CSC/ECE 778 – Optical Networks
Optical Burst Switching (OBS)

George N. Rouskas
Department of Computer Science
North Carolina State University
Outline

- OBS Paradigm and Motivation
- Burst Assembly
- Wavelength Reservation Schemes
- Contention Resolution
- OBS-Based Optical Internet
Switching Paradigms

- **Circuit Switching**
  - well-understood, extensively studied (RWA, reserv. protocols, etc)
  - current approach to optical networking
  - 2-way reservations → inefficient for bursty traffic
  - requires statistical multiplexing/grooming for high utilization

- **Packet Switching**
  - more appropriate for bursty data traffic
  - practical, cost-effective, scalable functions not available today
    - optical buffering
    - optical header processing
    - integrated optical devices (lasers, converters, amplifiers, etc)
Technical Compromise:

- no optical buffering or packet-level parsing
- more efficient when traffic volume less than full wavelength

Burst: unit of transmission/switching

- aggregation of client data (IP packets) → assembly/disassembly
- remains in optical form throughout OBS network
- cut-through switching → no buffering inside OBS network

Setup (burst-header control) Message:

- transmitted ahead of each burst → offset delay
- configures intermediate switches
Data vs. Control Plane in OBS

- **All-optical** data plane
  - optical transmission and switching of data bursts
  - intermediate switches do not parse/examine burst content

- **Electronic** control (signaling) plane
  - signaling messages (e.g., setup) undergo OEO processing at intermediate nodes
  - signaling is out-of-band (e.g., one control \( \lambda \) per fiber)
**Signaling engine:** implements signaling protocol, burst forwarding and related control functions, and configures the OXC
Fundamental Operation

- **Users**: transmit data (IP/GbE/ATM) to ingress OBS switch

- **Ingress switch**:
  - assembles user data packets into bursts
  - when a burst is ready to transmit:
    - selects a path/next-hop switch for the burst
    - determines the value of the `offset` delay
    - transmits the `setup` packet to the next-hop switch
    - waits for an amount of time equal to `offset`, then transmits burst
  - **does not wait** for confirmation that path has been set up
  - **one-way** reservation scheme
Setup message:
- contains info on burst’s destination, length, priority, etc.
- processed electronically at each intermediate switch

Intermediate switches:
- use setup msg info to configure OXC to optically switch the burst
- forward setup message to next-hop switch
- resolve output port contention (how? → later)

Egress switch:
- buffer arriving bursts
- extract individual packets from each burst and forward to users
Out-of-Band Signaling in OBS

Setup Message
Offset=T

Data Burst

Offset=T−d

OXC

Offset=T−2d

Signaling Engine
Out-of-Band Signaling (cont’d)

Ingress Switch       Switch 1       Switch 2       Egress Switch

Offset

Setup

Burst

Processing

Delay

Time

Setup

Setup
Offset Value

- Burst must arrive to egress switch just after setup message
- **Minimum** offset value:
  \[
  T_{\text{offset}}^{(\text{min})} = kT_{\text{setup}} + T_{\text{OX}}\]

- \(T_{\text{setup}}\): processing time of setup message at each switch
- \(T_{\text{OX}}\): OXC configuration time
- \(k\): number of hops in burst’s path

- Offset value may be larger than the minimum for:
  - alternate routing
  - service differentiation (later · · ·)
Burst Assembly

To the same egress

Data (IP, GbE)

Voice (SONET)

Ingress Node

Setup msg

E/O

To OBS core

Burst
Burst Disassembly

From OBS core

Burst

O/E

Setup msg

Egress Node

Data (IP, GbE)

Voice (SONET)

To local destinations
Burst Assembly Algorithms

1. Timer-based
2. Burst-length-based
3. Hybrid
Timer-Based Assembly Algorithms

- $T$: fixed size of each assembly cycle
- All packets arrive within a cycle are assembled in a single burst
- Central limit theorem $\rightarrow$ Gaussian burst length distribution
- Constant burst interarrival time; if sources get synchronized $\rightarrow$ persistent burst collisions
- Tradeoffs in selecting value of $T$:
  - large $T \rightarrow$ high packet delay
  - small $T \rightarrow$ many small bursts $\rightarrow$ high control overhead
- Depending on traffic arrivals, undesirable burst lengths may result
\( L \): minimum burst size

- Packets collected into a burst as long as total size is less than \( L \)
- Burst is transmitted once its size exceeds \( L \)
- (Almost) constant burst size

CLT: \( \rightarrow \) Gaussian burst interarrival times

- Imposes minimum burst size \( \rightarrow \) low control overhead
- No guarantee in terms of maximum packet delay
Hybrid Assembly Algorithms

- $T$: timer threshold
- $L$: burst length threshold
- Burst transmitted when either its length exceeds $L$ or timer expires
- Adaptive algorithms: dynamically adjust $T$ and $L$ according to real-time traffic measurements
  → better performance but higher complexity
Burst must be buffered for an amount of time equal to \textit{offset}.

How to handle bursts arriving during this time?

1. leave them for next burst $\rightarrow$ higher delay
2. minimize extra delay by performing burst length prediction:
   - $l$: length of burst when setup message sent
   - $f(t)$: predicted burst length for offset=$t$
   - setup message carries $l + f(t)$ as burst length
   - add new packets to this burst as long as length $\leq l + f(t)$
**Burst Reservation Protocols**

- **Main problem:** scheduling incoming bursts to use outgoing wavelengths

- **Two issues:**
  - when to reserve the output port/wavelength for an incoming burst
  - what scheduling algorithm to use

- **Classification:**
  1. Immediate reservation
  2. Delayed reservation
     - (a) without void filling
     - (b) with void filling
Just-in-Time (JIT) family of OBS signaling protocols

- an output wavelength is reserved for a burst immediately after the arrival of the corresponding setup message; if a wavelength cannot be reserved at that time, then the setup message is rejected and the corresponding burst is dropped.

- Simple, easy to implement, no burst scheduling required.
- Reserved wavelength remains idle for $T_{offset} - T_{setup} - T_{OXC}$ time.
- The idle time decreases as the burst travels towards the destination.
Immediate Wavelength Reservation

User A

Ingress Switch

Intermediate Switch

Egress Switch

User B

Initial Offset

\( T_{\text{OXC}} \)

\( T_{\text{setup}} \)

Burst

Time

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

User A User B

Initial Offset

\( T_{\text{OXC}} \)

\( T_{\text{OXC}} \)

\( T_{\text{OXC}} \)

Wavelength Reserved

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

\( \cdot \)

User A User B

Initial Offset

\( T_{\text{OXC}} \)

\( T_{\text{setup}} \)

\( \cdot \)

\( \cdot \)

Copyright © 2007 by George N. Rouskas – p.21

CSC/ECE 778 – Fall 2007: Optical Burst Switching
Immediate Wavelength Reservation (cont’d)

Setup Message Arrival (Burst i)

Setup Message Arrival (Burst i+1)

Free  Reserved  Free  Reserved  Free

\( t_1 \)  \( t_2 \)  \( t_3 \)  \( t_4 \)  \( t_5 \)  \( t_6 \)

Offset (Idle Time)

Optical Burst

Copyright © 2007 by George N. Rouskas
Delayed Reservation

Operation:

an output wavelength is reserved for a burst just before the arrival of the first bit of the burst; if, upon arrival of the setup message, it is determined that no wavelength can be reserved at the appropriate time, then the setup message is rejected and the corresponding burst is dropped.

Voids: idle times on a wavelength between successively scheduled bursts; may be used to carry bursts whose setup message arrives later.
Delayed Reservation (cont’d)

User A

Ingress Switch

Intermediate Switch

Egress Switch

User B

Initial Offset

OXC Configured

Wavelength Reserved

setup

T_{setup}

setup

setup

T_{OXC}

setup

T_{OXC}

void

void

void

void

Burst

Time

Void

Void

Void

Void

setup
Horizon reservation protocol:

an output wavelength is reserved for a burst only if the arrival time of the burst is later than the time horizon of the wavelength; if, upon arrival of the setup message, it is determined that the arrival time of the burst is earlier than the smallest time horizon of any wavelength, then the setup message is rejected and the corresponding burst dropped.

No attempt is made to fill voids on the wavelengths.
Horizon Operation

Setup Message Arrival (Burst i)
Setup Message Arrival (Burst i+1)

\[ t_1 \]
\[ t_2 \]
\[ t_3 \]
\[ t_4 \]
\[ t_5 \]
\[ t_6 \]
\[ t_7 \]

\[ T_{OXC} \]

Offset (Idle Time)
Optical Burst

Burst Interdeparture Time
Just-Enough-Time (JET) reservation protocol:

an output wavelength is reserved for a burst if the arrival time of the burst (1) is later than the time horizon of the wavelength, or (2) coincides with a void on the wavelength, and the end of the burst (plus the OXC configuration time $T_{OXC}$) occurs before the end of the void; if, upon arrival of the setup message, it is determined that none of these conditions are satisfied for any wavelength, then the setup message is rejected and the corresponding burst dropped.

Need scheduling algorithms for void filling.
JET Operation

- Setup Message Arrival (Burst B)
- Setup Message Arrival (Burst A)

Time

Offset

Optical Burst
Scheduling Algorithms for JET

- **LAUC-VF**: latest available unscheduled channel with void filling
- **LAUC-VF minimum ending void**: minimize new void generated between end of new reservation and an existing reservation
- **LAUC-VF best-fit**: minimize total length of starting and ending voids generated after the reservation
Scheduling Algorithms for JET (cont’d)

New burst

\[ t'' \]

\[ \lambda_1 \text{ LAUC–VF} \]

\[ t_1 \]

\[ t' \]

\[ t'' \]

\[ \lambda_2 \]

\[ t_2 \]

\[ t' \]

\[ t'' \]

\[ \lambda_3 \text{ LAUC/Horizon} \]

\[ t_3 \]

\[ t' \]

\[ t'' \]

\[ \lambda_4 \text{ Min–EV} \]

\[ t_4 \]

\[ t' \]

\[ t'' \]

\[ \lambda_5 \text{ Min–BF} \]

\[ t_5 \]

\[ t' \]

\[ t'' \]

\[ ts \]
JIT⁺ operation:

- An output wavelength is reserved for a burst if (1) the arrival time of the burst is later than the time horizon of the wavelength and (2) the wavelength has at most one other reservation.

- No void filling.

- Improves upon JIT, same complexity.
**JIT** Operation

**Setup Message Arrival (Burst i)**

**Setup Message Arrival (Burst i+1)**

**Setup Message Rejected**

**Burst i+2 Dropped**

**Offset (Idle Time)**

**Optical Burst**
Comparison: Space Complexity

\( W \): # of wavelengths

- **JIT**: vector of size \( W \) with finish times of reservation on each wavelength
- **JIT^+**: vector of size \( W \) with start/end values for each of two possible reservations on each wavelength
- **Horizon**: vector of size \( W \) with horizon of each wavelength
- **JET**: start/end values of all reservations on all wavelengths
$M$: max # of reservations on all channels

- **JIT, JIT**: $O(1)$, 1 memory lookup $\rightarrow$ amenable to hardware implementation
- **Horizon**: $O(W)$, many memory lookup operations
- **JIT**: $O(W \log M)$, many memory lookup operations
Performance Comparison: Example

Offset

Burst
Example: JIT
Example: Horizon
Example: JET

[Diagram showing time intervals and JET bursts]
But ... Wait A Minute ...
But ... Wait A Minute ...

Processing Times

OXC Configuration Time

Offset

Burst
Overhead Operations

- Processing Time
- OXC Configuration

- Signaling Engine
- OXC
- Burst Transmission

Time
A More Likely Scenario

Offset

Burst
A More Likely Scenario

- No large voids → no void-filling
- At most two reservations per wavelength → JIT$^+$
Results: Offset = 10 × Mean Burst Size
Results: Offset = Mean Burst Size

![Graph showing burst drop probability vs number of wavelengths for different models: JIT, analytical; JIT+, simulation; Horizon, analytical; Horizon, simulation; JET, analytical; JET, simulation. The x-axis represents the number of wavelengths, ranging from 8 to 128, and the y-axis represents burst drop probability, ranging from 0.001 to 1.0.]
Results: Offset = 0.1 \times \text{Mean Burst Size}
Discussion

- $T_{OXC} > kT_{setup} \Rightarrow$ No Void Filling $\Rightarrow$ JET $\equiv$ Horizon $\equiv$ JIT$^+$

- Minimum burst length $+T_{OXC} > kT_{setup} \Rightarrow$ No Void Filling $\Rightarrow$ JET $\equiv$ Horizon $\equiv$ JIT$^+$

- $T_{offset} =$ constant $\Rightarrow$ No Void Filling $\Rightarrow$ JET $\equiv$ Horizon

- $(1/\mu \gg T_{OXC} \text{ and } 1/\mu \gg T_{setup}) \Rightarrow$ JET $\approx$ Horizon $\approx$ JIT$^+$ $\approx$ JIT

- JET/Horizon offer benefit at edge nodes, not inside network where offset is dominated by $T_{OXC}$
Contention Resolution

1. One-way reservation → output port/λ contention inside network
2. Burst loss a serious issue
3. Contention resolution mechanisms:
   1. Deflection
      - wavelength domain
      - space domain
      - time domain
   2. Dropping
   3. Preemption
Deflection Techniques

\( \lambda \text{ domain:} \) contending burst sent on new \( \lambda \rightarrow \) wavelength conversion
- dramatic reduction in burst loss
- immature and expensive technology

\textbf{Space domain:} contending burst sent to different output port
- follow alternate route to destination \( \rightarrow \) deflection routing
- no extra hardware requirement
- out-of-sequence arrivals; possible instability; offset ?

\textbf{Time domain:} contending burst delayed for fixed time \( \rightarrow \) buffering using \textit{fiber delay lines} (FDL)
- conceptually simple; mature technology
- bulky FDLs; extra delay; more voids
Dropping and Preemption

If contending burst cannot be deflected → data loss unavoidable

Several approaches:
- drop burst with later setup msg → tail drop for buffer of size 1
- preempt existing burst based on priority or traffic profile → setup msg for preempted burst?
- burst segmentation: deflect/drop/preempt only overlapping segments of contending bursts → finer contention resolution, complicated control
Proactive Contention Resolution Techniques

- Wavelength assignment policies → when conversion not available
  - rank wavelengths based on performance statistics
  - customized allocation strategies for OBS

- Routing and traffic engineering
  - path optimization to balance the traffic
  - traffic isolation between contending source-destination pairs
Service Differentiation

**Objective:** provide QoS guarantees to different classes of traffic

Extensive research in the context of packet-switched networks

**But:** existing service disciplines/packet scheduling algorithms mandate use of buffers (RAM)

**FDL:**
- provides limited and deterministic (fixed) delay
- no buffer management functions equivalent to electronic RAM
1. **Offset-based techniques:** larger offset $\rightarrow$ lower loss, higher priority
   - works well only when offset difference is quite large $\rightarrow$ high delay
   - high buffer requirements at edge nodes

2. **Active dropping:** selective dropping of bursts based on loss measurements and/or traffic profile
   - similar to AQM/RED

3. **Burst segmentation:** QoS differentiation at packet level
   - high priority packets in middle of burst $\rightarrow$ drop/preempt head/tail segments first
TCP over OBS → how is TCP performance affected by
  assembly process and associated delay?
  lost burst carrying multiple packets?

IP/WDM multicast and OBS?

Compatibility between Internet and OBS QoS

Seamless integration → GMPLS and labeled OBS (LOBS)