Issues and Challenges in Optical Network Design

Biswanath Mukherjee
Distinguished Professor
Department of Computer Science
University of California, Davis
mukherje@cs.ucdavis.edu

Tutorial Talk
37th European Conference and Exhibition on Optical Communication, Geneva, Switzerland
September 21, 2011
Telecom Network Hierarchy

Long haul
- 100s-1000s km
- Mesh

Metro (interoffice)
- 10s of km
- Rings

Access
- a few km
- Hubbed rings, PONs

Users
Telecom Network Hierarchy

- **Long haul**
  - 100s-1000s km
  - Mesh

- **Metro (interoffice)**
  - 10s of km
  - Rings

- **Access**
  - a few km
  - Hubbed rings, PONs

**The “Last” Mile**
Telecom Network Hierarchy

Long haul
- 100s-1000s km
- Mesh

Metro (interoffice)
- 10s of km
- Rings

Access
- a few km
- Hubbed rings, PONs

The “Last” Mile
“First”
An Access Network (PON)
Basic PON Structure

- **Optical Line Terminal (OLT)**
- **Optical Network Units (ONU) / Optical Network Terminal (ONT)**

Long-Haul, Backbone Networks

Yesterday’s (SONET) Net

- Months to roll out new connections

Today’s Net

- Minutes (or seconds) to set up connections

Tomorrow’s Net
TE vs. NE vs. NP

- Traffic Engineering (TE)
- Network Engineering (NE)
- Network Planning (NP)
TE vs. NE vs. NP

• Traffic Engineering (TE)
  – “Put the traffic where the bandwidth is”

• Network Engineering (NE)

• Network Planning (NP)
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)
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 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
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
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

Cost ($$)$

Blocking probability
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
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering

Electronic (Computer) Engineering
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering

Electronics (Computer) Engineering

Software Engineering
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering

Electronic (Computer) Engineering

Software Engineering

System VALUE
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering

Electronic (Computer) Engineering

Software Engineering

System A

System VALUE
System/Network: Value Proposition

Current state of the art

Optical Science & Engineering

Electronic (Computer) Engineering

Software Engineering

"My" system + Yours too (?)

System A

System VALUE
Holistic Network Design: (Convergence Across “Layers”)

Applications
(“Customer” needs)

Network Architect
(many of us, ...)

Physical Layer
(optical comm. channel) --
materials, devices, subsystems
Holistic Network Design:
(Convergence Across “Layers”)

**Applications**
(“Customer” needs)

**Network Architect**
(many of us, ...)

- Network architectures to combat optical channel impairments
- **10/40/100G MLR network architectures**
- Breakthroughs needed in device technologies:
  - optical RAM, ultra-wideband amp, ...

**Physical Layer**
(optical comm. channel) --
materials, devices, subsystems
Holistic Network Design: (Convergence Across “Layers”)

Applications (“Customer” needs)

“Cloud” Services: (Virtual) Storage, Computing, ...

Differentiated Services: Bandwidth, Availability, Failure-Recovery Delay (“X ms”), ...

Network Economics: Pricing, SLA, ...

Sustainability: Energy, ...

Network Architect (many of us, ...) + Network architectures to combat optical channel impairments + 10/40/100G MLR network architectures + Breakthroughs needed in device technologies: - optical RAM, ultra-wideband amp, ...

Physical Layer (optical comm. channel) -- materials, devices, subsystems
Optical Network Design: Different Paradigms

- End-to-end ethernet
- Mixed-line-rate network design
- Dynamic optical circuit switching
- Hybrid dynamic circuit/packet switched network
- Robust network design
- Excess capacity management
- Energy-efficient network design
- Data-center network design
- Broadband access network design
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 → 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)
Telecom Nets: “End-to-End” Ethernet?
Transmission Rates and Mixed Line 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

- **Mixed Line Rates:**
  - 10 Gbit/s, 40 Gbit/s, 100 Gbit/s waves.
  - Need for (hierarchical) grooming.

- **Etherpath** = Lightpath carrying Ethernet frames.
Single Line Rate

![Graph showing single line rate with node connections and rates.](image-url)
Single Line Rate
Single Line Rate
Single Line Rate

40G REGENERATION
Single Line Rate
Single Line Rate

100G REGENERATION
Single Line Rate

100G REGENERATION
Mixing up Line Rates

[Graph with nodes labeled 1 to 13 and edges labeled with numbers 10G, 40G, and 100G]

09/21/2011
B. Mukherjee, UC Davis
Mixing up Line Rates
Mixing up Line Rates

Graph with nodes labeled 1 to 14 and edges with line rates indicated (10G, 40G, 100G).
Mixing up Line Rates

Graph showing network connections with line rates indicated by different colors:
- Red: 10G
- Blue: 40G
- Green: 100G

Nodes and edge labels:
- Node 1 connected to Node 2 with a line rate of 4800
- Node 2 connected to Node 3 with a line rate of 1500
- Node 3 connected to Node 4 with a line rate of 1500
- Node 4 connected to Node 5 with a line rate of 1200
- Node 5 connected to Node 6 with a line rate of 1200
- Node 6 connected to Node 7 with a line rate of 2400
- Node 7 connected to Node 8 with a line rate of 1500
- Node 8 connected to Node 9 with a line rate of 1500
- Node 9 connected to Node 10 with a line rate of 1500
- Node 10 connected to Node 11 with a line rate of 1500
- Node 11 connected to Node 12 with a line rate of 1500
- Node 12 connected to Node 13 with a line rate of 600
- Node 13 connected to Node 14 with a line rate of 300
- Node 14 connected to Node 1 with a line rate of 2100

Network connections and line rates are visually represented with arrows and numbers.
• Demand from 2→12: 10 Gbps paths only
• Demand from 2→7: 10 and 40 Gbps paths only
• Demand from 2→5: All of 10, 40 and 100 Gbps paths
Why Mixed Line Rate (MLR) Design?

• Is 100 Gbps always better?

• Tradeoffs:
  – Higher bit-rate
  – Volume discount
  \[ \text{VS} \]
  – Higher Cost
  – Shorter Reach

• MLR design
  – Promises to address these tradeoffs
  – Serves heterogeneity in traffic granularities
Mixed Line Rate (MLR) Networks

Mixed 10G, 40G, 100G

- Each node can have transponders of different line rates
- Same fiber can carry different line rates on different wavelengths
DCS: Dynamic Circuit Switching
DCS: Dynamic Circuit Switching

• 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
DCS: Dynamic Circuit Switching
DCS: Dynamic 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
  – DCS offers bandwidth-on-demand capabilities to applications (users can “dial” for bandwidth)
DCS: Dynamic 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
  – DCS 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)
DCS: Dynamic 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
  – DCS 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)

• Provision DDRs (Deadline-Driven Requests):
  – **Time dimension**: allows flexibility w.r.t. when to 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.
ESNet: Hybrid Dynamic Circuit/Packet Nets

ESnet IP switch/router hubs
ESnet IP core
ESnet Science Data Network (SDN) core (N X 10G)
International IP Connections

Layer 1 optical nodes
ESnet SDN switch hubs

09/21/2011
B. Mukherjee, UC Davis
Hybrid Dynamic Circuit/Packet Nets: Motivation

**Present Mode of Operation (PMO):**
- Fixed partition between SDN and packet networks

**Future Mode of Operation (FMO) (?):**
- Flexible partition between SDN and packet networks
Robust Network Design

• Multi-path routing
  – “Degraded service” vs. no service at all
Robust Network Design

• Multi-path routing
  – “Degraded service” vs. no service at all
Robust Network Design

• Multi-path routing
  – “Degraded service” vs. no service at all

• Fast reprovisioning
  – (upon network-state change)
Robust Network Design

• Multi-path routing
  – “Degraded service” vs. no service at all

• Fast reprovisioning
  – (upon network-state change)
Robust Network Design

• Multi-path routing
  – “Degraded service” vs. no service at all

• Fast reprovisioning
  – (upon network-state change)

• Data replication
  – across Disaster Protection Zones (DPZs)
Robust Network Design

- **Multi-path routing**
  - “Degraded service” vs. no service at all

- **Fast reprovisioning**
  - (upon network-state change)

- **Data replication**
  - across Disaster Protection Zones (DPZs)

- **Multicast support?**
  - for various services?
Excess Capacity Management

Any operational network usually has some spare capacity at a given time to accommodate variations of traffic demands and avoid early exhaustion.
Exploiting Excess Capacity to Improve Robustness

• Exploit excess capacity (EC) to improve:
  – Set up time
  – Protection switching time
  – Availability

• Pre-provisioning to decreases the connection setup time by reserving capacity for primary and backup paths in advance.

• Mixed-protection schemes that exploit EC to improve protection switching time and availability.
Energy Consumption of Telecom Networks

US Commercial Electricity Price

Source: US Energy Information Administration

Green Energy >20 cents

33%
Internet Traffic Profile

Source: AMS-IX traffic statistics
Network Domain Utilization

Network Domain Utilization

<50% Utilization

<30% Utilization

<15% Utilization

Energy-Efficient MLR/OFDM Network Design

• Adaptive link bandwidth allocation to save energy using MLR or orthogonal frequency division multiplexing (OFDM)

• Problem Statement:
  - Find optimum routing of lightpaths; and
  - Efficient allocation of flexible bandwidth
    • To optimize total energy consumption network

• Tradeoffs
  - Capacity vs. all-optical reach
  - Capacity vs. energy consumption
  - Energy consumption vs. all-optical reach
Data-Center Network (DCN)
Network Failure due to Disasters

30,000 km. of fiber optic cables
4,000 of telecom offices
Disaster Resilient Data-Center Network Design

- Protection against cascading and correlated multiple failures in communication network needs serious attention
  - Natural Disasters, WMD attacks, EMP attacks

- Cloud services based on datacenter networks are becoming very important. Needs to ensure-
  - Content protection
  - Path protection

- Content placement, routing and protection of paths and content are closely related to one another, so the interaction among these should be studied together
Energy-Efficient Content Distribution over Telecom Network Infrastructure

- Data center based solution
  - Centralized, over provisioned

- Distributed storage system
  - Content hosted across the network

- P2P based nano data center (NaDa)
  - Reduces power consumption for ISP
Problem Statement

- Given
  - Network with DC, Core, Metro, and Access
  - Set of content objects with known popularity
  - Storage availability for hosting nodes (can be infinite)
  - Requests of contents of the form (Client Node, Content, BW, [timestamp]) following Zipf distribution

- Goal: energy-efficient content placement while satisfying the requests
Long-Reach Broadband Access

ACK: ETRI Korea and NSF
Wireless Optical Broadband Access Network (WOBAN)
Summary

• Various paradigms of optical network design has been presented

• Networks of the future will be heterogeneous with convergence of various technologies

• Convergence among:
  – Access/metro/core network segments
  – Optical and wireless technologies
  – Physical, network, and services “layers”
  • including energy-conservation issues