Telecom

Optical Networks: The Road Ahead

Biswanath Mukherjee
Child Family Endowed Chair Professor
University of California, Davis, USA

mukherje@cs.ucdavis.edu

Keynote Talk Presented at:
4th Euro NGI Conference 2008
Krakow, Poland
April 28, 2008

• This is a reduced / redacted version of the presentation given at Euro NGI 2008.

• For more information on some of the materials, please consult my textbook:

Optical Networks: The Road Ahead

Broadband Access
- PONs
- WDM in PONs
- Long-Reach Broadband Access
- Hybrid wireless-optical access
- Metro: The vanishing breed?

Backbone Networks
- DOCS ("dial for bandwidth")
- Robust Network Design (multi-layer, multi-domain, multi-path, etc.)
- Ethernet Everywhere
- Network Engineering (NE) (vs. TE vs. NP)
- Higher-density switches!

Telecom Network Hierarchy
An Access Network (PON)

System/Network: Value Proposition
**Optical Network Architecture:**
Extending Our “Boundaries”

*Applications*

(*Customer* needs)

**Differentiated Services:**
- **Bandwidth:** OC-192, OC-48, ..., STS-1, VT1.5, ...
- **Failure-Recovery Delay:** The “50-ms myth!”
- **Network Economics:** Pricing, SLA, ...

**Network Architect**

(Piotr, I, ...)

+ routing protocols to combat optical channel impairments
+ breakthroughs needed in device technologies?
  - optical RAM, ultra-wideband amp, “tunable” AWG, ...

**Physical Layer**

(optical/wireless channel) --
materials, devices, subsystems

---

**Telecom**

Optical Networks: The Road Ahead

**Broadband Access**
- PONs
- WDM in PONs
- Long-Reach Broadband Access
- Hybrid wireless-optical access
- Metro: The vanishing breed?

**Backbone Networks**
- DOCS ("dial for bandwidth")
- Robust Network Design (multi-layer, multi-domain, multi-path, etc.)
- Ethernet Everywhere
- Network Engineering (NE) (vs. TE vs. NP)
- Higher-density switches!
**WDM in PONs (Evolutionary Upgrade)**

- Provision Wavelengths as Needed
  - Distinct growth phases
  - Upstream/downstream

- Non-Disruptive
  - Legacy users are unaffected
  - Infrastructure remains intact

- Maximum Utilization of Capacity
  - Users are still able to share wavelengths
  - Shift users requiring more bandwidth

ACK: NSF

---

**WDM in PONs (Evolutionary Upgrade)**

**Traffic Growth and Wavelength Allocation (Exponential)**

- Shifting (exponential model)
- $\lambda_0$ threshold times: 6, 12.5, 19, 25.5-months
Long-Reach Broadband Access

• Basic Problem
  – Reach of traditional PON = 10-20 km
  – Fiber capacity: much larger than data generated by users served by the limited geography in a traditional PON
  – Hence Long-Reach (LR) PON should be investigated
    • Increase distance to 100 km or more... may need power (so not exactly a “PON”... called “SuperPON” in literature), but use very few active elements... so we call them LR Access Networks

• Sampling of Current LR Access Efforts
  – ACTS-PLANET; BT Demos; SFI Demos

• Limitations of Current LR Access Methods
  – “One dimensional” (linear)
  – Users are located “two dimensionally” on the earth ... so “ring-and-spur” makes more sense for LR Access

Our LR Access Approach

ACK: ETRI Korea
Hybrid Wireless-Optical Broadband Access Network (WOBAN)

Optimization problem: how far to use fiber? at what point wireless takes over?

ACK: Nokia (NSN)

Example: WOBAN Front-End Mesh

Location:
Golden Gate Ave. to Pacific Ave. Divisadero to Van Ness Avenue
Area: 1 sq-mile (approx.)
Population: 15,000 (approx.)

A part of San Francisco WOBAN (SFNet)
### Current Deployment Status

<table>
<thead>
<tr>
<th>Area/Location</th>
<th>Architecture</th>
<th>Compatibility</th>
<th>Configuration</th>
<th>Operating Range</th>
<th>Player</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akron, OH</td>
<td>Flat AP infrastructure</td>
<td>WiFi</td>
<td>Multiple radio</td>
<td>2.4 GHz</td>
<td>MobilPico</td>
</tr>
<tr>
<td>Athens, GA</td>
<td>Multi-layered MAP</td>
<td>WiFi, WM3</td>
<td>Multiradio</td>
<td>2.4 GHz</td>
<td>Belair</td>
</tr>
<tr>
<td>Bristol, UK</td>
<td>Multi-layered deployment</td>
<td>WiFi, WM3, WM5</td>
<td>Multiradio multi-antenna</td>
<td>2.4-5 GHz</td>
<td>Belair</td>
</tr>
<tr>
<td>Chanhassen, MN</td>
<td>Flat deployment</td>
<td>WiFi</td>
<td>Single radio, omni-directional</td>
<td>2.4 GHz</td>
<td>Tempos, Pronto</td>
</tr>
<tr>
<td>Corpus Christi, TX</td>
<td>Flat deployment IP-compatible</td>
<td>WiFi</td>
<td>Single radio, omni-directional</td>
<td>2.4 GHz</td>
<td>Tempos, Pronto</td>
</tr>
<tr>
<td>College City, CA</td>
<td>Flat (wireless capable)</td>
<td>WiFi</td>
<td>Multiradio omni-directional</td>
<td>2.4-5 GHz</td>
<td>NetReach, Pronto</td>
</tr>
<tr>
<td>Farmers Branch, TX</td>
<td>Gateways with ODU ingress</td>
<td>WiFi, WM5</td>
<td>Multiple radio</td>
<td>2.4 GHz</td>
<td>NetReach, Pronto</td>
</tr>
<tr>
<td>guilty, AZ</td>
<td>Multi-layered deployment</td>
<td>WiFi, WM5</td>
<td>Multi-radio multi-antenna</td>
<td>2.4-5 GHz</td>
<td>Belair</td>
</tr>
<tr>
<td>Geneseo, Prince</td>
<td>Flat with mesh</td>
<td>WiFi</td>
<td>Multiradio omni-directional</td>
<td>2.4-5 GHz</td>
<td>Pirrucci</td>
</tr>
<tr>
<td>Ila Vista, CA</td>
<td>Flat deployment</td>
<td>WiFi</td>
<td>Multiradio</td>
<td>2.4-5 GHz</td>
<td>Pirrucci</td>
</tr>
<tr>
<td>Shingston, IA</td>
<td>3-tier hierarchical deployment</td>
<td>WiFi, WM5</td>
<td>Multiradio multi-antenna</td>
<td>2.4-5 GHz</td>
<td>Belair</td>
</tr>
<tr>
<td>Moorefield, NY</td>
<td>P2P with fiber-optic backbone</td>
<td>WiFi, WM5</td>
<td>Single radio omni-directional</td>
<td>2.4 GHz</td>
<td>Tempos</td>
</tr>
<tr>
<td>New Orleans, LA</td>
<td>WiFi routers with digital IP camera attached to IP backbone for video surveillance system</td>
<td>WiFi, WM5</td>
<td>Multi-radio multi-antenna</td>
<td>2.4-5 GHz</td>
<td>Tempos</td>
</tr>
<tr>
<td>Philadelphia, PA</td>
<td>Currently being deployed</td>
<td>WiFi</td>
<td>Must have dual radio</td>
<td>2.4-5 GHz</td>
<td>Belair, 4ever</td>
</tr>
<tr>
<td>Springfield, MO</td>
<td>Hierarchical (2.4 VLAN capable)</td>
<td>WiFi, WM5</td>
<td>Multiradio</td>
<td>2.4-5 GHz</td>
<td>Belair, 4ever</td>
</tr>
<tr>
<td>UC Montreal, Canada</td>
<td>Hierarchical deployment</td>
<td>WiFi, WM5</td>
<td>Multiradio</td>
<td>2.4-5 GHz</td>
<td>Belair, 4ever</td>
</tr>
<tr>
<td>Tempe, AZ</td>
<td>Gateways with ODU ingress</td>
<td>WiFi, WM5</td>
<td>Multiradio multi-antenna</td>
<td>2.4-5 GHz</td>
<td>NetReach, Pronto</td>
</tr>
</tbody>
</table>

*Wurton, Inc. is a new player with their 'spatially adaptive' MBBO-based routers having an antenna array and six radio transceivers.*

Ack: Suman Sarkar
Traceroute: to aland.bbn.com

1 169.237.4.254 (169.237.4.254) 0.973 ms 0.747 ms 0.475 ms
2 169.237.246.238 (169.237.246.238) 0.682 ms 0.797 ms 0.741 ms
3 area2-13--area2.ucdavis.edu (128.120.2.49) 0.880 ms 0.767 ms 0.746 ms
4 area2--area0.ucdavis.edu (128.120.0.133) 0.939 ms 1.295 ms 0.856 ms
5 area0--area14.ucdavis.edu (128.120.0.222) 1.063 ms 1.484 ms 0.887 ms
6 area14--a14.ucdavis.edu (128.120.9.142) 1.009 ms 0.950 ms 0.837 ms
7 a14--ucd-hpr.ucdavis.edu (128.120.9.138) 1.177 ms 1.577 ms 1.577 ms
8 * dc-oak-dc2--ucd-ge.cenic.net (137.164.24.225) 2.901 ms 2.694 ms
9 dc-sfo-dc1--oak-dc2-pos.cenic.net (137.164.22.32) 2.884 ms 2.751 ms 2.655 ms
10 dc-svl-dc1--sfo-dc1-pos.cenic.net (137.164.22.34) 3.923 ms 3.812 ms 3.674 ms
11 te2-3--480.tr01-plalca01.transitrail.net (137.164.131.253) 3.854 ms 4.300 ms 3.958 ms
12 bb1-g1-0.pxpaca.sbcglobal.net (198.32.176.112) 4.543 ms 4.463 ms 4.559 ms
13 151.164.93.249 (151.164.93.249) 68.968 ms 68.869 ms 69.440 ms
14 dist1-vlan30.lgtpmi.ameritech.net (65.42.245.97) 69.434 ms 69.213 ms 68.944 ms
15 rback1-g1-0.lgtpmi.sbcglobal.net (65.42.245.230) 68.845 ms 69.043 ms 69.145 ms
16 adsl-68-22-232-254.dsl.lgtpmi.ameritech.net (68.22.232.254) 823.731 ms 856.567 ms 822.425 ms
17 adsl-68-22-232-249.dsl.lgtpmi.ameritech.net (68.22.232.249) 858.838 ms 890.727 ms 859.053 ms

DOCS Architecture: Overview

ACK: NSF FIND Program, Dan Blumenthal, Nick McKeown, and John Bowers
Why DOCS? (or DCS)

• Emerging (video-enabled) applications:
  – Video downloads
  – Massively multiplayer games
  – Video collaborations
  – Telepresence
  – IPTV
  – Applications on a wire, etc.

• If you are happy with the PMO of our networks:
  – slow downloads
  – jittery streaming
  – unreliable audio
  then DOCS is not for you.

DOCS Architecture – Physical Aspects

• Circuit switches are suitable to optics (McKeown et al.)
  – Require no packet processing or buffering

• Circuit switches are simple
  – Requires about 90% less hardware than an equivalent packet switch
  – More attractive than equivalent packet switches:
    • About eight times the capacity of a packet switch
    • Five times the capacity per watt
    • Cost one quarter of the price of the equivalent packet switch.
  – Simplest way to build a circuit switch is to “Start with a packet switch and throw most of it away” [Pablo-HotNets-I 2002, IEEE MICRO 2002, JON 2003]

• Circuit switches perform well
  – Fast restoration times (less than 50 ms)
  – Contain less software than equivalent routers
  – Good service guarantees: no queuing delay (so predictable packet delays), and guaranteed bandwidth
DOCS: Basic Premise

- **DOCS** = Dynamic Optical Circuit Switching

- **Approach:**
  - Bursty (packet) traffic generated by users/applications
  - Aggregate traffic at the network edge
  - Establish high-bandwidth pipes between edge nodes through the network core
  - **DOCS** offers bandwidth-on-demand capabilities to applications (users can “dial” for bandwidth)

- **Example Applications:**
  - Real-time download (say within 5 sec)
  - Database/website backup (say between 1 am – 3 am, and not to exceed a 15-min duration)

New Trends for Future Telecom Backbone Networks

- **Emergence of new bandwidth-hungry applications:**
  - Scientific grid-computing applications
  - Consumer applications: Video-on-Demand (VoD), IPTV, ...

- **Emerging services that require dynamism and flexibility:**
  - Flexible bandwidth reservation, e.g., Bandwidth on Demand (dial for bandwidth), Advanced reservations, ...
  - Service guarantees (guaranteed bandwidth, fast switching/restoration times, SLA, availability)

- **Increasing gap between optical and electronic transmission and processing speeds**
**DOCS Example Applications - DDRs**

- **DDR = Deadline-Driven Request**
- **Real-time file download**
  - maximum transfer time a person can tolerate (perhaps 5 sec)
- **Database/website/server backup:** scheduled at any time, perhaps during night, but with a fixed deadline (completion time)
  - flexible application requiring high bandwidth, but not necessarily instantaneously.
- **Our research:** Provision DDRs in a DOCS network from two perspectives:
  - **Time dimension:** allows flexibility w.r.t. when to schedule the request.
  - **Transmission bandwidth (rate) dimension:** allows flexibility (adjustable to network state) w.r.t. data rate to be allocated to the request.

---

**Ethernet Everywhere**

- **Ethernet is a success story in Local Area Networks (LAN)**
  - About 90% of LANs use Ethernet.
- **Extending its reach from LAN into Metro Area Networks (MAN) has already been established.**
- **Focus now is to extend Ethernet into carrier core networks.**
  - Future mode of operation: Ethernet over WDM $\rightarrow$ native Ethernet frames directly over WDM.
    - Elimination of several layers of other technologies.
    - CapEx and OpEx savings.
  - Connection-oriented Ethernet.
  - Forwarding: VLAN-XC, Provider Backbone Transport (PBT), T-MPLS.
- **Following requirements must be taken into account:**
  - High resilience.
  - Long reach: 1500-4000 km.
  - Rates of up to 100 Gbit/s Ethernet (GbE).
  - High degree of mesh.

ACK: Siemens (NSN)
Transmission Rates

- **Ethernet Rates:**
  - 100 Gbit/s Ethernet.
  - Max possible CapEx savings.

- **Constraint: Signal transmission range for a certain rate**
  - Signal’s quality depends on the physical impairments.
  - Transmission Range = Signal traveled distance after which signal quality degrades to a level that it needs regeneration.

- **Transmission Ranges:**
  - Range of 10 Gbit/s signal = 3000 Km.
  - Range of 100 Gbit/s signal = 500 Km.

- **Etherpath** = Lightpath carrying Ethernet frames.

---

Telecom Nets: “End-to-End” Ethernet?

![Telecom Nets Diagram](image-url)
TE vs. NE vs. NP

- **Traffic Engineering (TE)**
  - “Put the traffic where the bandwidth is”

- **Network Engineering (NE)**
  - “Put the bandwidth where the traffic is”

- **Network Planning (NP)**
  - “Put the bandwidth where the traffic is forecasted to be”

- **TE** – online, dynamic, provisioning problem, ms time scale
- **NE** – intermediate problem, months time scale
- **NP** – offline, static, dimensioning problem, 5-yr time scale

Robust Network Design

**Network Events:**

- Connection arrival
- Connection departure
- Fault occurs
- Fault fixed

**Network Architecture Needs:**

- More dynamism / flexibility / agility / automation ...
- Handle multiple (near-simultaneous) faults efficiently
- Reprovision (backup) capacity when “network state” changes
  - Efficiency tradeoff: bandwidth vs. implementation
- Is a service path “ahead” or “behind” the contract (SLA)?
  - Reprovision, if necessary; role of penalty and economics in SLA
- Utilize “excess capacity”? 
Further Reading

**Broadband Access**
- PONs
- WDM in PONs – *IEEE Photonics in Switching Conf. (PIS)*, Aug. 07
- Long-Reach Broadband Access – stay tuned
- Hybrid wireless-optical access – *JLT, Nov. 07; OFC’07; ICC’07, ...*
- Metro: The vanishing breed?

**Backbone Networks**
- DOCS (“dial for bandwidth”) – *PIS, Aug. 07*
- Robust Network Design – *Lots of literature*
- Ethernet Everywhere – *OFC’07 PD paper; GB’07; JLT, Jan. 08*

Extra Slides
What Is An Optical Network?

- It is **NOT NECESSARILY** all optical
  - packet switched

- **Characteristics of an optical network**
  - Transmission: optical
  - Switching: could be optical, could be electronic, could be hybrid
    - could be circuit, could be packet, could be burst

- **Most Promising Approach Today**
  - Electronic circuit switching with sub-lambda granularity (STS-1, STS-3, ...)

- **Example Utility for IP Networking**
  - Connect any two IP routers (geographically far apart) with a direct
    - ("virtual") bandwidth pipe... of whatever capacity (STS-1, ..., STS-192)
  - Increase (or decrease or delete) the capacity on demand
  - Dynamically control the "topology" connecting the IP routers
  - Create a "separated control network" (of whatever bandwidth)

---

Resilient Mesh Net --Overview

- **Multi-path routing**
  - How many paths?
  - State-dependent path capacities
  - Virtual concatenation (VC)
    - ("traffic grooming")

- **Survivable Routing**
  - Primary + backup path(s)
  - Precomputed backup (?)
  - Shared backup (?)
  - Need reliable path (five-9’s)
    - Component/link availability
  - Need fast recovery
Data over (Next-Gen) SONET/SDH

- Inverse multiplexing
- Relaxing time-slot contiguity/alignment
- Reassembly and jitter control

Next-Gen SONET Chip Availability

<table>
<thead>
<tr>
<th>Company</th>
<th>B/W (b/s)</th>
<th>Trans. GFP</th>
<th>VC</th>
<th>VC Groups</th>
<th>Differential Delay</th>
<th>LCAS</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agere</td>
<td>622M</td>
<td>No</td>
<td>High</td>
<td>16</td>
<td>±62.5 ms</td>
<td>N/S</td>
<td>$347</td>
</tr>
<tr>
<td>Agere</td>
<td>1.2G</td>
<td>No</td>
<td>High</td>
<td>16</td>
<td>±62.5 ms</td>
<td>N/S</td>
<td>$379</td>
</tr>
<tr>
<td>Agere</td>
<td>2.5G</td>
<td>No</td>
<td>High</td>
<td>16</td>
<td>±62.5 ms</td>
<td>N/S</td>
<td>$522</td>
</tr>
<tr>
<td>Agilent</td>
<td>2.5G</td>
<td>No</td>
<td>High</td>
<td>4</td>
<td>±80 ms</td>
<td>N/S</td>
<td>$475</td>
</tr>
<tr>
<td>Cypress</td>
<td>2.3G</td>
<td>Yes</td>
<td>High</td>
<td>16</td>
<td>±16 ms</td>
<td>N/S</td>
<td>$375</td>
</tr>
<tr>
<td>Galazar</td>
<td>2.3G</td>
<td>No</td>
<td>Low/High</td>
<td>24</td>
<td>±1.25 ms</td>
<td>H/W</td>
<td>$500</td>
</tr>
<tr>
<td>MultiLink</td>
<td>10G</td>
<td>No</td>
<td>High</td>
<td>64</td>
<td>±1.28 ms</td>
<td>RISC</td>
<td>-</td>
</tr>
<tr>
<td>PMC-Sierra</td>
<td>2.5G</td>
<td>No</td>
<td>High</td>
<td>2</td>
<td>±25 ms</td>
<td>N/S</td>
<td>$340</td>
</tr>
<tr>
<td>Transwitch</td>
<td>155M</td>
<td>No</td>
<td>Low/High</td>
<td>8</td>
<td>±32 ms</td>
<td>H/W</td>
<td>$125</td>
</tr>
<tr>
<td>Transwitch</td>
<td>2.5G</td>
<td>Yes</td>
<td>High</td>
<td>48</td>
<td>±32 ms</td>
<td>RISC</td>
<td>$550</td>
</tr>
<tr>
<td>Vitesse</td>
<td>2.5G</td>
<td>Yes</td>
<td>High</td>
<td>48</td>
<td>±54 ms</td>
<td>H/W</td>
<td>$240</td>
</tr>
<tr>
<td>Vitesse</td>
<td>10G</td>
<td>Yes</td>
<td>High</td>
<td>64</td>
<td>±54 ms</td>
<td>H/W</td>
<td>$536</td>
</tr>
<tr>
<td>West Bay</td>
<td>2.3G</td>
<td>No</td>
<td>High</td>
<td>48</td>
<td>±24 ms</td>
<td>H/W</td>
<td>$400</td>
</tr>
</tbody>
</table>

Source: Light reading
Problem Statement

• **Given:**
  - Network topology $G (V, E, C(e, w))$
  - Basic time-slot component, e.g., VT1.5, STS-1, STM-1, etc.
  - A high-speed request $R$ with bandwidth requirement $B$.
  - An inverse-multiplexing control parameter $K$ for maximal allowed path number.

• **Find:**
  - $m$ distinct paths, where $1 \leq m \leq K$,
    such that the aggregated capacity they offer $\geq B$.

Implication to Service Resilience

• Route splitting enables *degraded services*
  - Ethernet service: peak rate 600 Mbps, guaranteed rate 300 Mbps (against single network-element failure)
  - Need LCAS
Motivation:
Critical Infrastructure Survivability

- **Overloads, Attacks, and Failures:**
  - What’s the difference? It’s a “Race against Time”

- “Our society depends on critical infrastructures for delivery of essential services.
- “Telecom, banking, power supply, public transport, etc. increasingly rely on information infrastructures
  ... not only for management and control
  ... but also for monitoring outages and recovery.
- “Combinations of wireless, ad hoc, and fixed networks are becoming a reality in many domains.”

- Tomorrow's networks face the survivability challenge:
  - how to deliver critical services in a timely manner in presence of overloads, attacks, and failures?

---

Modeling “Disasters”

- Define SRGs (shared-risk groups)
  - for (potentially) correlated failing components
  - e.g., links + nodes in close geographical proximity

- SRLG (shared-risk link group)
  - Special case of SRG
  - Quite common since fibers are laid in ducts, and duct cuts are more frequent than we wish

- Extend link models (or develop new models) to apply to SRGs
Which Method?

- **PREV: Provisioning fast REstorable VCG**
  - One backup path per node pair

- **PIVM: Protecting Individual VCG Member**
  - One working VCG per connection
  - One backup VCG per working VCG member