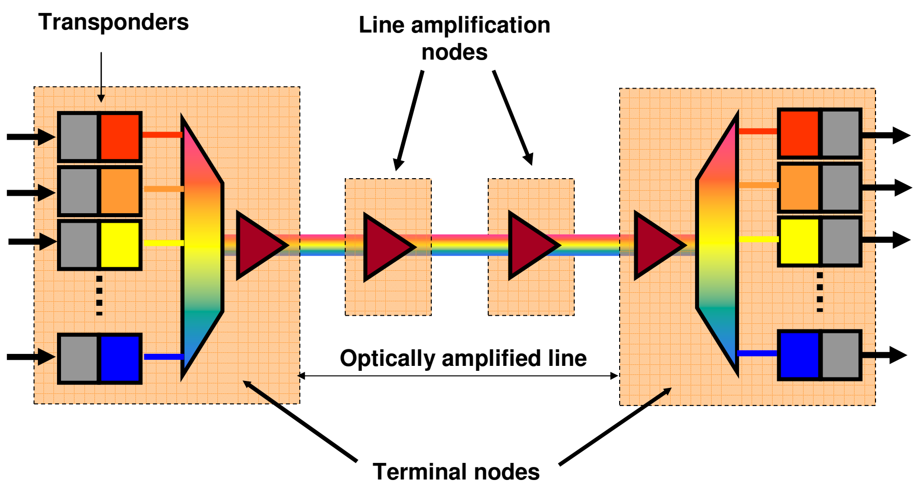
- Transmission (long haul)
  - 80 λs (1530*nm* to 1565*nm*) now, and additional 80 λs (1570*nm* to 1610*nm*) soon
  - OC-48 (2.5 Gbps) per λ (separated by 0.4 *nm)* and OC-192 (10 Gbps) (separated by 0.8 *nm*)
  - 40 Gbps per λ also on the way (>1 Tbps per fiber)
- Cross-connecting and Switching
  - Up to 1000 x 1000 optical cross-connects (MEMS)
  - 64 x 64 packet switches (switching time < 1 *ns*)

Principle of WDM multiplexing

- Wavelength Division Multiplexing consists of transmitting several signals over one fibre using optical carriers at different wavelengths.
- Hence it is just optical FDM.

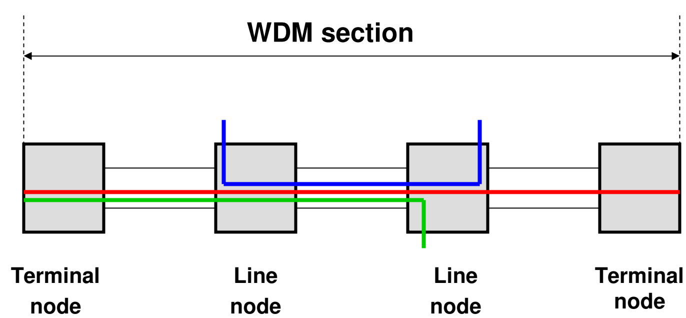


Architecture of a DWDM system



Switching in a WDM system

- To transports different client traffics over the same fibre using different optical channels at different wavelengths :
  - From terminal node to terminal node
  - From terminal node to line site
  - From line site to line site



Example: IST TOPRATE

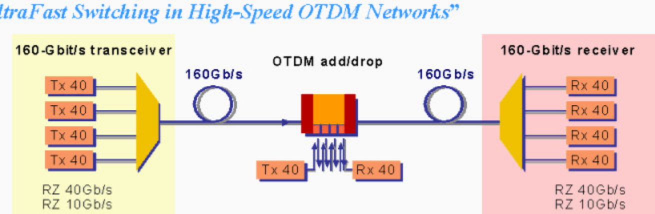
IST TOPRATE: Achievements/ Results

- Optical Time Division Multiplexing (OTDM) technology
  - ⇨ to the next bitrate hierarchy of 160 Gbit/s and beyond
  - ⇨ full use of 40 Gbit/s ETDM technology
- Dense Wavelength Division Multiplexing (DWDM)
  - ⇨ Nx160 Gbit/s DWDM / use of AWG demultiplexers
- Fibre transmission
  - ⇨ Dispersion management optimization
  - ⇨ 160 Gbit/s single channel transmission over 2x100 km
  - ⇨ 4x160 Gbit/s DWDM transmission over 3x80 km
- Many novel techniques investigated / used for
  - ⇨ PMD compensation: tunable planar devices (PLC)
  - ⇨ electronic 40Gbit/s eye-monitoring
  - ⇨ optical demultiplexing: novel photonic crystal fibre
  - ⇨ optical clock recovery: 40 GHz optical clock operated in 160 Gbit/s receiver

Example: IST FASHION

IST FASHION

"UltraFast Switching in High-Speed OTDM Networks"



- Transmission of 160 Gb/s OTDM data (16x10 Gb/s and 4x40 Gb/s) over 500 km of fiber to demonstrate the usability in ultra high speed core networks
- Adding to and dropping 10 Gb/s and 40 Gb/s data from a 160 Gb/s OTDM data stream to demonstrate the flexibility of OTDM technology for usage in flexible optical networks
- 160 Gb/s OTDM field trial over installed fibers to study the influence of environmental conditions

Ethernet evolution

- 10 Gbit/s Ethernet is widely deployed to interconnect IP routers and Ethernet switches
- Proliferation of 10 Gbit/s services calls for the next factor of 10: 100 Gbit/s Ethernet
- 100 Gbit/ s Ethernet standard is under development by the IEEE
- The IEEE 802.3 Higher Speed Study Group (HSSG) has adopted several objectives
  - 100 GbE Optical fiber Ethernet Standards at least 100 meters and 10 Km
  - Full duplex operation
  - Current frame format and size standards

How to switch in next generation networks

- High speed switching can be implemented using either electronic cross-connect switches or optical cross-connect switches
- Cross connect switching in SDH infrastructures is performed in the electronic domain
  - It represents a reference benchmark
- Optical cross connect based on MEMS technology and operating at millisecond switching speeds are now available and suitable for slow switch reconfiguration time

Electronic vs Optical Switching

- Data transmission is carried out in the optical domain today in WANs (Wide Area Networks) and MANs (Metropolitan Area Networks)
  - switching is mostly done in the electronic domain
- Electronic switching uses electronic switching fabrics
  - Converts data from optical to electronic for switching purposes, and then from electronic back to optical for transmission.
- Optical switching uses optical switching fabrics
  - Payload stays in the optical domain
  - Control plane is in the electronic domain

Caveats of OEO Switching

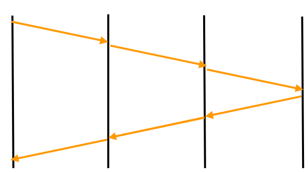
- Internet traffic doubles every 6 months (1997-2008)
- Semiconductor performance doubles every 18 months which is known as the Moore's Law
- The first time in history that improvements have been required faster than the improvement rate for semiconductors, Moore's Law.
  - Complex operations are needed at a OEO router's line card e. g. processing the packet header, longest prefix match, packet buffering, etc.
- The cost of OEO at OC-48 (2.5Gbps) and at OC-192 (10 Gbps) is relatively high

Optical switching alternatives

- Optical Circuit Switching (OCS)
- Optical Packet Switching (OPS)
- Optical Burst Switching (OBS)

Optical Circuit Switching

- Two-way process with request and acknowledge
  - Round Trip Time = tens of *ms* therefore long setup delays
  - Suitable for smooth traffic and QoS guarantees due to fixed bandwidth allocation
  - Bandwidth inefficient for bursty (data) traffic
    - Wasted bandwidth during off/low-traffic periods
    - Overhead due to frequent set-up/release

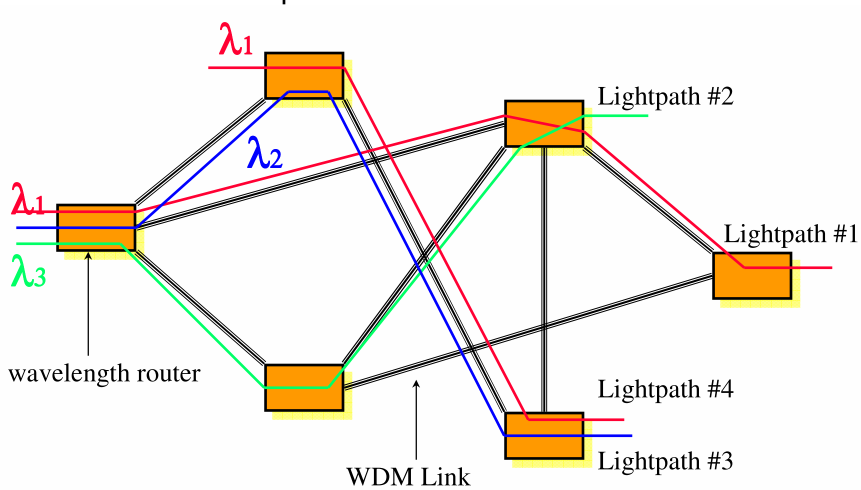


Wavelength Routing

- Setting up a circuit means setting up a lightpath (or λ path)
  - A wavelength, or a concatenation of wavelengths, is allocated for the connection from source to destination
- λ-path specific pros and cons:
  - \+ Mature OXC technology (*msec* switching time)
  - – Very coarse granularity (OC-48 and above)
  - – Limited # of wavelengths (thus # of lightpaths)
  - – No aggregation (merge of λs) inside the core
    - traffic grooming at the edge can be complex/inflexible
- Current state of the art

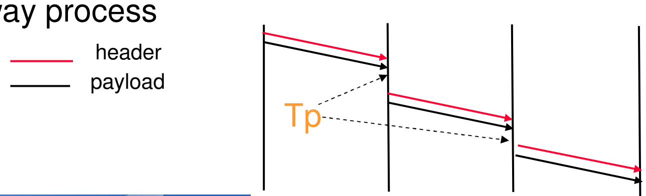
Wavelength Routing: Lightpaths

Point-to-point link concatenation



Optical Packet Switching

- A packet contains a header (e.g., addresses) and a payload (variable or fixed length)
  - Can be sent without circuit set-up delay
  - Statistic sharing of link bandwidth among packets with different source/destination
- Store-and-forward at each node
  - Buffers a packet, processes its header, and sends it to the next hop
  - Packet header is today processed in the electronic domain or all-optically in the future at each node and switched to the next hop
- One-way process



In band header

- Optical header at pre-defined bit rate
- Header signal spilled at switch input and converted to electronics
- Guard band between header and payload to cope with optical devices switching speed
- Header and payload bit rates may be different

Packet format alternatives

- Synchronous fixed length packets (ATM like scenario)
  - ☺ Node design and operation
  - ☺ Queuing performance
  - ☹ Synchronization
  - ☹ Interworking with variable length data
- Asynchronous variable length packets (Internet like scenario)
  - ☺ No synchronization
  - ☺ Interworking with any format
  - ☹ Node operation
  - ☹ Queuing performance

Optical Packet Switching

+ Statistical multiplexing of data
+ Suitable for bursty traffic
– Requires fast switching speeds (nanoseconds)
– Stringent synchronization requirements
– Queuing requirement inside the switch
– Still viewed as a long term solution

Motivations for OBS

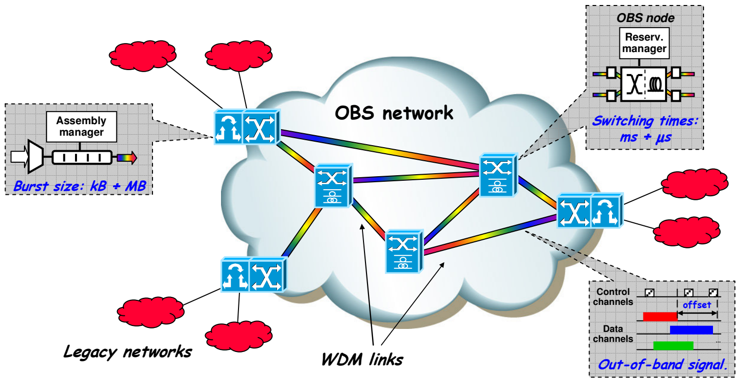
- New traffic profiles
  - P2P file downloading
  - multimedia streaming
  - grid networking
- Problems in wavelength routed networks
  - low network utilization and flexibility
- Problems in optical packet switched networks
  - lack of optical buffering
  - need for fast packet switching and header processing
- Need of graceful migration from wavelength routing networks

Optical Burst Switching

- Main design objectives
  - decreasing complexity of OPS with still employed statistical multiplexing in optical domain
  - building a buffer-less network
  - user data travels transparently as an optical signal and cuts through the switches at very high rates
- Solution
  - sending a header to temporarily reserve a wavelength path
  - Sending then an optical burst (a block of IP packets) through the network
- OBS (one-way reservation) can be viewed as lying between OPS (no reservation) and WS networks (two-way reservation)

OBS Network Architecture

- Control and data information travel **separately** on different channels
- Data coming from legacy networks are aggregated into **a burst unit** in edge node
- **The control packet** is sent first in order to reserve the resources in intermediate nodes
- The burst follows the control packet with some **offset time**, and it crosses the nodes remaining in **the optical domain**

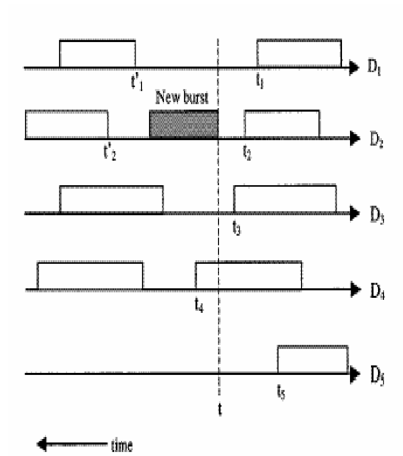


OBS Principles

- Variable-length packets, named bursts
- Asynchronous node operation
- Strong separation between the control and data planes
  - Control burst (with control information) transmitted on dedicated control channel and processed electronically
  - Data burst transmitted and switched all-optical way

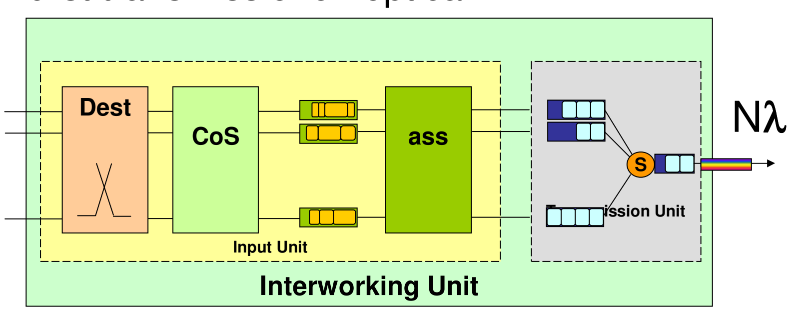
Channel Scheduling

- Problem of assigning a burst to a channel when it gets information about when it will arrive.
- Ideally bursts are assigned to channels that become free just before the bursts arrive.
- To minimize idle time (voids) and to help for scheduling later by maintaining maximum flexibility for later bursts.



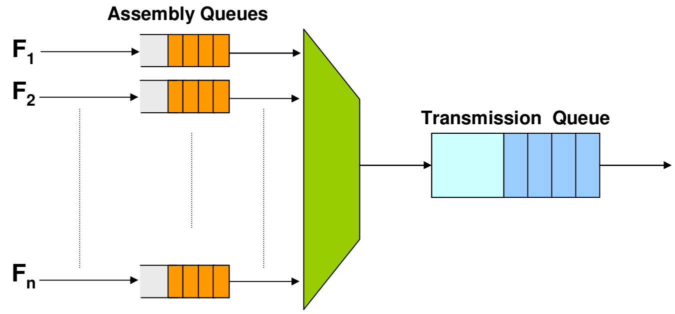
Burst assembly at edge node

- Legacy network interfacing
- Burst classification (address, QoS, …)
- Burst assembly (per flow, mixed flow…)
- Burst transmissionon optical



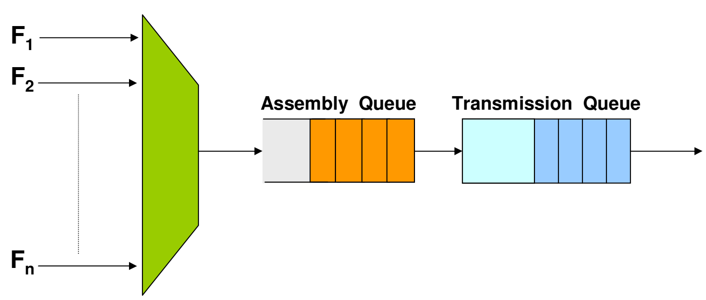
Per-flow aggregation at network edge

- Aggregation is needed both in OPS and OBS data planes
- Ingress per-flow queuing
- Optical packet assembled with segments of the same flow
- An assembly time-out for each active flow is needed

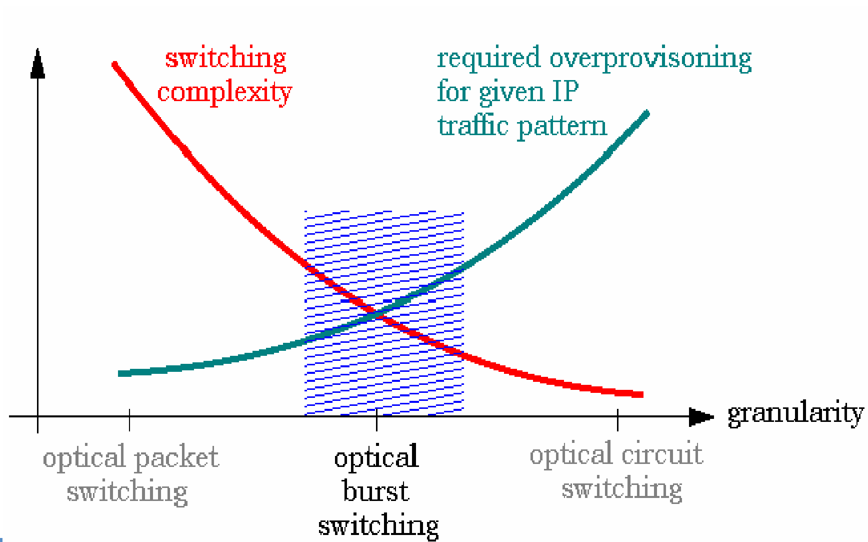


Mixed-flow aggregation

- TCP segments from different flows and with the same optical destination address aggregated in the same optical burst
- Only one assembly time-out is needed
- Lower complexity of the assembly mechanism



Comparison among OBS/OPS/OCS



Part II:
Switch architectures

Outline

- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- Buffered architectures
- Conclusions

Outline

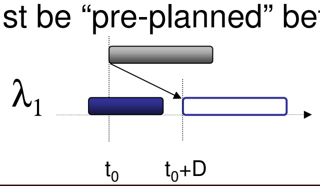
- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- Buffered architectures
- Conclusions

Contention Related Issues

- Contention arises when information transfers contend for the same resources
  - It happens at different time scales for OCS, OBS and OPS
- In OPS/OBS statistical multiplexing is applied
  - In optics no RAMs available
  - Queuing approach based on Fiber Delay Lines or Slow Light
- Alternative solutions
  - Exploitation of different domains jointly considered
    - Time
    - Space
    - wavelength

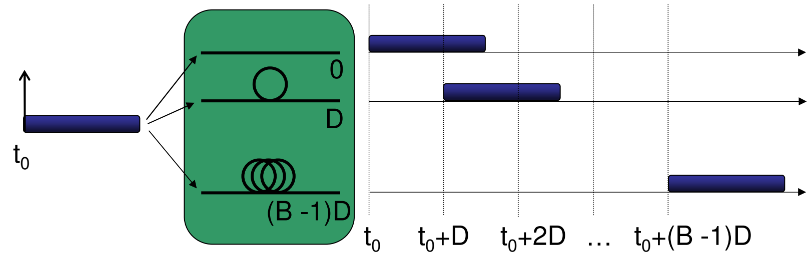
Contention Resolution in **Time Domain**

- Time domain: contending packets delayed by fiber delay lines (FDLs)
  - FDL introduces propagation delay
  - with FDLs, a packet can be delayed of fixed propagation delay, typically multiple of a fixed quantity D: D, 2D, … , ND
  - D is the granularity
- Time to transmit must be “pre-planned” before packet payload arrival time



Packet is lost if the wavelength is busy when the maximum delay is reached

FDL optical buffer



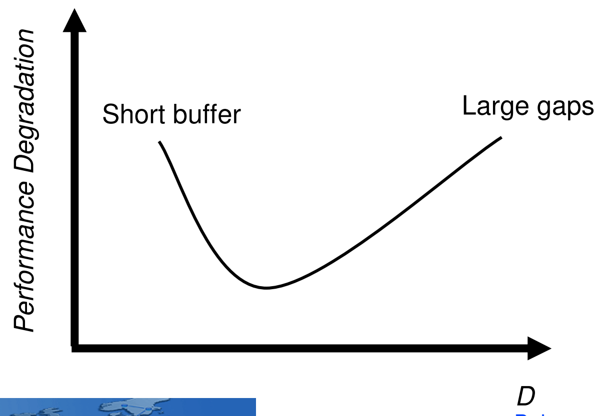
- Realized with B **Fiber Delay Lines** (FDL):
  - packets are delayed until the output wavelength is available
  - available delays are typically consecutive multiples of the **delay unit** D (different choices are also possible)
  - packets are lost when the buffer is full, i.e. the required delay is larger than the maximum delay achievable D_M = (B -1)D

Choosing the Buffer Delay Unit

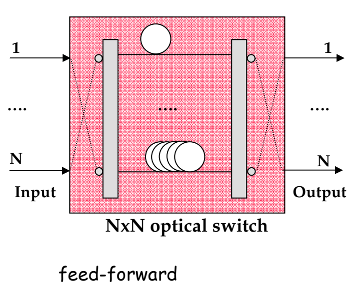
- D is directly related to
  - time resolution of the delay buffer
  - maximum delay achievable (buffer size)
- For a given number of delay lines (B):
  - decreasing D
    - + the time resolution improves and the average void size decreases
    - – the buffering capacity decreases
  - increasing D
    - + the buffering capacity increases
    - – the time resolution decreases and the void size increases

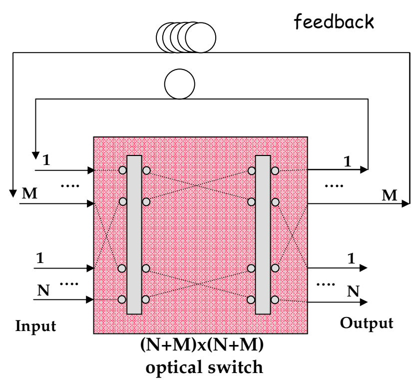
Choosing the Buffer Delay Unit

- D is directly related to
  - time resolution of the delay buffer
  - maximum delay achievable (buffer size)



FDL reference architectures

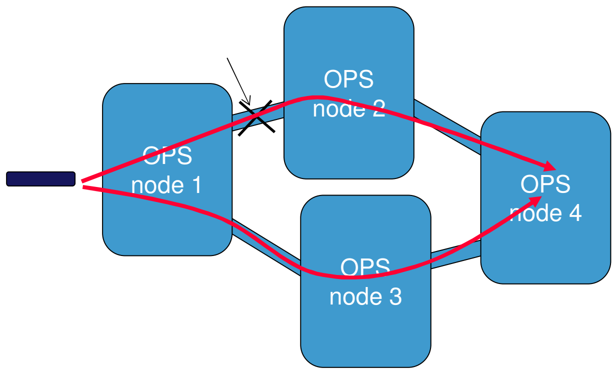




- In the feed-forward method, packets are fed into fiber delay lines of different lengths and when they come out, they have to be switched out.
- In the feedback scheme, a packet may re-circulate as long as there is a bandwidth shortage at the output ports.

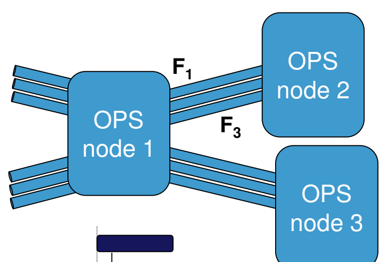
Contention Resolution in **Space Domain**

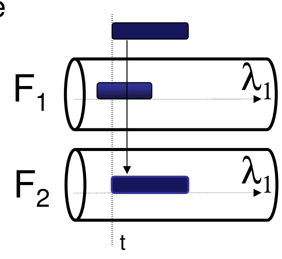
- Route-based approach
- When a packet cannot be forwarded on the first-choice path (output fiber congested), alternative routing paths can be considered



Contention Resolution in **Space Domain**

- Multi-fiber approach
- Multiple fibers per input/output interface
- Next hop-based routing
  - all fibers and wavelengths on the same output interface are equivalent for routing purposes
  - Wavelength re-use on the same interface
  - Contending packets transmitted on different fibers

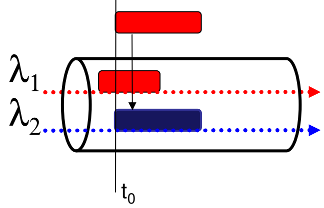




Packets dropped when all channels are busy

Contention Resolution in **Wavelength Domain**

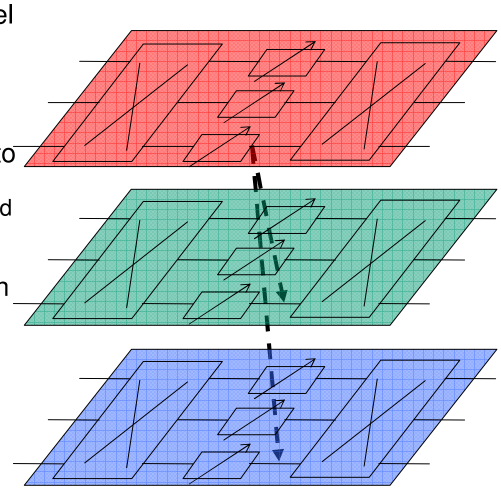
- Packets competing for the same output wavelength
  - One sent directly
  - The others converted to different wavelengths



Packets dropped when:
all wavelengths are busy *or*
no internal devices for wavelength shifting are available

Wavelength Conversion

- DWDM technology provides parallel planes
  - packets on different wavelengths do not collide
- WCs allow to shift from one plane to another one
  - packets can be wavelength converted to solve contentions
- Equivalent to multi-plane solution in electronic domain
  - space equivalent of optical switches
- Scheduling procedure needed to decide
  - when conversion is needed
  - the wavelength (plane) to convert to



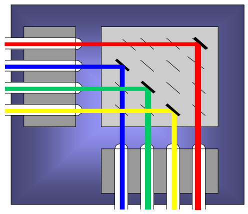
Wavelength Converters

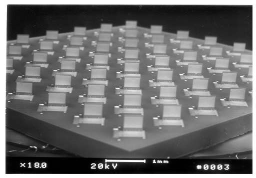
- Key components for contention resolution directly in the optical domain (avoiding O/E conversion – processing – E/O conversion)
- Very difficult to be implemented, with any kind of technology (most of them exploit non-linear effects of optical devices)
  - complex and costly devices
- Due to the complexity and cost of these devices, schemes allowing to share them have been defined
  - trade-off between the number of WCs and other optical devices and performance
- Different kind of WCs, according to the required functionality and technology
  - **tunable-input/tunable-output (TWCs)**: convert any wavelength to any other
  - **fixed-input/tunable-output (FTWCs)**: fixed wavelength on input converted to any other
  - **tunable-input/fixed-output (FWCs)**: any wavelength converted to a fixed wavelength
  - **limited range (LWCs)**: able to convert a sub-set of the wavelength range
- FWCs are the easiest to be implemented, but they are less flexible

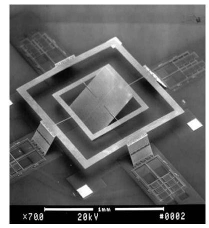
Optical switching: enabling technologies

- AWG – Arrayed Waveguide Gratings
  - Generalization of the Mach–Zehnder Interferometer
  - Two multiport couplers interconnected by an array of waveguides
  - Several copies of the same signal shifted in phase
  - The output port is selected depending on the wavelength used
- MEMS – Micro-Electro-Mechanical Systems
  - Miniature movable mirrors made in silicon
  - Mirrors are deflected from one position to another using a variety of electronic actuation techniques
  - Depending on the mirror position, the optical signal is transmitted or deflected (switching time: ca. 100 µs)
- SOA – Semiconductor Optical Amplifier
  - Based on the principle of stimulated emission (same as LASER)
  - It may be used as an ON/OFF switch (switching time: ca. 1 ns)

MEMs Switch Matrix

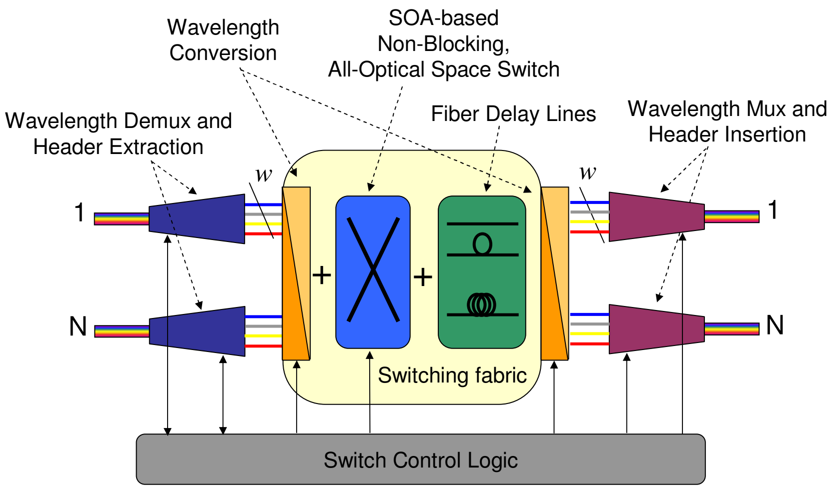




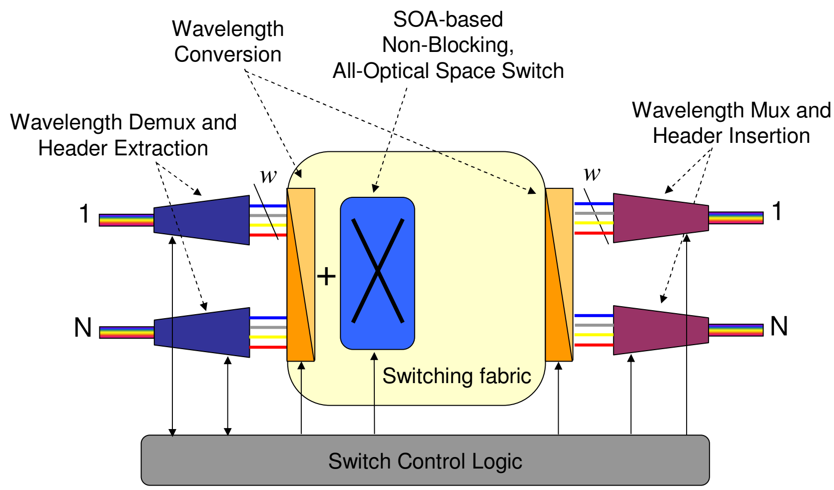


4x4 switch = 16 mirrors. Each mirror can be moved independently

General view of node elements



Buffer-less node



Outline

- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- Buffered architectures
- Conclusions

Buffer-less architectures

- They are employed with OCS and OPS
  - Cross-connect principle
- Interconnection of N fibers with M channels
  - Channel switching
- Contention resolution in wavelength domain
  - Wavelength shift is achieved by wavelength conversion
- Optical switches with Wavelength converters sharing schemes proposed to limit switch cost

Wavelength Converters Sharing Schemes

- Shared-Per-Link (SPL)
  - WCs shared among the packets directed to the same output link
- Shared-Per-Node (SPN)
  - WCs shared among all the arriving packets
- Shared-Per-Input-Wavelength (SPIW)
  - WCs shared among the packets arriving on the input channels related to the same wavelength
- Shared-Per-Output-Wavelength (SPOW)
  - WCs shared among the packets forwarded to the output channels related to the same wavelength

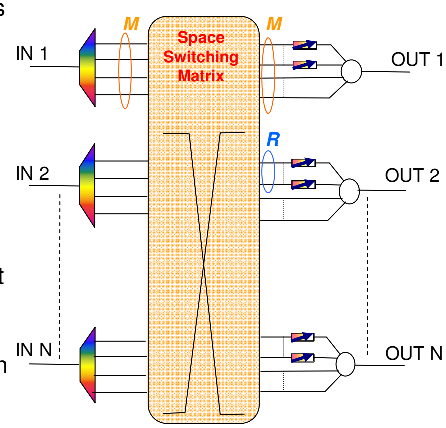
Different sharing schemes require different kinds of WCs

Contention in optical switch with shared wc

- Channel blocking
  - it is a consequence of overload on a wavelength channel on output link
  - more than one packet require the same output channel
  - it can be **resolved** by finding a different channel on the same fiber
  - it requires wavelength conversion
- Internal blocking
  - it is a consequence of resource limitation inside the switch
  - mostly related to WC unavailability or limited range of WCs
  - it leads to **packet loss** (no buffer)
- Output blocking
  - it is a consequence of overload on output link
  - exceeding packets require the same output link
  - it leads to **packet loss** (no buffer)

Shared-Per-Link (SPL) Scheme

- Optical switch with **N** input/output fibers carrying a WDM signal with **M** wavelengths
- Each output link is equipped with a dedicated pool of ***R*** WCs
  - ***R*** < M partially equipped
  - ***R*** = M fully equipped
- Input packets firstly forwarded on the same wavelength they come from (without wavelength conversion)
- In case of channel contention, wavelength conversion is performed



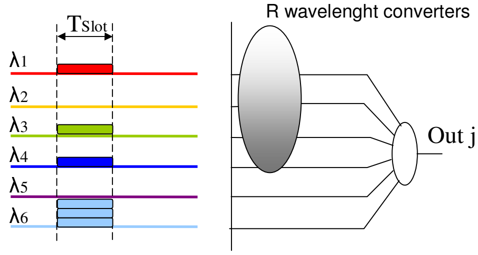
Packet Loss: Case 1

G: number of different wavelengths sending at least 1 packet to output interface j

Case 1: G > M – R

- M – R packets are transmitted to the output channels without WC
  - All channels without WC are exploited
- R packets are transmitted to the output channels with WC
- Total packets transmitted: M, remaining packets are lost
- Number of transmitted packets is the same as in the full conversion case

Example:
M=6
R=3
G = 4 > M – R



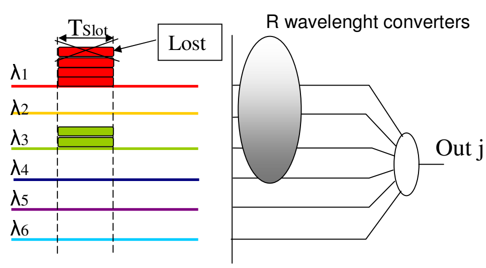
Packet Loss: Case 2

Case 2: G < M – R

- G packets are transmitted to the output channels without WC
- R packets are transmitted to the output channels with WC
- Total packets transmitted: G + R < M, remaining packets are lost

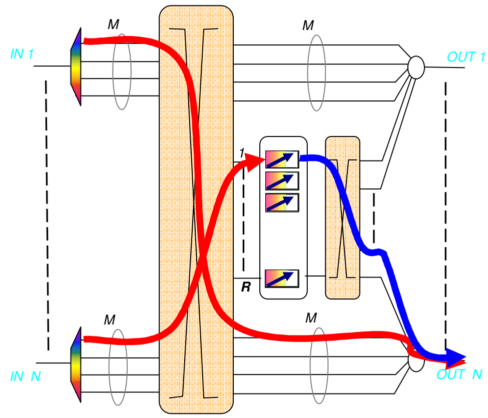
- Not all channels without WC can be exploited to transmit packets
  - **loss is higher than in the full wavelength conversion case**

Example:
M=6
R =3
G = 2 < M – R



Shared-Per-Node (SPN) Scheme

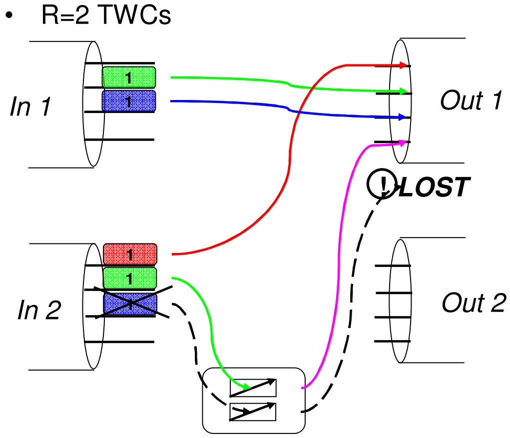
- Optical switch with N input/output fibers with M wavelength channels
- ***R*** (< NM) TWCs shared by all input channels
- Only the packets requiring conversion are sent to the WC pool
- Further optical switching stage required to reach the target output link



Strictly Non-Blocking (SNB) Space Switching Matrix

Case 1: Packet Loss Due To Output Blocking

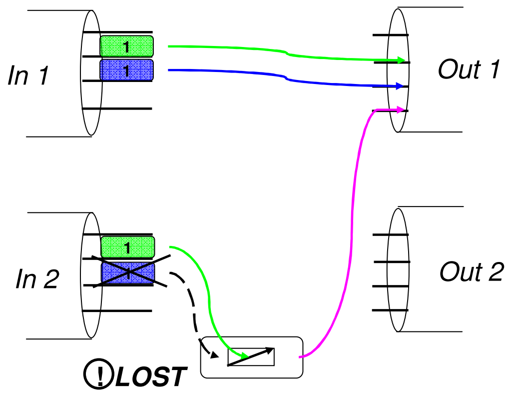
- N=2 input/output fibers
- M=4 wavelengths per fiber
- R=2 TWCs



- 5 packet arrivals directed to output fiber 1 in a time slot
- One packet from each different wavelength is firstly sent without conversion
- Further packets are sent after wavelength conversion
- If destination output fiber is congested, packet is lost
- *If destination output fiber is congested, packet is not sent to TWC pool*

Case 2: Packet Loss Due To Lack of WCs

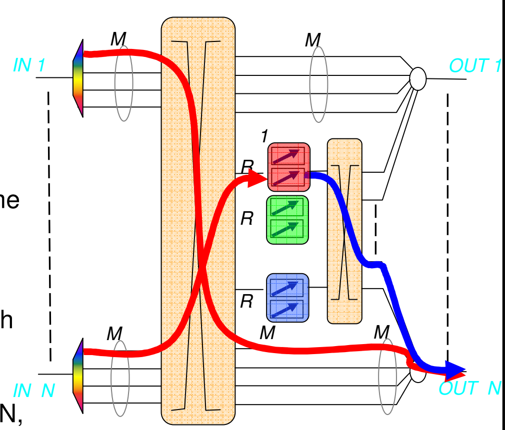
- N=2 input/output fibers
- M=4 wavelengths per fiber
- R_N=1 TWC



- 4 packet arrivals directed to output fiber 1 in a time slot
- First, one packet from each different wavelength is sent without conversion
- Then, other packets are sent exploiting wavelength conversion
- If no TWC is available, packet is lost even if there are available wavelengths on the output fiber

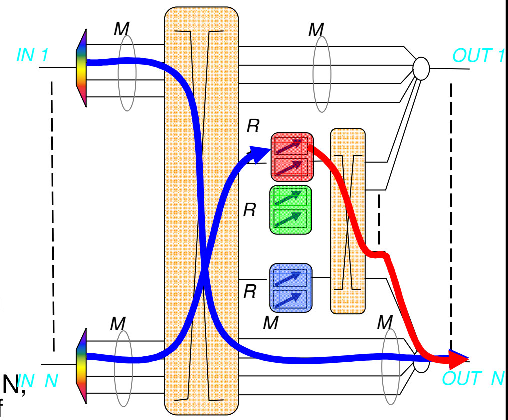
Shared-Per-Input-Wavelength (SPIW) Scheme

- Optical switch with N input/output fibers with M wavelength channels
- WCs partitioned among the input wavelengths
  - M groups of WCs
- ***R*** (< N) WCs shared among the same input wavelength
  - MR in total
- **FTWCs** can be used, given that each WC has a fixed input wavelength
- Same switching fabrics as in the SPN, whose size depends on the number of WCs



Shared-Per-Onput-Wavelength (SPOW) Scheme

- Optical switch with N input/output fibers with M wavelength channels
- WCs partitioned among the output wavelengths
  - M pools of WCs
- ***R*** (< N) WCs shared among the packets forwarded to the same wavelength
  - MR WCs in total
- **FWCs** can be used, given that each WC has a fixed output wavelength
- Same switching fabrics as in the SPN, their size depends on the number of WCs



Outline

- Contention Resolution in Optical Switches
- Buffer-less architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- Buffered architectures
- Conclusions

A Simple Analytical Model of SPN loss

- Hypothesis:
  - Synchronous environment, packet length equal to time slot duration
  - Bernoulli independent arrivals on input wavelengths, with load ***p***
  - Packet arrivals uniformly addressed to the N output fibers (probability 1/N)
- Variables:
  - ***p***: arrival probability on an input channel in a time slot
  - P_u: probability that the output fiber is congested and packet discarded (output blocking)
  - P_b: probability that the packet is blocked on its wavelength (wavelength blocking)
  - A_{wc}: traffic offered to WC pool from each wavelength
  - P_{bwc}: probability that packet lost due to WC unavailability (internal blocking)
  - P_{loss}: overall packet loss probability

Expression of Packet Loss Probability

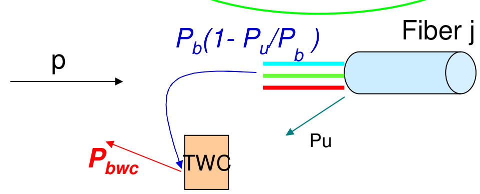
Packet loss due to output contention

$$P_{loss} = P_u + P_b\left(1 - \frac{P_u}{P_b}\right) P_{bwc}$$

Packet loss due to lack of WCs

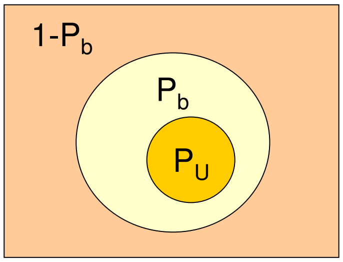
Probability that a packet requires conversion

- Probability that a packet requires conversion, joint probability of:
  - Pb, wavelength blocking
  - 1-Pu/Pb, packet not blocked on the output fiber given that it is blocked on its wavelength



Considerations on Pu and Pb

- Pu: probability that the output fiber is congested and packet discarded (output blocking)
- Pb: probability that the packet is blocked on its wavelength (wavelength blocking)



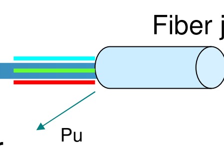
$$P_U \subset P_b \qquad P(U,b) = P_U$$

$$P(U,b) = P(U|b)P_b = P(b|U)P_U$$

$$P(b|U) = 1$$

$$P(U|b) = \frac{P_U}{P_b}$$

Expression of P_u



- P_u is evaluated on destination output fiber
- P_u is calculated assuming full wavelength conversion capability
- Up to MN packet arrivals directed to that output fiber, only M are sent
- Packet loss occurs when there are h > M arrivals, and the tagged packet is one of those discarded

$$P_u = \sum_{h=M+1}^{N \cdot M} \left(1 - \frac{M}{h}\right) \binom{N \cdot M - 1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{N \cdot M - h}$$

- Probability of h arrivals is evaluated as the probability of h – 1 arrivals on MN – 1 input wavelengths other than the tagged
- Simple combinatorial formula gives accurate expression of packet loss

Expression of P_b

- P_b is evaluated considering a wavelength "k" on target output fiber
- Up to N packet arrivals carried by wavelength "k" and directed to the target output fiber
- Packet blocked on wavelength "k" when there are h > 1 arrivals and the tagged packet is not the one forwarded without conversion

$$P_b = \sum_{h=2}^{N} \left(1 - \frac{1}{h}\right) \binom{N-1}{h-1} \left(\frac{p}{N}\right)^{h-1} \left(1 - \frac{p}{N}\right)^{N-h}$$

- Probability of h arrivals evaluated as probability of h – 1 arrivals on wavelength "k" in the N – 1 input fibers other than the tagged

Traffic on WC Pool A_{WC}

- It is necessary to evaluate the traffic offered to the WC pool from each wavelength channel
  - Probability that a packet is sent to WC pool: $P_b \cdot \left(1 - \frac{P_u}{P_b}\right)$
  - Load per input wavelength: p
- Traffic on WC pool: $A_{wc} = p \cdot P_b \cdot \left(1 - \frac{P_u}{P_b}\right)$

Expression of P_{bwc}

- Assuming **Bernoulli independent arrivals** on input of WC pool, there are up to MN possible arrivals, each one with probability A_{wc}
- There are R < MN WCs in the bank
- Packet loss occurs when there are h > R arrivals and the tagged packet is one of those discarded

$$P_{bwc} = \sum_{h=R+1}^{N \cdot M} \left(1 - \frac{R}{h}\right) \binom{N \cdot M - 1}{h - 1} (A_{wc})^{h-1} (1 - A_{wc})^{N \cdot M - h}$$

- Probability of h arrivals is evaluated as probability of h – 1 arrivals on MN – 1 output wavelengths other than the tagged
- Hypothesis of independent arrivals leads to overestimation of the packet loss
  - correlation among the number of packets forwarded in different fibers is neglected
  - packets already forwarded in a fiber, means less packets forwarded to other fibers, this is not considered
- anyway, the approximation obtained is good when NM is high

Packet Loss Probability: Special Cases

- Full wavelength conversion (R=MN):
  - no packet loss on TWC bank

$P_{bwc}=0 \Rightarrow P_{loss} = P_u$

- No wavelength conversion (R=0):
  - packets requiring conversion are lost

$P_{bwc}=1 \Rightarrow P_{loss} = P_u+P_b(1-P_u/P_b)P_{bwc} = P_b$

Considerations on the model

- Analytical model proposed for SPN architecture is very flexible
- The expression of the overall packet loss is valid in general
- The sharing scheme only influences the loss at the WC pool(s), P_{bwc}
- The model can be used to evaluate packet loss in other sharing schemes, updating the expression of P_{bwc}
- Here, the analytical model is used to evaluate loss performance of SPIW scheme

Expression of P_{BWC} for SPIW Scheme

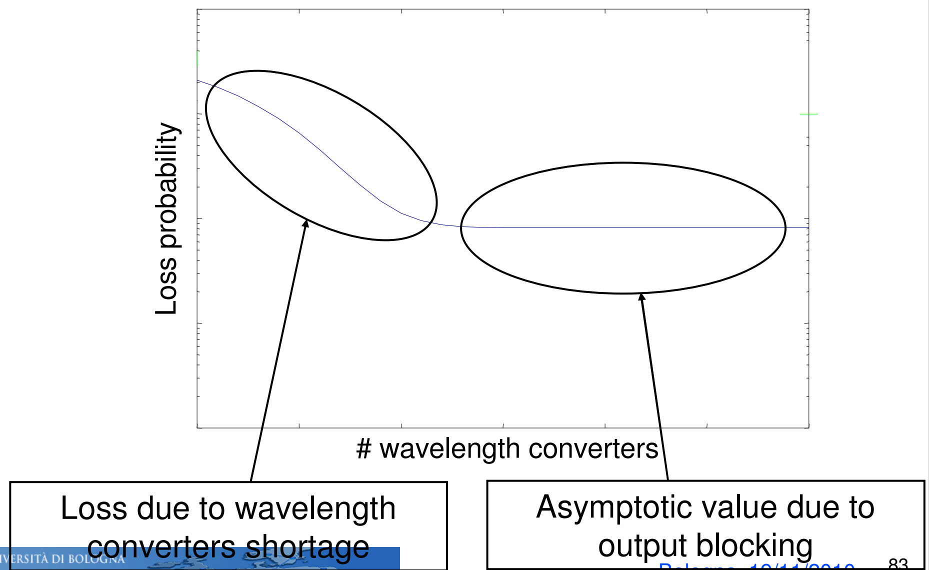
- Considering SPIW, only P_{bwc} has to be changed according to the new sharing scheme
- In SPIW, up to N packets carried by same wavelength contend among each other for only R FTWCs
- At a given WC pool, loss occurs when there are h > R arrivals and the tagged packet is one of those discarded

$$P_{BWC} = \sum_{h=R+1}^{N} \left(1-\frac{R}{h}\right)\binom{N-1}{h-1}(A_{WC})^{h-1}(1-A_{WC})^{N-h}$$

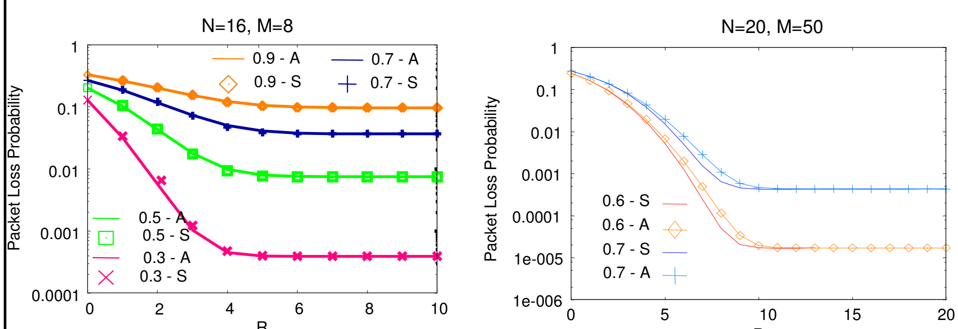
Up to N packets contending for R FTWCs

- In this case, the hypothesis of independent arrivals leads to less precise results, due to the lower number of channels considered (N instead of NM)

Loss probability behavior

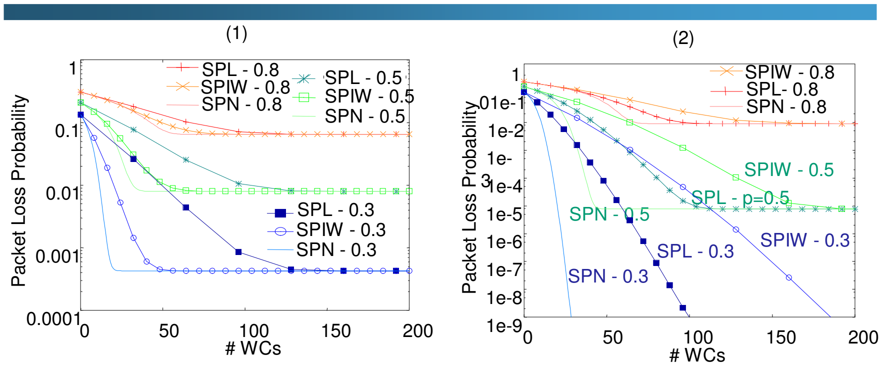


Model Validation for SPIW



- Comparison among analysis (A) and simulation (S) for SPIW
- Analytical model provides very good results for N high and M slow, while the model slightly overestimates the packet loss when N low and M high
  - due to hypothesis of independent arrivals at WC pools

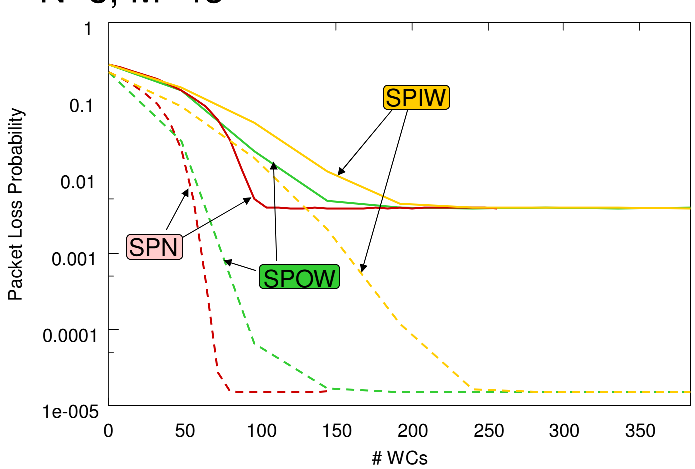
Comparison among SPN, SPL, SPIW



- Example: packet loss for SPN, SPL, SPIW as a function of the number of WCs varying load, as in the case:
  - 1) N=32, M=8
  - 2) N=8, M=32
- Same asymptotic value of the packet loss due to output blocking
- Relative performance between SPL and SPIW depends on switch configuration

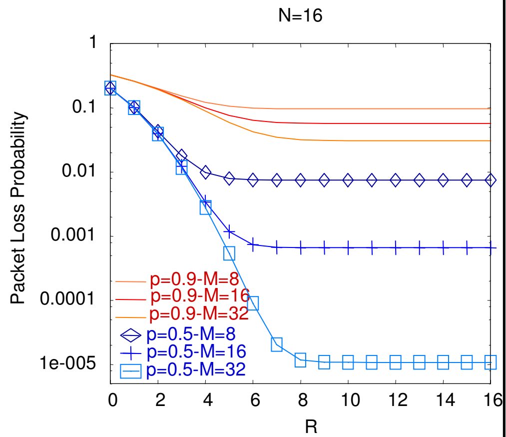
Further comparisons

- N=8, M=48



Effect of M on SPIW packet loss

- Set-up: N=16, M=8, 16, 32
- Packet loss probability as a function of number of WC per wavelength varying the number of wavelengths per fiber
- Packet loss greatly improves by increasing the number of wavelengths



Outline

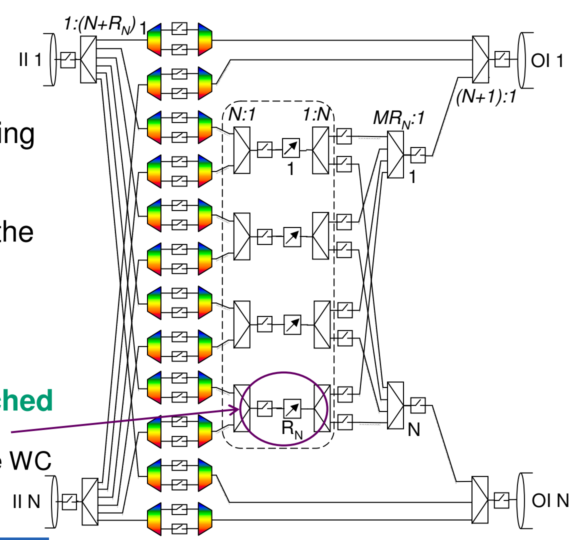
- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- **Practical architectures**
- Physical path analysis
- Buffered architectures
- Conclusions

Practical architectures

- The proposed sharing schemes must be realized taking available optical technology into account
  - spitters/coupler, MUX/DEMUX, switching gates like Semiconductor Optical Amplifiers (SOA) switch…
- Different implementations can be proposed, based on broadcast-&-Select (B&S), wavelength routing (using Arrayed Waveguide Gratings, AWGs), space diversity and so on
- B&S solutions based on optical gates (SOA or MEMS technologies as example) are presented

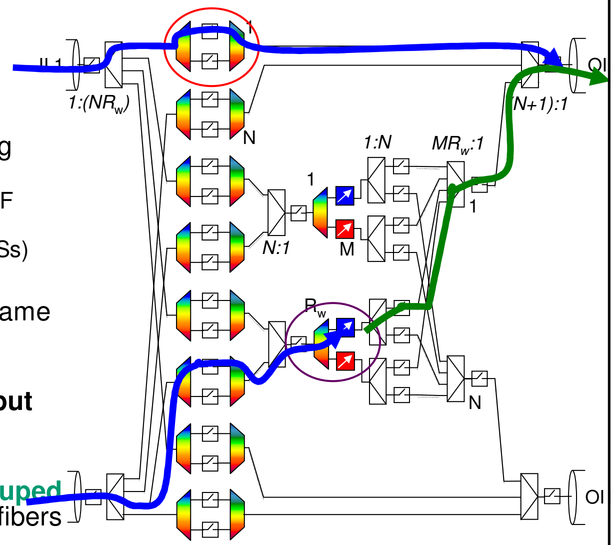
SPN Architecture

- **N** input/output fibers, **M** wavelengths
- B&S SOA-based switching matrix
- WCs shared among all the input channels
- **R** TWCs
- Each **WC must be reached by all the input fibers**
  - N WSs to reach a single WC



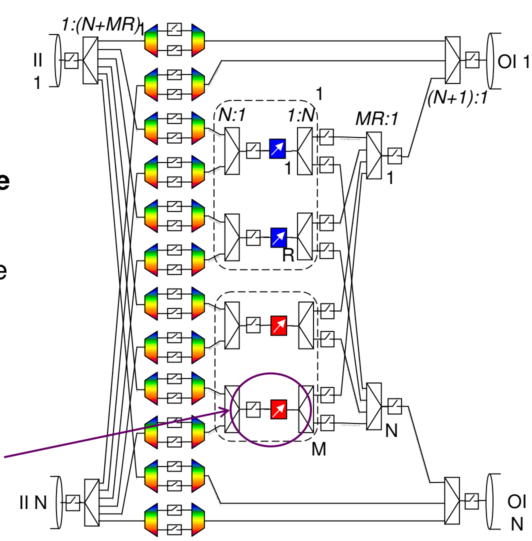
SPIW Architecture

- **N** input/output fibers, **M** wavelengths
- B&S SOA-based switching matrix
  - SOA employed as ON/OFF gates
  - Wavelength Selectors (WSs)
- **R** WCs dedicated to the same input wavelength
- **Fixed-input/tunable output** WCs
- Allows **M WCs to be grouped** and reached by the input fibers in a simple way
  - N WSs to reach M WCs



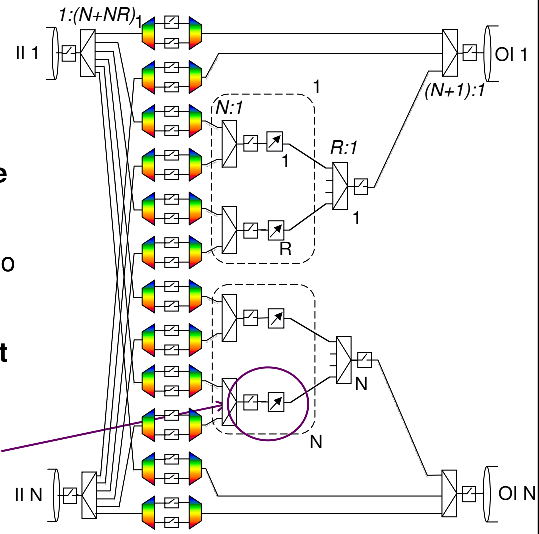
SPOW architecture

- **N** input/output fibers, **M** wavelengths
- SOA-based switching matrix
- WCs **shared among the same output wavelength**
- **R$_w$** WCs dedicated to the same output wavelength
- **Tunable-input/fixed-output** WCs
  - simpler devices
- Each WC must be reached by all the input fibers
  - N WSs to reach a single WC



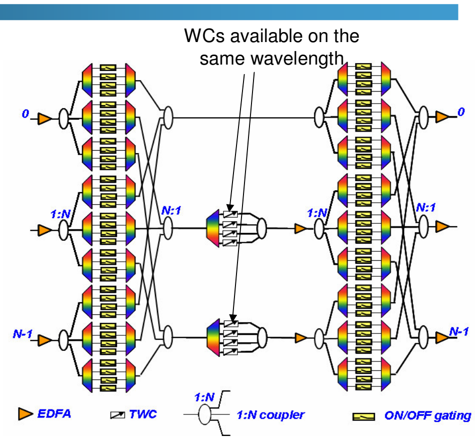
SPL architecture

- **N** input/output fibers, **M** wavelengths
- SOA-based switching matrix
- WCs **shared among the same output link**
- **R** (0 ≤ R ≤ M) WCs dedicated to the same output link
- **Tunable-input/tunable-output** WCs
- Each WC must be reached by all the input fibers
  - N WSs to reach a single WC



Multi-stage SPIW

- To improve scalability
- Same logical performance as SPIW
- B groups of wavelength converters
- N-B direct fibers
- (N+1) links for splitters and couplers



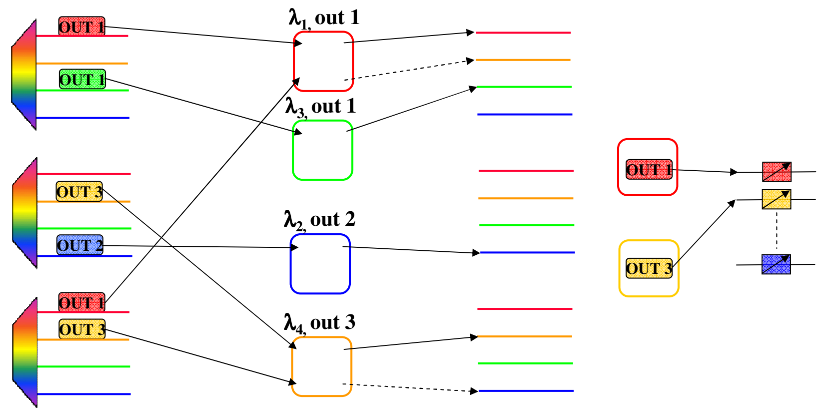
Scheduling Algorithms

- Sharing schemes require proper scheduling algorithms to solve contentions
  - in both synchronous and asynchronous scenarios
- In asynchronous scenario, when a packet arrives the scheduling needs to assign resources to that packet
- In synchronous scenario, the scheduling algorithm must assign the resources to all incoming packets in that time slot
  - must be executed in a (fraction of) time slot duration
  - computational complexity as low as possible (possible parallelization over fibers/wavelengths
- <u>OPTIMAL</u> scheduling algorithm: able to forward the maximum number of packets
  - minimum packet loss

Heuristic Scheduling Algorithms

- Synchronous context
- 3 different phases (executed in each time slot):
  - 1) packets carried by the same wavelength and directed to the same fiber are grouped in a common set
    - packets in different sets are contention free, packets in the same set contend for the same output channel
  - 2) packets that do not need wavelength conversion are directly sent to the output fibers
    - one packet per set, randomly chosen, is sent to the related output channel
    - packets exceeding the capacity of the output fibers (M channels) are discarded due to OUTPUT BLOCKING
  - 3) packets that need conversion are sent to the output fibers:
    - according to the WC availability, the remaining packets in the sets are sent to the proper WC pool
    - packets blocked due to WC unavailability are discarded due to INTERNAL BLOCKING
- The first two phases are common to all sharing schemes, the **third phase** is strictly **related to the WC sharing** strategy

Scheduling Algorithm for SPIW Scheme



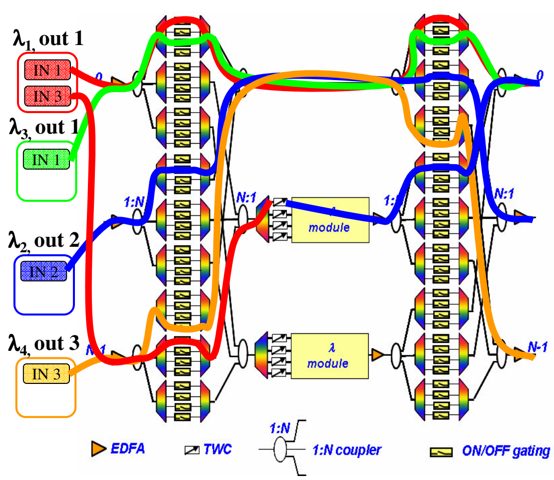
N=3, M=4

PHASE 1: O(NM)

PHASE 2: O(NM)

PHASE 3: O(N+MB)

Sample scheduling in multi-stage SPIW



Complexity (Number of Optical Devices)

- Each sharing scheme requires a different organization of the optical devices
- Different architectures must be compared in terms of complex (and expensive) optical devices employed
- Here, we consider WCs and SOAs, which are active components
- Complexity in terms of WCs and SOAs employed

	total # WCs	total # SOAs (switching purposes)
SPN	$C_{SPN} = R$	$G_{SPN} = N(NM + (M + 1)C_{SPN})$
SPIW	$C_{SPIW} = MR$	$G_{SPIW} = N(NM + 2C_{SPIW})$
SPOW	$C_{SPOW} = MR$	$G_{SPOW} = N(NM + MC_{SPOW})$
SPL	$C_{SPL} = NR$	$G_{SPL} = N(NM + MC_{SPL})$

SPIW requires less SOAs when architectures equipped with the same number of WCs

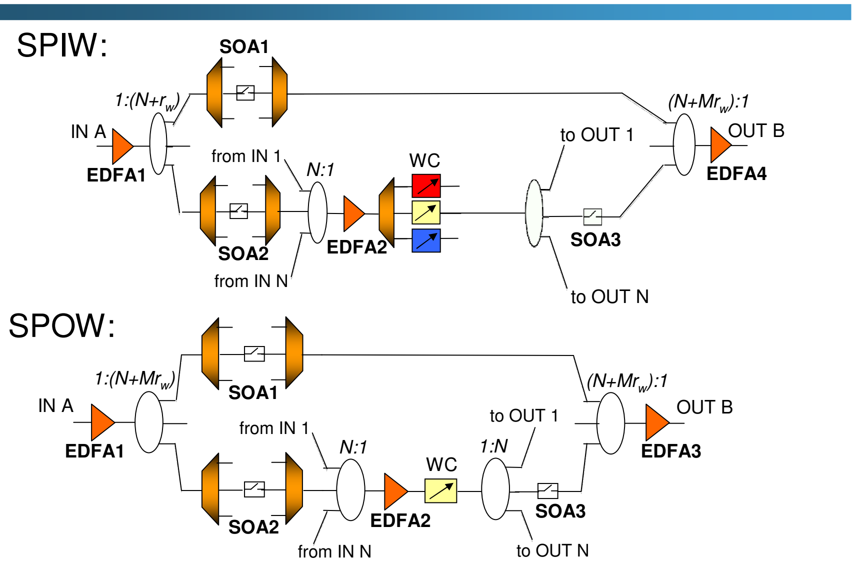
Advantages and Disadvantages of Sharing Architectures

	# WCs	*Kind of WCs*	*#SOAs*	*performance/ complexity trade-off*
SPN	lowest	TWCs (complex)	highest	fair
SPL	high (especially when N high, M low)	TWCs (complex)	near to SPN with some save	not good
SPIW	good (when N high, M low) fair in other cases	FTWCs (easier than TWCs)	lowest	good (when N high, M low)
SPOW	low (near to SPN)	FWCs (easiest)	near to SPN with some save	good

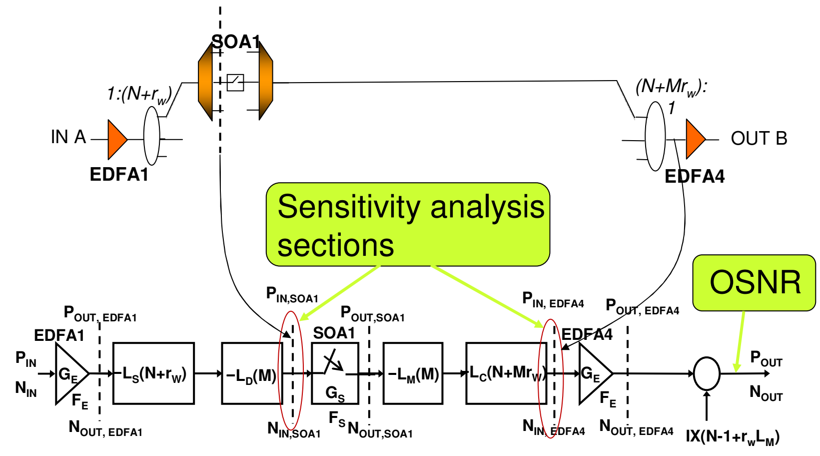
Outline

- Contention Resolution in Optical Packet Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- **Physical path analysis**
- Buffered architectures
- Conclusions

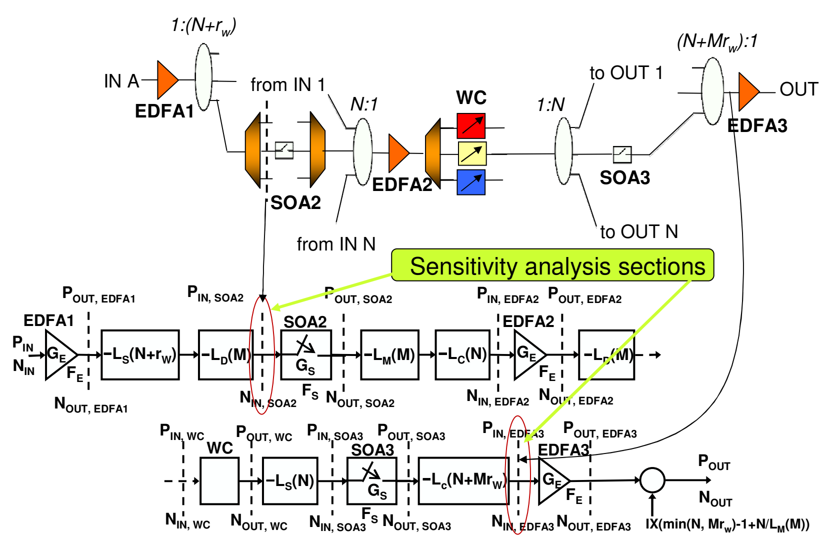
SPIW and SPOW paths



SPIW path without wavelength conversion



SPIW path with wavelength conversion



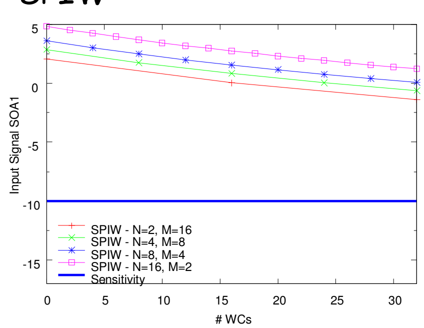
Numerical results

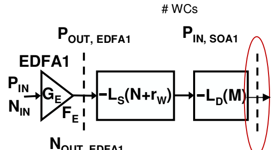
- Reference parameters:
  - EDFA: $\mathbf{P}_{OUT,\,E}$ = 25 dBm
  - SOA: $\mathbf{P}_{OUT,\,S}$ = 8 dBm, $\mathbf{S}_S$ = -10 dBm, $\mathbf{F}_S$ = 7 dB, $\mathbf{E}_R$ = 35 dB
  - WC: $\mathbf{P}_{OUT,WC}$ = 3dBm, $\mathbf{S}_{WC}$ = 0 dBm
  - $OSNR_T$ at the receiver: 20 dB
- Single channel analysis
- Amplifier requirement
  - input power > sensitivity
- Receiver requirement
  - $OSNR > OSNR_T$
- Interfering sources are accounted for as additional noise at the receiver input

Sensitivity analisys: NM=32

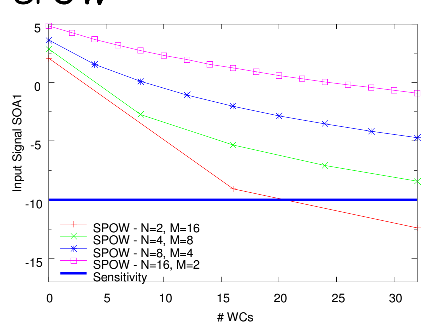
Power at SOA1 input decreases with the number of wc

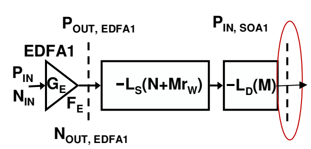
•SPIW





•SPOW

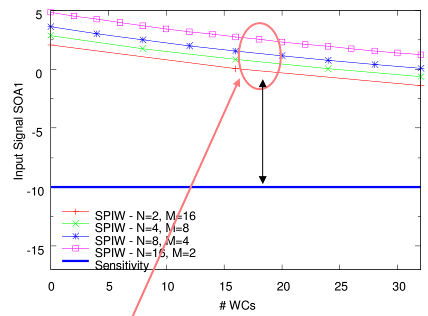




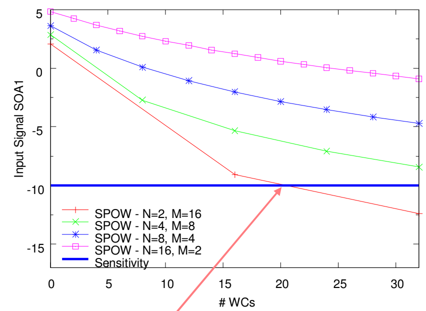
Sensitivity analisys: NM=32

Power at SOA1 input decreases with the number of wc

- SPIW



- SPOW



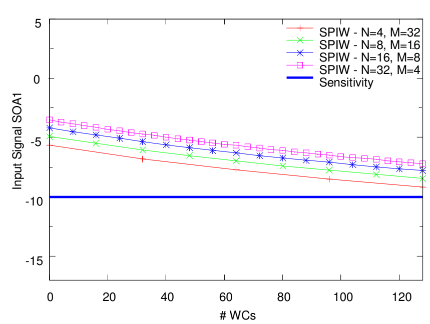
Input power is too low

SPIW always match sensitivity constraints

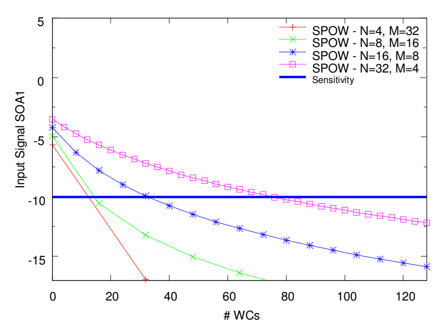
Sensitivity analysis: NM=128

-

- SPIW



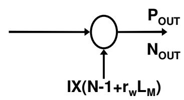
SPOW

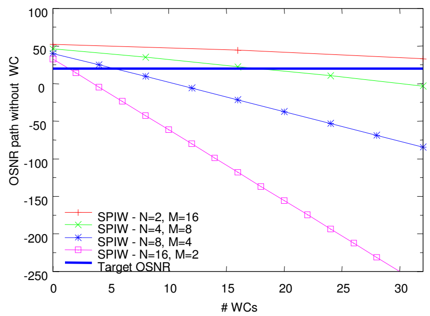


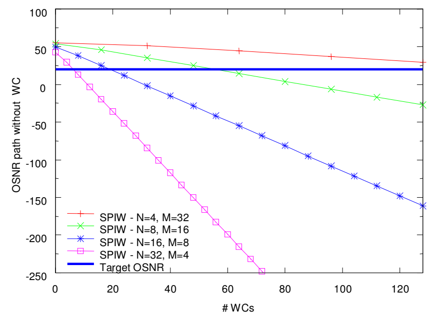
- Major limitations as the total number of channel increases
- Poor contention resolution capability in SPOW

OSNR analysis: path without wc

N=32 N=128

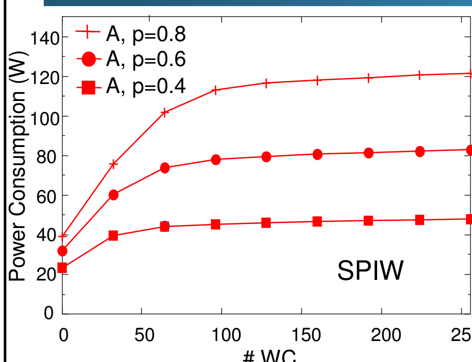


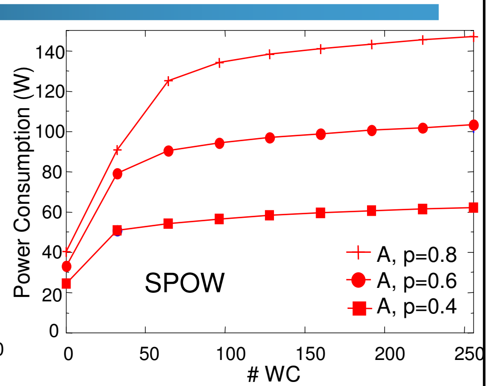




- Strong dependence on the number of wavelength converters when N is high due to interference effects

Power consumption in SPIW and SPOW



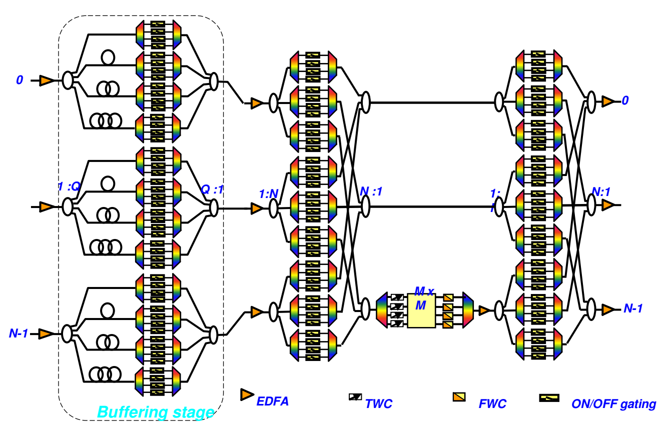


- Considering actual power loss in passive devices, the SPOW consumes more power than SPIW (when equipped with the same number of WCs and high load)
- The dimensioning process should take not only loss but also power consumption (and scalability analysis) into account

Outline

- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- **Buffered architectures**
- Conclusions

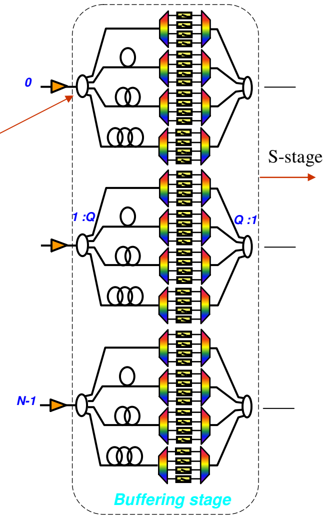
The multi-stage architecture with input buffers



- A queuing stage based on the broadcast and select principle is added
- The behaviour of the S-λ-S sub-system is the same as before

The queueing stage

- A power coupler generates Q copies of the multi-wavelength bundle of channels
  - D = k T, k=0..Q-1; T is the packet time
- Each FDL is followed by a wavelength selector
  - Optical packets are available at the first S-stage with all possible delays D
  - Input queuing scheduling can be applied



Queue scheduling

- Slot by slot operation
- Packets on the same input channel are organized in a list
- FIFO
  - Lists are served on a FIFO basis
  - Round robin service among lists
- Window
  - Allow to overcome the HOL phenomenon
  - If packets are present in the list, within a window of size W, which belong to a contention-free path, they are considered for forwarding

Scheduling of packet forwarding through the switch aims at minimizing wavelength conversion

Queue Scheduling pseudo code (FIFO)

```
step 1:
for (k=0; k < N; k++)
  for (j=0; j < M; j++)
    if (there is a packet carried by λj on input fiber k)
      the packet is inserted on the tail of list Lk^j;

step 2:
for (k=0; k < N; k++)
   for (j=0; j < M; j++)   {
     in=(k+RRF) mod N;
     λ=(j+RRW) mod M;
     if (Lin^λ != NULL)   {
        out=output fiber of the packet on the head of  Lin^λ;
        tx=sel_lambda(λ, out);
        if (tx>0)
          the packet is forwarded and removed from the head of  Lin^λ;
        if ( (tx== -1) && (T - Cin^λ == Q -1) )
          the packet is removed from the head of Lin^λ;
     }
   }
RRF=(RRF+1) mod N;
if (RRF==0)      RRW=(RRW+1) mod M;
```

Step 1: lists are formed for each channel

Step 2:

a packet is extracted from the head of each list (FIFO)

RRF and RRW are counters that assure fairness

Sel_lambda is a procedure to find a path through the switch

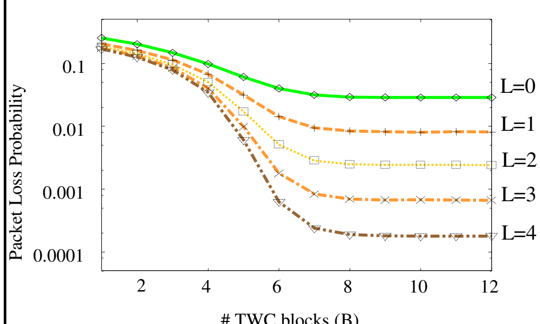
Quality of service

- Input queuing allows also quality of service differentiation in the node
- A simple QoS algorithm is considered to manage two QoS classes
  - The scheduling algorithm considers high priority packets first, if present in the list
  - Low priority packets are then searched for, otherwise

Simulation set up

- Bernoulli independent arrivals on input channels
- Uniform addressing scheme
- FIFO-RR and W-RR scheduling
- Confidence interval at 95%
  - Less than or equal to the 5% of the mean

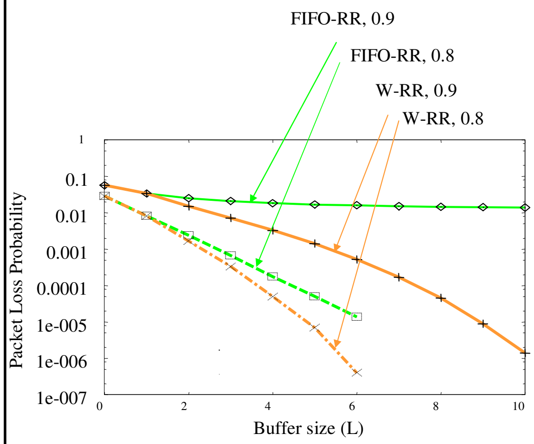
Packet loss improvement



- N=16, M=16
- FIFO-RR
- p=0.8
- L buffer size
- Remarkable improvement when internal block is overcome (asymptotic region)

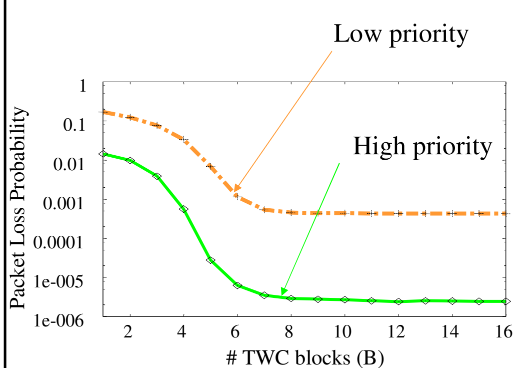
Output contention resolution in time domain

Effects of the queue scheduling



- N=16, M=16
- B=N
- FIFO-RR,W-RR
- p = 0.8 and 0.9
- The benefit becomes more evident as buffer length increases

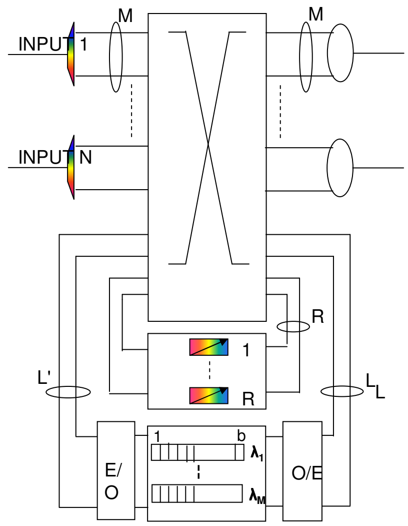
Quality of service aspects



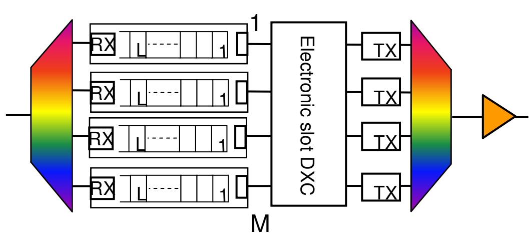
- N=16, M=16
- L=5
- p=0.8
- Remarkable QoS differentiation even with low buffer values

Single stage with recirculation shared buffer

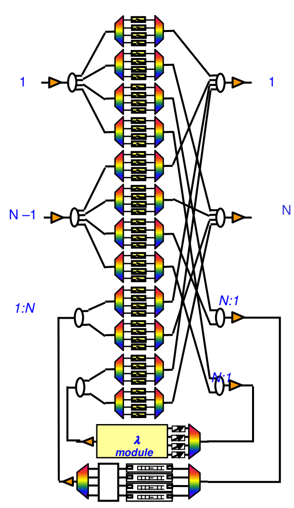
- N input/output fibers
- M wavelengths per fiber
- Strictly non-blocking space switching fabric
  - (NM+R+L') x (NM+R+L)
- R Full Range Tunable Wavelength Converters (TWCs) shared per node
- Queuing stage with M queues (one per wavelength), with size b
  - FDLs
  - ***Electronic buffers***



Practical implementation of the hybrid switch

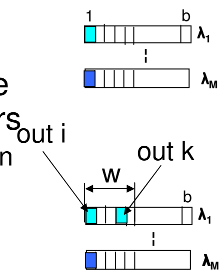


- Each buffer is used by the same wavelength

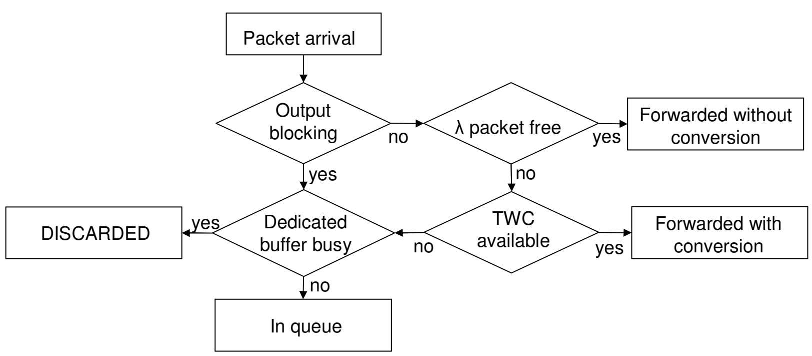


Queuing policies

- Hp: In each time slot, packets in the queues are served before packets from input fibers
- Two policies for buffer management:
  - FIFO: only the Head Of Line (HOL) packets (up to M) are served
  - Windowing: w packets in each queue can be forwarded, if directed to different output fibers
    - Window size: w=b, up to M*max(b, N) packets can be forwarded



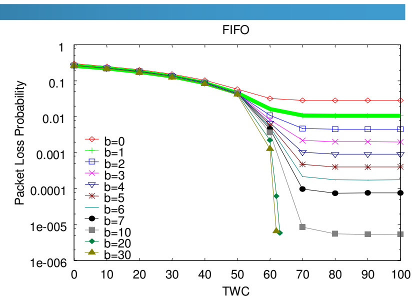
Packet forwarding

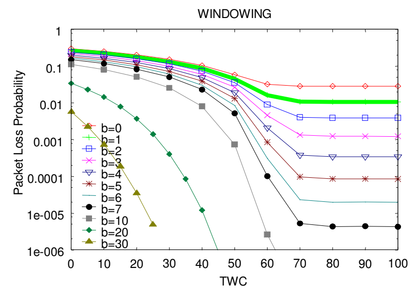


- The packet is sent without conversion, if possible
- Otherwise the packet is sent with conversion
- If output fiber is congested or no TWC is available, the packet is stored in the queue
- The packet is lost when its queue is busy

Packet loss probability

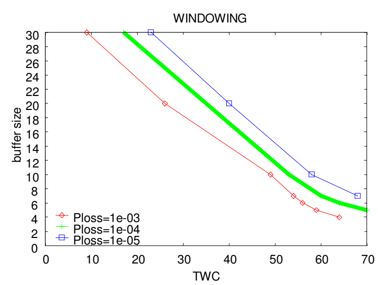
- Ploss = *f*(TWC), varying buffer size b,
- N=16, M=16, L=L'=256, p=0.8
- FIFO: improvement when buffer size increases only if enough TWCs are available
  - Due to the saturation of the queues
- Windowing: relevant improvement of the throughput in all regions
  - Small number of TWCs can be enough to assure low packet loss

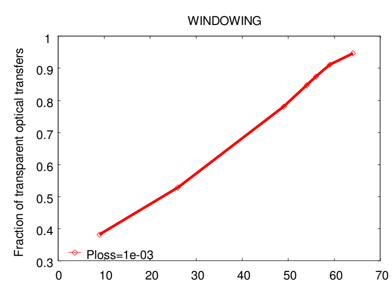




Switch dimensioning

- b=*f*(TWC), to obtain a target PLP
- N=16, M=16, L=256, p=0.8
- Different couples (b, TWC) can be chosen to obtain the same packet loss
- The percentage of packets forwarded without O/E conversion decreases as the number of TWCs decreases
- If a certain percentage of packets must be forwarded in optics, a certain number of TWCs must be employed





Outline

- Contention Resolution in Optical Switches
- Buffer-less Architectures
- Logical performance evaluation
- Practical architectures
- Physical path analysis
- Buffered architectures
- Conclusions

Conclusions

- Different sharing schemes lead to different technology to be employed
  - different complexity
  - different cost
- Difficult to find an architecture which is the best solution for all switch dimensioning
  - the best architecture must be selected according to N and M and traffic context
- In general, SPIW can (in most cases) be implemented with a lower number of SOAs and provides less power consumption
- SPOW is equipped with FWCs, which are easier to be implemented
- For this reason, they represent a good alternative to the well studied SPL and SPN architectures

Open points and research opportunities

- Optical switching is feasible but
  - Limitation on the number of channels
  - SOAs are costly and power consuming
  - Very limited storage capability
  - Extremely high capacity and overall costs

Possible solutions:

- Hybrid technology optical switching
  - Fast and slow optical switching subsystems
  - Electrical buffers
  - NEED of managing/programmable capability to match technology performance with transport techniques and quality of service
- Hardware resource sharing
  - network transport technologies
  - service providers
  - custumers

↓

- Node programmability is the emerging concept
- Modular node design
- Data,control,management capabilities

Events in Bologna: February 2011

- **BONE Closing session**
  - Farnese Chapel, City hall (Palazzo d'Accursio), February 7th
- **Optical network Design and Modeling (ONDM 2011)**
  - A peer-reviewed international conference
  - Faculty of engineering, February 8th-10th
- **Workshops**
  - **Workshop 1** - Building the Future Optical Network in Europe: Key final outcomes of the EU BONE project, February 8th, Faculty of engineering, 14-16 p.m
  - **Workshop 2** - Control plane evolution in metro and core networks, February 9th, Faculty of engineering, 16-18 p.m
- **Information available at** http://www.ondm2011.unibo.it/
  - To participate to any of these events write an e-mail to
  - carla.raffaelli@unibo.it
  - Registration of students of Master Degree in Telecommunications is free!

Part III:

Towards new network concepts: programmable router and experimental activities

Prof. Walter Cerroni

Motivations

- New transport paradigms are needed for Future Internet
  - User- , service- and content-centric
  - Integrating heterogeneous virtualized resources on top of a common physical infrastructure
- Current network-centric approach provide
  - Technology-dependent transport
  - Semi-static service provisioning
  - Limited knowledge of service requirements
- Need for **built-in** network functionalities capable of **dynamically** providing **on-demand virtual** communication resources based on **high-level service needs**

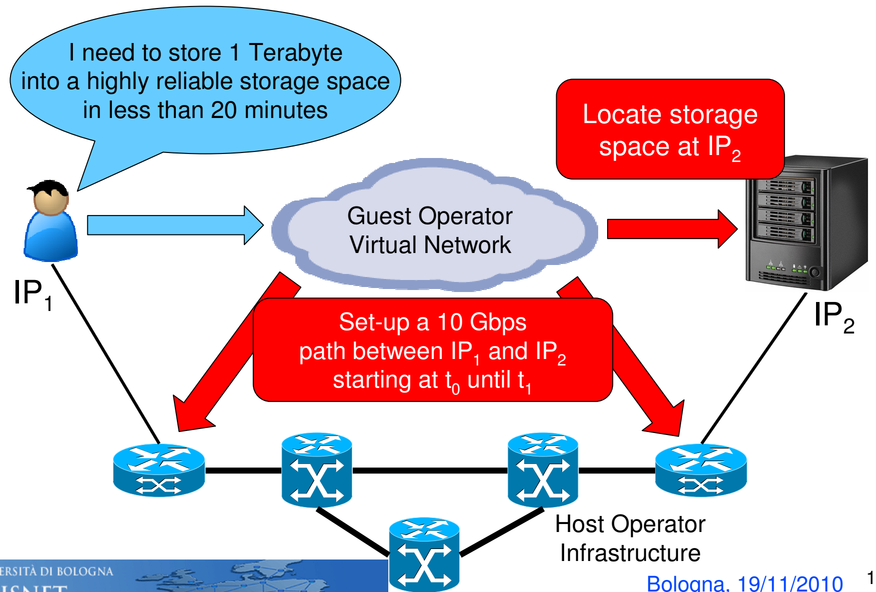
Key features

- **Node and Network Programmability**
  - Open and accessible control of switching/routing facilities and other network functions
  - Infrastructure owned by a **Host Operator (HO)** and dynamically configured by multiple **Guest Operators (GOs)** to offer different on-demand connectivity services to their customers
  - Scalable and cost-effective resource sharing solution
  - Security and reliability issues

Network programmability background

- Original idea from the late '90s
  - oriented to traditional IP networks
  - *IEEE Communications Magazine, Vol. 36, No. 10, October 1998*
  - *IEEE Networks, Vol. 12, No. 3, May/June 1998*
- Two approaches
  - **Programmable networks**
    - set of open and standardized programming interfaces that allowed customer applications to activate and manage services by reconfiguring low-level routing and switching resources
  - **Active networks**
    - more radical approach by allowing each packet to carry not only the traditional IP header information (needed for routing and forwarding purposes), but also code fragments to be executed on the nodes in order to customize specific networking functions
- Standardization attempts
  - IEEE P1520 project (never became standard…)

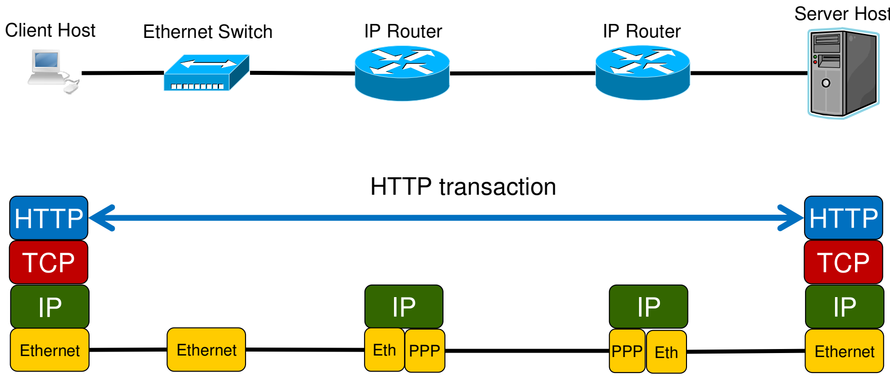
Example of network programmability



Major players

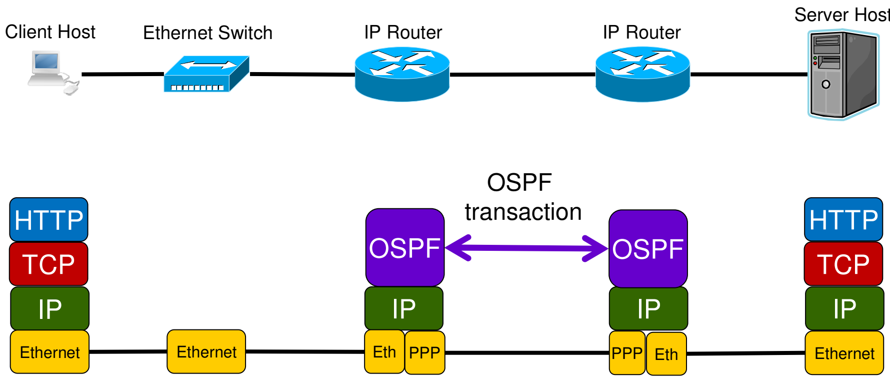
- **Optical Networks**
  - Key role in the physical infrastructure
    - high bandwidth, reduced power consumption, small footprint, etc.
  - Capable of multiple switching granularity based on service needs
    - Circuit (OCS)
    - Packet (OPS)
    - Burst (OBS)
- **Programmable Node Architectures**
  - Key role in the network infrastructure control and management planes

Network planes: data/forwarding plane



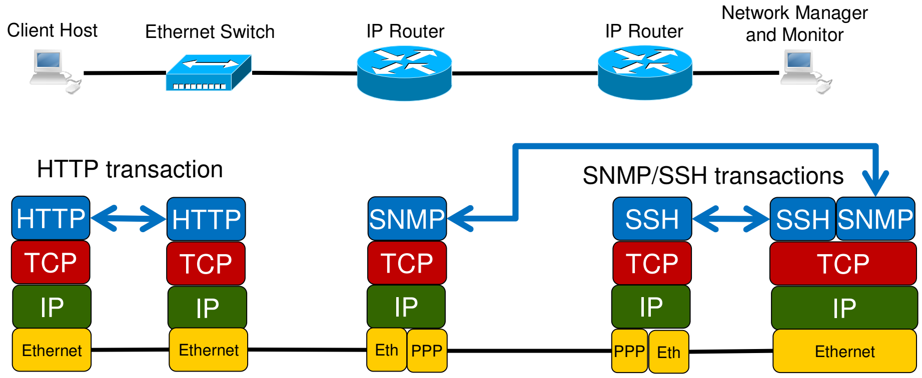
- User data are forwarded between client and server hosts across the IP network
- Network equipment performs lookup and switching or forwarding operations

Network planes: control plane



- Routing information is exchanged between IP routers to build routing tables needed to perform packet forwarding
- Network equipment performs signaling and computation operations

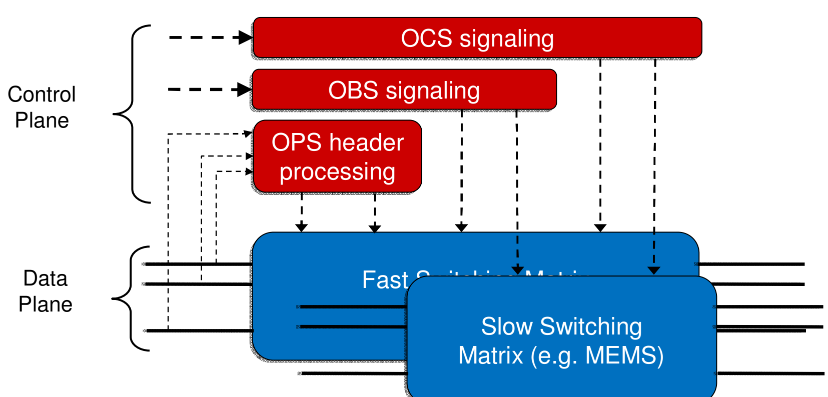
Network planes: management plane



- Management information is exchanged between hosts and network equipment to monitor or configure the network using different management protocols
- Network equipment performs host-like operations

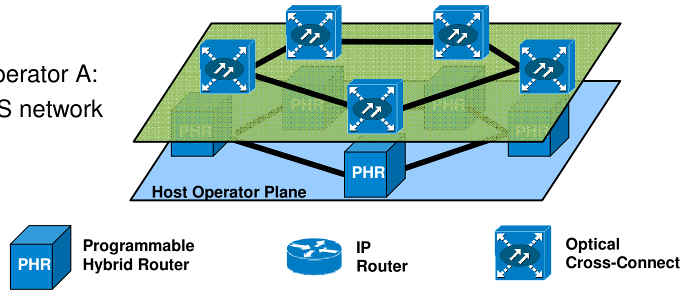
Programmable hybrid optical router

- Multiple switching granularity available on demand
  - Multiple Guest Operators offer different switching services tailored to the customer needs
  - Co-existence of OPS, OBS, OCS within the same infrastructure



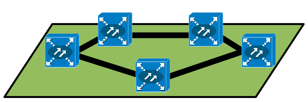
Programmable OPS/OCS hybrid network

Guest Operator A:
Pure OCS network

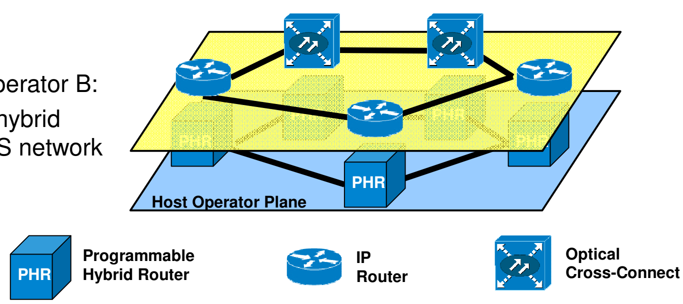


Programmable OPS/OCS hybrid network

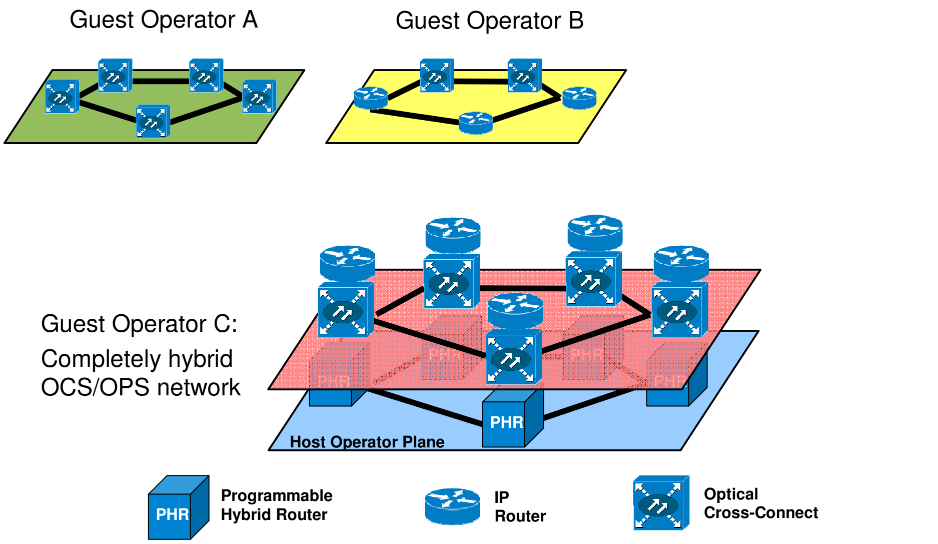
Guest Operator A



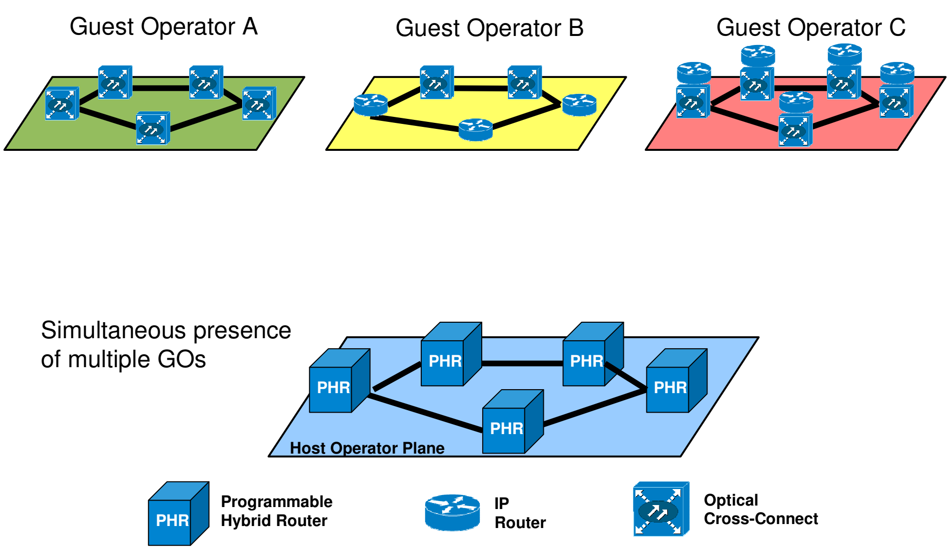
Guest Operator B:
Partially hybrid OCS/OPS network



Programmable OPS/OCS hybrid network



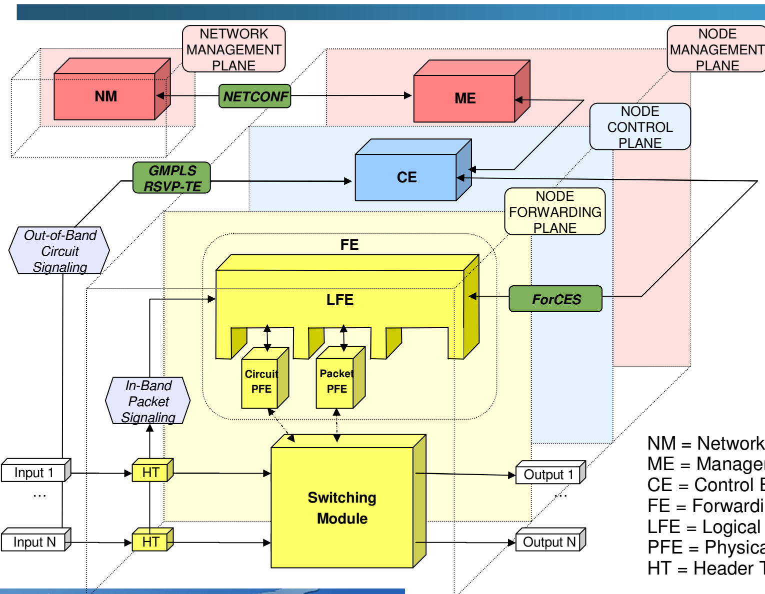
Programmable OPS/OCS hybrid network



New programmable node design approach

- Key concept:
  Separation of control and forwarding functions
  – Based on IETF ForCES framework (RFCs 3746 and 5810)
    - network boxes as multi-vendor systems where control and forwarding subsystems can be developed and can evolve independently
  – Provides the required modular architecture to implement programmable node functions
- Extended with a further separation between
  – High-level, logical forwarding functions
  – Low-level, hardware-dependent device configuration tasks
- Resulting in a modular architecture capable of resource virtualization

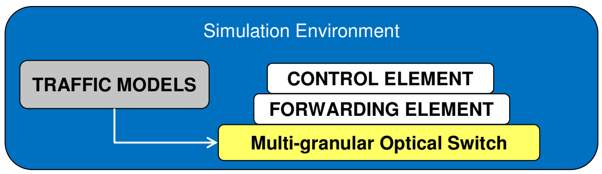
Programmable hybrid router architecture



NM = Network Manager
ME = Management Element
CE = Control Element
FE = Forwarding Element
LFE = Logical FE
PFE = Physical FE
HT = Header Tap

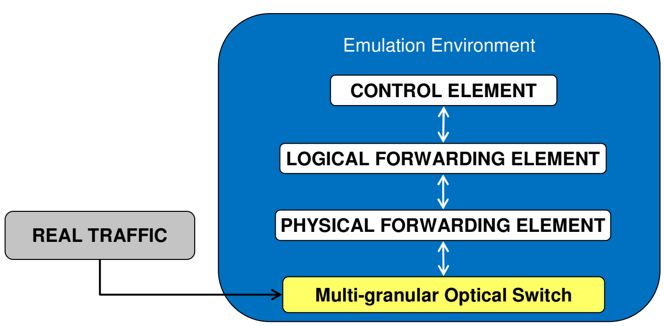
How do we test it?

- Different testing approaches to characterize the system architecture from logical and physical perspectives
  - *Analytical evaluation*: too difficult for complex systems and often based on approximations
  - *Physical implementation*: usually expensive and sometimes unfeasible due to technology limitations
  - *Simulation*: typically adopts a simplified model and provides an abstracted representation of the system
    - e.g. node control design: real interactions between control plane and forwarding plane are often neglected



Why software emulation?

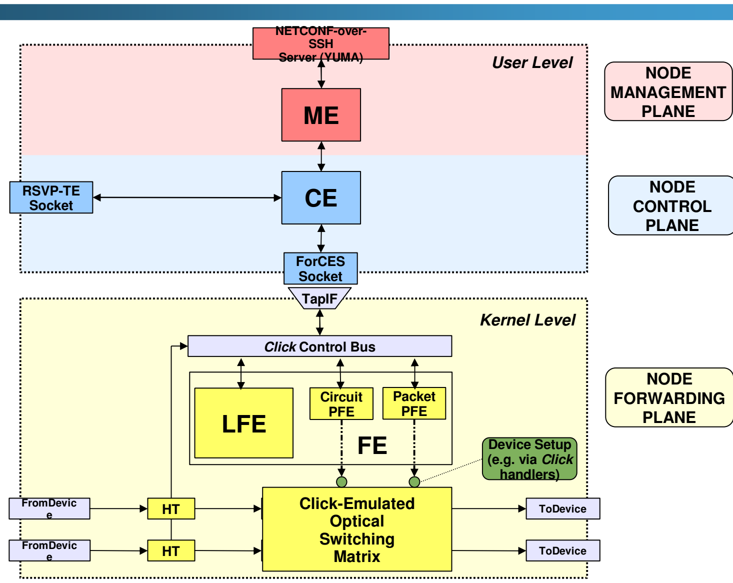
- To go further than typical simulations
  - Implementation of the software components of the control and forwarding planes
  - Emulation of the physical components of the data plane
  - Implementation of their real interacting functions
  - Accounting for as many feasibility aspects as possible



Software router emulation environment

- **Click Modular Router**: http://read.cs.ucla.edu/click/
  - Flexible, modular, fine-grained architecture for implementing software-based routing with full control of the packet flow
  - Build your own IP router: a number of basic elements to perform switching and routing functions are provided
  - Add your own customized features: elements can be extended and new ones can be designed
- Click emulation of programmable optical routers provides
  - Cheap and fast prototyping of all the control and forwarding plane features to be implemented in the real system
  - Modular flexibility to enforce node programmability
  - Emulation of the physical switching operations

Software router configuration



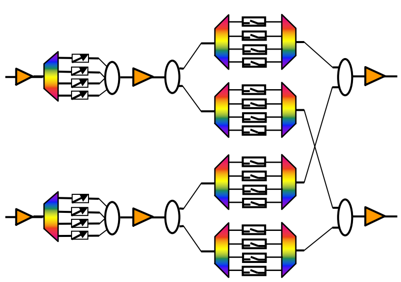
Logical FE general procedure

```
if schedule == 0 then {//new request (from CE)}
    if out_counter[n] < W then {//wavelength free on n}
        select_main_FM(need);
        send_request_to_FM(n, w, m);
    else {//no wavelength free on n}
        drop packet/flow;
    end if
end if
if schedule == 1 then {//request successfully processed}
    out_counter[n]++;
end if
if schedule == 2 then {//req. unsuccessfully processed}
    if alternative_FM == true then {//alt. FMs available}
        select_alternative_FM(need);
        send_request_to_FM(n, w, m);
    else {//alternative FMs not available}
        drop packet/flow;
    end if
end if
```

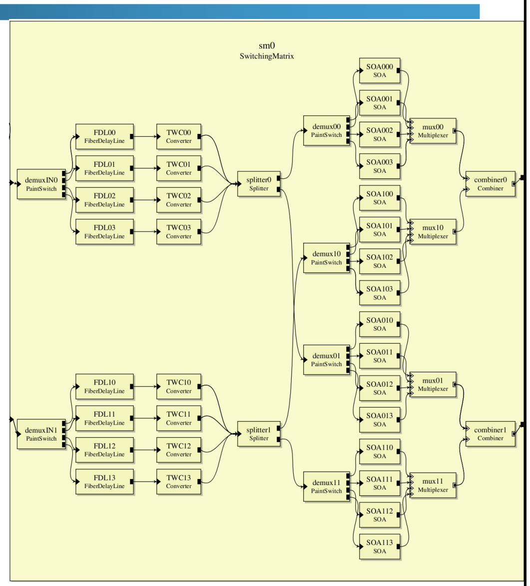
Logical scheduling operations

Notify Physical FEs for device set-up

Example of emulated switching matrix

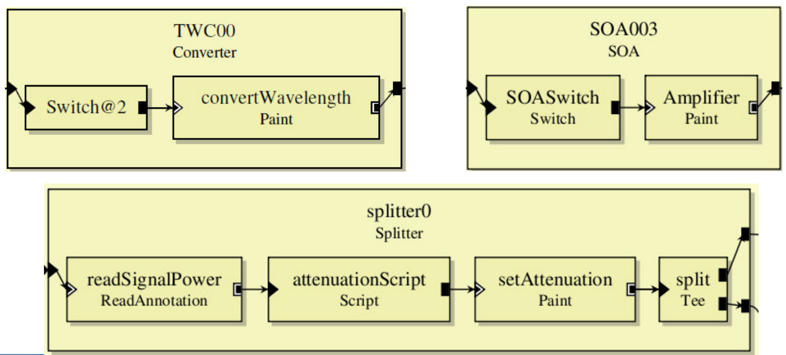


- Broadcast & Select Switch
- SOA-based space switching
- Input tunable wavelength converters



Emulation of the physical layer

- Wavelength multiplexing and signal propagation emulated using Click's *Paint* functions
  - Incoming data packets are marked by the *OpticalSource* elements according to the input wavelength they are supposed to be received on
  - Signal power level and OSNR value are associated to each data packet and modified by each device traversed

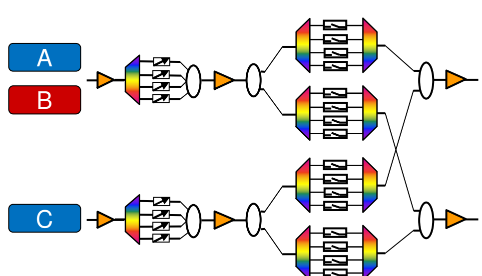


Testing the physical layer emulation

A
```
01 InputFiber(0) - Counter Value: 1, 1264763261.187402
02 InputFiber(0) : Paint=00, DestFiber=0, Power=13
```

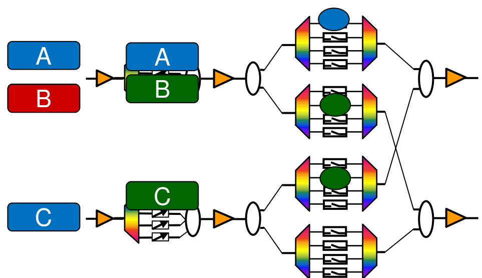
B
```
03 InputFiber(0) - Counter Value: 2, 1264763261.190988
04 InputFiber(0) : Paint=02, DestFiber=1, Power=10
```

C
```
05 InputFiber(1) - Counter Value: 1, 1264763261.192472
06 InputFiber(1) : Paint=00, DestFiber=0, Power=12
```



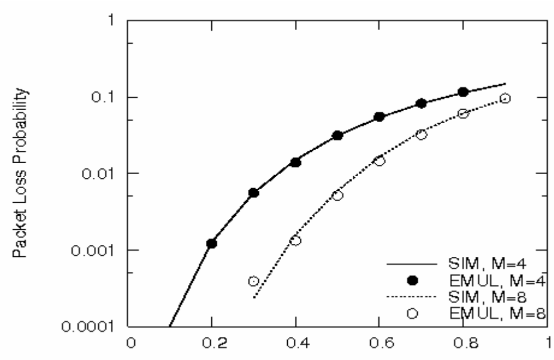
Testing the physical layer emulation

A
```
08 TWC(00)to(0)   1264763262.187860: Paint=00 Power=13
09 SOA(000)       1264763262.187860: Paint=00 Power=0f
10 OutputFiber(0) - Counter Value: 1
```
B
```
11 TWC(02)to(1)   1264763262.190989: Paint=01 Power=10
12 SOA(011)       1264763262.190989: Paint=01 Power=0f
13 OutputFiber(1) - Counter Value: 1
```
C
```
14 TWC(10)to(1)   1264763262.192473: Paint=01 Power=12
15 SOA(101)       1264763262.192473: Paint=01 Power=0f
16 OutputFiber(0) - Counter Value: 2
```



Testing the logical performance

- Synchronous OPS only
- N = 4 I/O fibers, M wavelengths/fiber
- Packet Loss Rate vs. input traffic (N×M Bernoulli sources)
- Comparison with well-established node simulator
- **The emulator correctly schedules the packets**



Testing the multi-granular switching

- Capture of control and management plane traffic during router programming operations

Not authorized

Granted

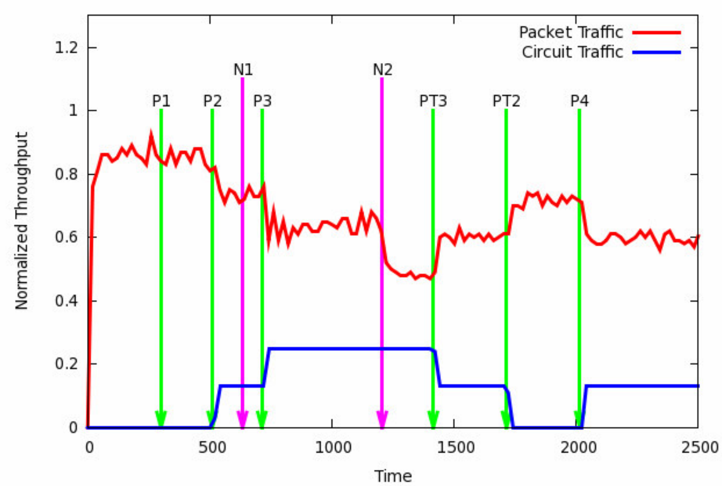
NETCONF commands

Granted

No.	Time	Source	Destination	Protocol	Info
2	300	172.16.2.1	192.168.102.1	RSVP	PATH Message. SESSION: IPv4-LSP,
3	510	172.16.5.1	192.168.105.1	RSVP	PATH Message. SESSION: IPv4-LSP,
4	513	10.50.51.2	10.50.51.1	RSVP	RESV Message. SESSION: IPv4-LSP,
5	634	192.168.10.209	192.168.10.69	TCP	53294 > netconf-ssh [PSH, ACK] S
6	634	192.168.10.69	192.168.10.209	TCP	netconf-ssh > 53294 [PSH, ACK] S
7	634	192.168.10.209	192.168.10.69	TCP	53294 > netconf-ssh [ACK] Seq=34
8	713	172.16.2.1	192.168.102.1	RSVP	PATH Message. SESSION: IPv4-LSP,
9	716	10.50.50.2	10.50.50.1	RSVP	RESV Message. SESSION: IPv4-LSP,
10	1206	192.168.10.209	192.168.10.69	TCP	53294 > netconf-ssh [PSH, ACK] S
11	1206	192.168.10.69	192.168.10.209	TCP	netconf-ssh > 53294 [PSH, ACK] S
12	1206	192.168.10.209	192.168.10.69	TCP	53294 > netconf-ssh [ACK] Seq=68
13	1415	172.16.2.1	192.168.102.1	RSVP	PATH TEAR Message. SESSION: IPv4
14	1715	172.16.5.1	192.168.105.1	RSVP	PATH TEAR Message. SESSION: IPv4
15	2015	172.16.1.1	192.168.101.1	RSVP	PATH Message. SESSION: IPv4-LSP,
16	2018	10.50.50.2	10.50.50.1	RSVP	RESV Message. SESSION: IPv4-LSP,

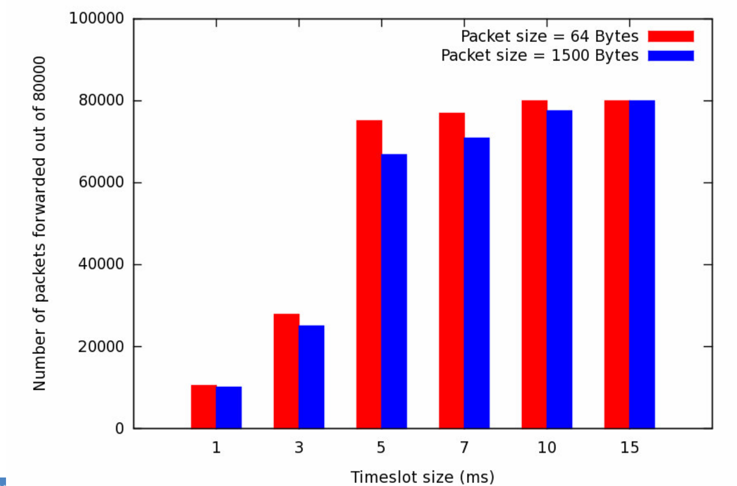
Measuring the throughput

- OPS (Bernoulli sources) + OCS (CBR sources)
- N = 2 I/O fibers, M = 4 wavelengths/fiber



Testing the forwarding scalability

Number of packets being correctly forwarded in a *bar state* configuration on standard PC hardware



Security and reliability issues

- Enforcing security is a two-fold problem
  1. Issues when a GO requests the activation of a given programmable function on a HO node
     GO vs. HO Transactions
  2. Issues when one of the customers of a trusted GO requests to use one of the programmable functions available
     Customer vs. GO/HO Transactions
- Major aspects to be covered
  - Authentication
  - Authorization
  - Integrity
  - Confidentiality
  - Protection and Availability
  - Accounting (not considered here)

Authentication

1. GO vs. HO
   - The GO must authenticate itself before the HO allows it to use its network infrastructure
   - The GO must be sure to talk to the desired HO
   - Reciprocal authentication needed (e.g. using PKI certificates)
   - A **Guest Operator Service Level Agreement (GO-SLA)** must be negotiated and established between each GO and the HO
2. Customer vs. GO/HO
   - Customers of a trusted GO must authenticate themselves when requesting a given service to the GO/HO
   - Service-specific solutions based on the signaling protocols
   - e.g. GMPLS-based multi-service optical network may use standard authentication mechanisms provided by RSVP-TE
     - RSVP-TE messages carry an *Integrity Object* including a sequence number and a SHA-1 message digest with secret keys shared between neighbors

Authorization

1. GO vs. HO
   - A trusted GO is allowed to activate a programmable function only if it is specified in its GO-SLA
   - Possible solution: white list approach, including the trusted GOs and the list of programmable functions available in their GO-SLAs
2. Customer vs. GO/HO
   - An authenticated customer is allowed to request only services complying with the GO-SLA of its GO
   - In addition, different customers of the same GO may have different service profiles, e.g. allowing different numbers of instances of a given service
   - Possible solution: a second-level white list, including the type of service and number of instances of each customer service profile defined in each GO-SLA

Integrity

1. GO vs. HO
   – Accidental or intentional alterations of the messages exchanged between GOs and HOs must be detected
   – Possible solution: to digitally sign each transaction, adopting a mechanism to be agreed upon during the authentication phase
2. Customer vs. GO/HO
   – Message integrity during exchanges with authenticated customers must be kept as well
   – Service-specific solutions based on the signaling protocols
   – e.g. GMPLS-based multi-service optical network may use standard integrity check mechanisms provided by RSVP-TE
     • The *Integrity Object* within a RSVP-TE message includes a keyed hash function of the entire message

Confidentiality

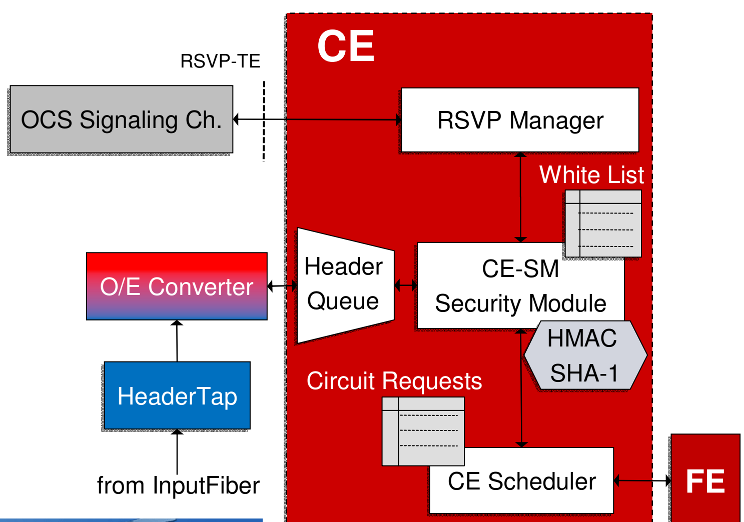
• Any critical information exchange must be encrypted using a robust method
• Control plane
  – use of PKI solutions
    (e.g. when distributing the shared keys to the neighbors for RSVP-TE authentication and integrity check)
• Data plane
  – left to end-user applications
    (e.g. using SSL/TLS)
  – activation of specific programmable network functions
    (e.g. secure tunnel based on IPsec)

Protection and Availability

- The services offered by the HO must be available 24-7
- The HO must enforce protection mechanisms to minimize the impact of any accidental service interruption
  - e.g. equipment failures, cable cuts, natural disasters
  - redundancy and backup resource allocation should be planned, e.g. using GMPLS protection and restoration techniques
- Malicious denial-of-service attacks may also compromise service availability
  - intrusion detection/prevention solutions should be enforced

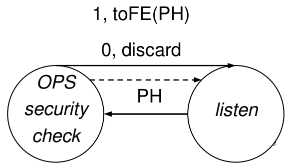
Enforcing security in the node control plane

Customer vs. GO/HO transactions OPS + OCS



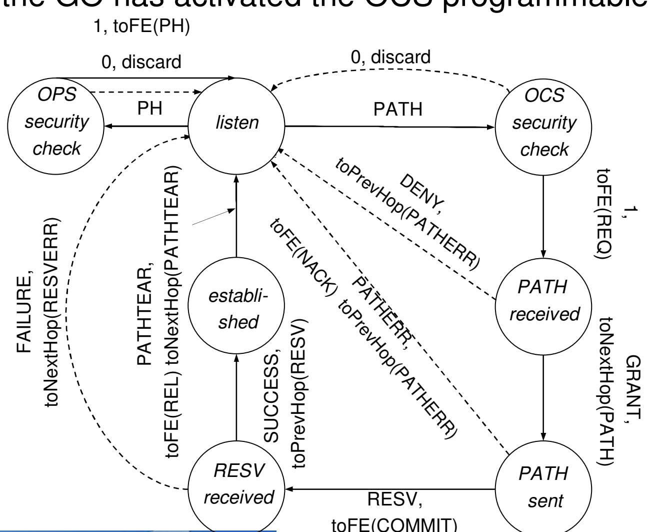
CE finite state machine w.r.t a given GO

Before the GO activates the OCS programmable function



CE finite state machine w.r.t a given GO

After the GO has activated the OCS programmable function



Conclusion and further developments

- Programmable optical router architecture as a key solution for future flexible network service provisioning
- Security and reliability issues to be tackled
- Software-router emulation as a flexible, inexpensive and fully-functional test platform
- Preliminary tests show promising results
- Further work currently under development
  - Extensive performance assessment of the programmable multi-service architecture
  - Implementation of standard ForCES protocol
  - Implementation of the OBS control plane
  - Improvement of optical signal propagation emulation
  - Further benchmarking of kernel-level processing speed