Telecom ~~Optical~~ Networks: The Road Ahead

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

mukherje@cs.ucdavis.edu
http://networks.cs.ucdavis.edu/~mukherje/

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:

  Biswanath Mukherjee, "Optical WDM Networks," Springer, 2006.

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

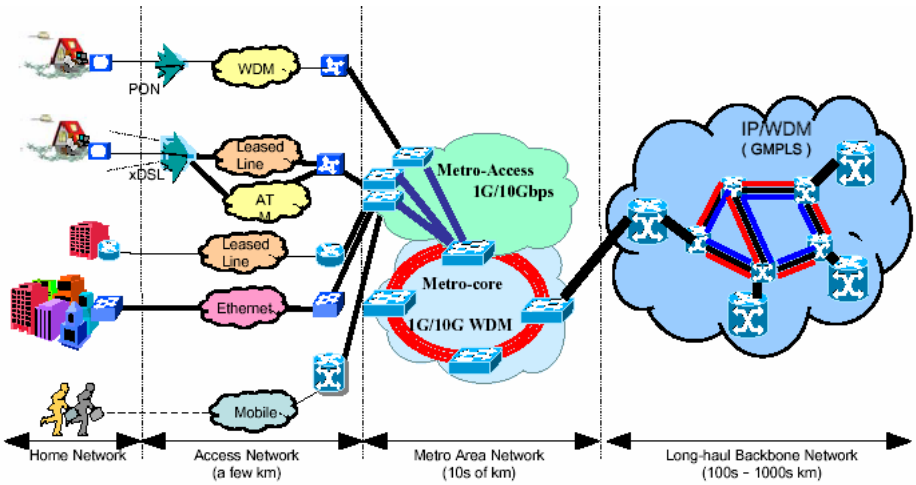
- DOCS ("dial for bandwidth")

(✗) Robust Network Design (multi-layer, multi-domain, multi-path, etc.)
- Ethernet Everywhere

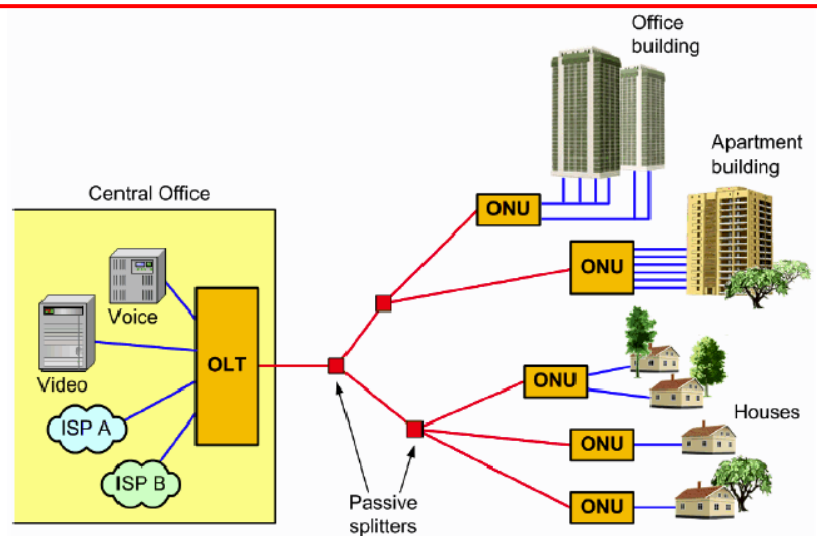
- Network Engineering (NE) (vs. TE vs. NP)

(✗) Higher-density switches!

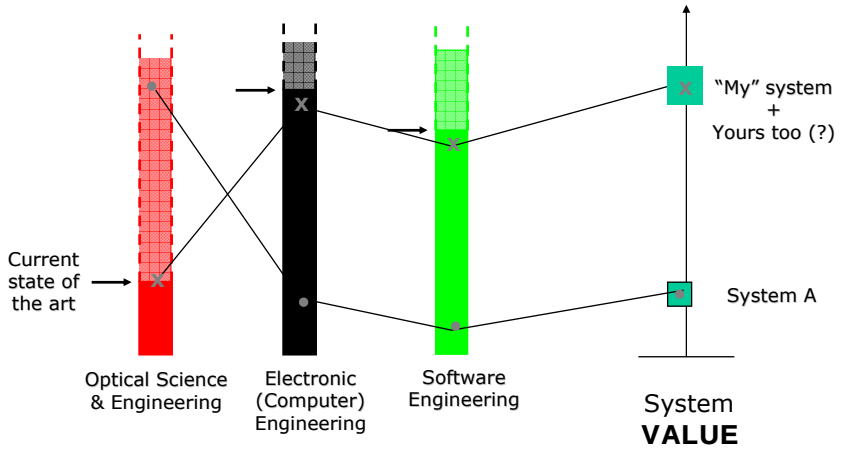
Telecom Network Hierarchy



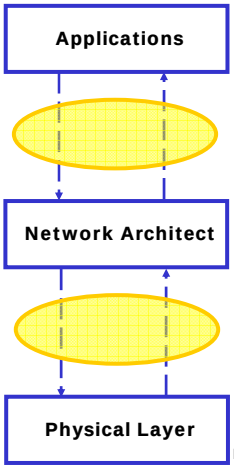
An Access Network (PON)



System / Network: Value Proposition



Optical Network Architecture:
Extending Our "Boundaries"



("Customer" needs)

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

(Piotr, I, …)

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

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

Telecom ~~Optical~~ Networks: The Road Ahead

Broadband Access

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

Backbone Networks

- DOCS ("dial for bandwidth")
- (X) Robust Network Design (multi-layer, multi-domain, multi-path, etc.)
- Ethernet Everywhere

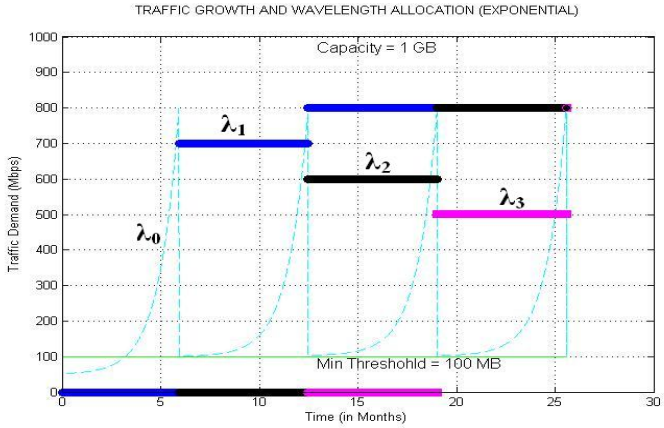
- Network Engineering (NE) (vs. TE vs. NP)
- (X) Higher-density switches!

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)

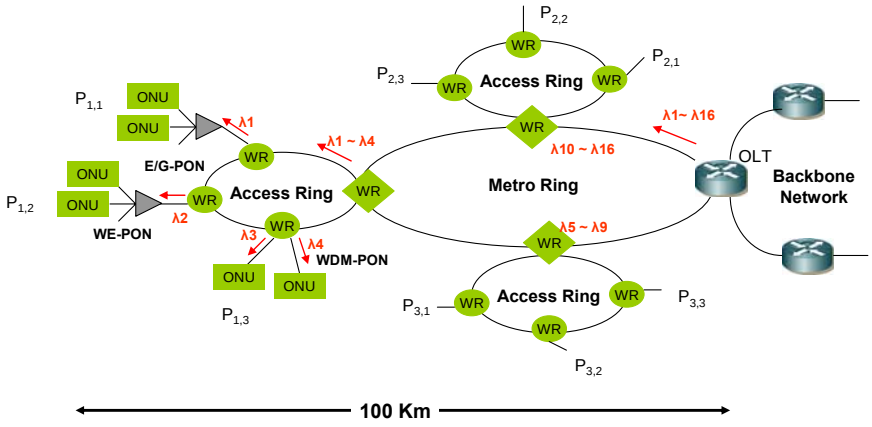


- Shifting (exponential model)
- λ_0 threshold times: 6, 12.5, 19, 25.5-months

Long-Reach Broadband Access

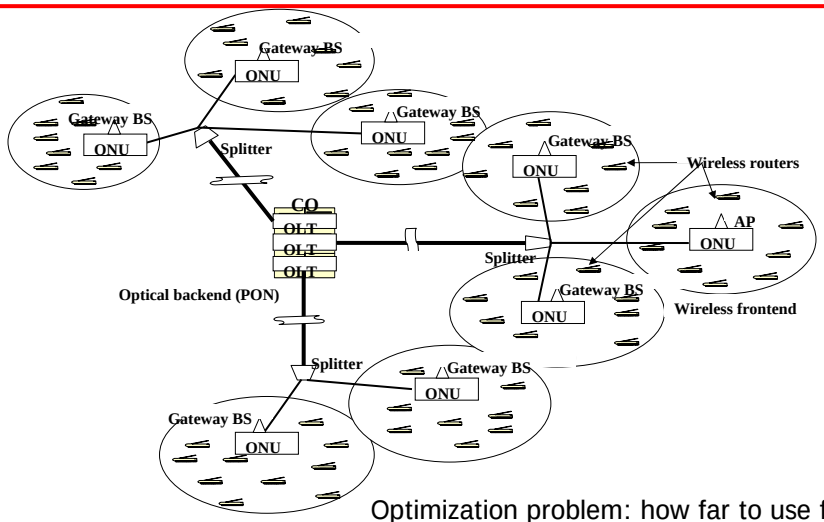
- Basic Problem
  - Reach of traditional PON = 10-20 km
  - Fiber capacity: much larger than data generated by users served by the limited geography in a traditional PON
  - Hence Long-Reach (LR) PON should be investigated
    - Increase distance to 100 km or more… may need power (so not exactly a "PON"… 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

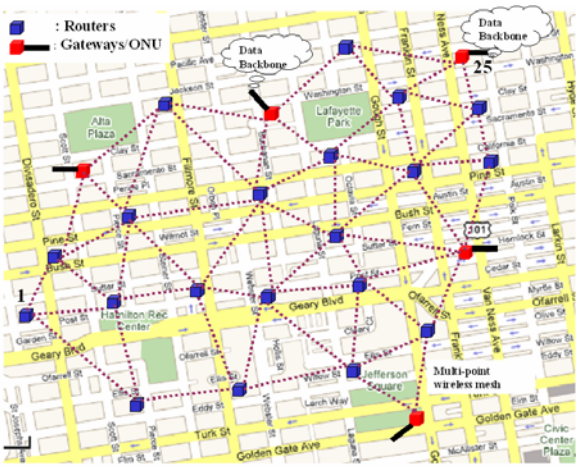
Hybrid Wireless-Optical Broadband Access Network (WOBAN)



ACK: Nokia (NSN)

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

Example: WOBAN Front-End Mesh



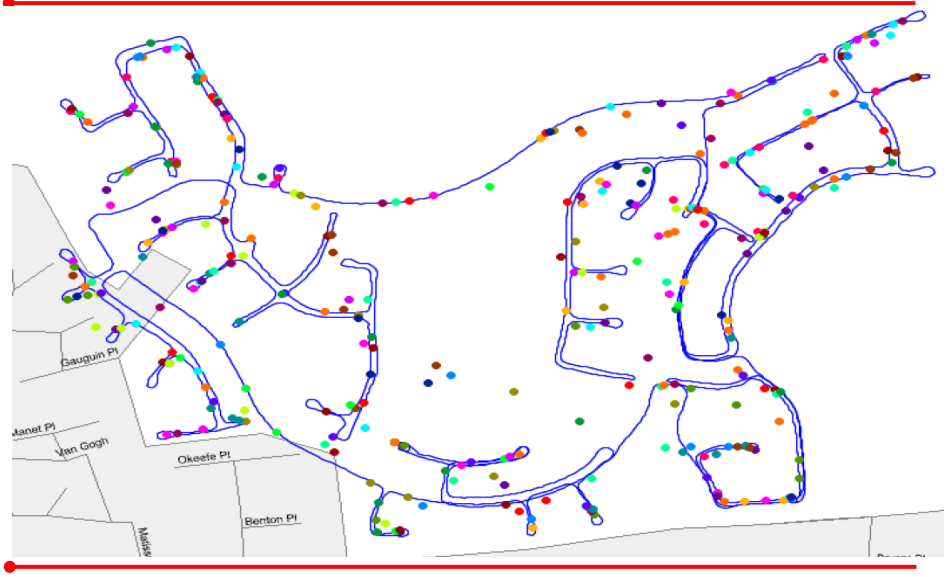
Location:
Golden Gate Ave. to Pacific Ave.
Divisadero to Van Ness Avenue

Area: 1 sq-mile (approx.)

Population: 15,000 (approx.)

A part of San Francisco WOBAN (SFNet)

Wildhorse, Davis, CA: Wireless Routers



Current Deployment Status

Area/Location	Architecture	Compatibility		Configuration	Operating Range	Player
		Present	Future			
Akron, OH	Flat AP infrastructure	WiFi	WiFi	Multiple radio	2.4 GHz	MobilePro
Athens, GA	Multi-layered MP2P	WiFi	WiFi, WiMax	Multi-radio	2.4, 5 GHz	Belair
Bristol, UK	Multi-layered deployment	WiFi	WiFi, WiMax	Multi-radio multi-antenna	2.4, 5 GHz	Belair
Chaska, MN	Flat deployment	WiFi	WiFi, WiMax	Single radio (omni-directional)	2.4 GHz	Tropos, Pronto
Corpus Christi, TX	Flat deployment (GPS-compatible)	WiFi	WiFi	Single radio (omni-directional)	2.4 GHz	Tropos, Pronto
Culver City, CA	Flat (intermesh capable)	WiFi	WiFi	Multi-radio omni-directional	2.4, 5 GHz	Firetide
Farmers Branch, TX	Gateways with OC-3 ingress	WiFi	WiFi, WiMax	Multiple radio	2.4, GHz	NeoReach, Pronto
Galt, CA	Multi-layered deployment	WiFi	WiFi, WiMax	Multi-radio multi-antenna	2.4, 5 GHz	Belair
Gilbert, AZ	Flat AP infrastructure	WiFi	WiFi	Multiple radio	2.4 GHz	MobilePro
Gordes, France	Flat with intermesh	WiFi	WiFi	Multi-radio omni-directional	2.4, 5 GHz	Firetide
Isla Vista, CA	Flat deployment	WiFi	WiFi	Multi-radio	2.4, 5 GHz	Firetide
Islington, UK	3-tier hierarchical deployment	WiFi	WiFi	Multi-radio multi-antenna	2.4, 5 GHz	Belair
Moorehead, MN	P2MP with fiber optic backbone	WiFi	WiFi	Single radio (omni-directional)	2.4 GHz	Tropos
New Orleans, LA	WiFi routers with digital IP cameras attached to IP-backbone for video surveillance system					Tropos
Philadelphia, PA	Currently being deployed					Earthlink
San Francisco, CA	Currently being deployed					Earthlink, Google
Springfiled, MO	Hierarchical (L2 VLAN capable)	WiFi	WiFi	Multi-radio	2.4, 5 GHz	Belair
St. Maarten, Carribean	Hierarchical deployment	WiFi	WiFi	Multi-radio	2.4, 5 GHz	Belair, Lucent
Tempe, AZ	Gateways with OC-3 ingress	WiFi	WiFi, WiMax	Multi-radio multi-antenna	2.4, 5 GHz	Strix, NeoReach
Wavion, Inc. is a new player with their "spatially adaptive" MIMO-based routers having an antenna array and six radio transreceivers.						

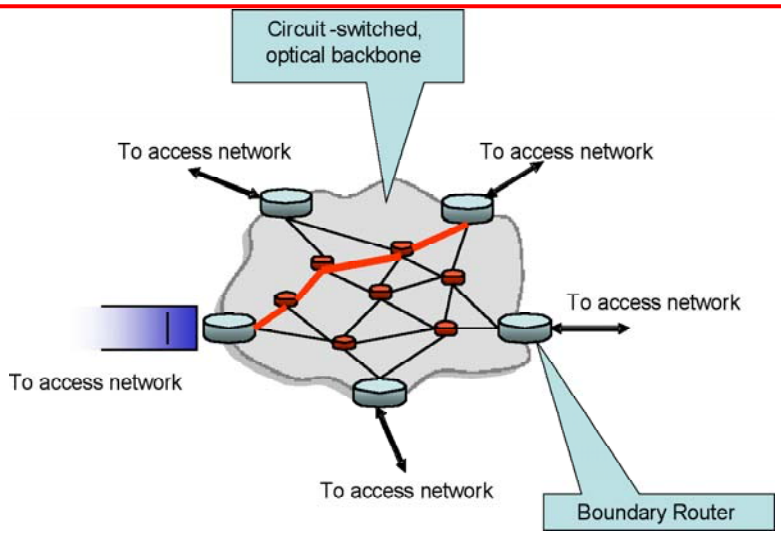
Ack: Suman Sarkar

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 m[illegible] 1.57[illegible]
8 * dc-oak-dc2--ucd-ge.cenic.net [illegible]
9 dc-sfo-dc1--oak-dc2-pos.cenic.n[illegible] [illegible] m[illegible] [illegible]55 ms
10 dc-svl-dc1--sfo-dc1-pos.cenic.n[illegible] [illegible]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

What is This?

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 <u>not</u> 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 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

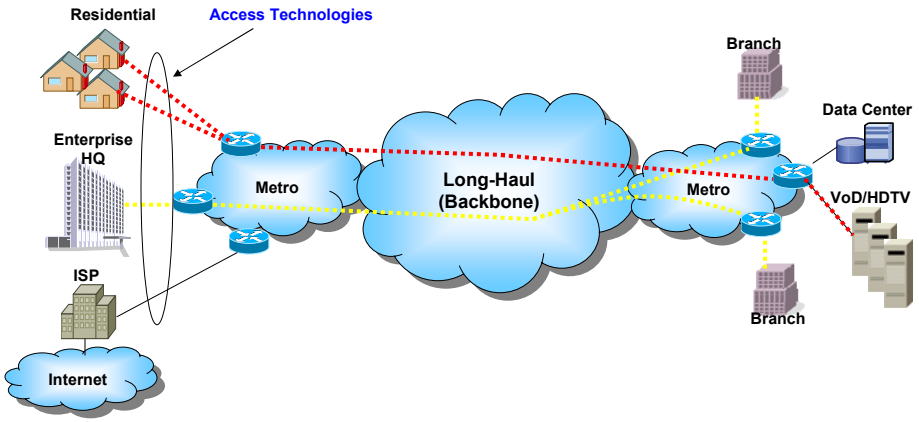
- **Ethernet is a success story in Local Area Networks (LAN)**
  - About 90% of LANs use Ethernet.
- **Extending its reach from LAN into Metro Area Networks (MAN) has already been established.**
- **Focus now is to extend Ethernet into carrier core networks.**
  - **Future mode of operation: Ethernet over WDM → 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?



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*"

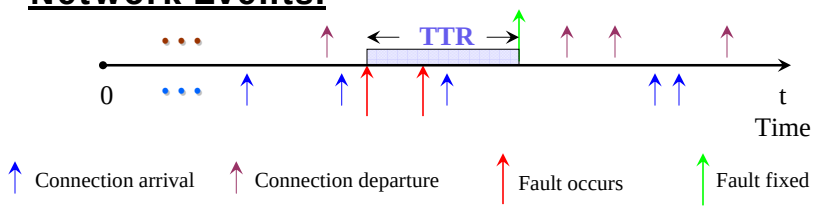
Blocking probability

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

Cost [$$$]

Robust Network Design

Network Events:



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

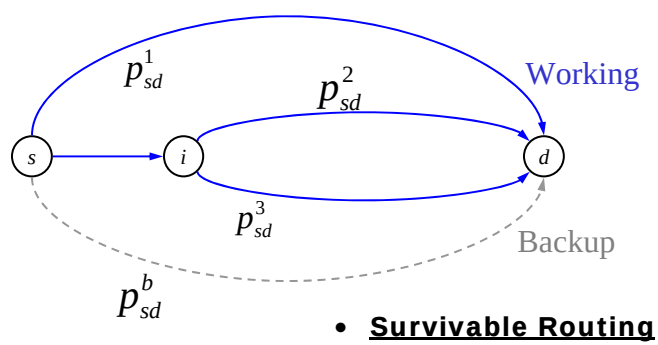
- DOCS ("dial for bandwidth") – PIS, Aug. 07
- Robust Network Design – Lots of literature
- Ethernet Everywhere – OFC'07 PD paper; GB'07; JLT, Jan. 08

Extra Slides

What Is An Optical Network?

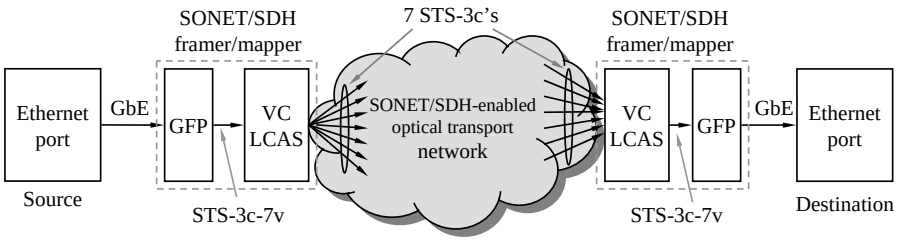
- It is **NOT NECESSARILY** all optical
  " " " " packet switched
- **Characteristics of an optical network**
  - **Transmission: optical**
  - **Switching:** could be optical, could be electronic, could be hybrid
    could be circuit, could be packet, could be burst
- **Most Promising Approach Today**
  - 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

Company	B/W (b/s)	Trans. GFP	VC	VC Groups	Differential Delay	LCAS	Price
Agere	622M	No	High	16	±62.5 ms	N/S	$347
Agere	1.2G	No	High	16	±62.5 ms	N/S	$379
Agere	2.5G	No	High	16	±62.5 ms	N/S	$522
Agilent	2.5G	No	High	4	±80 ms	N/S	$475
Cypress	2.5G	Yes	High	16	±16 ms	N/S	$375
Galazar	2.5G	No	Low/High	24	±125 ms	H/W	$500
Multilink	10G	No	High	64	±128 ms	RISC	-
PMC-Sierra	2.5G	No	High	2	±25 ms	N/S	$340
Transwitch	155M	No	Low/High	8	±32 ms	H/W	$125
Transwitch	2.5G	Yes	High	48	±32 ms	RISC	$550
Vitesse	2.5G	Yes	High	48	±54 ms	H/W	$240
Vitesse	10G	Yes	High	64	±54 ms	H/W	$536
West Bay	2.5G	No	High	48	±24 ms	H/W	$400

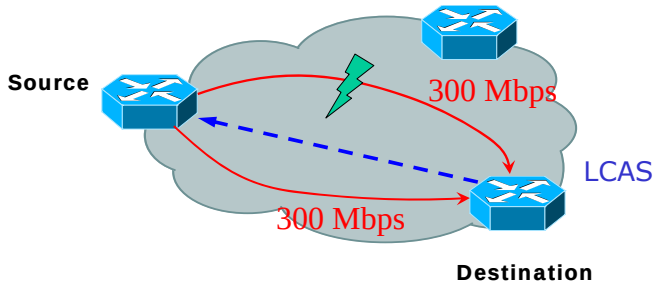
Source: Light reading

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?

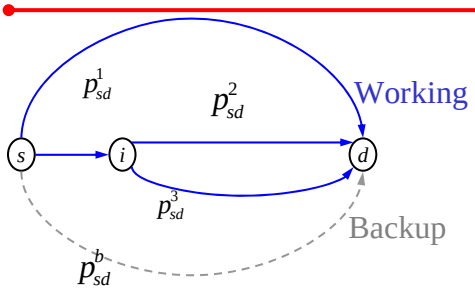
Modeling "Disasters"

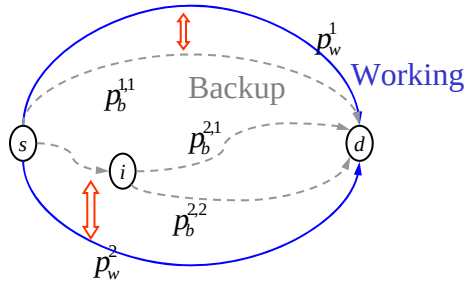
- Define SRGs (shared-risk groups)
  - for (potentially) correlated failing components
  - e.g., links + nodes in close geographical proximity

- 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