Optical Packet Switching

DEISNet
Gruppo Reti di Telecomunicazioni
http://deisnet.deis.unibo.it

WDM Optical Network

Legacy Networks
Edge Systems
WDM Links
Core Nodes

\( \lambda_1 \), \( \lambda_2 \), \( \lambda_3 \), \( \lambda_4 \)
Optical Packet Switching

- Long-term alternative to the optical circuit switching techniques currently under development (e.g. Wavelength Routing, MP\(\lambda\)S)
- Availability of the optical resource at packet level efficient use of the bandwidth
- Optical Packet Format:

  ![Optical Packet Format Diagram](image)

  - No O/E/O payload conversion needed at the core nodes, only header conversion may be performed

Optical Packet Router Architecture

![Optical Packet Router Architecture Diagram](image)
Contention resolution in OPS networks

• Typical problem of packet-level switching
• Resolution techniques available in OPS networks:
  – **wavelength domain**: contending packets transmitted on separate wavelengths of the same WDM link

![Wavelength Domain Diagram]

– **time domain**: contending packets delayed by optical buffers

![Time Domain Diagram]
Contestion resolution in OPS networks

• Typical problem of packet-level switching
• Resolution techniques available in OPS networks:
  – **wavelength domain**: contending packets transmitted on separate wavelengths of the same WDM link
  – **time domain**: contending packets delayed by optical buffers
  – **space domain**: contending packets forwarded to different links, according to a given adaptive, multi-path routing strategy

![Diagram](image)

Contention resolution in OPS networks

• Typical problem of packet-level switching
• Resolution techniques available in OPS networks:
  – **wavelength domain**: contending packets transmitted on separate wavelengths of the same WDM link
  – **time domain**: contending packets delayed by optical buffers
  – **space domain**: contending packets forwarded to different links, according to a given adaptive, multi-path routing strategy

• Performance of OPS contention resolution schemes traditionally evaluated in terms of
  – packet loss rate
  – packet latency

• Effects on the packet stream are also important
  – packet sequence
  – delay jitter
Optical Buffer

- Realized with B Fiber Delay Lines (FDL):
  - the delay must be chosen at packet arrival
  - packets are delayed until the output wavelength is available
  - available delays are 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 \( D_m = (B-1)D \)

The FDL-gap problem

- Using asynchronous, variable-length optical packets is efficient to carry one or more IP datagrams or packets from heterogeneous legacy networks
- Assumption: FIFO queuing, i.e. packets cannot overtake one another
- Due to the packets variable length and the finite delay granularity, when packets are queued a time gap between the end of a packet and the beginning of the following one is present \( (0 < \text{gap size} < D) \)
The Excess Load

- During the gap:
  - the server is idle
  - the queued packet must wait for the delay to expire
  - the result is some waste of the available output bandwidth

The gap can be seen as an artificial increase of the length of the previous packet, it results in an excess load at the output.

Choosing the Buffer Delay Unit

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

- Simulation VS. approximate Markov model assuming a finite queue

![Graph showing packet loss probability against D (normalized to the average packet length)]

- Too many FDLs
- Wavelength MUX required

Wavelength and Delay Selection (WDS)

- At packet arrival
  - Given the output fiber (from routing table lookup)
  - Forwarding algorithm must determine
    - the output wavelength
    - the required delay

Example:
- 4 delays
- 4 wavelengths

- WDS problem:
  how to assign the best wavelength to the packet?
Queuing Model of a WDM Optical Buffer

• For each output fiber:
  – $n$ identical servers represented by the $n$ wavelengths
  – $n$ logical FDL buffer in parallel, one for each wavelength (realized with a single set of FDLs operated in WDM)

WDS policies

• Gap-filling techniques may be too complex for the optical packet switching time-scale
WDS policies

• Gap-filling techniques may be too complex for the optical packet switching time-scale
• Simpler policies based on wavelength availability, with increasing intelligence
• Pursuing an analytical approach seems very complex • simulation study

\[ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \]

\[ t_0, t_0+D, t_0+2D, t_0+3D, t_0+4D \]

WDS policies

• **D-type**: delay oriented, aiming at minimizing the queuing time
  – choice of the first available wavelength

• **G-type**: gap oriented, aiming at minimizing the gaps between queued packets
  – choice of the wavelength with the closest queued packet

\[ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \]

\[ t_0, t_0+D, t_0+2D, t_0+3D, t_0+4D \]
Single-node WDS performance

Packet Loss Probability

D-type vs. G-type

Role of time and wavelength domains

D (normalized to the average packet length)
Synchronous network, slotted operation

• Asynchronous, variable-length legacy packets are split and inserted into a number of slots (slot size = $T$)
• Alternatives for the optical packet format:
  – Fixed-Length Packet (FLP): each slot is routed independently
  – Slotted Variable-Length Packet (SVLP): each train of slots is routed altogether

Incoming traffic

<table>
<thead>
<tr>
<th>FLP</th>
<th>SVLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

- net load = $\rho$
- real load = $\rho_F > \rho$
- real load = $\rho_S < \rho_F$

SVLP – performance

Slot size normalized to the average packet length

$D$ (normalized to the average packet length)
QoS differentiation in the OPS node

- Due to FDL buffering constraints, traditional priority queuing and scheduling techniques are not feasible
- QoS differentiation at the OPS node level possible through **resource partitioning**
  - any K wavelengths are reserved to HP traffic based on the actual occupancy
    - e.g. K=2 → LP packets are allowed as long as more than 2 wavelengths are available

Routing strategies in OPS networks

- Path computation algorithms provide:
  - a **default path**, i.e. the shortest path (e.g. in terms of number of hops)
  - a number of **alternative paths**, with equal or higher hop count than the default one
- The routing strategy may use:
  - **Shortest-Path Routing (SPR)** only
    - static routing tables
    - not using any alternative path
  - **Multi-Path Routing (MPR)**
    - dynamic routing tables
    - alternative paths used to relieve congestion on the default one
- Decisions to be taken by MPR strategy
  - how many alternative paths should be considered
  - how they should be dealt with by WDS policies
Dynamic MPR strategies (1)

- Strategies applying WDS on one of the alternative paths only when the default path is congested:
  - a path is considered congested when, on the corresponding output link, all the wavelengths are busy and there are no places left in the optical buffer

- SAP (Shortest Alternative Paths)
  - beside the default link, an alternative set of routes is considered, including any other shortest path different from the default one

- n-SAP (n-Shortest Alternative Paths)
  - besides the default link, n alternative sets of routes are considered, corresponding to paths with up to n–1 hops more than the shortest one

Dynamic MPR: 2-SAP

Alternative paths available toward the packet destination
Dynamic MPR: 2-SAP

Packets are sent through the default path, as long as this one is not congested.

Dynamic MPR: 2-SAP

In case the default path is congested, the best wavelength is chosen by the WDS policy on one of the alternative paths.
Dynamic MPR: 2-SAP

In case the default path is congested, the best wavelength is chosen by the WDS policy on one of the alternative paths.

Dynamic MPR strategies (2)

- Strategies applying WDS not on a single link, but on an entire set of links
  - sharing the wavelengths belonging to different paths

- **SSP (Shared Shortest Paths)**
  - WDS is performed over all the wavelengths on any shortest path link, including the default one

- **n-SSP (n-Shared Shortest Paths)**
  - WDS is performed over all the wavelengths on any link corresponding to paths with up to n–1 hops more than the shortest one
Dynamic MPR: SSP
WDS performed on any shortest path link

Dynamic MPR: 2-SSP
WDS performed on any shortest and 2nd-shortest path link
Dynamic MPR strategies performance

Simulation of European network (15 nodes, 24 links)
Dynamic MPR effective when non-shortest paths are used

QoS differentiation at the routing level

• Integration of QoS management into dynamic MPR strategies
  – aggregate QoS classes (sort of DiffServ approach)
  – simple set-up: 2 priority classes
• High-Priority (HP) traffic: always routed along the shortest path (SPR) using node-level resource partitioning
  – limited packet loss
  – limited delay and packet jitter
• Low-Priority (LP) traffic: two options
  – always routed along the shortest path (SPR) using spare resources
  – overflow traffic re-routed to alternative paths (MPR)
Simple test topology

Average node degree: $E = 2.4$

16 wavelengths per link, each loaded with 0.8 → traffic matrix generated accordingly

QoS performance

- **SPR/MPR** is the routing policy adopted for LP traffic
  - HP traffic (20%) always uses SPR
- Accurate dimensioning gives a good degree of traffic differentiation
  - LP routing policy does not affect HP
  - LP performance is slightly affected by HP resource dimensioning (within the range considered)
QoS performance

- **SPR/MPR** is the routing policy adopted for LP traffic
  - HP traffic always uses SPR
  - $K = 3$

Reference topology 1

Average node degree: $E = 3.25$
Topology 1: uniform matrix, balanced load

Reference topology 2

Average node degree: $E = 5.75$
Topology 2: uniform matrix, balanced load

Topologies 1 & 2: no. of hops distribution
Link failure recovery in OPS networks

- Connectionless approach
- Single link failure
- Link failure detection technique
  - notified by physical layer
  - time-based
  - signaling-based
- During failure detection, loss of optical packets supposed to be transmitted on the failed link
- After failure detection, the recovery procedure is called and a MPR alternative path is used

Packet loss during failure detection

- Analytical model for the packet loss probability as a function of detection time
  - $p$ is the loss probability in a failure free scenario
  - $t_f$ is the failure time
  - $d$ is the failure detection delay

$$P_L(t) = \begin{cases} 
  p & t \leq t_f \\
  1 - (1 - p) \frac{t_f}{t} & t_f < t < t_f + d \\
  1 - (1 - p) \frac{t_f}{t_f + d} & t \geq t_f + d 
\end{cases}$$
Packet loss during failure detection

![Diagram showing packet loss during failure detection](image)

Adding resources for protection

- The MPR-based protection scheme needs additional resources to be effective
- First, the network is dimensioned to have a given average load per wavelength (e.g. 0.7) with relation to the input traffic matrix
- Then, further wavelengths are added to each fiber so that each node sees all its output fibers with the same capacity
- Additional cost due to protection is 73% of the initial cost in terms of number of wavelengths
Impact on throughput

![Graph showing impact on throughput]

- **Network Throughput [bps]**
- **Simulation Time [sec]**

<table>
<thead>
<tr>
<th>Simulation Time</th>
<th>Network Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.25e+12</td>
</tr>
<tr>
<td>0.02</td>
<td>3.30e+12</td>
</tr>
<tr>
<td>0.04</td>
<td>3.35e+12</td>
</tr>
<tr>
<td>0.06</td>
<td>3.40e+12</td>
</tr>
<tr>
<td>0.08</td>
<td>3.45e+12</td>
</tr>
<tr>
<td>0.1</td>
<td>3.50e+12</td>
</tr>
<tr>
<td>0.12</td>
<td>3.55e+12</td>
</tr>
</tbody>
</table>

Out-of-sequence packet delivery

- A consequence of dynamic resource allocation techniques deployed to reduce congestion

- Implications:
  - complex reordering operations at the optical network edges
    - due to the huge bit rate of the optical channels
  - throughput degradation at the transport/application level
    - TCP congestion control highly affected by unordered segments
    - real-time, UDP-based traffic requirements impaired by excessive delay due to unordered packets and/or reordering process

- Possible solutions:
  - WDS policy with some time constraints in order to keep the correct packet sequence
  - dynamic multi-path routing limited to delay-equivalent paths

- But what do we intend with **correct packet sequence**?
Alternatives for packet sequence

ARRIVAL

\[ t_n \quad P_n \quad t_{n+1} \quad P_{n+1} \]

DEPARTURE

1. strictly in-sequence
2. strictly in-sequence (transparent)
3. strictly in-sequence
4. loosely in-sequence
5. loosely in- OR out-of-sequence
6. \( \Delta t_n \)
7. \( \Delta s_n \)

Jitter distribution

G-type WDS

G-type WDS with sequence constraint
Impact on loss

Conclusions

• OPS from the network perspective: issues
  – dynamic multi-path routing for contention resolution
  – QoS differentiation through resource partitioning and differentiated routing strategy
  – link failure protection strategy through multi-path routing and additional resources provisioning
  – packet sequence issues

• Further ongoing activities:
  – effective low-complexity implementation of void filling WDS
  – study of dynamic routing through ant colony optimization
  – study of the impact of packet sequence break on higher layer protocols
  – study of the impact of limited wavelength conversion on node and network performance