Important Notes

- OBS is a promising switch paradigm that offers many advantages over the existing technologies but is not likely to be the be-all, end-all solution.
- OBS has several variations, and adopting OBS will be an evoluzional process during which heterogeneous technologies are expected to co-exist for a quite while.
Overview

- Part I
  - Background
  - Optical Switching Paradigms
  - Basic OBS Concepts
  - IP/WDM Integration

- Summary

Electronic vs Optical Switching

- Data is transmitted optically (in WANs, MANs and even some LANs)
- Electronic switching: uses digital (electronic) switching fabrics; converts data from O to E for switching, and then from E to O for transmission.
- Optical (photonic) switching: uses optical switching fabrics; keeps data in the optical domain
Why Not Status-Quo (OEO)?

- data traffic growth still doubling every year
  - pure electronic processing and switching can hardly keep up (Moore’s Law)
  - though the cost of OEO at OC-48 (2.5Gbps) is going down, the overall cost (including WDM systems at OC-48) is still a dominant factor
    - electronic mux/demux, space/power consumption, heat dissipation, no transparency (future proof)
  - OEO at OC-192 (and higher in the future) will still be a dominant cost factor

Optical/Photonic (OOO) Switching

- Pros:
  - low-cost (no OEO), and high-capacity
  - transparency (bit-rate, format, protocol)
  - synergetic to optical transmission and future-proof

- Caveats
  - opaque (OEO) switches are more mature/reliable
  - still need some electronic processing/control
  - optical 3R/performance monitoring are hard
Part I
- Background
- Optical Switching Paradigms: historically, circuit switching is for voice and packet switching is for data
- Basic OBS Concepts
- IP/WDM Integration

Summary

2/7/2005

Circuit Switching

- long circuit set-up (a 2-way process with Req and Ack): $RTT = \text{tens of } ms$
- pros: good for smooth traffic and QoS guarantee due to fixed BW reservation;
- cons: BW inefficient for bursty (data) traffic
  - either wasted BW during off/low-traffic periods
  - or too much overhead (e.g., delay) due to frequent set-up/release (for every burst)
Wavelength Routing

- setting up a lightpath (or \( \lambda \) path) is like setting up a circuit (same pros and cons)

- \( \lambda \)-path specific pros and cons:
  - very coarse granularity (OC-48 and above)
  - limited # of wavelengths (thus # of lightpaths)
  - no aggregation (merge of \( \lambda \)s) inside the core
    - traffic grooming at edge can be complex/inflexible
  - mature OXC technology (\( msec \) switching time)

Packet (Cell) Switching

- A packet contains a header (e.g., addresses) and the payload (variable or fixed length)
  - can be sent without circuit set-up delay
  - statistic sharing of link BW 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
Optical Core: Circuit or Packet?

- five src/dest pairs
  - circuit-switching (wavelength routing)
    - 3 λs if without λ conversion
    - only 2 λs otherwise
  - if data is sporadic
    - packet-switching
      - only 1 λ needed with statistical muxing
      - λ conversion helps too

Packet Core: A Historical View (hints from electronic networks)

- optical access/metropolitan networks (LAN/MAN)
  - optical buses, passive star couplers (Ethernet)
  - SONET/WDM rings (token rings)
  - switched networks? (Gigabit Ethernet)
- optical core (WAN)
  - λ-routed virtual topology (circuits/leased lines)
  - dynamic λ provisioning (circuits on-demand)
  - optical burst (packet/flow) switching (IP)
Packet Core: Technology Drivers

- explosive traffic growth
- bursty traffic pattern
- to increase bandwidth efficiency
- to make the core more flexible
- to simplify network control & management by *making the core more intelligent*

Self-Similar (or Bursty) Traffic

- Left:
  - Poisson traffic (voice)
  - smooth at large time scales and mux degrees
- Right:
  - data (IP) traffic
  - bursty at all time scales and large mux degrees
  - circuit-switching not efficient \((\text{max} \gg \text{avg})\)
To Be or Not to Be BW Efficient?
(don’t we have enough BW to throw at problems?)

- users’ point of view:
  - with more available BW, new BW intensive (or hungry) applications will be introduced
    - high BW is an addictive drug, can’t have too much!
  - carriers’ and venders’ point of view:
    - expenditure rate higher than revenue growth
    - longer term, equipment investment cannot keep up with the traffic explosion
    - need BW-efficient solutions to be competitive

Optical Packet Switching: Holy Grail

- No.1 problem: lack of optical buffer (RAM)
- fiber delay lines (FDLs) are bulky and provide only limited & deterministic delays
  - store-n-forward (with feed-back FDLs) leads to fixed packet length and synchronous switching
- tight coupling of header and payload
  - requires stringent synchronization, and fast processing and switching (ns or less)
Impacts on Components

(a) Cross-Connect (1000 by 1000, ms switching time)
(b) Packet-Switch (64x64, with ns switching time)

Optical Burst Switching (OBS)

- a burst has a long, variable length payload
  - low amortized overhead, no fragmentation
- a control packet is sent out-of-band ($\lambda_{\text{control}}$)
  - reserves BW ($\lambda_{\text{data}}$) and configures switches
- a burst is sent after an offset time
  - arrives at a switch after it has been configured so no buffering needed [OBS-1'97, OBS-2'97]
Packet (a) vs. Burst (b) Switching

Optical Burst Switching Node

Multiple data channels share one control channel. Data bursts remain in optical domain while CPs go through O/E/O conversions.
Optical Packet Switching Node

All-optical processing is not practical yet (will ever competitive?). Need O/E/O conversion of header on every $\lambda$ (hundreds of them in each fiber). Also not scalable and cost-effective.

2/7/2005

OEO approach

All traffic goes through O/E/O conversions (for sub-$\lambda$ granularity). However, as transmission speed goes higher, this approach is neither scalable nor cost-effective (heat, power, footprint)

2/7/2005
Optical Circuit Switching (Wavelength Routing) node

Bandwidth is assigned at the wavelength (λ) granularity after lightpath is set up. No statistical multiplexing gain and high overhead for bursty traffic.

Part I
- Background
- Optical Switching Paradigms
- Basic OBS Concepts
- IP/WDM Integration

Summary
Burst Switching – Time Line

- Burst Switching For Digital Voice
- Resv Just-In-Time (RIT) Protocol
- Tell-and-Wait (TAW) Protocol
- Tell-and-Go (TAG) Protocol
- Just-Enough-Time (JET) for OBS

- 1983
- 1990
- 1995
- 1997

Burst Signaling for TDM Networks

- Tell-and-Wait (TAW) = Connect-Confirmation (CC)
  - Send REQ first to make reservation; Transmit the burst after ACK is received (hop-by-hop distributed control)

- Reservation Just-In-Time (RIT)
  - Similar to TAW/CC; But switching fabric configured just before burst arrival; Burst transmitted at a time specified by ACK (centralized control or with global knowledge/synchronization)

- Tell-and-Go (TAG)
  - Send REQ and then burst (before receiving ACK); Delay burst at intermediate nodes to wait for REQ processing and switch configuration (hop-by-hop distributed control)
Optical Burst Switching (OBS) Protocols for WDM Networks

- Just-Enough-Time (JET)
  - Qiao, Yoo, 08/97 (IEEE/LEOS, NSF Proposal, DARPA Workshop), SPIE’98, JHSN’99, JSAC’00 [1-5]
- Terabit Burst Switching (based on TAG)
  - J. S. Turner, 12/97 (Tech. Rep) [6], 1999 (JHSN) [7]
- Just-In-Time (hop-by-hop RIT)
  - Wei, Tsai, McFarland et, al., SPIE’98, IFIP’00 [8,9]
  - Xu et al. IEEE ComMag’01, Baldine et al ComMag’02
- Wavelength Routed-OBS (centralized TAW/RIT)
  - Düser and Bayvel, JLT’02

OBS Publications in Major Conferences and Journals

- conservative figures for year 2001 and beyond
- dedicated sessions on OBS at OFC’03, Infocom’03, ICC’03 etc.

![Bar chart showing OBS publications by year](chart.png)
OBS Basic Concepts

- Burst Assembly (and Disassembly) at Edge
  - client data (e.g., IP packets) assembled into bursts

- Burst Switching/Reservation Protocol
  - Control packet (CP) sent an offset time $t$ ahead of burst
  - Dedicated control channel (out-of-band signaling) for CP
  - No fiber delay lines (FDLs) nor O/E/O conversions for burst at any intermediate (core) nodes

- Photonic Burst Switching Fabric inside Core
  - Leverages the best of optics (for burst switching) and electronics (for CP processing and fabric control)

---

Burst Assembly

- Assembly queues for different egress nodes
- Time or length threshold is reached

---

ATM Cell
IP Packet
SONET Frame
Burst Assembly

Assembly queues for different egress nodes

A CP is generated and sent out

Control channel

Data channel

ATM Cell
IP Packet
SONET Frame

2/7/2005
Fiber Delay Line (FDLs)

- Feed-forward (above) or Feed-backward (Loop)
- No optical RAM for store-and-forward
- FDLs provide only limited delay and cannot perform most of useful buffer functions
- FDL units are bulky, affect signal quality etc.

Just-Enough-Time (JET)

- An offset time between CP and burst
  - No fiber delay line (FDL) required to delay the burst when CP is processed and switch fabric is configured.
- CP carries the burst length information
  - Facilitates delayed reservation (DR) for intelligent, efficient allocation of BW and FDL (if any), including look-ahead scheduling.
  - Later adopted by TBS [7], JIT [10,11] and others (OPS)
JET with Offset Time $T$

- Offset $= T$
  - CP arrives the OEO node at time $t_1$

- CP goes through O/E conversion and configures switch fabric

2/7/2005
When burst arrives at the intermediate node, the switch fabric is already configured.
Without any delay, the burst goes through the optical switch fabric.

Reduce Offset Time and Tolerate Switch Setting Delay (better than packet switching)

- control packet can leave right after $\delta = \Delta - s$ (s is the switch setting time)
Delayed Reservation (DR)

DR leads to efficient allocation of BW and any available FDLs (though not shown). Without DR, 2nd burst will be dropped in both cases (and FDLs will be wasted in Case 2).

Burst scheduling

- Which output channel to use?
  - If none is available, which FDL (if any) to use?
- Two categories of scheduling algorithms
  - Without void (closed interval) filling
    - Only use open interval (also called Horizon/LAUC) [Turner’99]
  - With void filling
    - Can minimizes the starting void (Min-SV or LAUC-VF) or the ending void (Min-EV) etc., [Xu et.al. Infocom’03]
Scheduling Algorithms

Min-SV [Infocom’03] achieves the best performance in terms of computational complexity and the bandwidth utilization.

Statistical Multiplexing in OBS

Burst level transmission granularity and delayed reservation makes statistical multiplexing possible in OBS network.
Statistical Multiplexing in OBS

Burst level transmission granularity and delayed reservation makes statistical multiplexing possible in OBS network
Sub-\(\lambda\) Switching Capability

By-pass traffic is treated the same as add/drop traffic and both are switched all-optically.

2/7/2005
Sub-\(\lambda\) Switching Capability

By-pass traffic is treated the same as add/drop traffic and both are switched all-optically.

Contestation Resolution

- When multiple bursts compete for the same output channel, how to avoid/reduce burst loss?
- Three major strategies
  - Deflection in space, time and wavelength
  - Preemption of an existing reservation
  - Segmentation of a burst into smaller pieces
Contention Resolution

- **Deflection** [Yoo, Qiao, Dixit, SPIE'00]
  - Space domain: applying deflection routing
  - Wavelength domain: use a different wavelength via wavelength conversion
  - Time domain: wait using a fiber delay line

- **Segmentation**
  - Drops, deflects or preempts one or more segments instead of an entire burst [Qiao NSF’97, Deti et al ’02 and Vokkarane&Jue ’02]

Part I

- Background
- Optical Switching Paradigms
- Basic OBS Concepts
- IP/WDM Integration

Summary
Network Architectures

- today: IP over (ATM/SONET) over WDM
- trend: Integrated IP/WDM (with optical switching)
- goal: ubiquitous, scalable and future-proof

Evolutionary Trend

Removing the transport layers

Traditional
- IP
- ATM
- SONET
- Optical

Today
- IP Over ATM
- POS

Tomorrow
- IP-OG
- Optical

Lower Cost, Complexity, & Overhead
Internet Protocol (IP)

- main functions
  - break data (email, file) into (IP) packets
  - add network (IP) addresses to each packet
  - figure out the (current) topology and maintains a routing table at each router
  - find a match for the destination address of a packet, and forward it to the next hop
    - a link to a popular server site may be congested
Asynchronous Transfer Mode

- break data (e.g., an IP packet) into smaller ATM cells, each having 48+5 = 53 bytes
- a virtual circuit (VC) from point A to point B needs be pre-established before sending cells.
- support Quality-of-Service (QoS), e.g., bounded delay, jitter and cell loss rate
- basic rate: between 155 and 622 Mbps
  - just start to talk 10 Gbps (too late?)

ATM: Legacy

- interest in ATM diminished
  - a high cell tax, and segmentation/re-assembly and signaling overhead
  - failed to reach desktops (& take over the world)
  - on-going effort in providing QoS by IP (e.g., IPv6 & Multi-protocol Label Switching or MPLS)
Benefit of VC (as in ATM)

- faster and more efficient forwarding
  - an exact match is quicker to find than a longest sub-string match (with a destination as done in IP)
- facilitates traffic engineering
  - paths can be explicitly specified for achieving e.g., network-wide load-balance
  - packets with the same destination address (but different VCI’s) can now be treated differently

IP-over-ATM

- IP routers interconnected via ATM switches
- breaks each packet into cells for switching
- Multi-protocol over ATM (MPOA)
  - ATM-specific signaling to establish an ATM VC between source/destination IP routers
  - segmentation and re-assembly overhead
- a flow: packets with the same source or destination (slightly differs from a burst)
Multi-Protocol Label Switching

- A control plane integrating network-layer (routing) and data-link layer (switching)
  - packet-switched networks with VC’s
- LSP: label switched path (VC’s)
  - identified with a sequence of labels (or VCIs)
  - set up between label switched routers (LSRs)
- Each packet is augmented with a shim containing a label, and switched over a LSP

SONET/SDH

- standard for TDM transmissions over fibers
  - basic rate of OC-3 (155 Mbps) based on 64 kbps PCM channels (primarily voice traffic)
  - expensive electronic Add-Drop Muxers (ADM) @ OC-192 (or 10 Gbps) and above
  - many functions *not* necessary/meaningful for data traffic (e.g., bidirectional/symmetric links)
  - use predominantly rings: not BW efficient, but quick protection/restoration (<= 50 ms)
Wavelength Division Multiplex

- up to 50 THz (or about 50 Tbps) per fiber (low loss range is now 1335nm to 1625nm)
- mature WDM components
  - mux/demux, amplifier (EDFA), transceiver (fixed-tuned), add-drop mux, static λ-router,
- still developing
  - tunable transceiver, all-optical λ-conversion and cross-connect/switches, Raman amplifiers

Advance in WDM Networking

- Transmission (long haul)
  - 80 λ.s (1530nm to 1565nm) now, and additional 80 λ.s (1570nm to 1610nm) soon
  - OC-48 (2.5 Gbps) per λ. (separated by 0.4 nm) and OC-192 (separated by 0.8 nm)
  - 40 Gbps per λ. also coming (>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)
IP over WDM Architectures

I. IP routers interconnected with WDM links
   - with or without built-in WDM transceivers
II. An optical cloud (core) accessed by IP routers at the edge
   - pros: provide fat and easy-to-provision pipes
   - either transparent (i.e., OOO) or opaque (i.e., O-E-O) cross-connects (circuit-switches)
   - proprietary control and non-IP based routing

IP over Optical: Network Model

Router Network

Optical Network

Optical Path

End-to-end path (LSP)
Tune IP layer topology to changing traffics

- IP Layer Traffic Patterns Change
  - Step 1 - Add new $\lambda$ from A to B
  - Step 2 - Delete $\lambda$ from A to C

Integrated IP/WDM

- IP and GMPLS on top of every optical circuit or packet switch:
  - IP-based addressing/routing (electronics), but data is optically switched (circuit or packet)
  - GMPLS-based provisioning, traffic engineering and protection/restoration
  - peer-to-peer, overlay or hybrid models
Why IP over WDM

- IP: the unifying/convergence network layer
- IP traffic is (& will remain) *predominant*
  - annual % increase in voice traffic is in the teens
- IP/WDM the choice if start from scratch
  - ATM/SONET were primarily for voice traffic
  - should optimize for pre-dominant IP traffic
- IP routers’ port speed reaches OC-48
  - no need for multiplexing by ATM/SONET

Why IP/WDM

- IP is resilient (albeit *rerouting* may be slow)
- Having a WDM layer (with optical switches) provides fast restoration (not just WDM links for transmission only)
- no need to re-invent routing and signaling protocols for the WDM layers and corresponding interfaces
- facilitates traffic engineering and inter-operability
Evolutionary Trend

Removing the transport layers

Traditional
IP
ATM
SONET
Optical

Today
IP Over ATM
IP
SONET
Optical

Tomorrow
POS
IP
IP-OG
Optical

Lower Cost, Complexity, & Overhead

Observation

- IP over WDM has evolved:
  - from WDM links, to WDM clouds (with static virtual topology and then dynamic \( \lambda \) services),
  - and now integrated IP/WDM with MP\( \lambda \)S
- to be truly ubiquitous, scalable and future-proof, a WDM optical core should also be
  - capable of OOO packet/burst-switching, and basic QoS support (e.g., with LOBS control)
MPLS-variants: MP\(\lambda\)S and LOBS

- optical core: circuit- or packet- switched?
- circuit-switched WDM layer
  - OXC’s (e.g., wavelength routers) can be controlled by MPLambdaS (or MP\(\lambda\)S)
- optical burst switched WDM layer
  - optical switches controlled by Labeled Optical Burst Switching (LOBS)

Labeled OBS (LOBS) [Qiao, 2000]

- Extends G-MPLS to OBS networks,
  - where CPs carry additional label information
- Differs from MP\(\lambda\)S:
  - Associate \(\lambda\) with a label on the time scale of a burst
  - Support sub-\(\lambda\) granularity and statistical multiplexing
- Opens many traffic engineering issues
  - Routing and wavelength assignment of LOBS paths
  - Protection and restoration
  - Periodical transmission support
Labeled Optical Burst Switching

Physical Layer performs functions for burst switching, wavelength conversion, burst delay/buffering, optical amplification, etc. LOBS (MPLS) layer provisions OBS services. This includes burst assembly, WDM topology and resource dissemination, survivability, etc.

IP layer performs layer three functions (e.g., addressing, routing)

Electronic layer

Opto-electronic layer

Optical layer

Monitoring layer (optional)

Physical Layer

Similar to MPLS (e.g., different LOBS paths can share the same λ)

Unique LOBS issues: assembly (offset time), QoS in bufferless core, routing & λ-assignment, contention resolution, light-splitting (for WDM multicast) [LOBS’00]

Potentials of OBS from Business/Economic point of view

- Networks are adopting MPLS
- OBS + MPLS/G-MPLS = LOBS
- LOBS mixes persistent & burst length circuits
- Sources of Economic Benefit:
  - Unified control plane – single layer, universal service
  - Transparent transport of packets & frames
  - Statistical multiplexing – even between circuit frames
  - Delivers best effort and deterministic QoS
  - More efficient than packet switching
LOBS Vs Today’s Networks

Today’s Networks:
• Multiple services = multiple layers
• Faster speeds = all new electronics.
• More wavelengths = more electronics
• Multiple Trade Craftsmen = OPEX$
• Many OEO Conversions = CAPEX$ (highest cost component of network)

A LOBS Network
• All services in a single layer
• Faster speeds & more wavelengths cost nothing.
• Fewer Trade Craftsmen = < OPEX$
• Few OEO Conversions = <<CAPEX$
• Less Power & Space = < OPEX$

Comparing Costs

IP Core Mesh
• 24 x 40 Line Cards @ $125K ea.
• CAPEX = $120 Million + Misc.
• (40 λ.s go through OADMs)

Equivalent LOBS Core Mesh
• 1 control λ., 79 data λ.’s
• 24 Line Cards @ $125K ea.
• CAPEX = $7.2 Million + Misc.

$113 Million of CAPEX Avoided + Reduced OPEX + Power & Space + Simpler network control
**OBS/LOBS in the Value Chain**

- **Component Vendors**
  - Signaling Processors
  - Burst Processors
  - Fast Optical Switching Fabrics

- **Equipment Vendors**
  - Control Plane Extensions to GMPLS

- **OSS Vendors**
  - LOBS technology
  - Core and Edge Switch/Routers
  - One box –
  - More efficient –
  - Much lower cost –
  - Legacy compatible –
  - Service Transparent –

- **Carriers**
  - Converged services
  - Capex reduction
  - Opex efficiency
  - Improved economics
  - Link-by-link adoption
  - Maintain legacy contracts

---

**OBS – A Future Proof Solution**

<table>
<thead>
<tr>
<th>Optical switching paradigms</th>
<th>Bandwidth Utilization</th>
<th>Latency (setup)</th>
<th>Optical Buffer</th>
<th>Proc./Sync. Overhead (per unit data)</th>
<th>Adaptivity (traffic &amp; fault)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit</td>
<td>Low</td>
<td>High</td>
<td>Not required</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Packet/Cell</td>
<td>High</td>
<td>Low</td>
<td>Required</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>OBS</td>
<td>High</td>
<td>Low</td>
<td>Not required</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

OBS combines the best of the two while avoiding their shortcomings
Summary of Research Topics

- Burst scheduling algorithms for legacy support (SONET, GigaE, ATM etc)
- Contention resolution and avoidance strategies
- Node and switching fabric architectures
- IP/WDM Mcast and Tree-Shared Mcast in OBS
- Labeled OBS (LOBS) and GMPLS extension

Select Pre-2000 OBS Publications

Select Pre-2000 OBS Publications

11. See http://www.cse.buffalo.edu/~yangchen/OBS_Pub_year.html for more references published in and after 2000
12. See http://www.cse.buffalo.edu/~qiao/wobs for recent Workshops on OBS held in conjunction with Opticomm’03 and Globecom’03