Optical Networking

Charlie Catlett
Senior Fellow, Computation Institute
catlett@mcs.anl.gov

Director, Illinois Wired/Wireless Infrastructure for Research and Education
(www.i-wire.org)

Executive Director, TeraGrid Project
(www.teragrid.org)

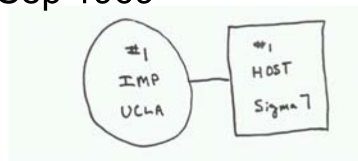
Chair, Global Grid Forum
(www.gridforum.org)

- "High Performance Nets"
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

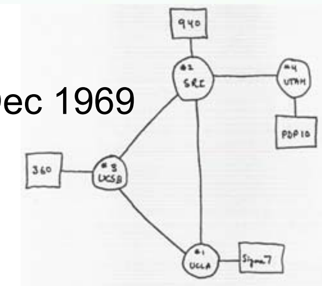
February 2003

The Internet: 1969 through Today

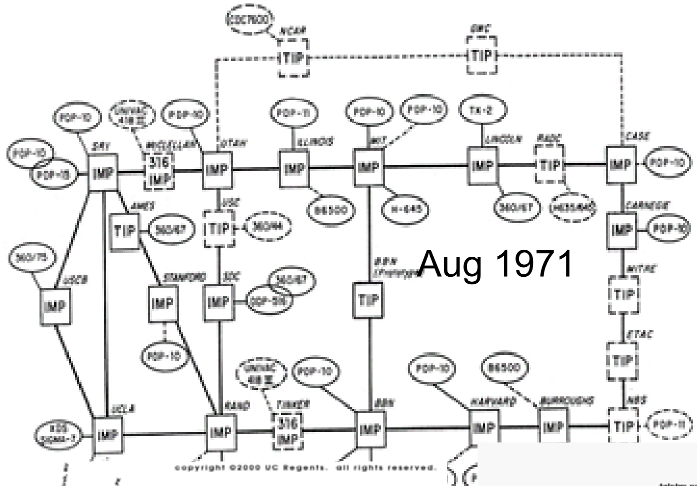
Sep 1969



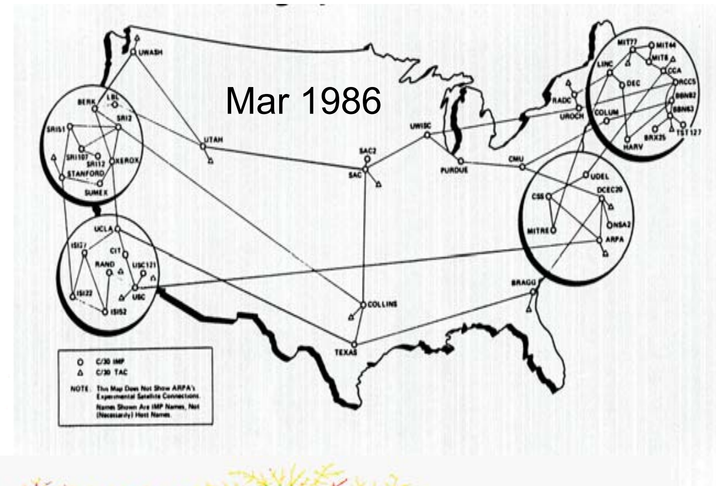
Dec 1969

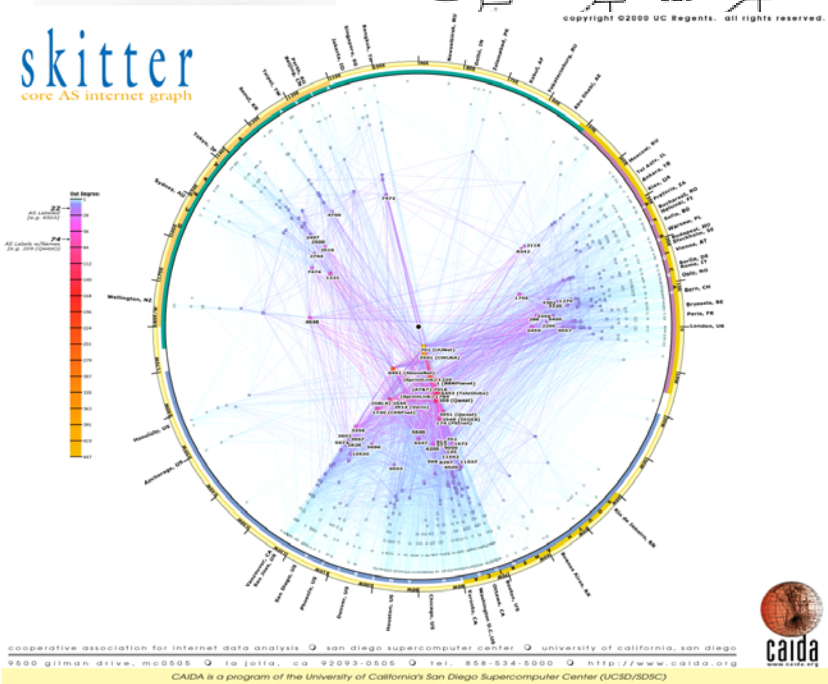


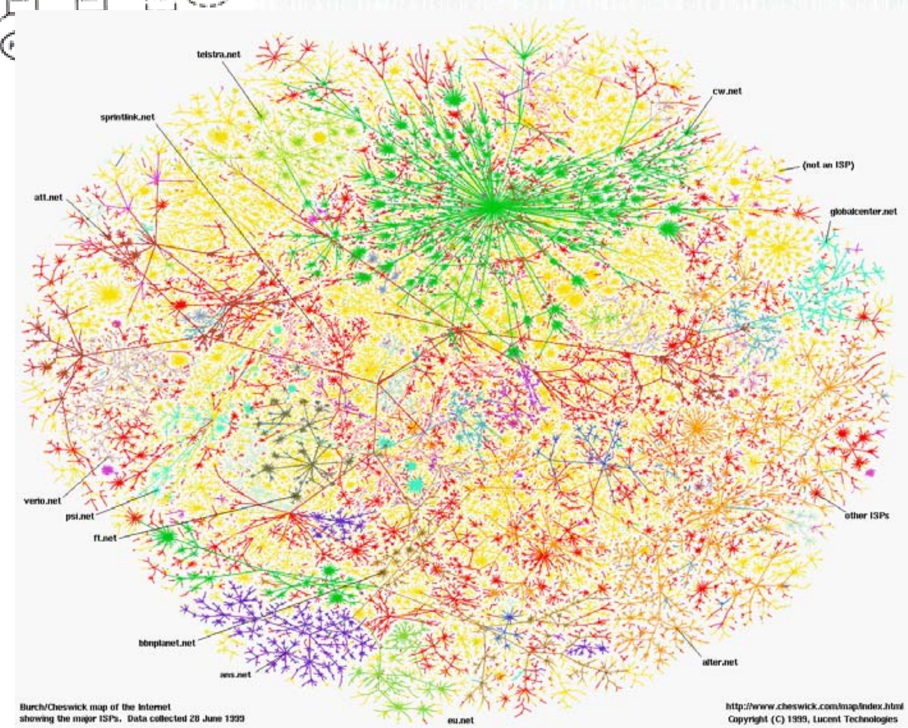
Aug 1971



Mar 1986

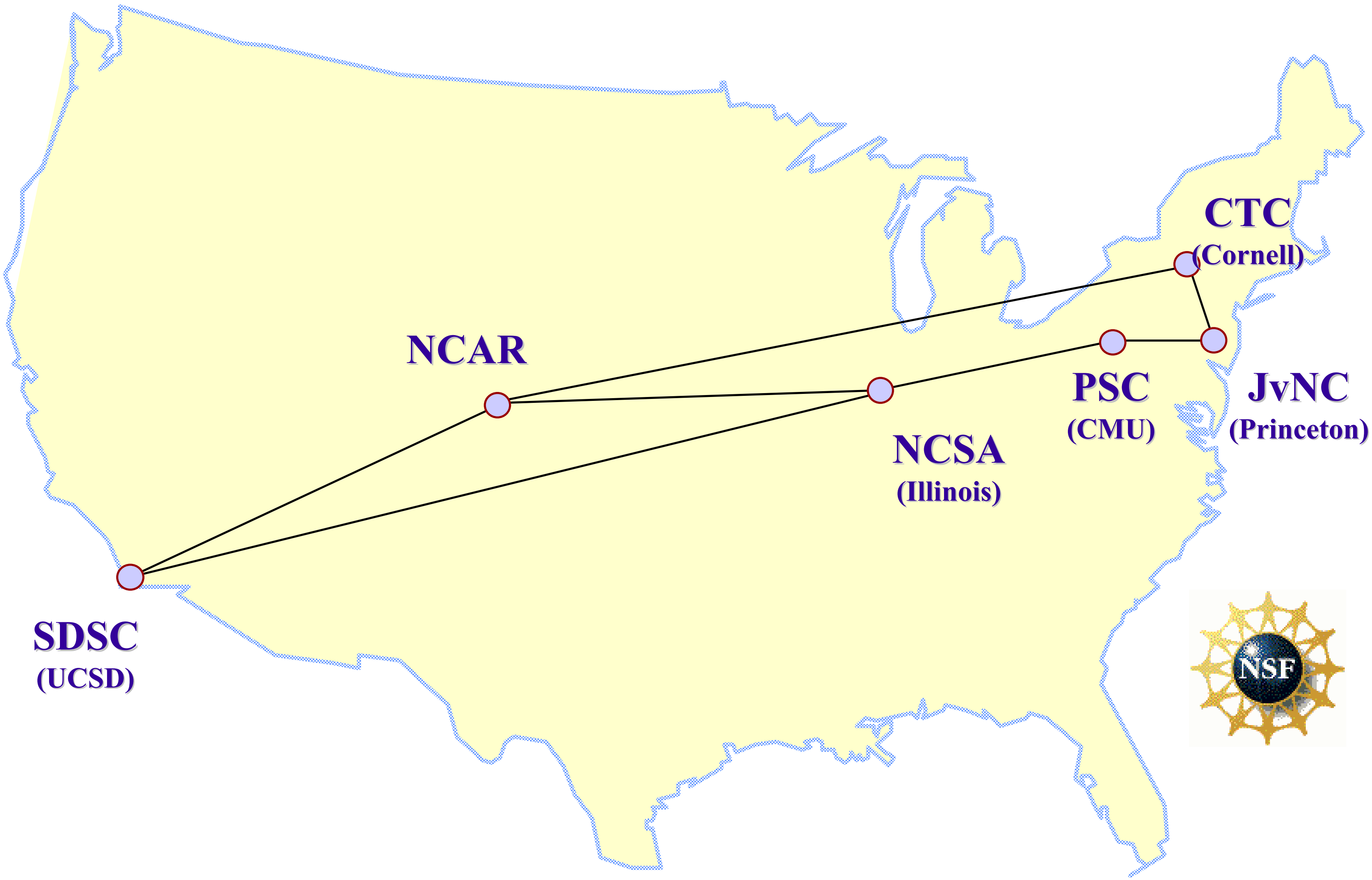






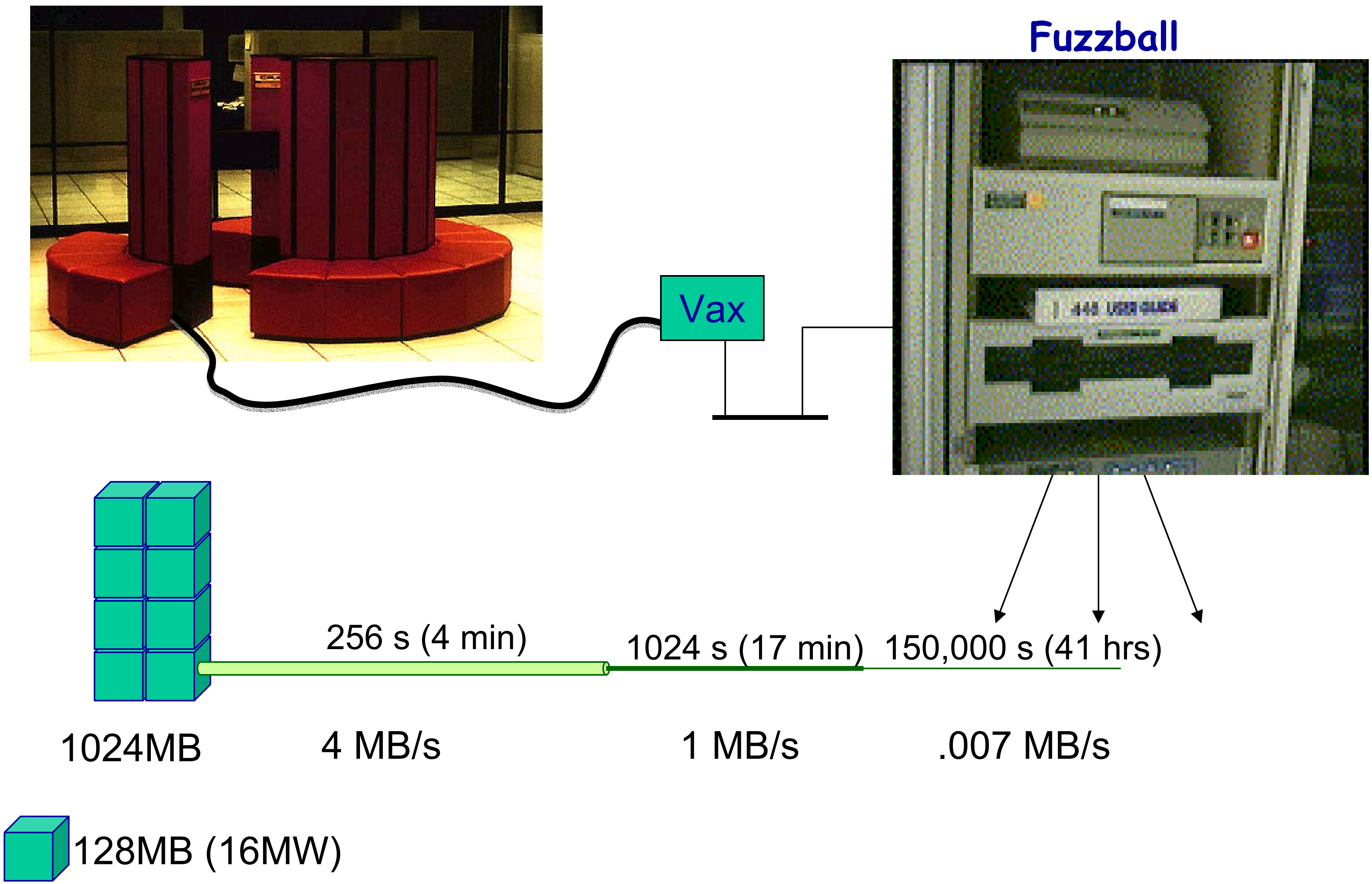
Sources: SIGCOMM Review, CAIDA/UCSD, AT&T Labs

NSFNET 56 Kb/s Backbone (1986-8)

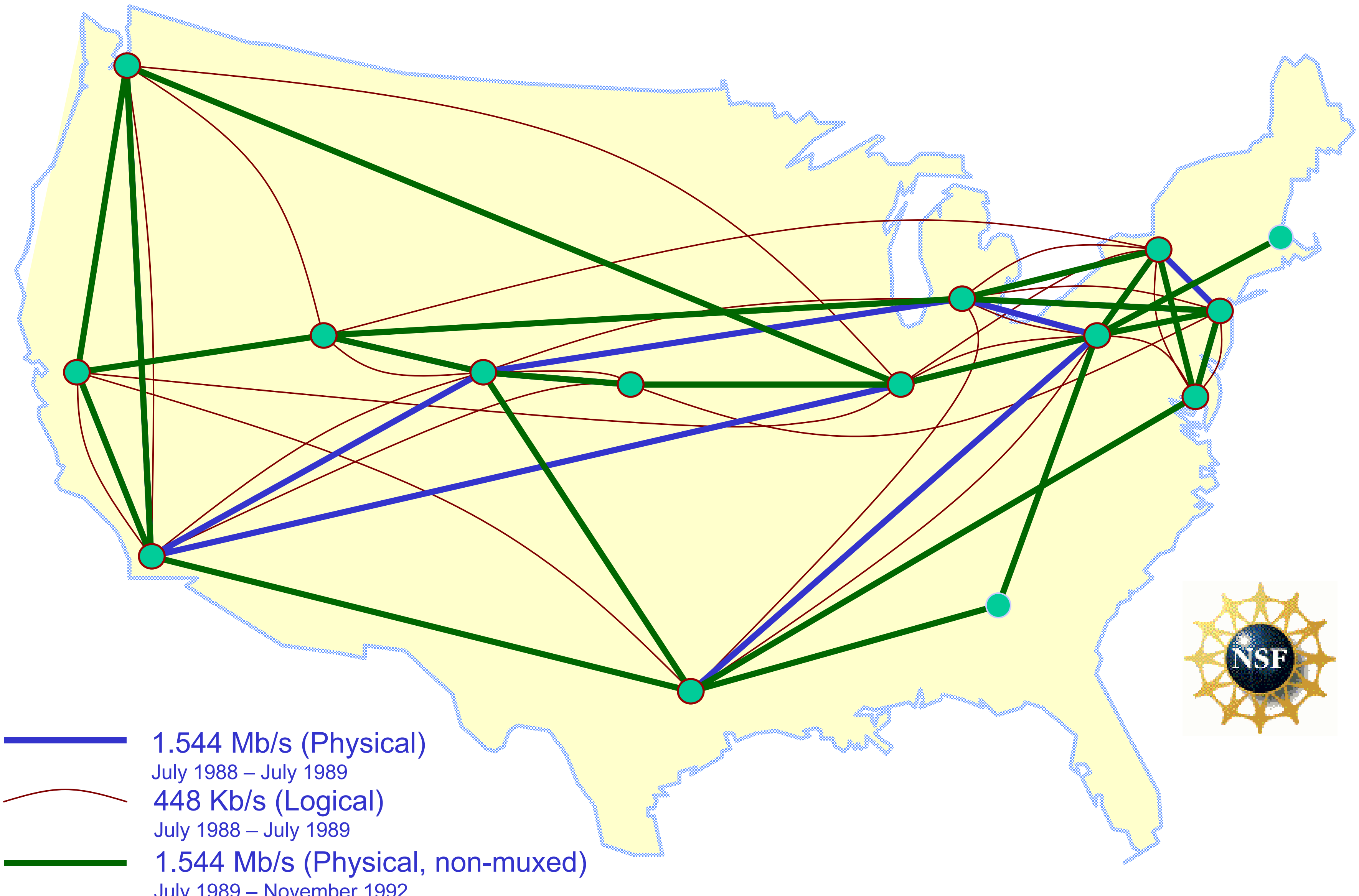


Charlie Catlett (catlett@mcs.anl.gov)

NSFNET 56 Kb/s Site Architecture



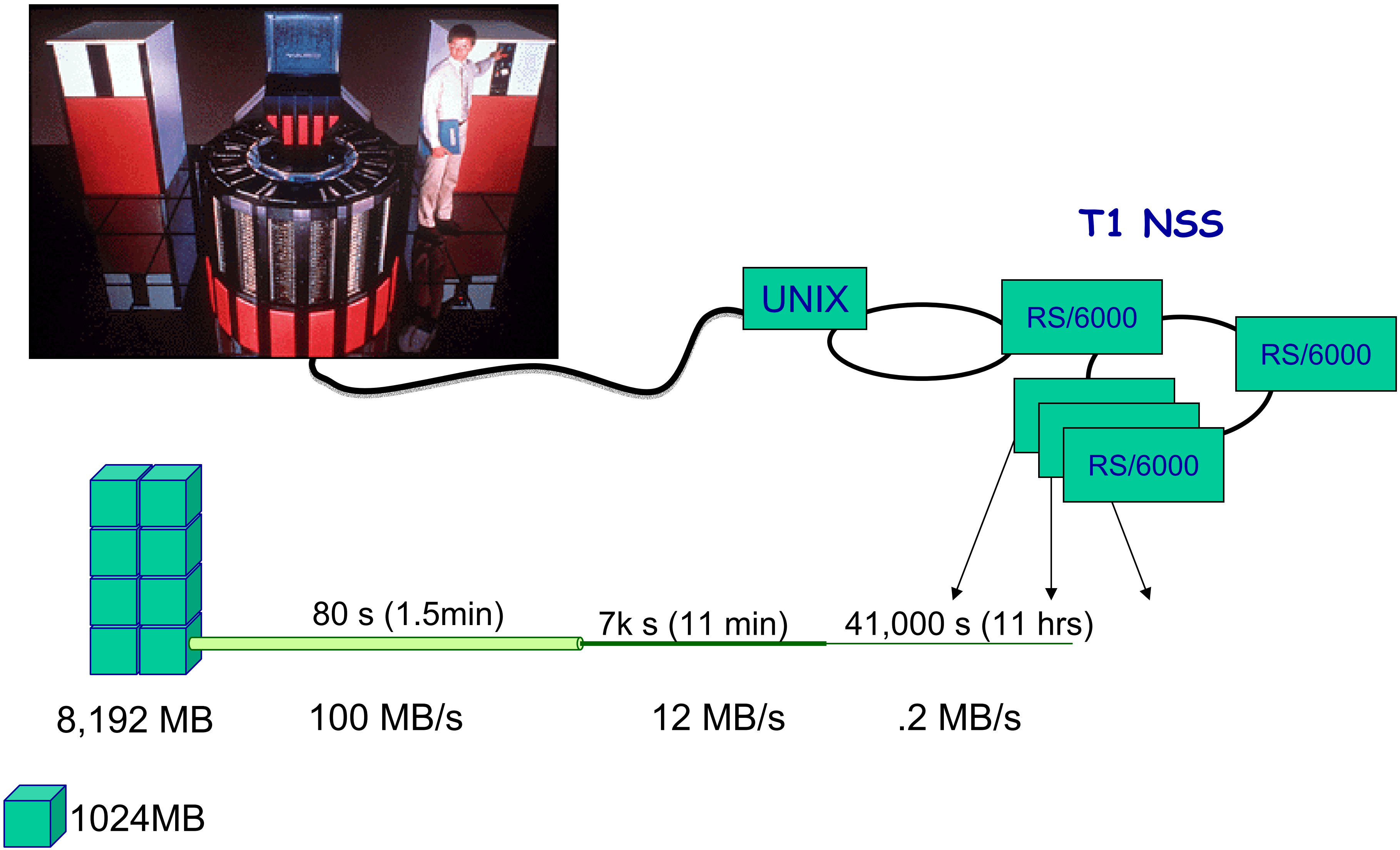
NSFNET T1 Backbone



Charlie Catlett (catlett@mcs.anl.gov)

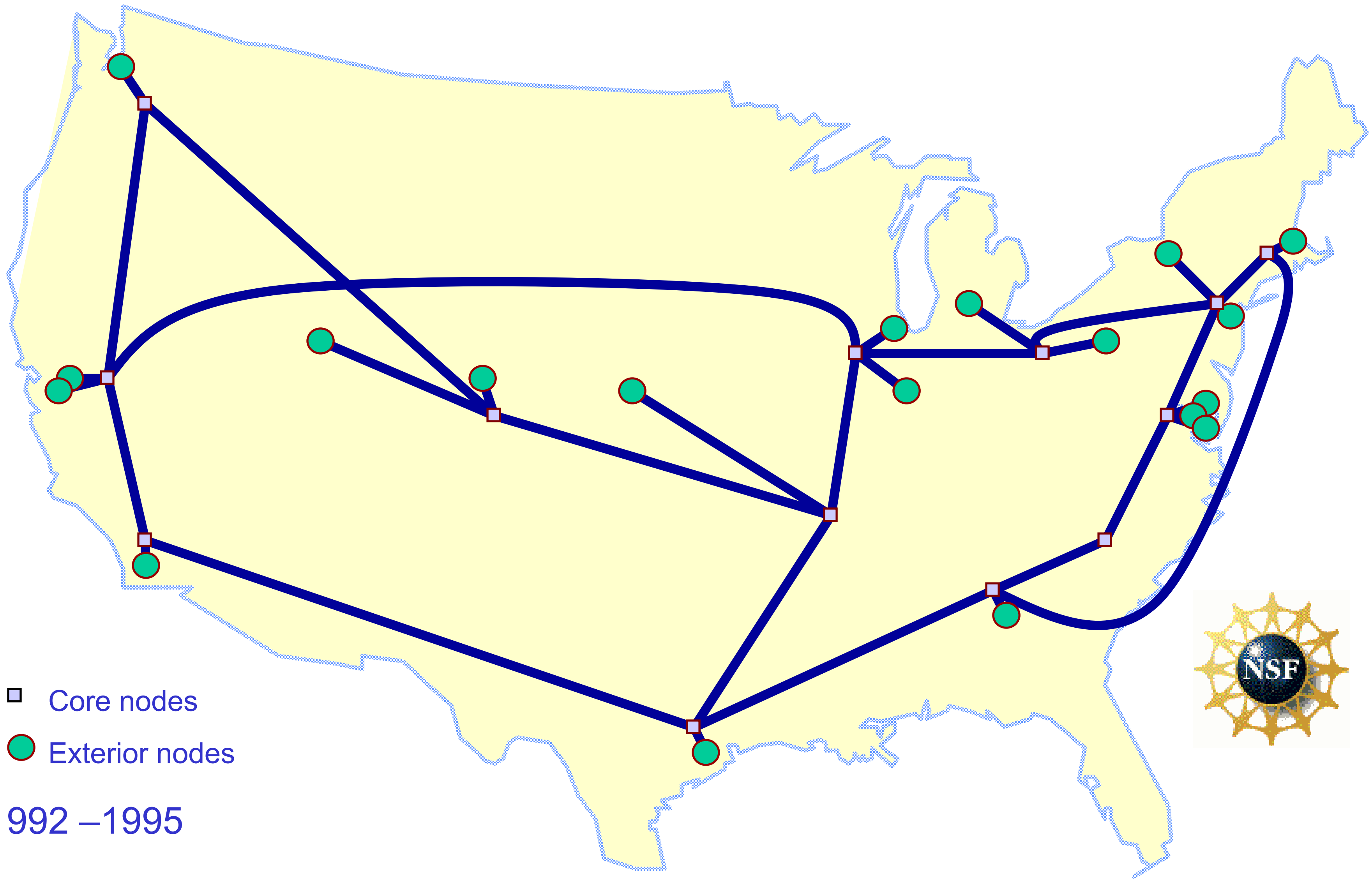
Topology data from NLANR MOAT (www.nlanr.net)

T1 Site Architecture



Charlie Catlett (catlett@mcs.anl.gov)

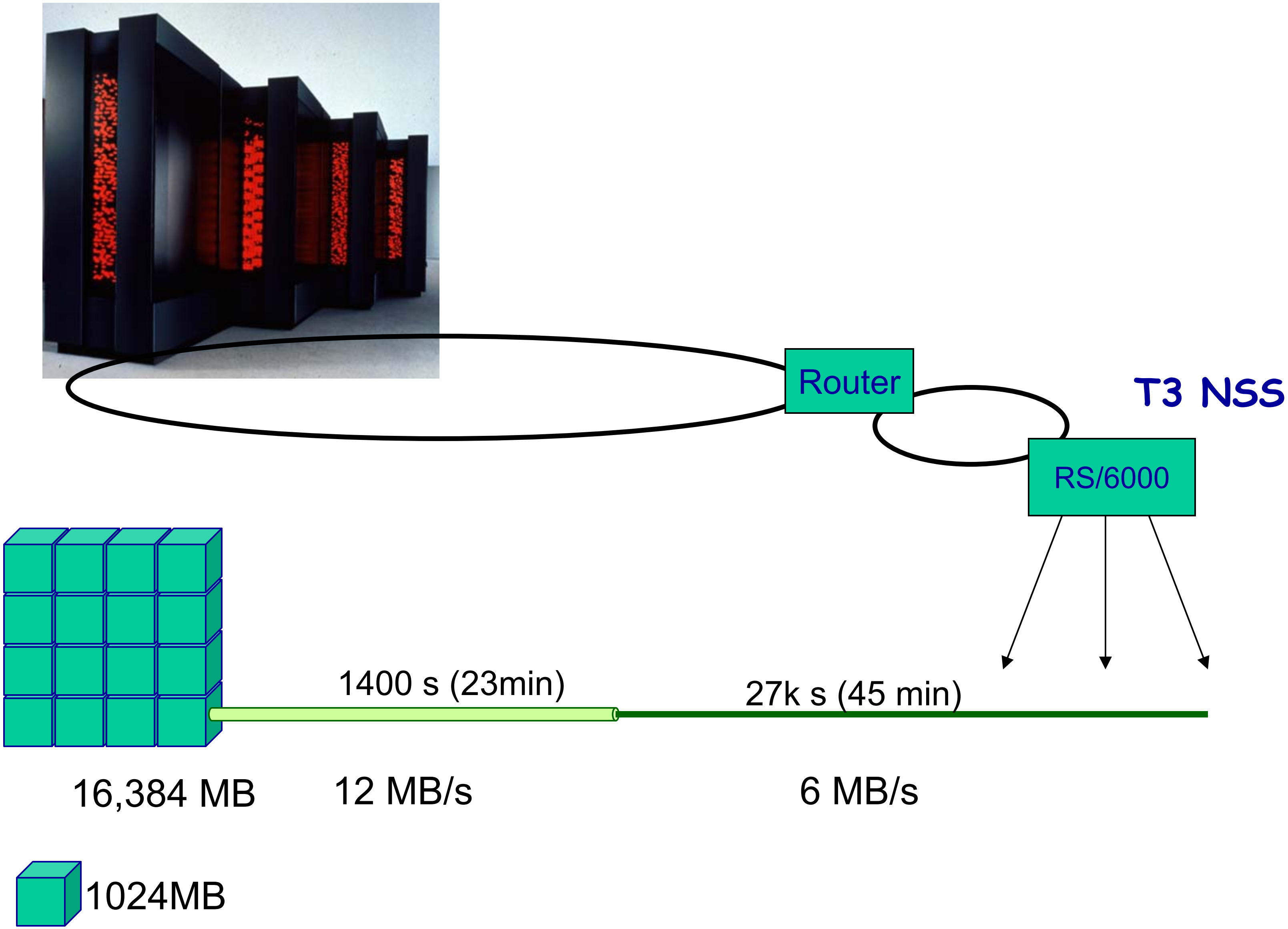
NSFNET T3 Backbone



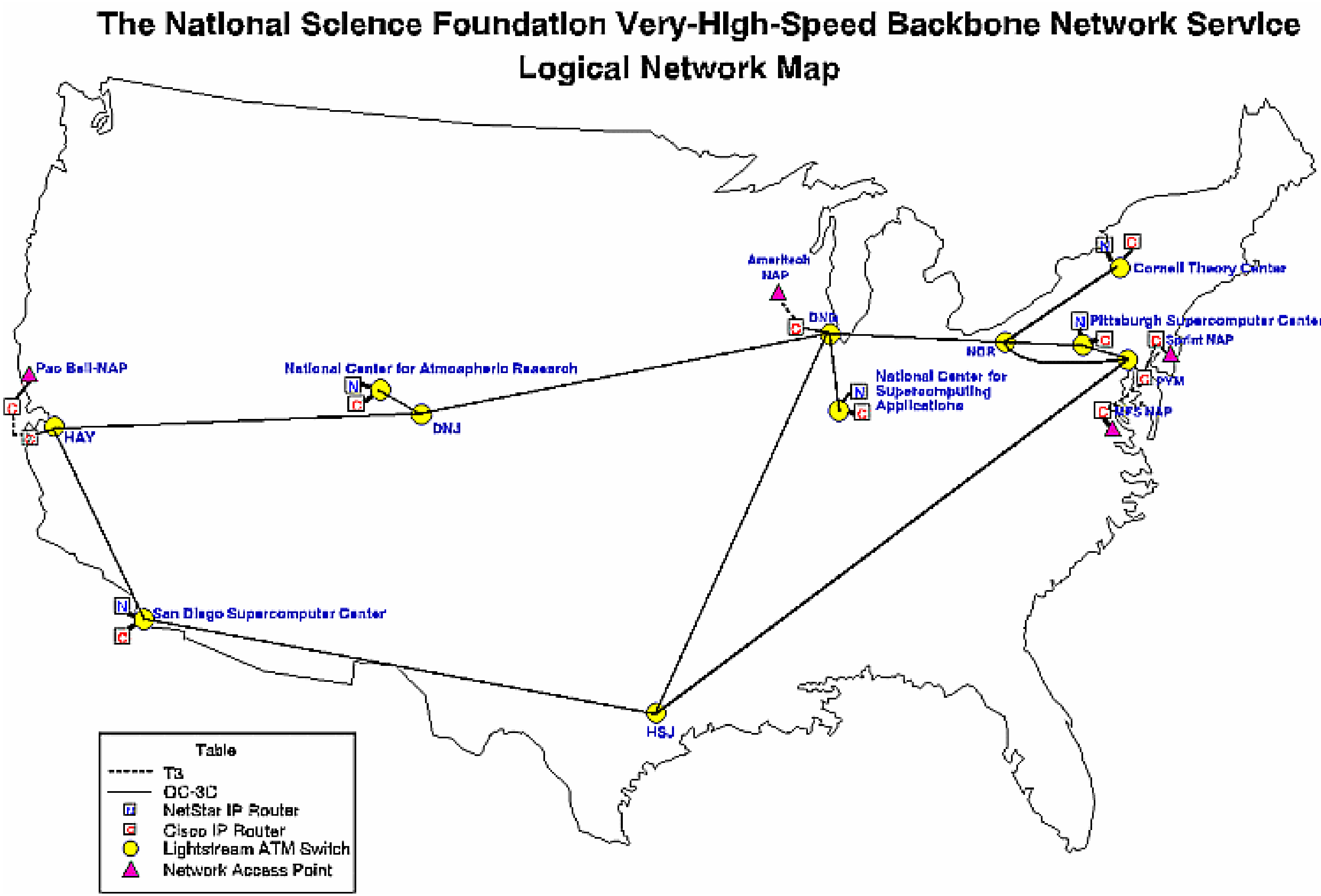
Charlie Catlett (catlett@mcs.anl.gov)

Topology data from NLANR MOAT (www.nlanr.net)

T3 Site Architecture

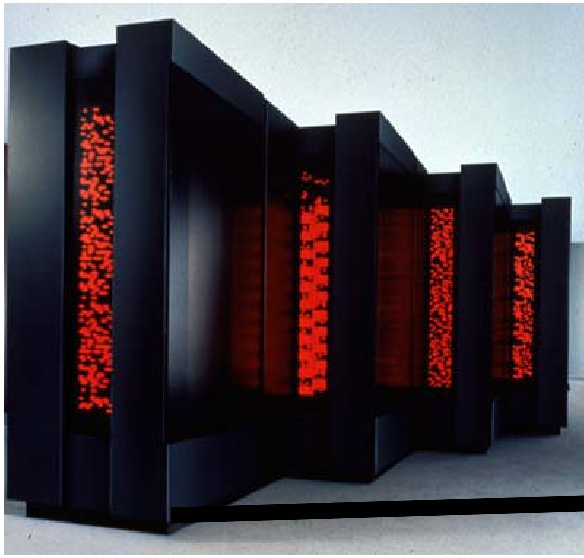


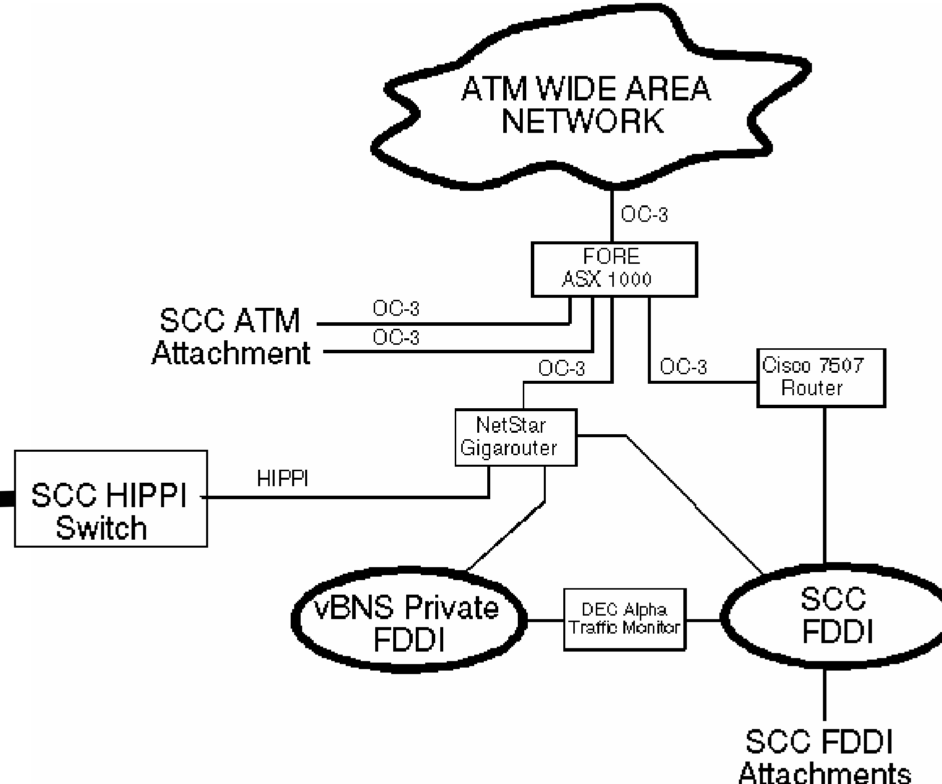
vBNS Logical Map

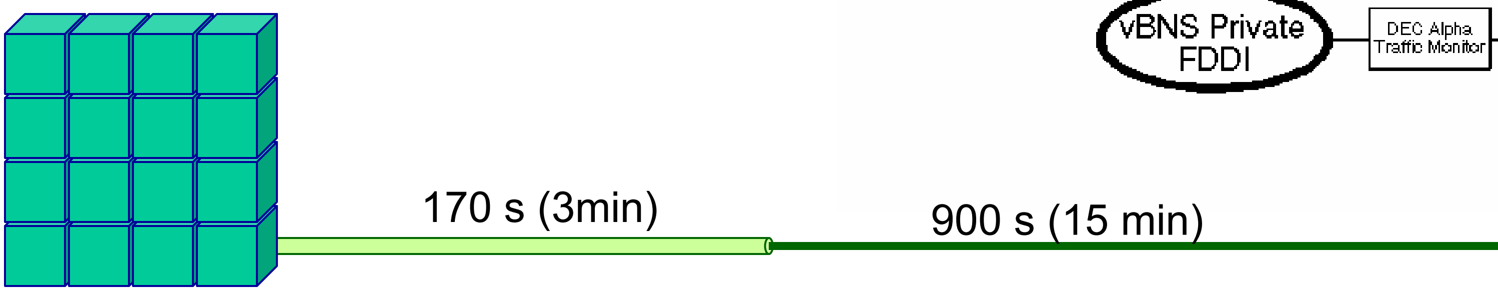


Source: MCI vBNS project (www.vbns.net)

vBNS Site Configuration







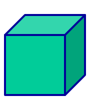
170 s (3min)

900 s (15 min)

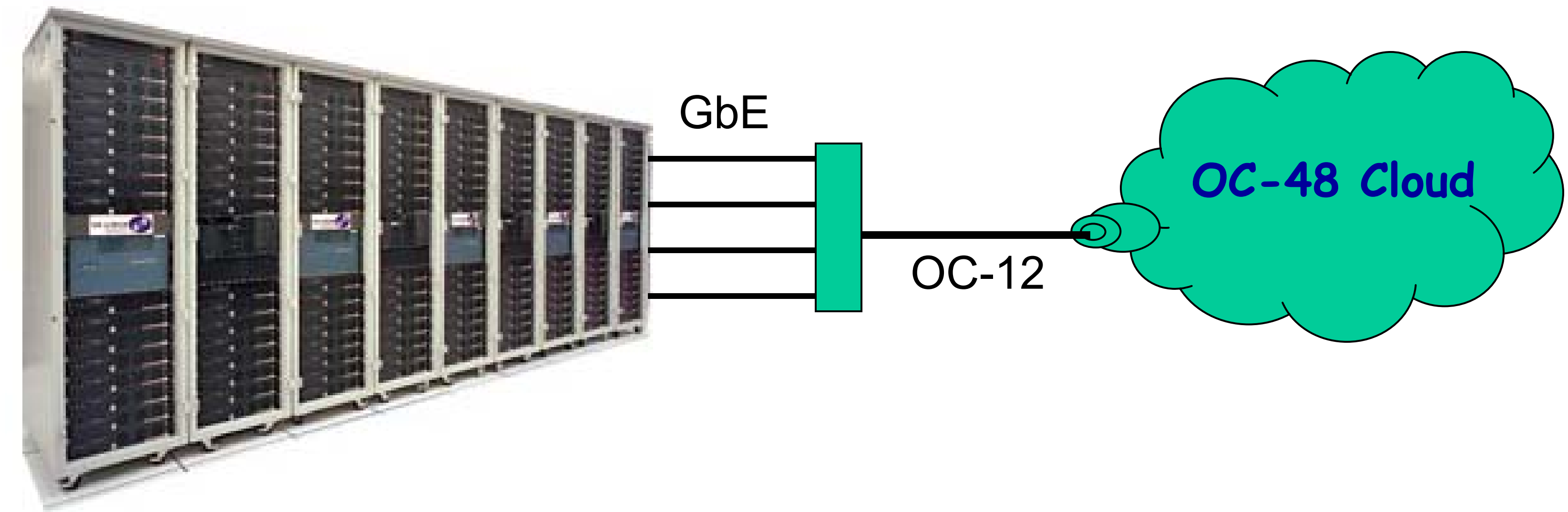
16,384 MB

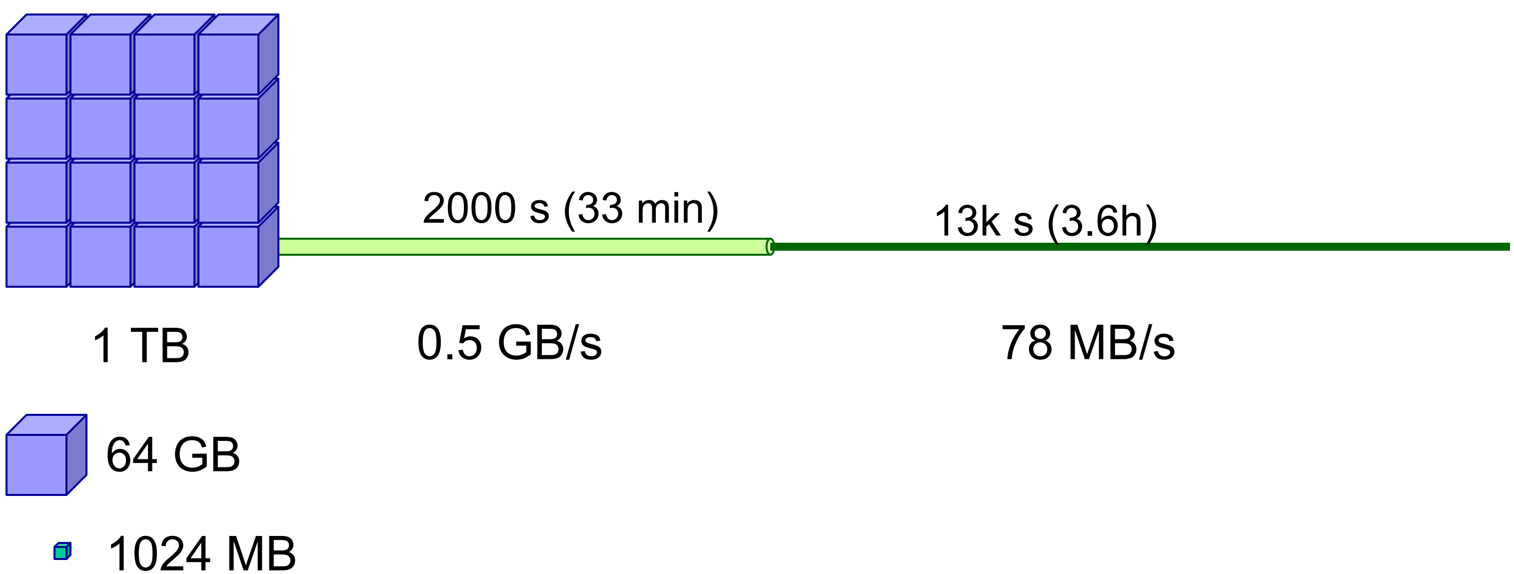
100 MB/s

18 MB/s

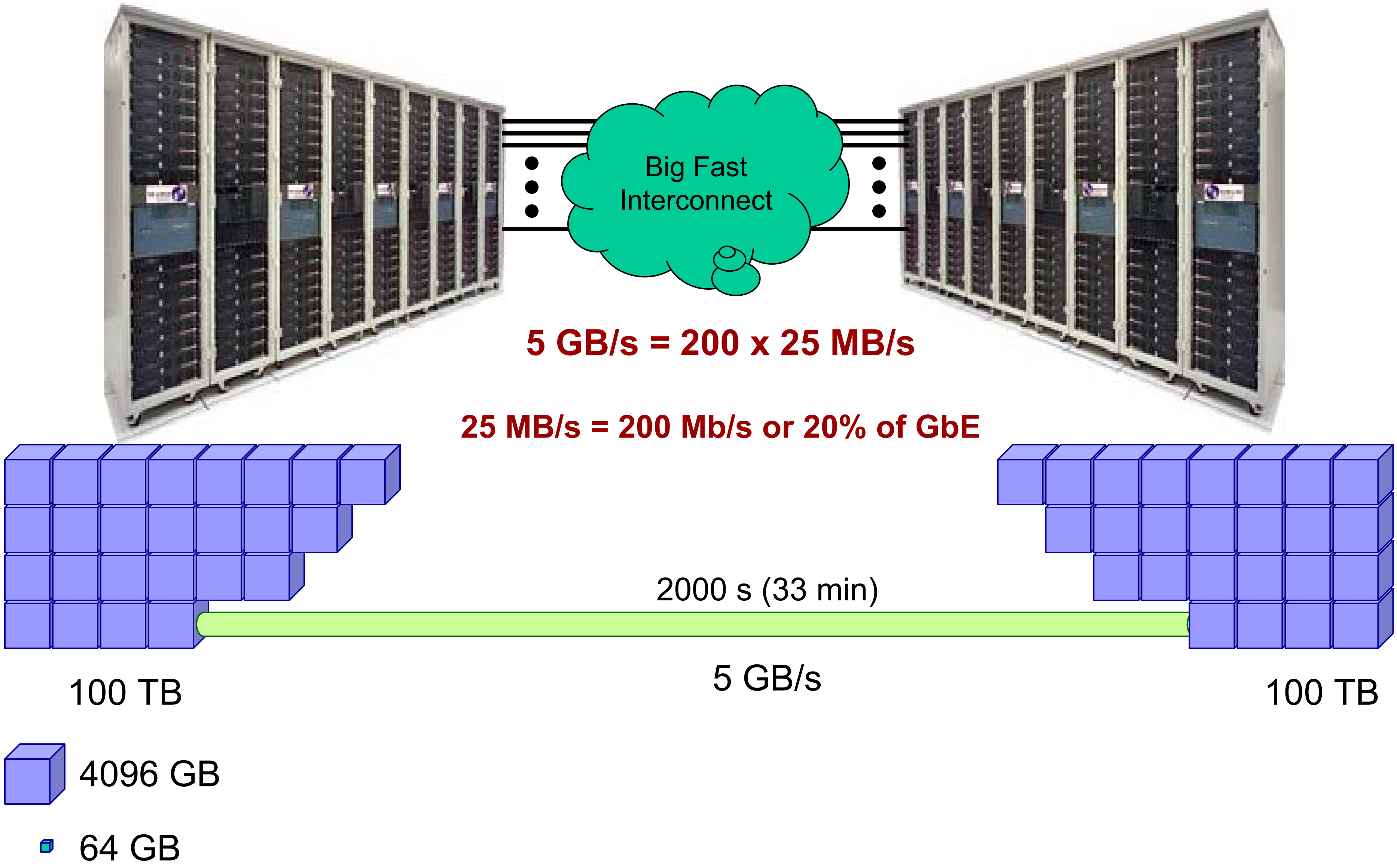

1024MB

Today's Architecture





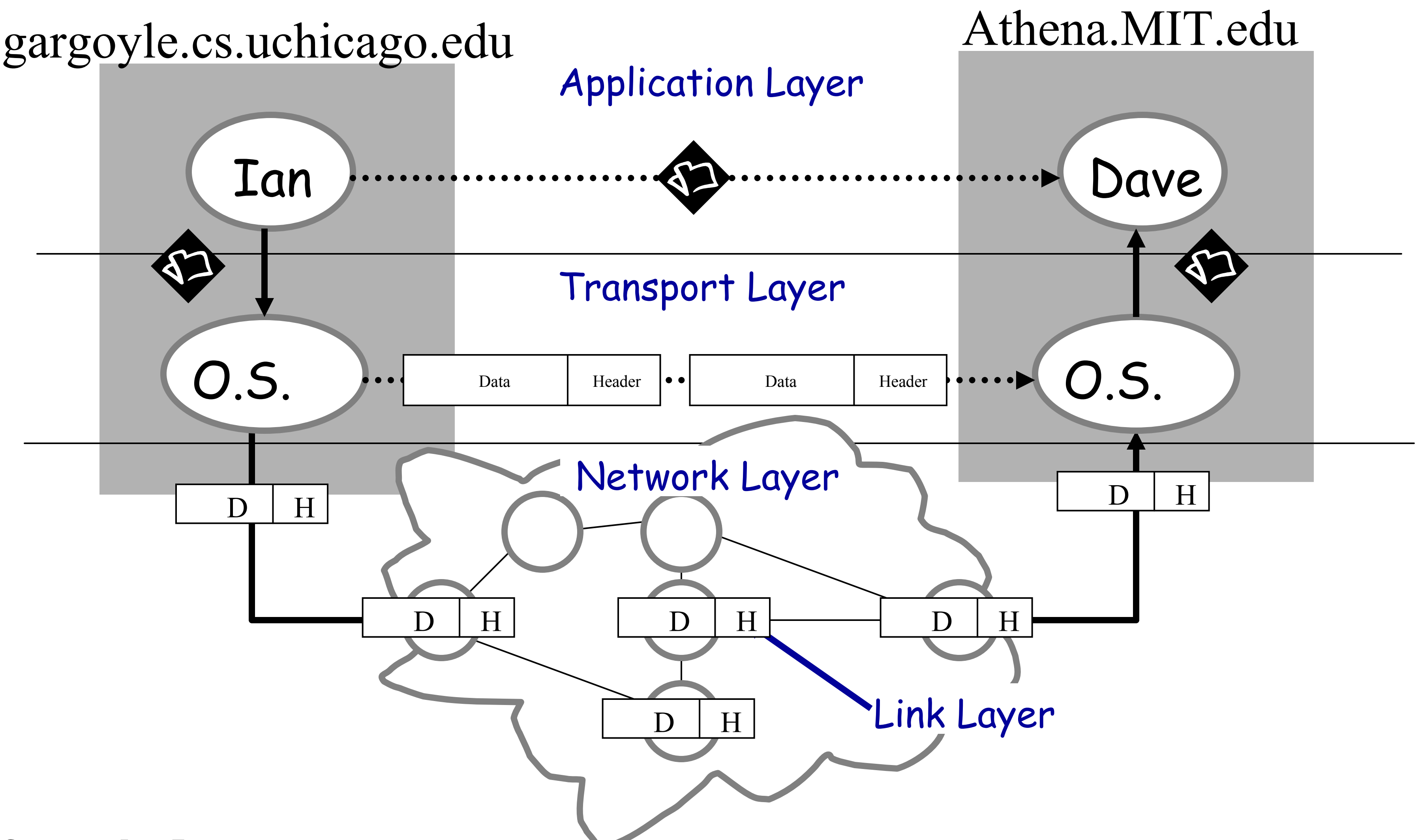
To Build a Distributed Terascale Cluster…



Optical Networks

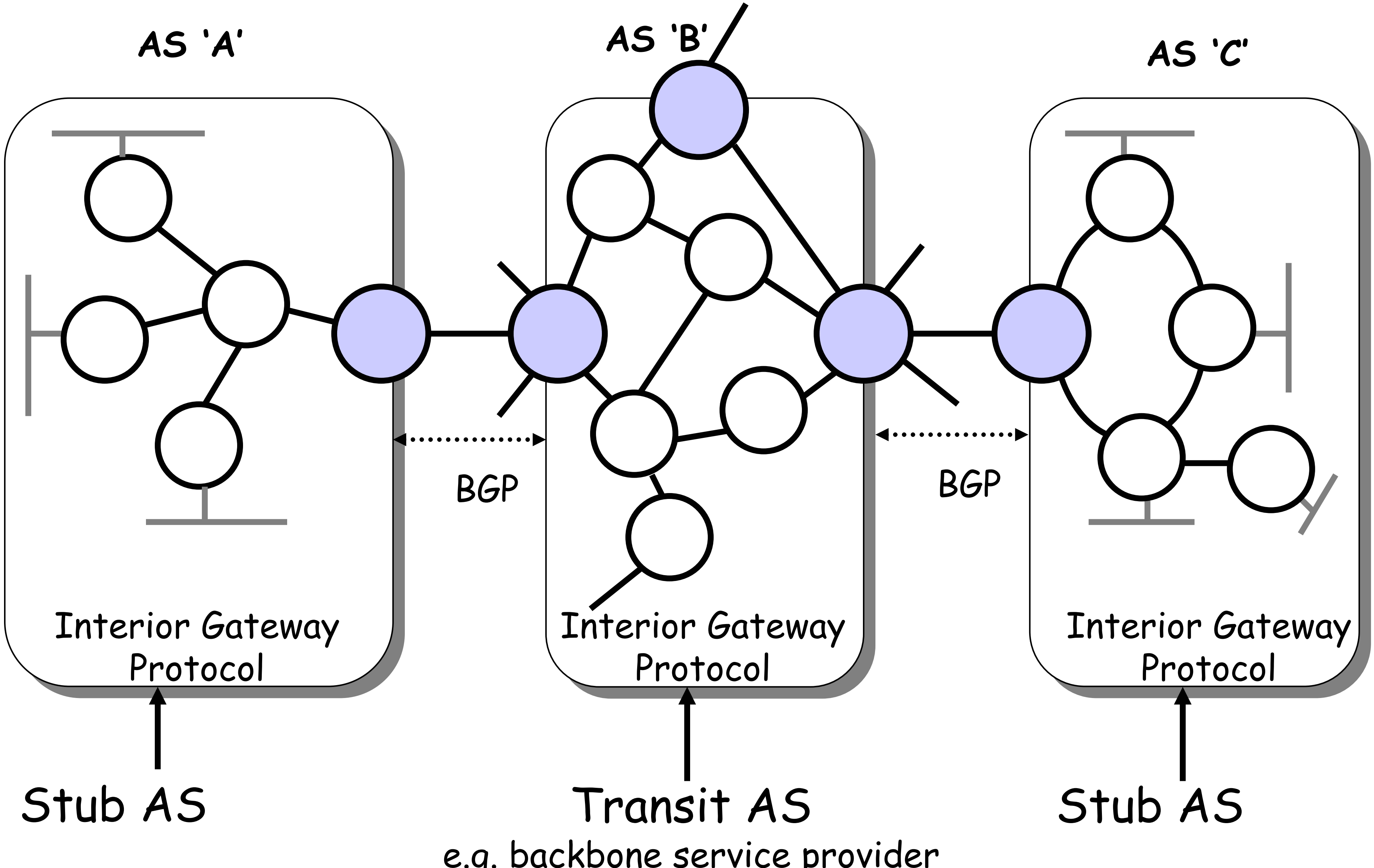
- “High Performance Nets”
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

Recap: Layered Architecture



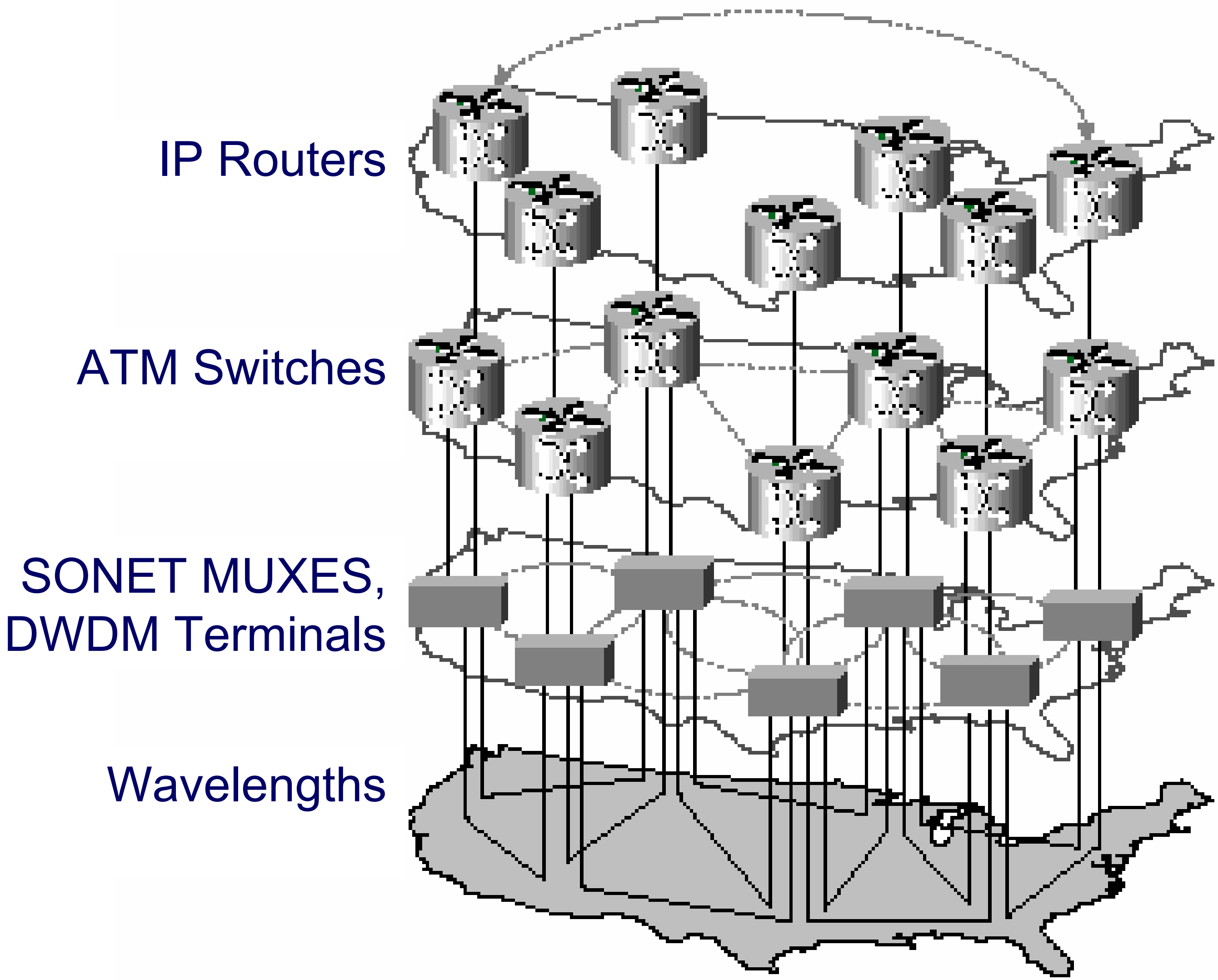
Source: Ian Foster

Recap: The Internet Routes Packets



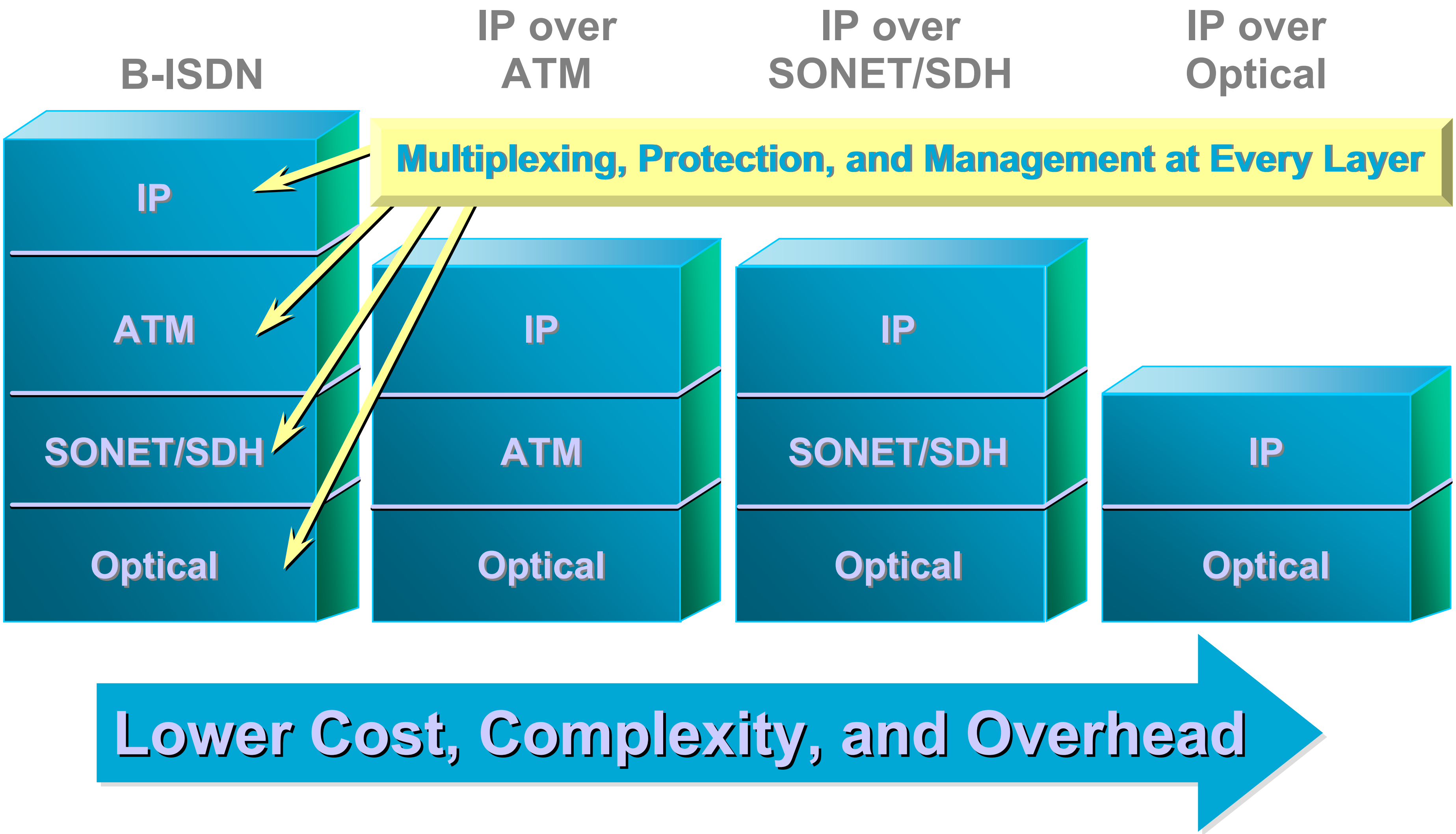
Source: Ian Foster

Layered Network View



Source: Cisco

IP Transport Alternatives

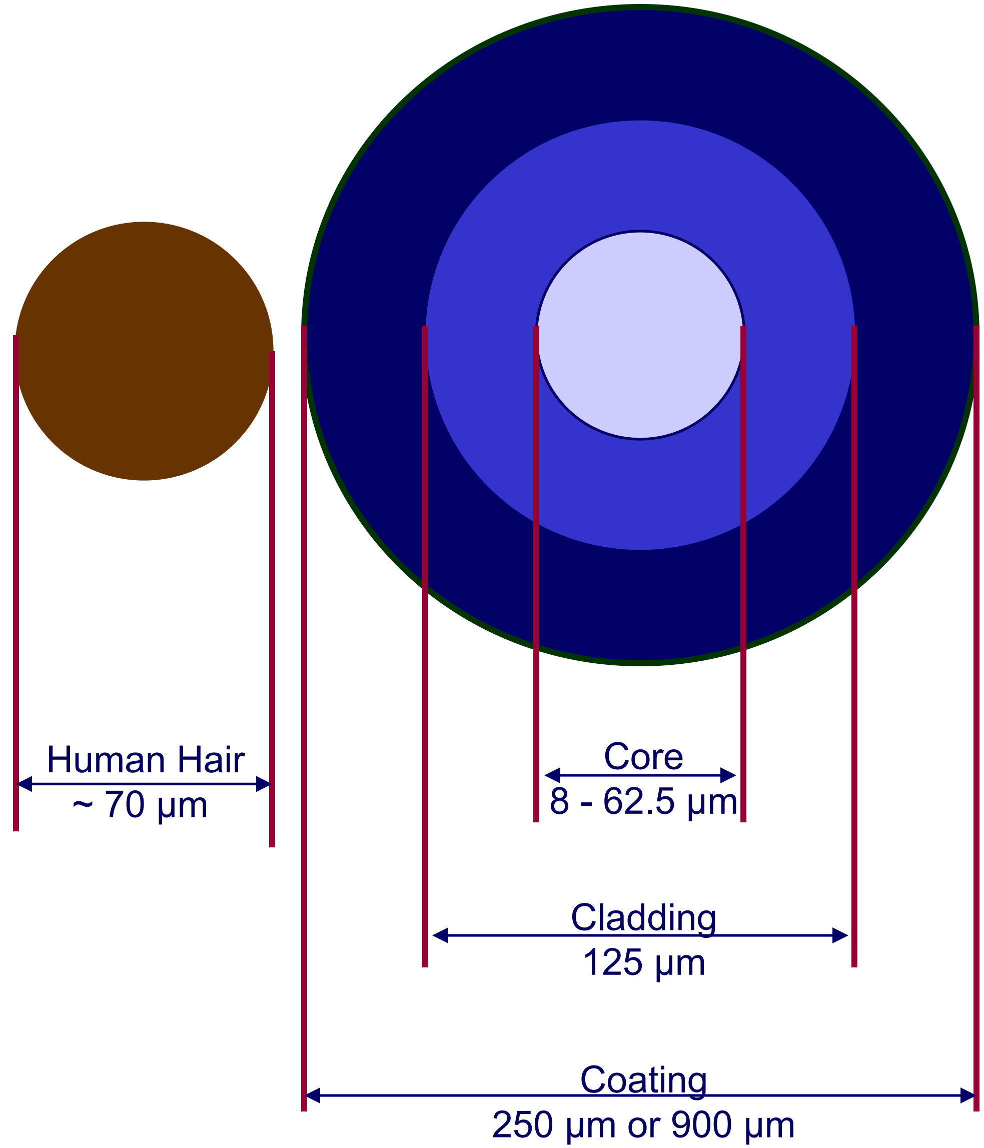


Source: Cisco

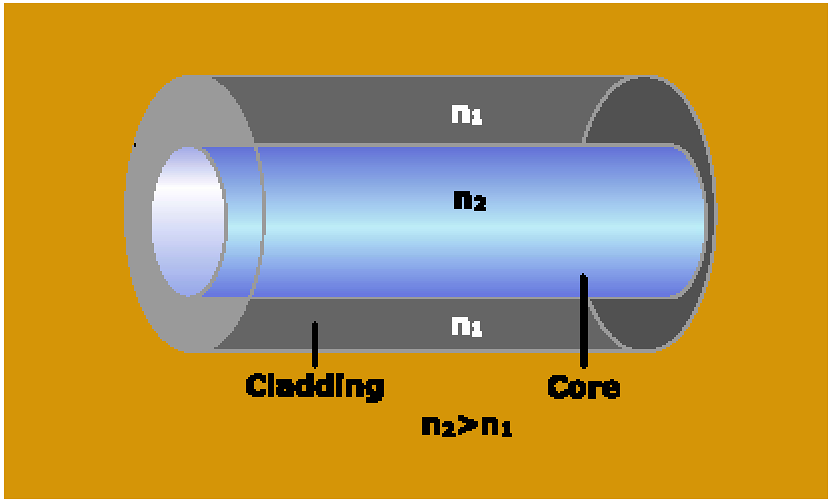
Optical Networking Building Blocks

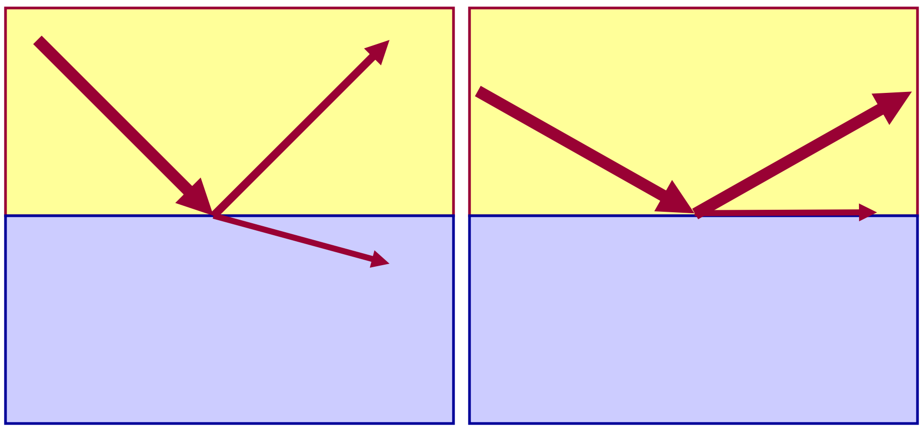
- **Optical Fiber**
  - Types of fiber
  - Fiber characteristics
- **Attenuation and Dispersion**
- **Amplification**
- **Filters, Multiplexors, Demultiplexors**
- **Transmission Systems**

- **Vocabulary…**
  - Wavelength = channel = lambda (λ)

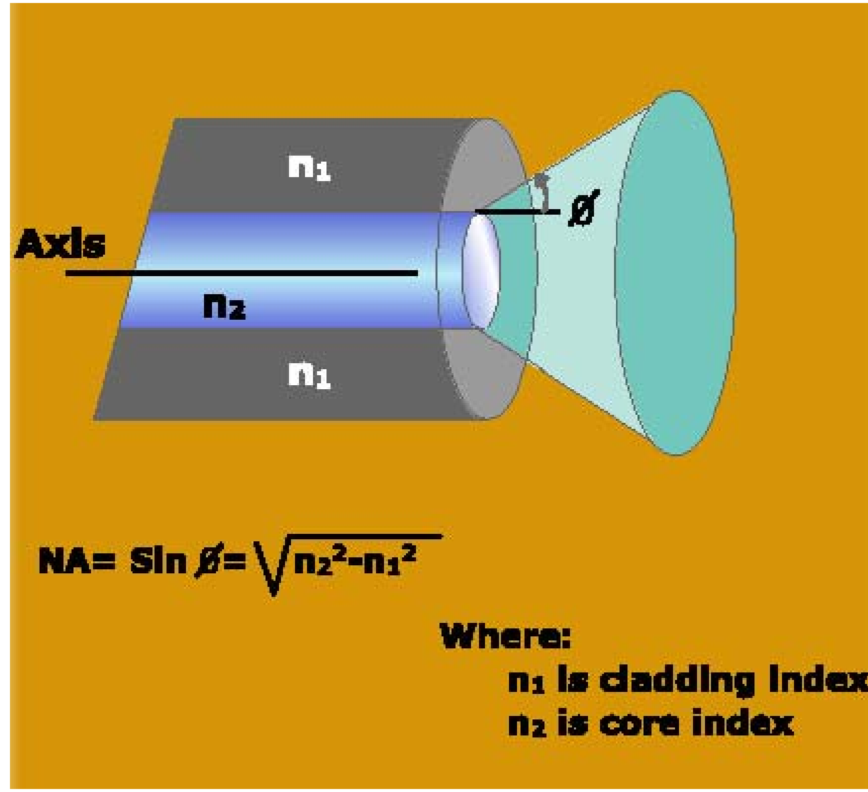


Total Internal Reflection & Numerical Aperture





- Total internal reflection allows light to propagate down fiber

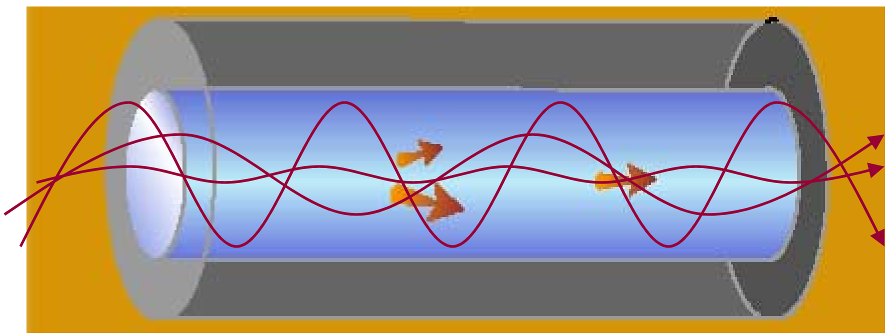


- *NA* defines envelope within which light will enter the core and propagate
- Light must be "launched" into the core within the *NA*
- Light detectors must take into account *NA* as well

Source: Corning

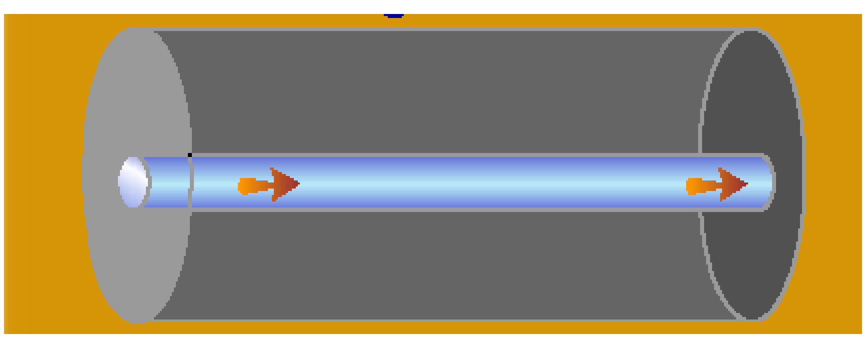
Fiber Types

Multimode Fiber



- Different components of the light signal entering at different launch angles → "modes"
- Each mode travels a different path through the fiber, thus a different path length
- <u>Disadvantage</u>: modes arrive at different times – "modal dispersion" (more on dispersion later)
- <u>Advantage</u>: easier coupling to launch signal into fiber → cheaper components

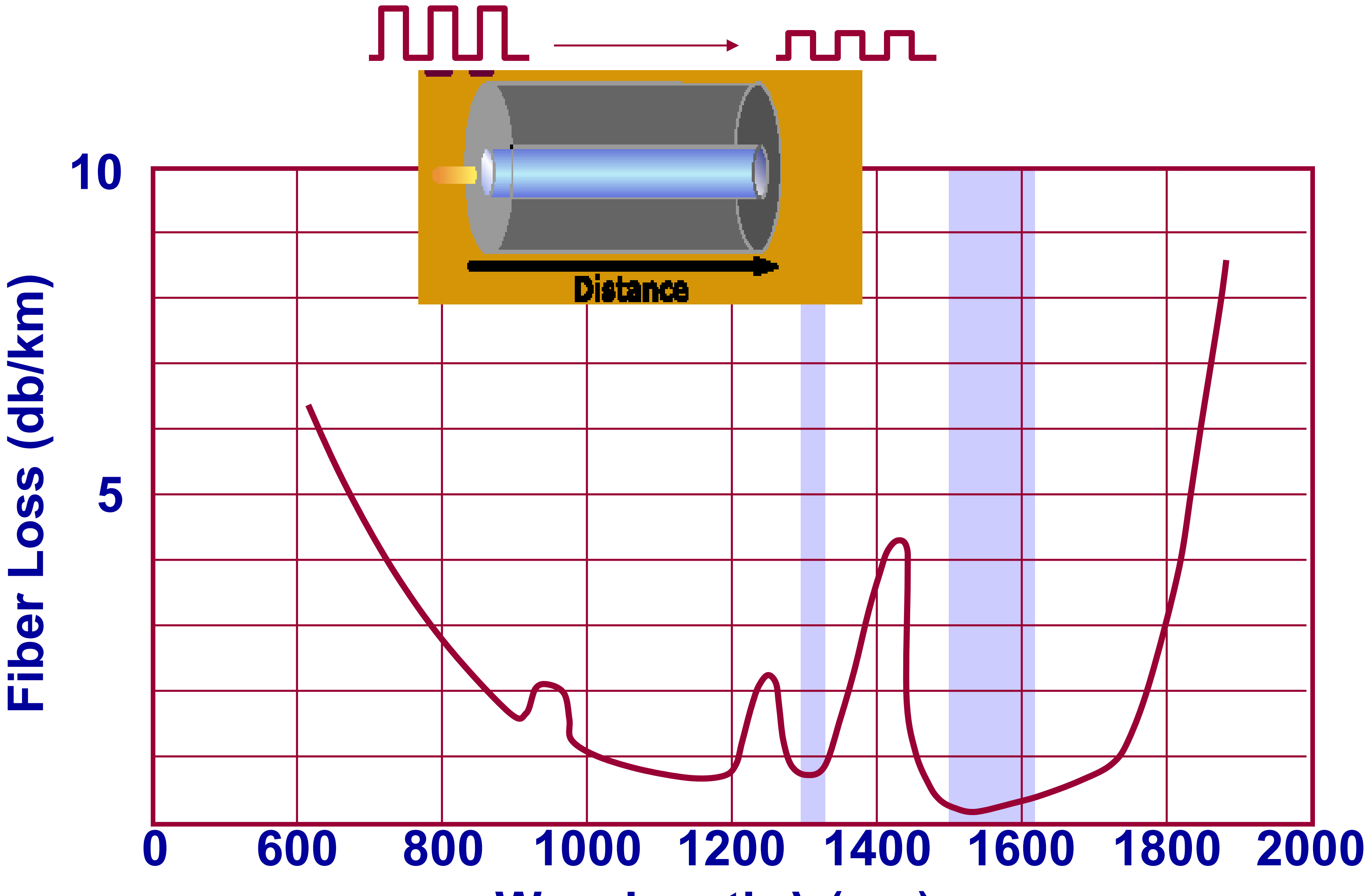
Single mode Fiber



- Only allows one mode
- <u>Disadvantage</u>: tighter tolerances for coupling
- <u>Advantage</u>: no modal dispersion!

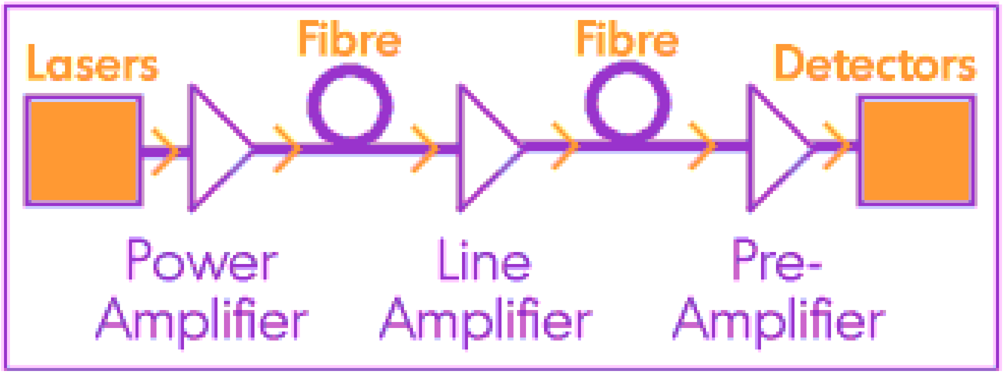
Source: Corning

Attenuation: Loss in Optical Fiber

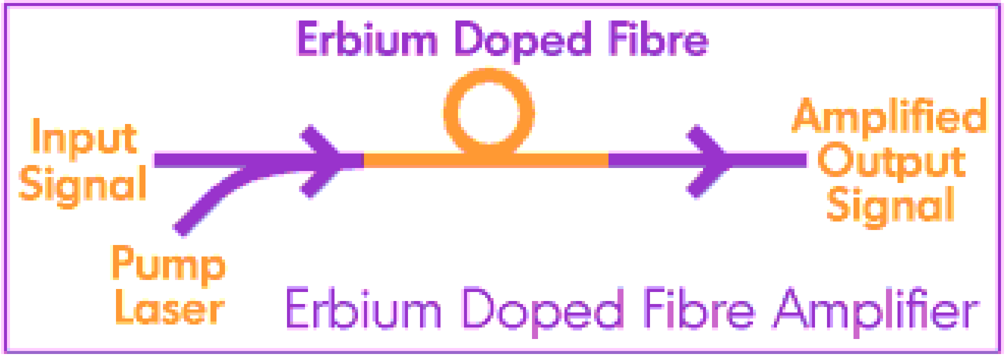


Source: Corning

Amplification



- Laser signal at X db
- Receiver threshold at Y db (Y << X)
- Maximum distance before amplification determined by fiber characteristics (db loss per km at signal wavelength)
- More on this later!

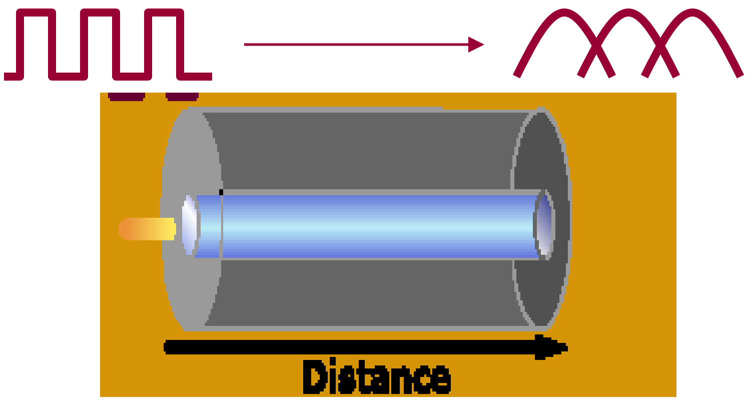


- Amplification requires no conversion to electronics
- Amplifiers tuned to a range of wavelengths – on amplifier per strand

Source: www.lightreading.com

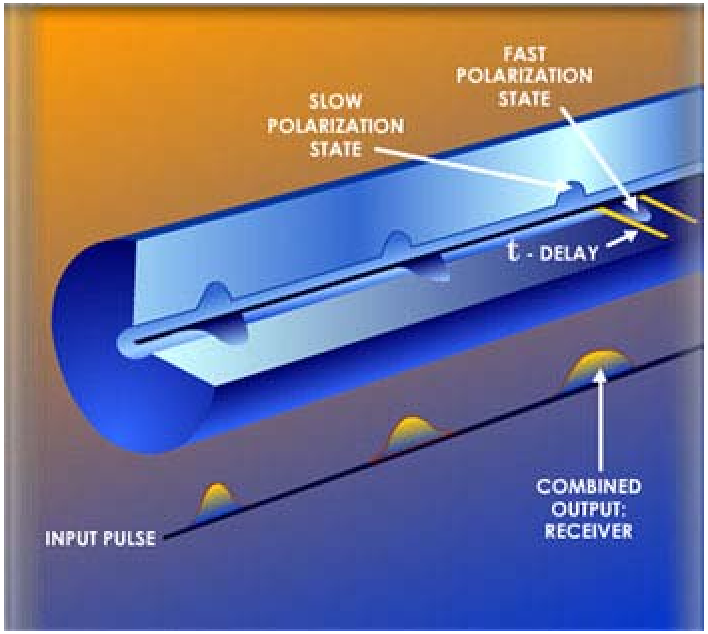
Dispersion

Chromatic Dispersion



Units: ps / nm * km

Polarization Mode Dispersion



Bitrate	Allowable Dispersion (DL)
2.5 Gb/s	12-16,000 ps/nm
10 Gb/s	800-1,000 ps/nm
40 Gb/s	60-100 ps/nm

Source: Corning

Dispersion at 40 Gb/s

40 Gb/s signal at transmitter (top) and after 80 km of NZ-DSF fiber without dispersion compensation.

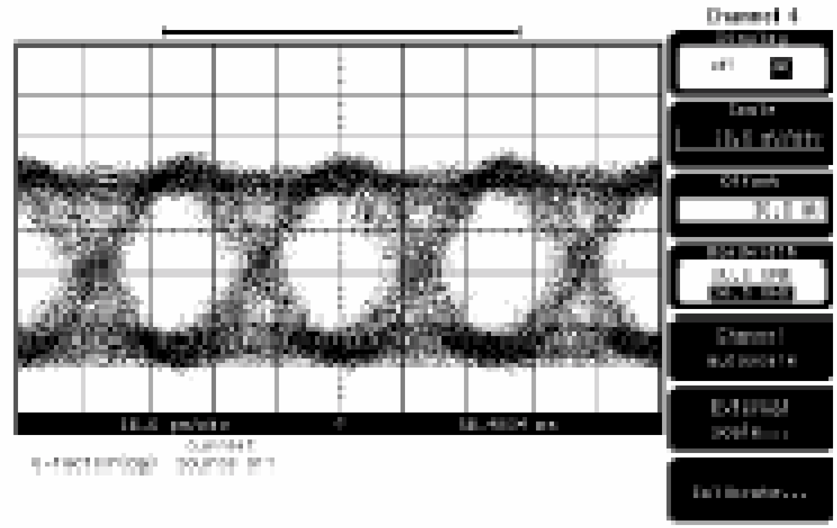


Figure 2. 40 Gb/s signal at Transmitter

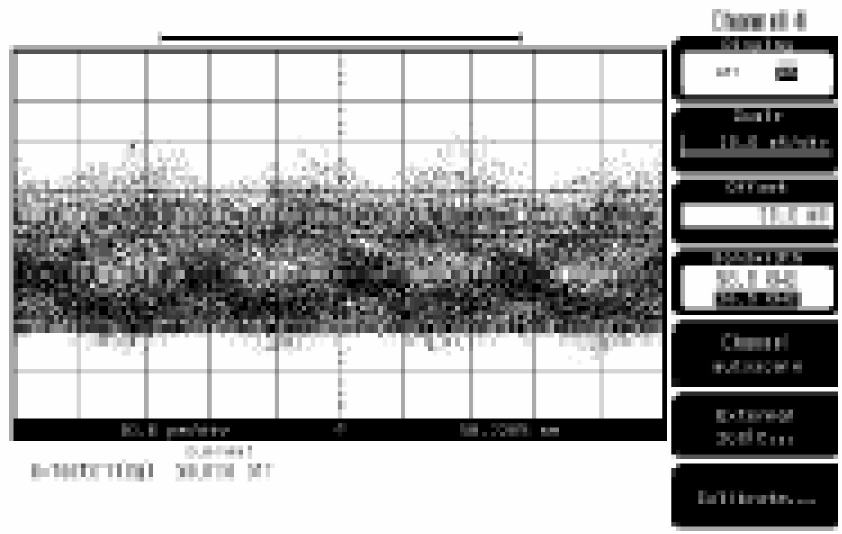
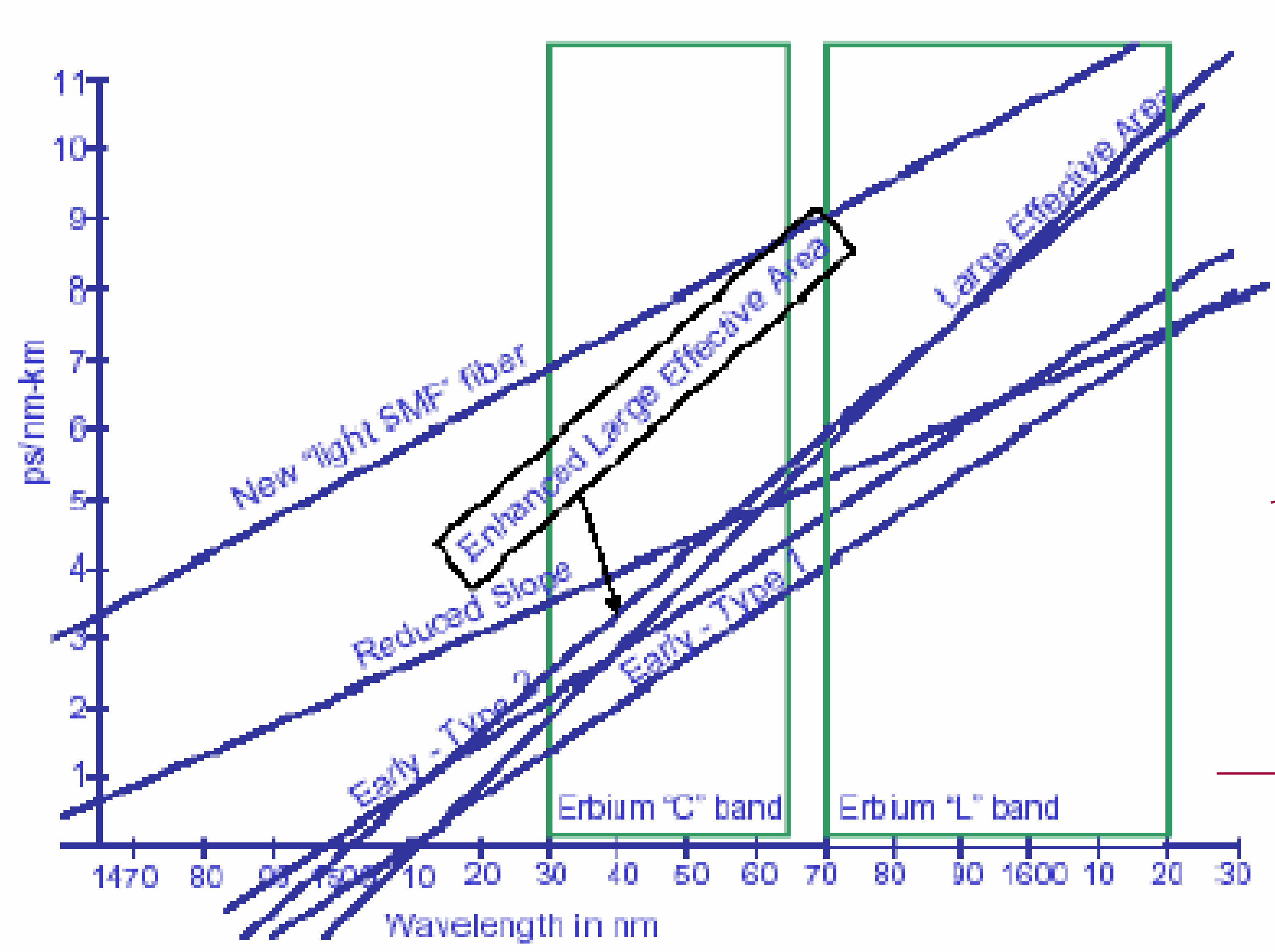
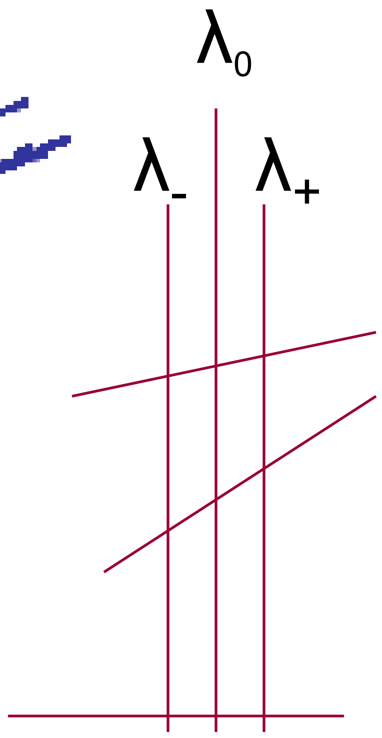


Figure 3. 40 Gb/s signal after traversing 80 km of NZDSF fiber without compensation

Source: Lasercom, Inc

Dispersion-Shifted Fiber

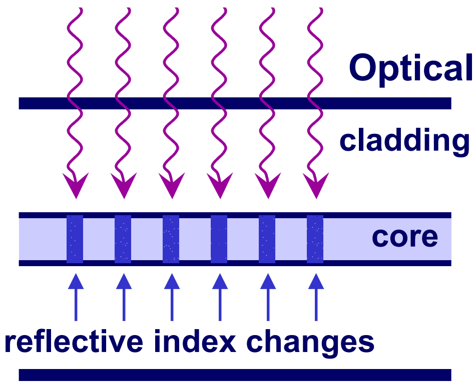




Source: Lasercom, Inc

Filtering, Muxing, Demuxing

Fiber Bragg Grating

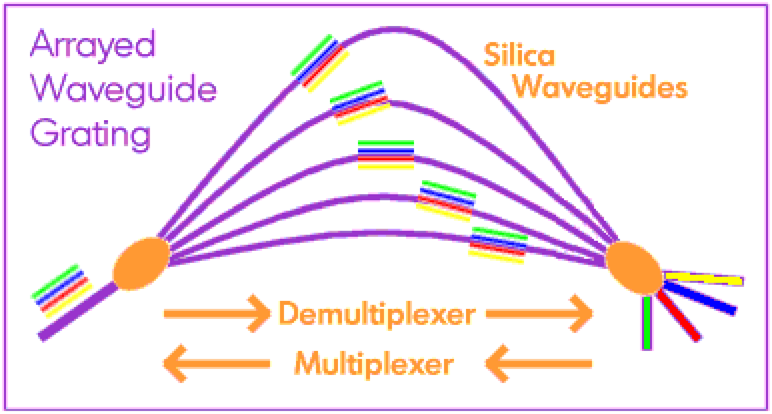


Light at ***Bragg wavelength*** is reflected.

Bragg wavelength is f **(spacing, Δindex)**

Array Waveguide Grating

(Optical Phased Array)

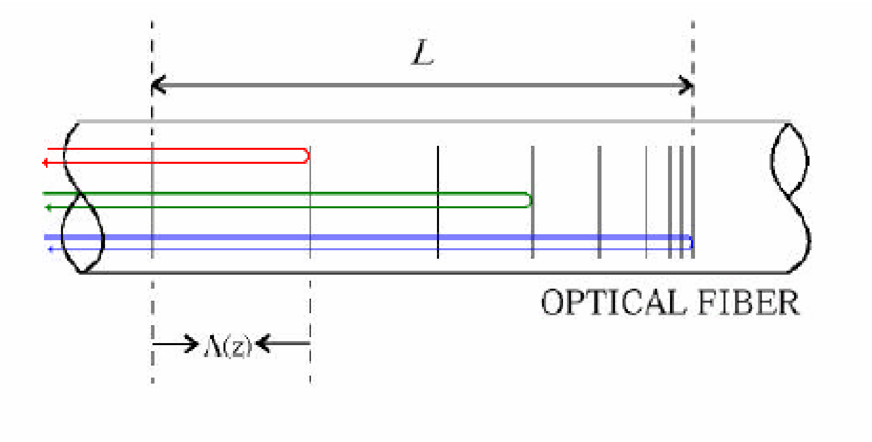


- **Multiple waveguides at fixed length intervals**
- **All wavelengths pass through each waveguide, arrive slightly out of phase**
- **Interference "focuses" individual wavelengths at different points on receiver surface, effectively splitting out individual wavelengths**
- **Same principle as phased array antenna**

Source: Corning

Dispersion Compensation

With Bragg Grating



- Use a "loopback" section of fiber with series of Bragg filters, sending "fast" wavelengths on a longer loop than "slow" wavelengths.

Insert section of Negative NZ-DSF Fiber

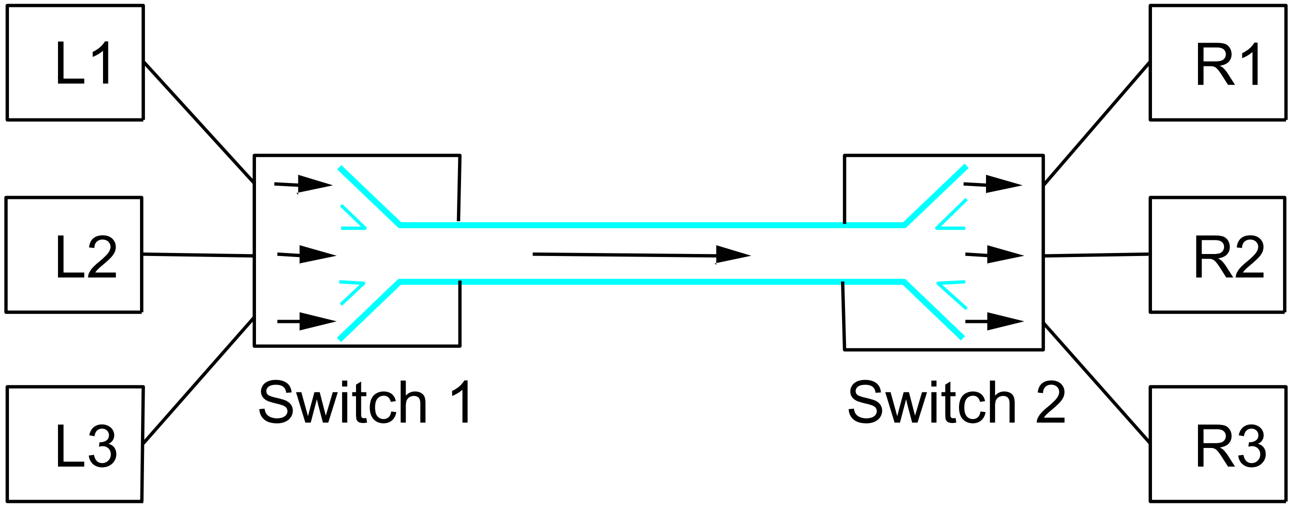
Source: Corning

Optical Networks

- "High Performance Nets"
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

Recap: Multiplexing

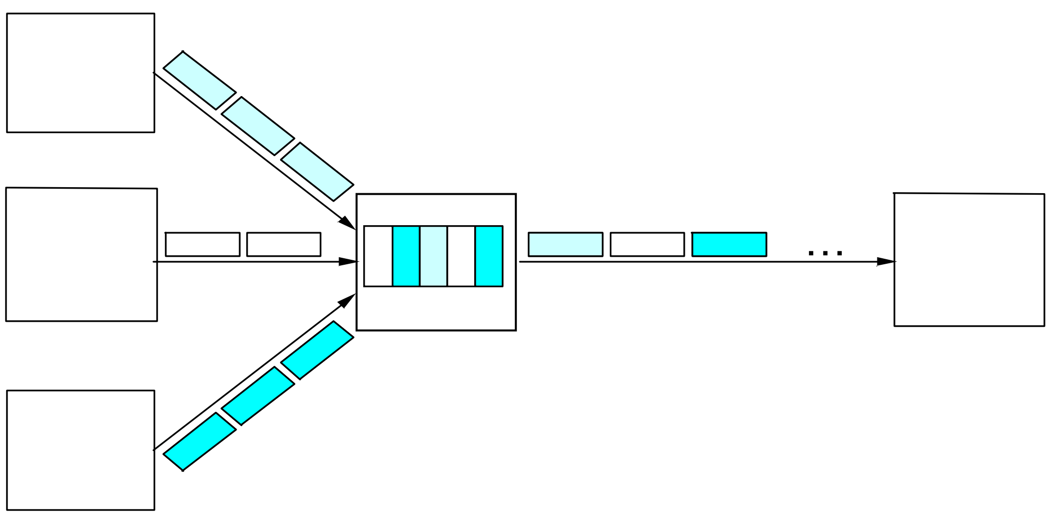
- **Time-Division Multiplexing (TDM)**
- **Frequency-Division Multiplexing (FDM)**



Source: Ian Foster

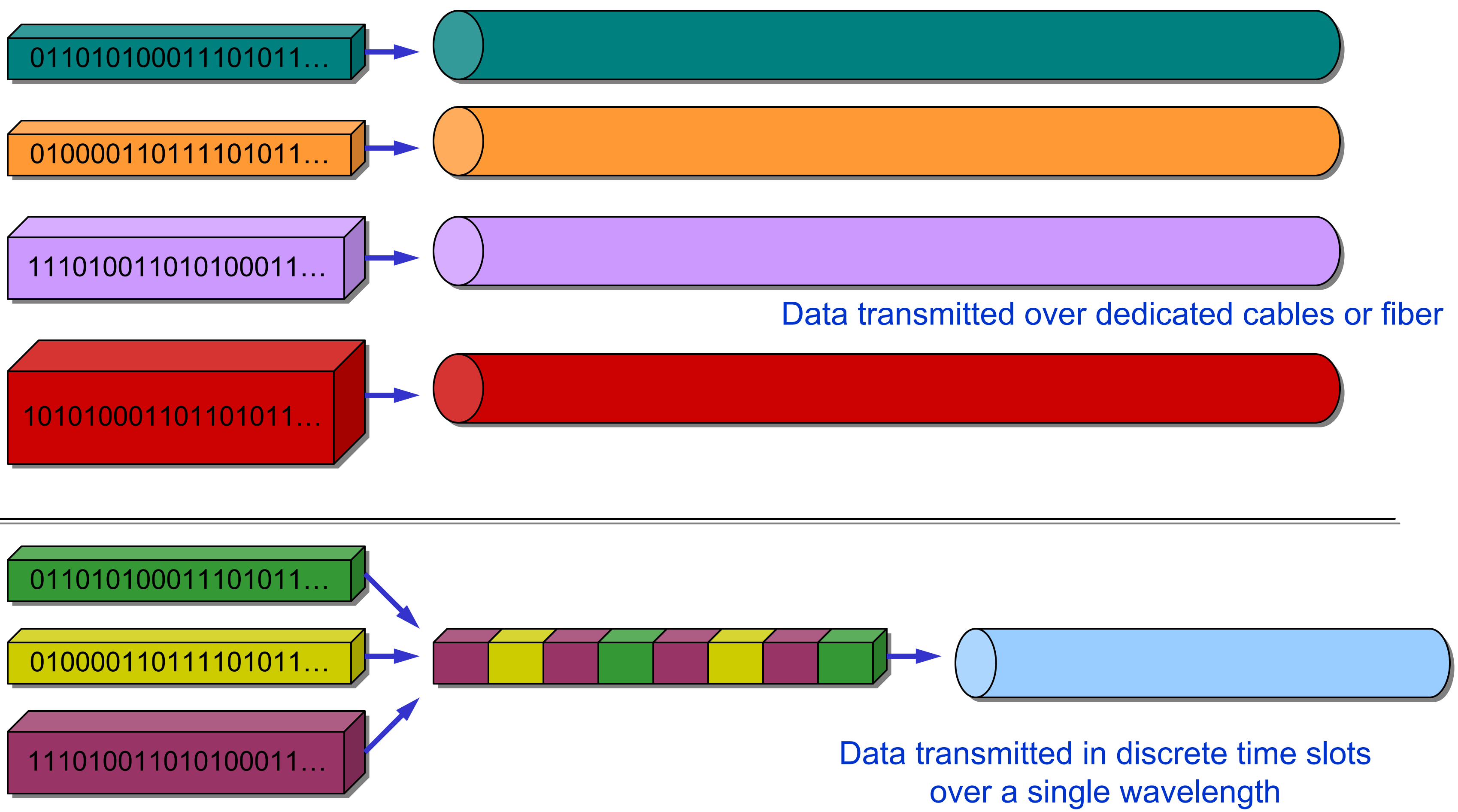
Recap: Statistical Multiplexing

- On-demand time-division
- Schedule link on a per-packet basis
- Packets from different sources interleaved on link
- Buffer packets that are *contending* for the link
- Buffer (queue) overflow is called *congestion*



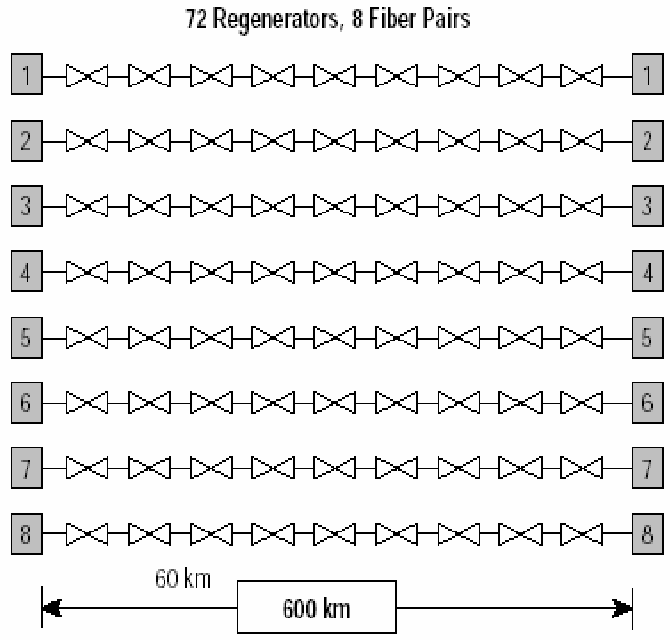
Source: Ian Foster

Traditional Transmission Systems



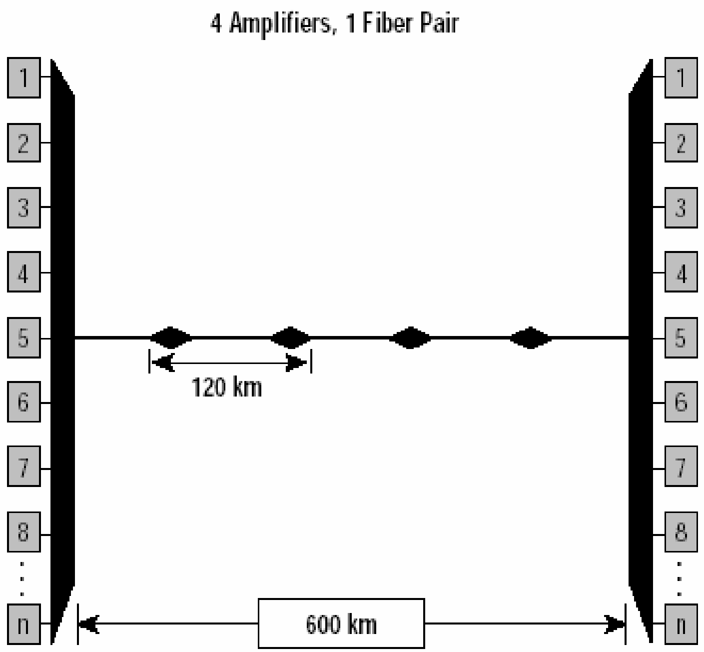
Two Changes: WDM and Optical Amplifiers

Then



- One channel per pair of fiber
- Electrical regeneration every ~60km

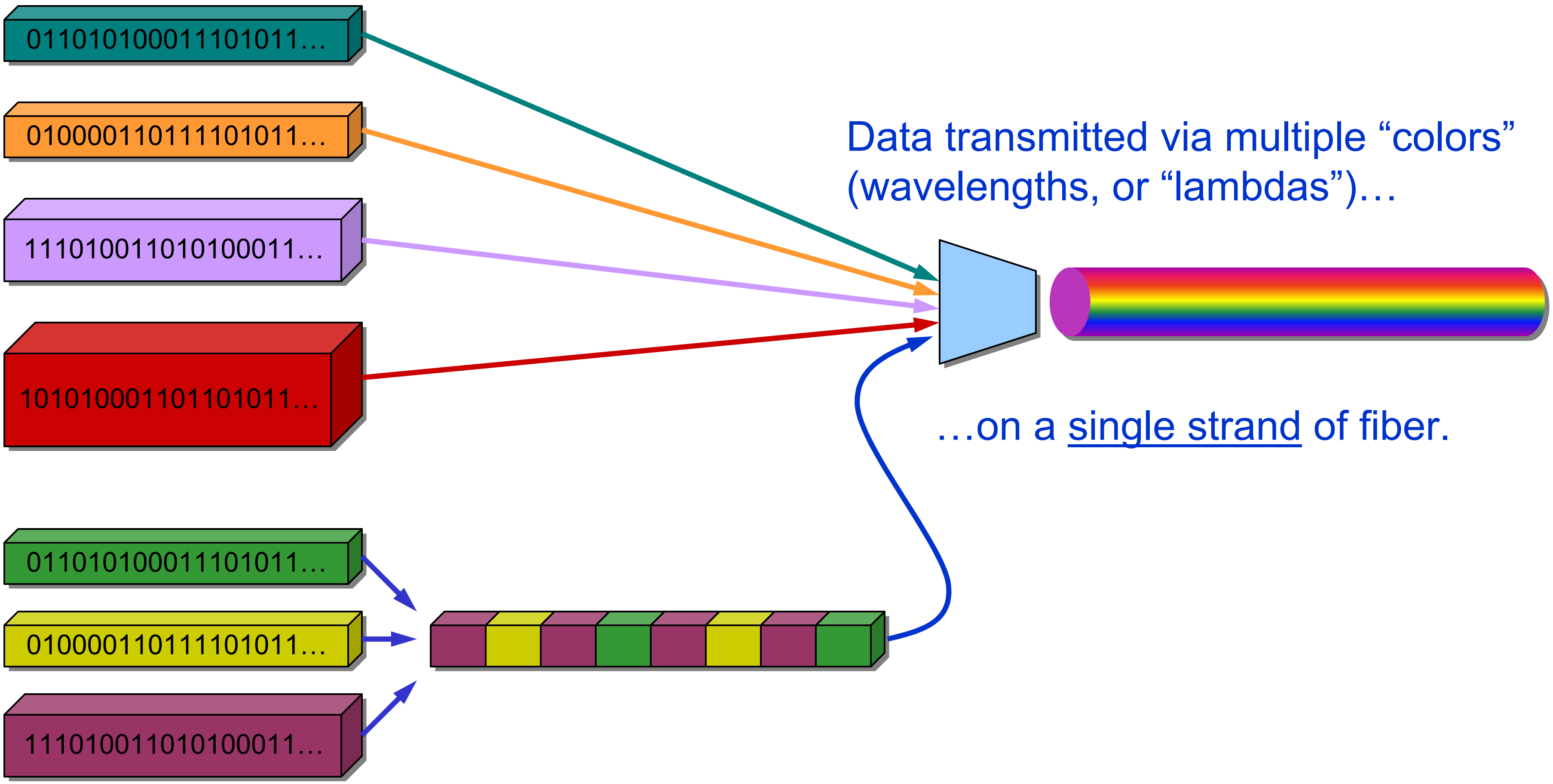
Now



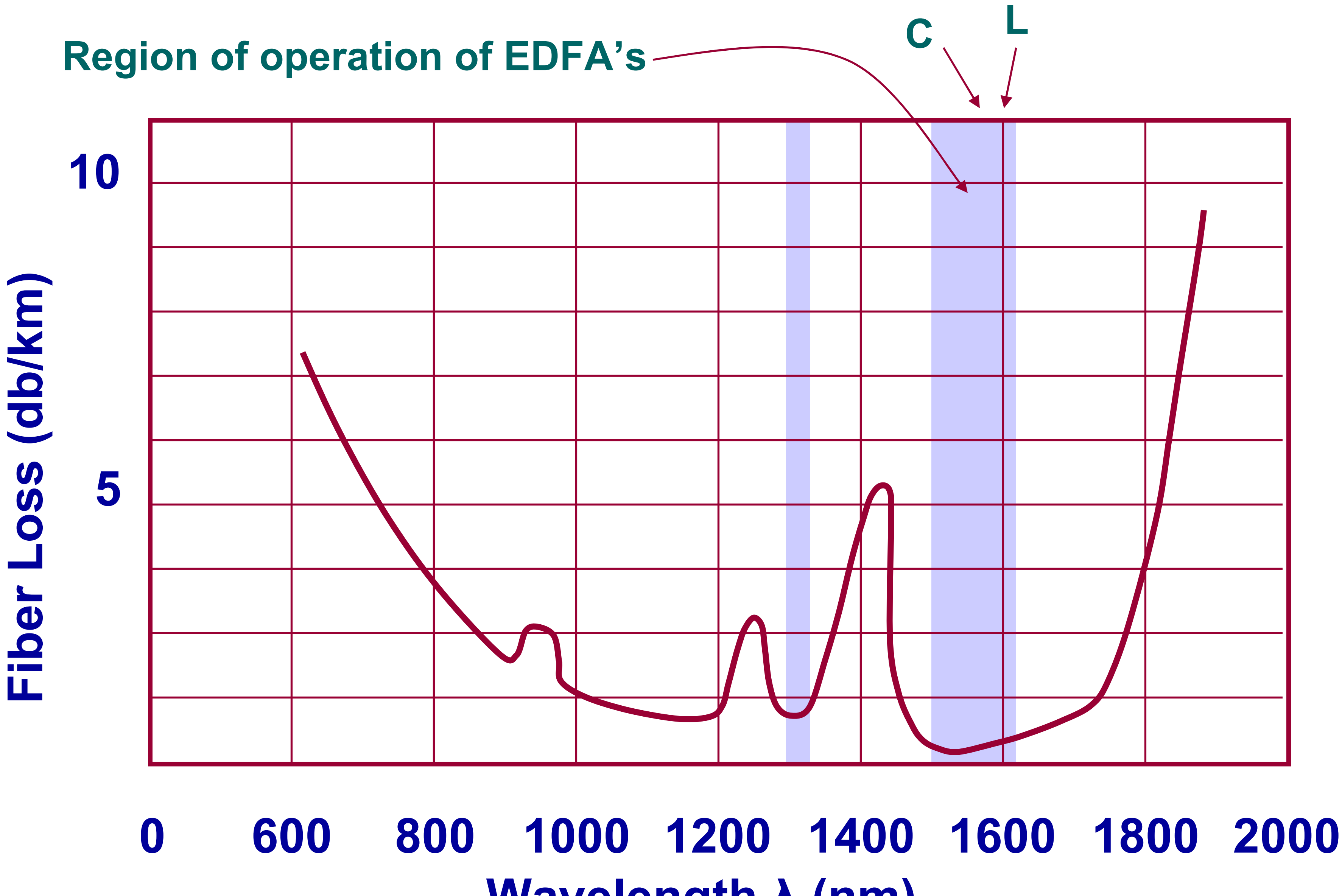
- Wave Division Multiplexing allows all 8 channels to share one fiber pair
- Optical amplifiers allow for greater distances, less equipment in line

Source: Cisco

Wave Division Multiplexing (WDM)



Attenuation: Loss in Optical Fiber



Source: Corning

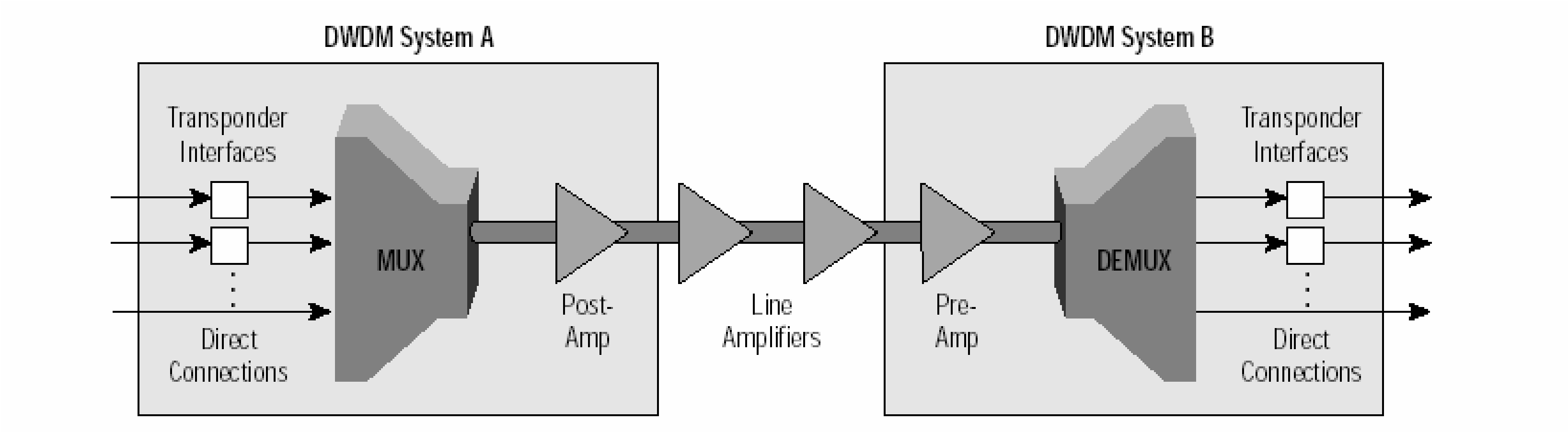
Dense Wave Division Multiplexing: ITU Grid

	L-Band (1565-1620nm)				C-Band (1530-1565nm)				S-Band (1490 – 1530nm)			
	100 GHz Grid		50 GHz Offset		100 GHz Grid		50 GHz Offset		100 GHz Grid		50 GHz Offset	
	THz	nm	THz	nm	THz	nm	THz	nm	THz	nm	THz	nm
1	186.00	1611.79	186.05	1611.35	191.00	1569.59	191.05	1569.18	196.00	1529.55	196.05	1529.16
2	186.10	1610.92	186.15	1610.49	191.10	1568.77	191.15	1568.36	196.10	1528.77	196.15	1528.38
3	186.20	1610.06	186.25	1609.62	191.20	1567.95	191.25	1567.54	196.20	1527.99	196.25	1527.60
4	186.30	1609.19	186.35	1608.76	191.30	1567.13	191.35	1566.72	196.30	1527.22	196.35	1526.83
5	186.40	1608.33	186.45	1607.90	191.40	1566.31	191.45	1565.90	196.40	1526.44	196.45	1526.05
44	190.30	1575.37	190.35	1574.95	195.30	1535.04	195.35	1534.64	200.30	1496.72	200.35	1496.34
45	190.40	1574.54	190.45	1574.13	195.40	1534.25	195.45	1533.86	200.40	1495.97	200.45	1495.60
46	190.50	1573.71	190.55	1573.30	195.50	1533.47	195.55	1533.07	200.50	1495.22	200.55	1494.85
47	190.60	1572.89	190.65	1572.48	195.60	1532.68	195.65	1532.29	200.60	1494.48	200.65	1494.11
48	190.70	1572.06	190.75	1571.65	195.70	1531.90	195.75	1531.51	200.70	1493.73	200.75	1493.36
49	190.80	1571.24	190.85	1570.83	195.80	1531.12	195.85	1530.72	200.80	1492.99	200.85	1492.62
50	190.90	1570.42	190.95	1570.01	195.90	1530.33	195.95	1529.94	200.90	1492.25	200.95	1491.88

0.8 nm spacing **0.4 nm spacing**

Charlie Catlett (catlett@mcs.anl.gov)

Long Haul Transmission with WDM



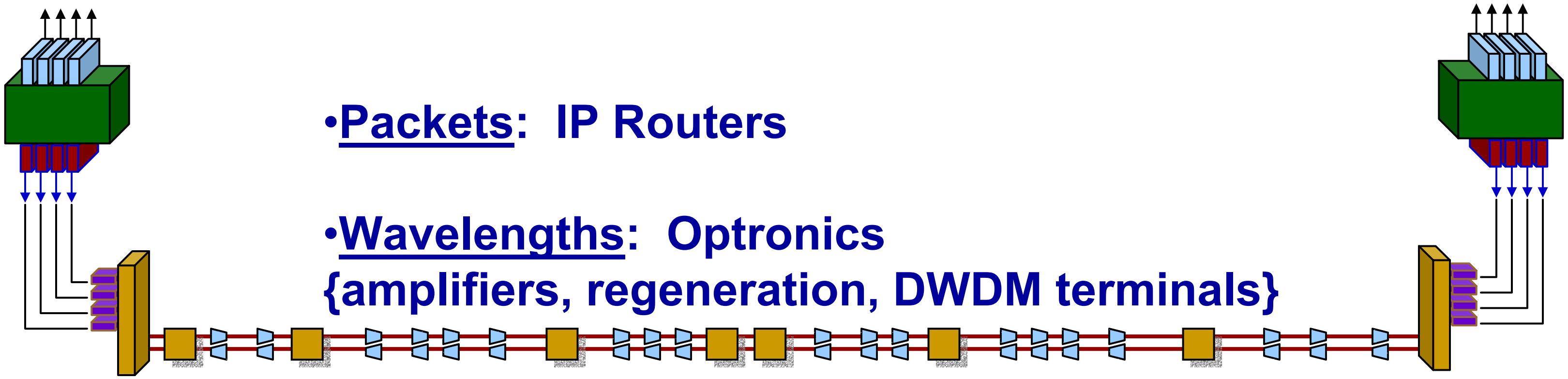
- **Course Wave Division Multiplexing (~8 channels per fiber)**
- **Dense Wave Division Multiplexing (>>16 channels per fiber)**

- **Optical regeneration every ~60-100 km (20-25 dB actually).**

- **Full electrical "3R" (Re-amplification, Re-shaping, Re-timing) every 4-5 spans or ~ 3-400 km. Requires O-E-O.**

Source: Cisco

Optical Network Cost Components



•**Packets**: IP Routers

•**Wavelengths**: Optronics
{amplifiers, regeneration, DWDM terminals}

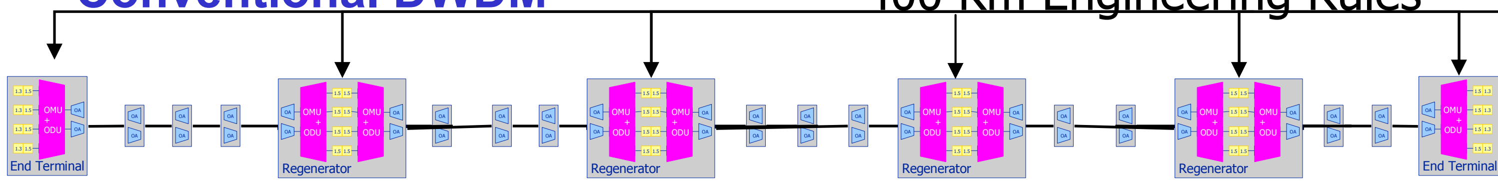
•**Medium**: Fiber and facilities

- Often vendors will do back-to-back terminals in major hubs to increase port capacity. This increases cost, as transponders are required at each point.
- Largest cost component: interfaces and transponders

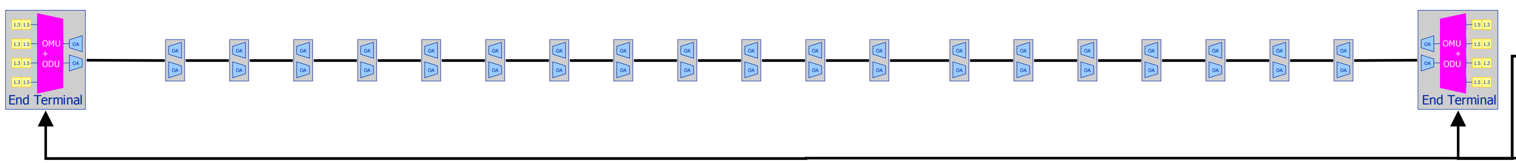
Charlie Catlett (catlett@mcs.anl.gov)

Ultra-Long Haul DWDM

Conventional DWDM 400 Km Engineering Rules

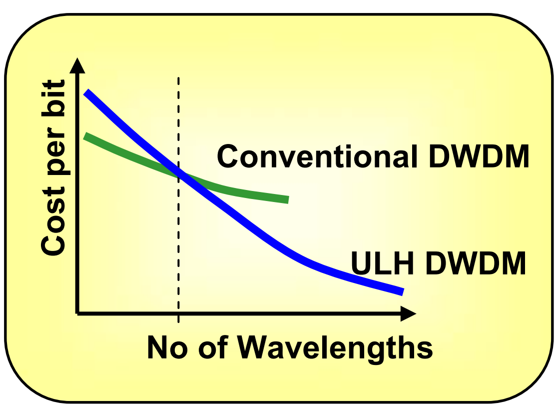


4 Regens & **2** End Terminals



2 End Terminals

Ultra-Long Haul DWDM 2000+ Km Engineering Rules



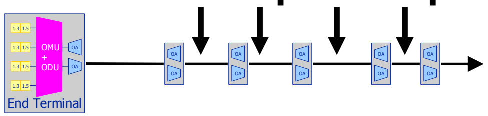
ULH DWDM reduces or removes regenerators saving ~ 30% of the overall equipment costs

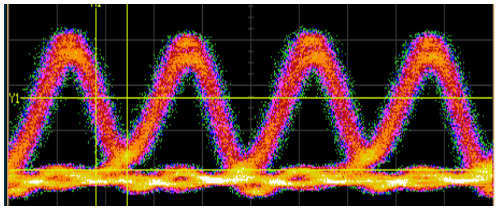
Source: Qwest

ULH DWDM Technical Challenges

Ultra-Long Haul DWDM + Fiber = System performance

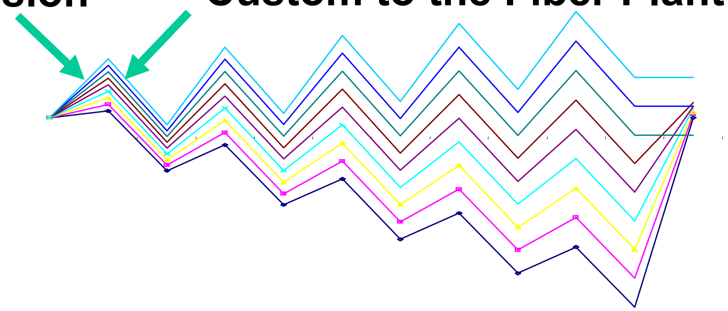
Uneven amplifier spacing





Fiber Dispersion

Dispersion Compensation Custom to the Fiber Plant



Other Noise Sources:
Four Wave Mixing
Polarization Mode Dispersion
Splice reflectances
Amplified Spontaneous Emission

ULH DWDM and Fiber Integration

- Not Plug and Play; Analog link not Digital
- Each fiber optic link is a custom engineered
- In an optically transparent network each λ (W + P) path is custom
- Early deployments at half the distances had unexpected results

Source: Qwest

Optical Networks

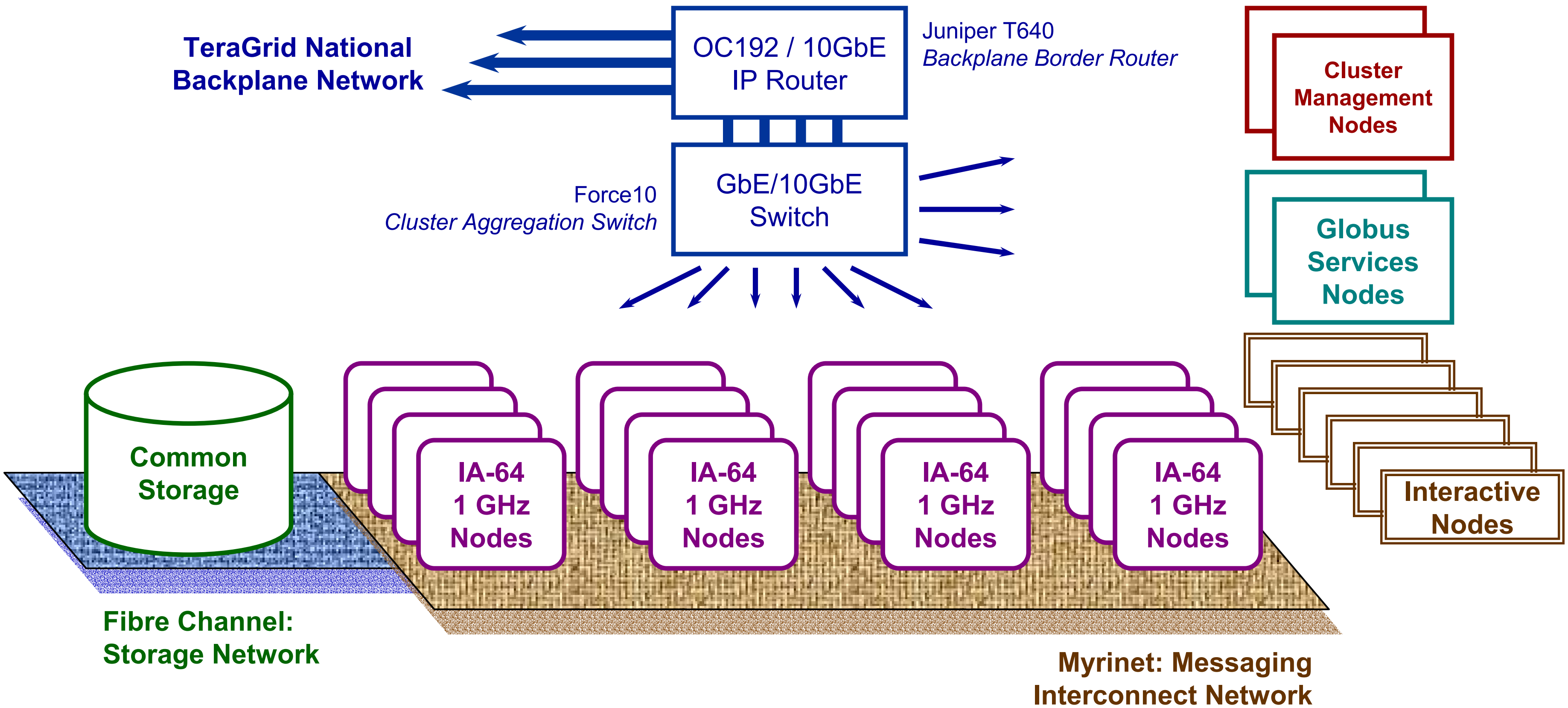
- "High Performance Nets"
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

Case Study: TeraGrid

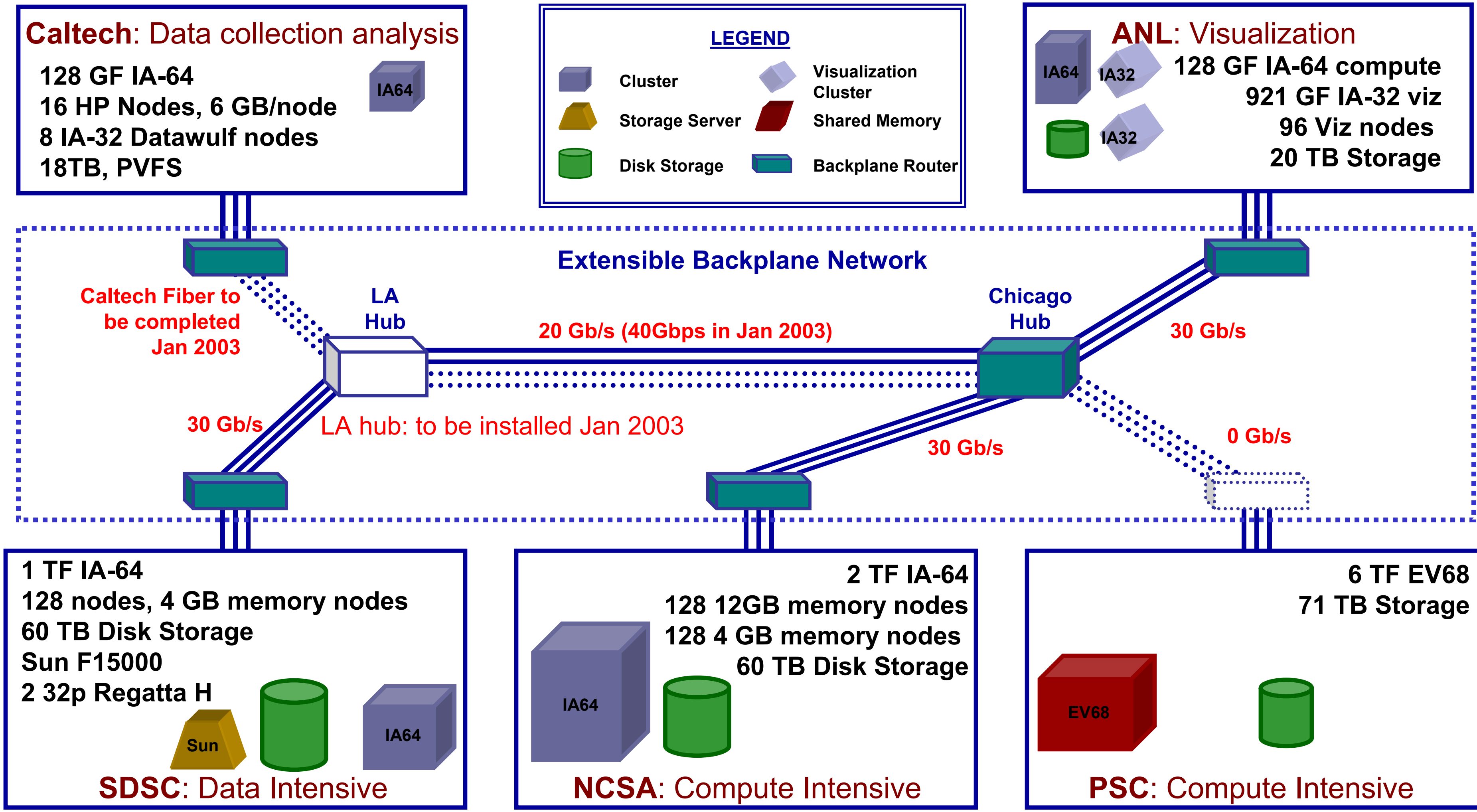
For more information: www.teragrid.org

Source: NCSA

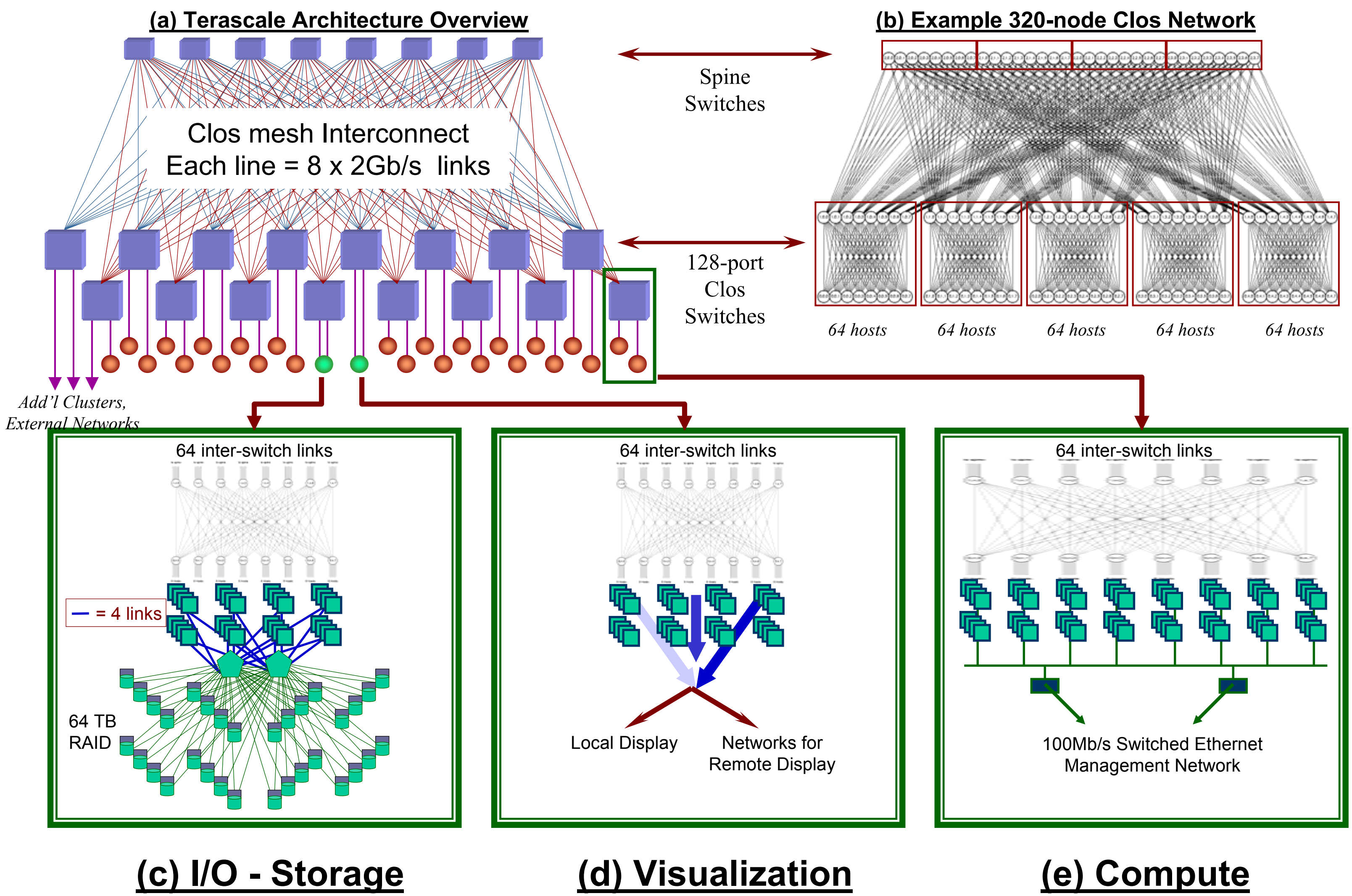
TeraGrid Clusters



TeraGrid Circa January 2003



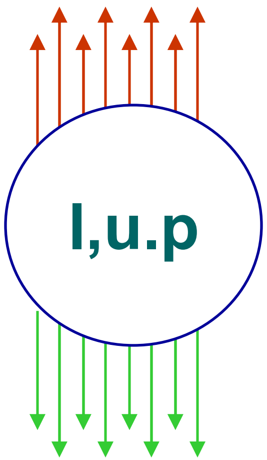
864-node Terascale Architecture



Charlie Catlett (catlett@mcs.anl.gov)

Example: 320-node Clos Network

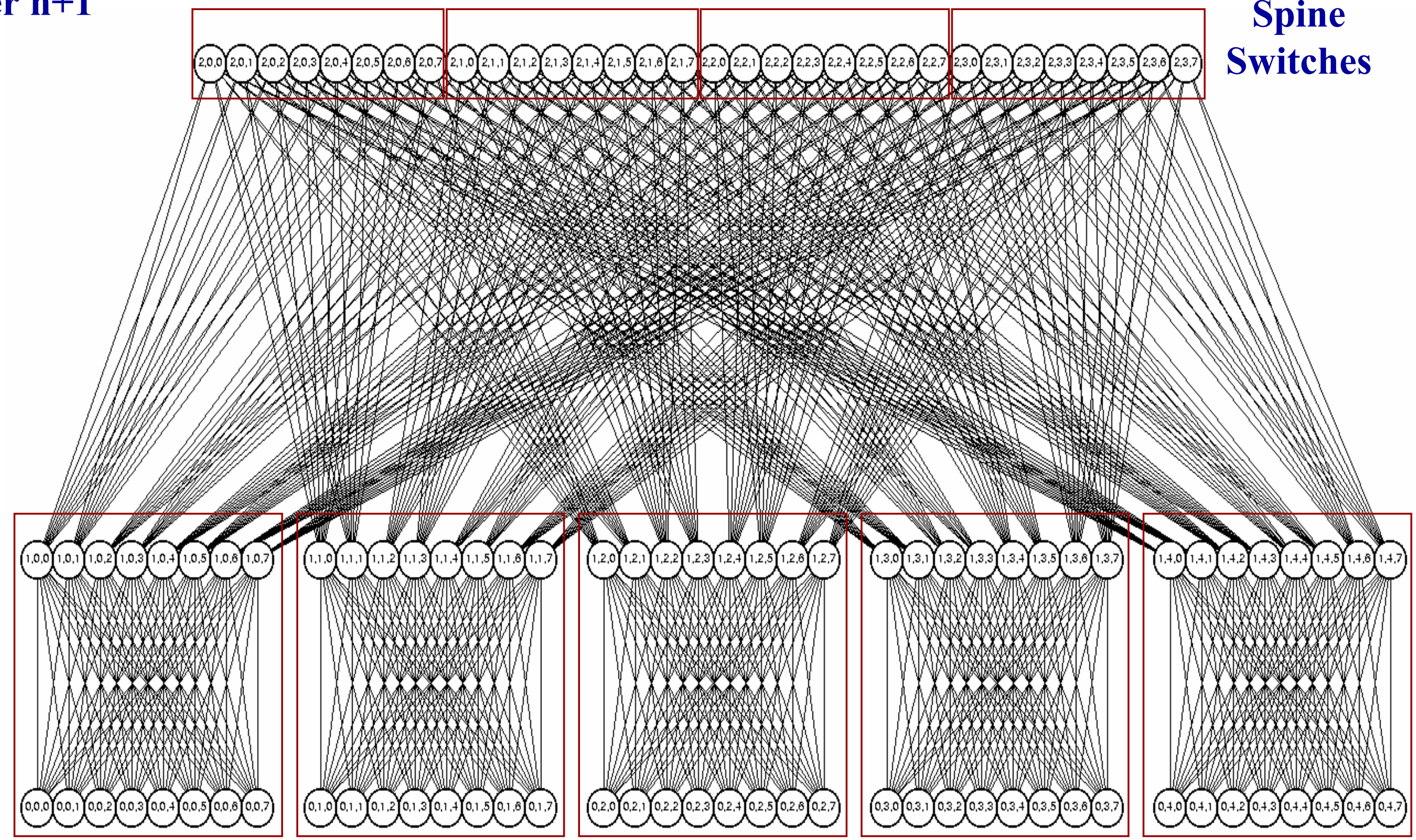
All 8 units at layer n+1



8 hosts, or…
8 units at
layer n-1

l = layer
u = unit
p = port

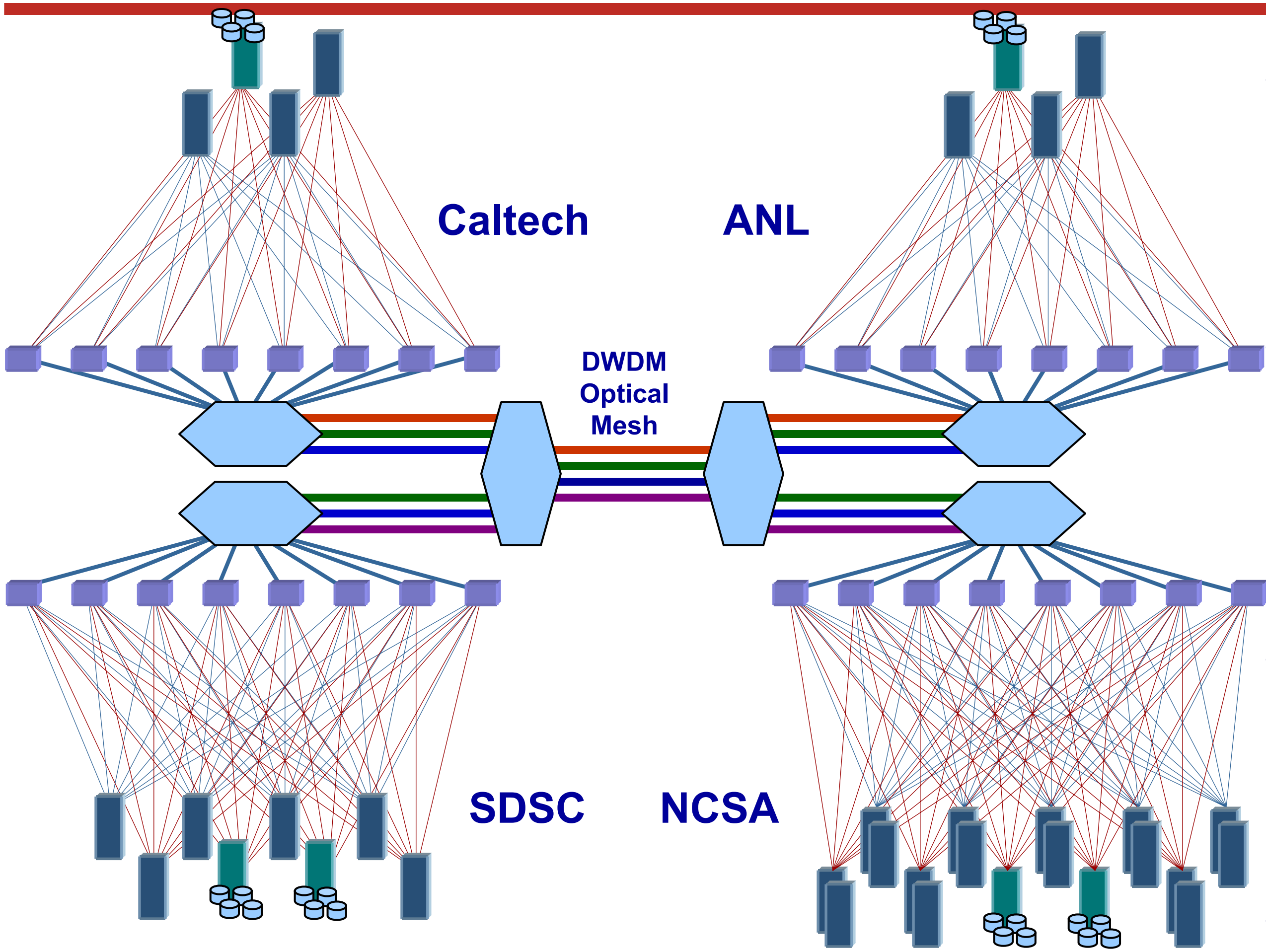
Spine Switches



64 hosts *64 hosts* *64 hosts* *64 hosts* *64 hosts*

128-port Clos Switches

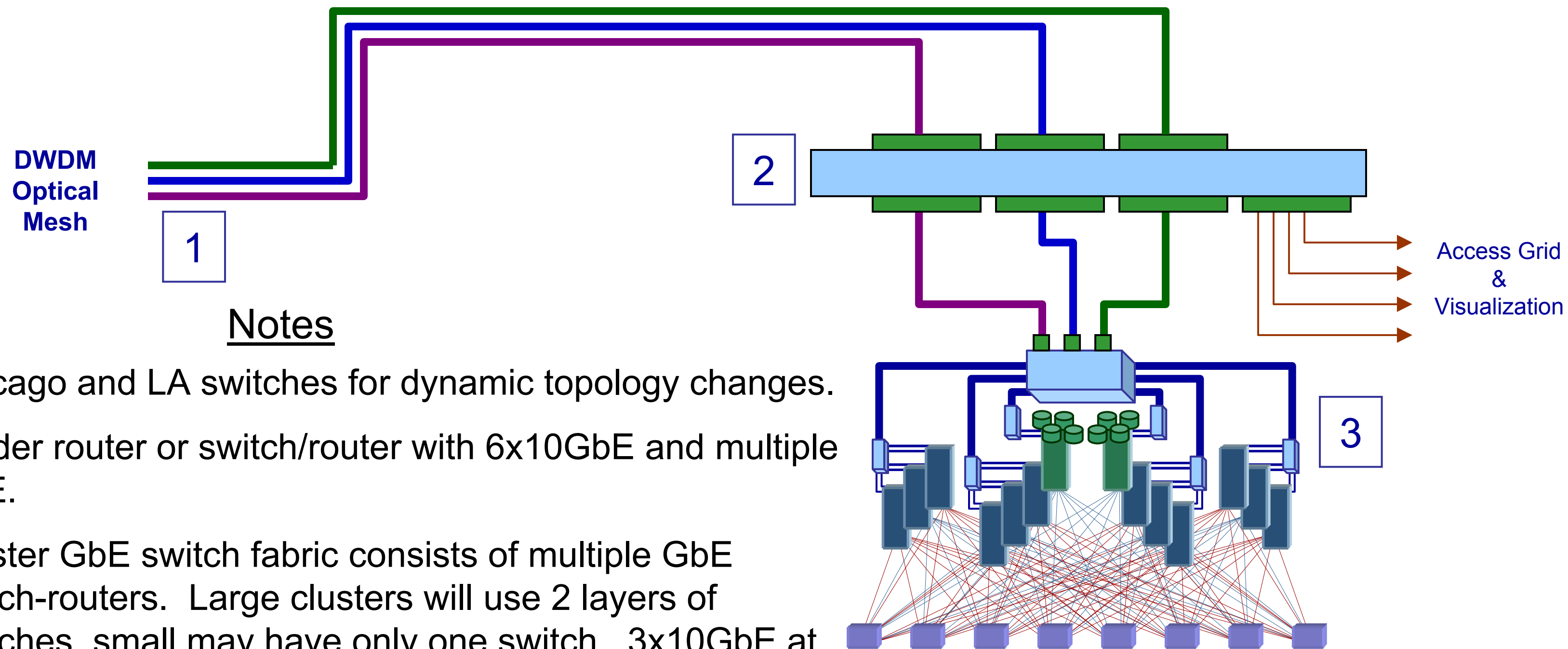
Initial Distributed Terascale Facility Design



- GbE interfaces in Myrinet root switches
  - •Myrinet scaling, GbE interface feasibility
- Layer 2 topology
  - •It's just a big bunch of GbE nodes…
  - •Requires spanning tree: sub-optimal routing here
- Lambda switching
  - •Immature technology
  - •Makes less sense with small number of λ's
- Fixed Lambda Mesh
  - •Very inefficient, inflexible bandwidth allocation

Charlie Catlett (catlett@mcs.anl.gov)

DTF Network Configuration Options



Notes

(1) Chicago and LA switches for dynamic topology changes.

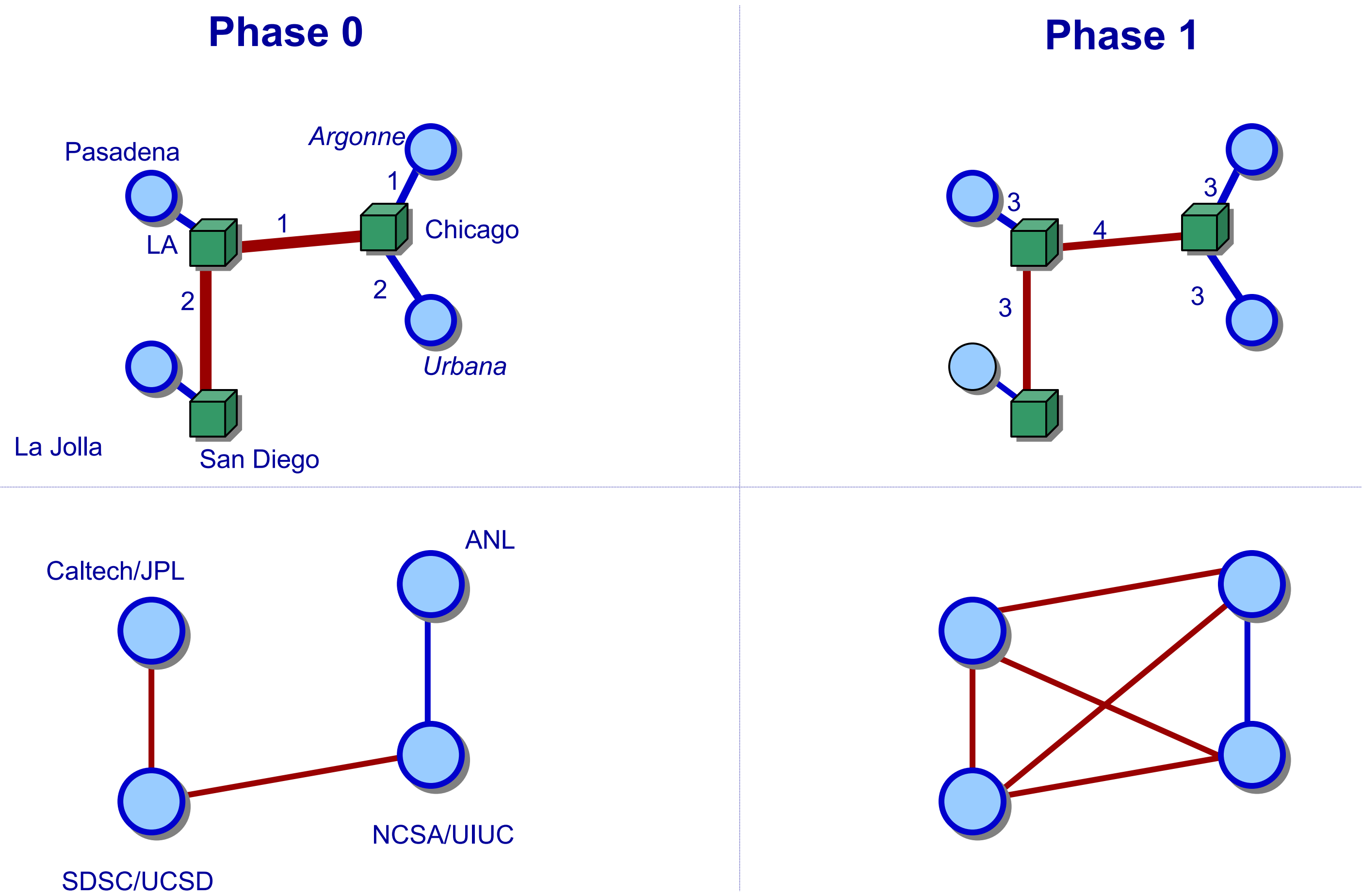
(2) Border router or switch/router with 6x10GbE and multiple GbE.

(3) Cluster GbE switch fabric consists of multiple GbE switch-routers. Large clusters will use 2 layers of switches, small may have only one switch. 3x10GbE at the top, nx10GbE at the bottom going out to n bottom layer switch/routers. Bottom layer switch/routers have 10GbE at the top and nxGbE at the bottom, connecting to individual cluster servers.

Physical to Logical Topology for TeraGrid

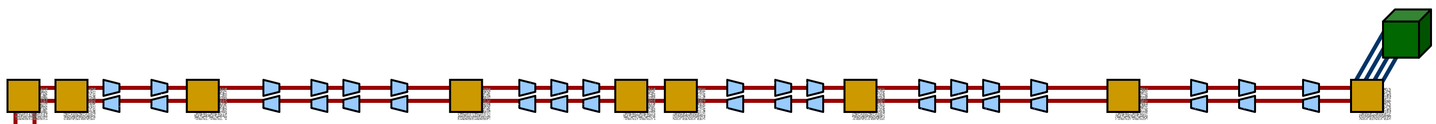
Physical
# denotes λ count

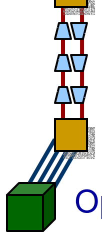
Light Paths (Logical)



Charlie Catlett (catlett@mcs.anl.gov)

Example Costs: 2000-mile Optical Network





Optical Transport System

Capital Costs

One Pair Fiber (2000 mi x 700/mi x 2) = $2.80M
Optical Amplifiers (every 50mi, 2000 mi, $100k/pair) = $4.00M
DWDM Base System (10 regen at $100k) = $1.00M
Four 10Gb/s Lambda ($75k/transponder x 18 x 4) = $5.40M
Total for 4-lambda system = $13.2M

Recurring Costs

Fiber Maintenance (@$200/mi) = $0.20M
Equipment space & power ($10k, every 50 mi) = $0.40M
Amplifier and DWDM Maintenance (15%) = $1.15M
Total for 4-lambda system = $1.75M

Integration with LANs

Capital Costs

IP Routers (2, $400k) = $0.80M
WAN interfaces (8, $200k) = $1.60M
LAN interfaces (8, $200k) = $1.60M
Total for 4-lambda system = $4.00M

Recurring Costs

15% of capital = $0.60M

Optical Transport System

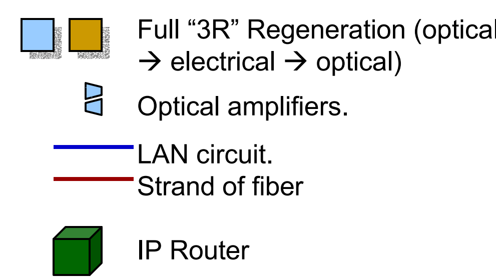
Capital Costs: $13.2M
Recurring Costs: $1.75M

Integration with LANs

Capital Costs: $4.00M
Recurring Costs: $0.60M

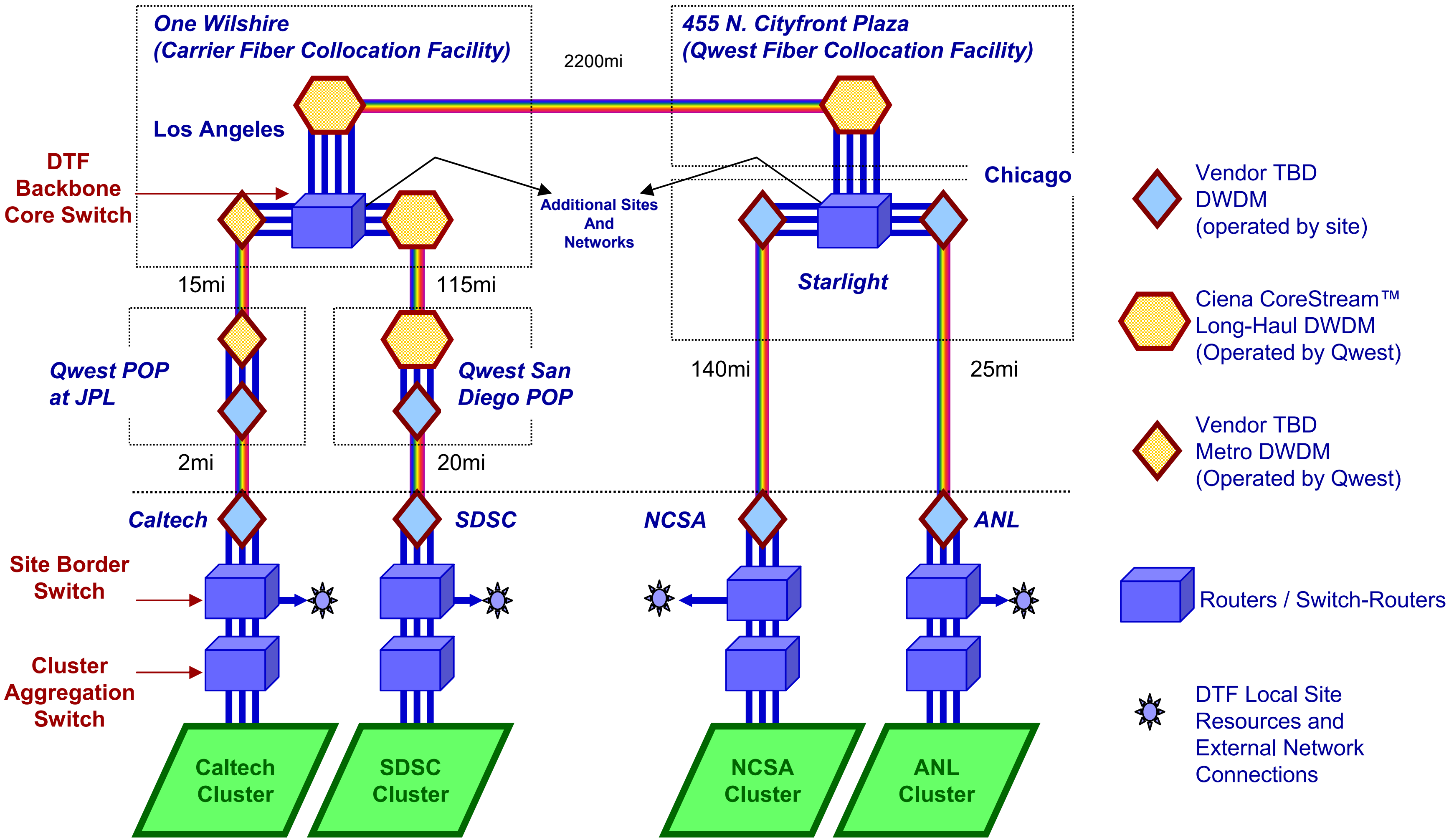
Total 4-Lambda Network

$17.2M, $2.35M/yr



Charlie Catlett (catlett@mcs.anl.gov)

DTF Network Architecture Options



Charlie Catlett (catlett@mcs.anl.gov)

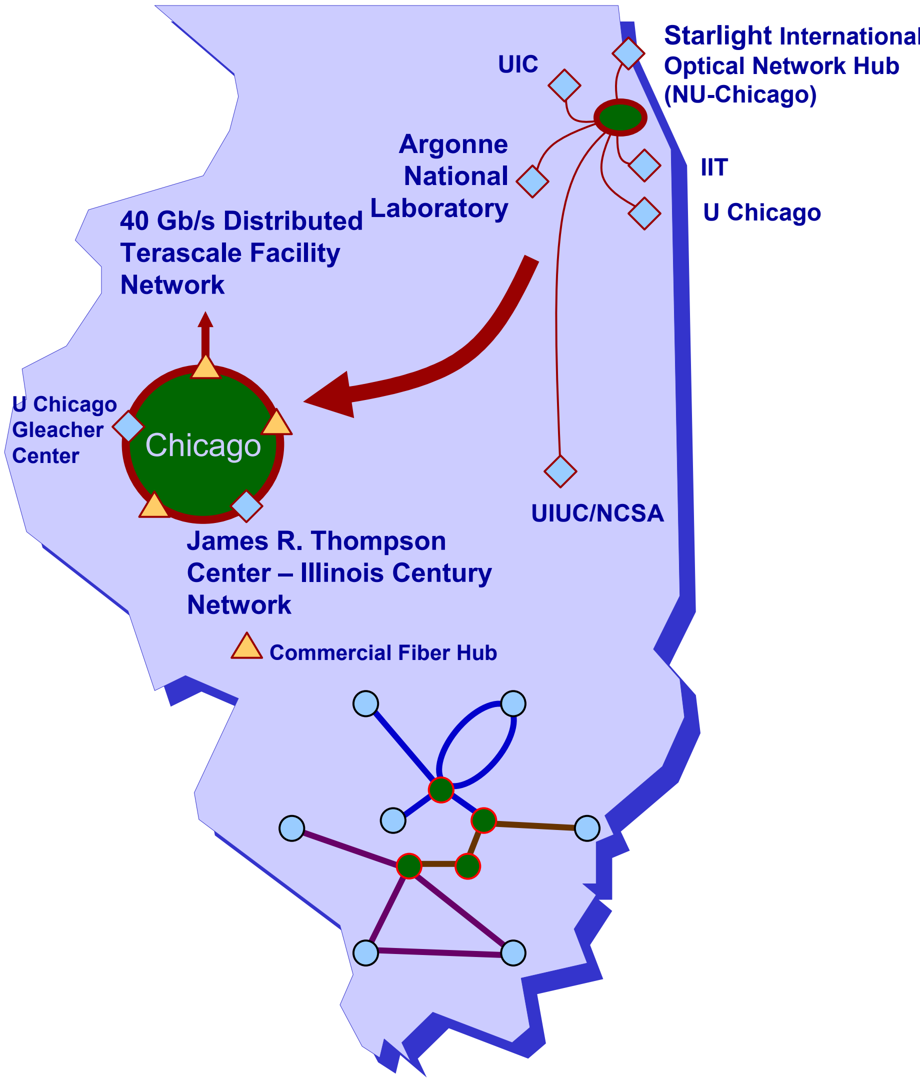
Optical Networks

- "High Performance Nets"
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

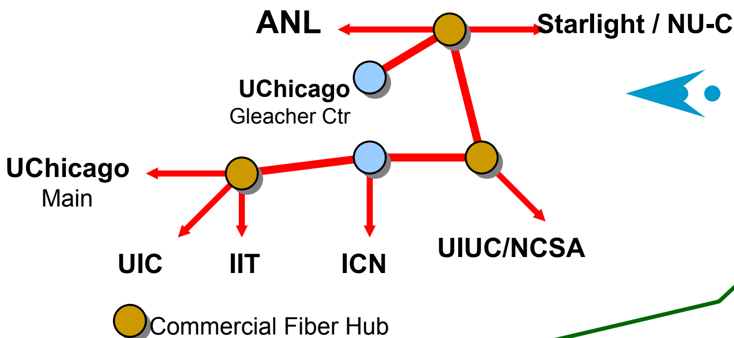
Case Study: I-WIRE

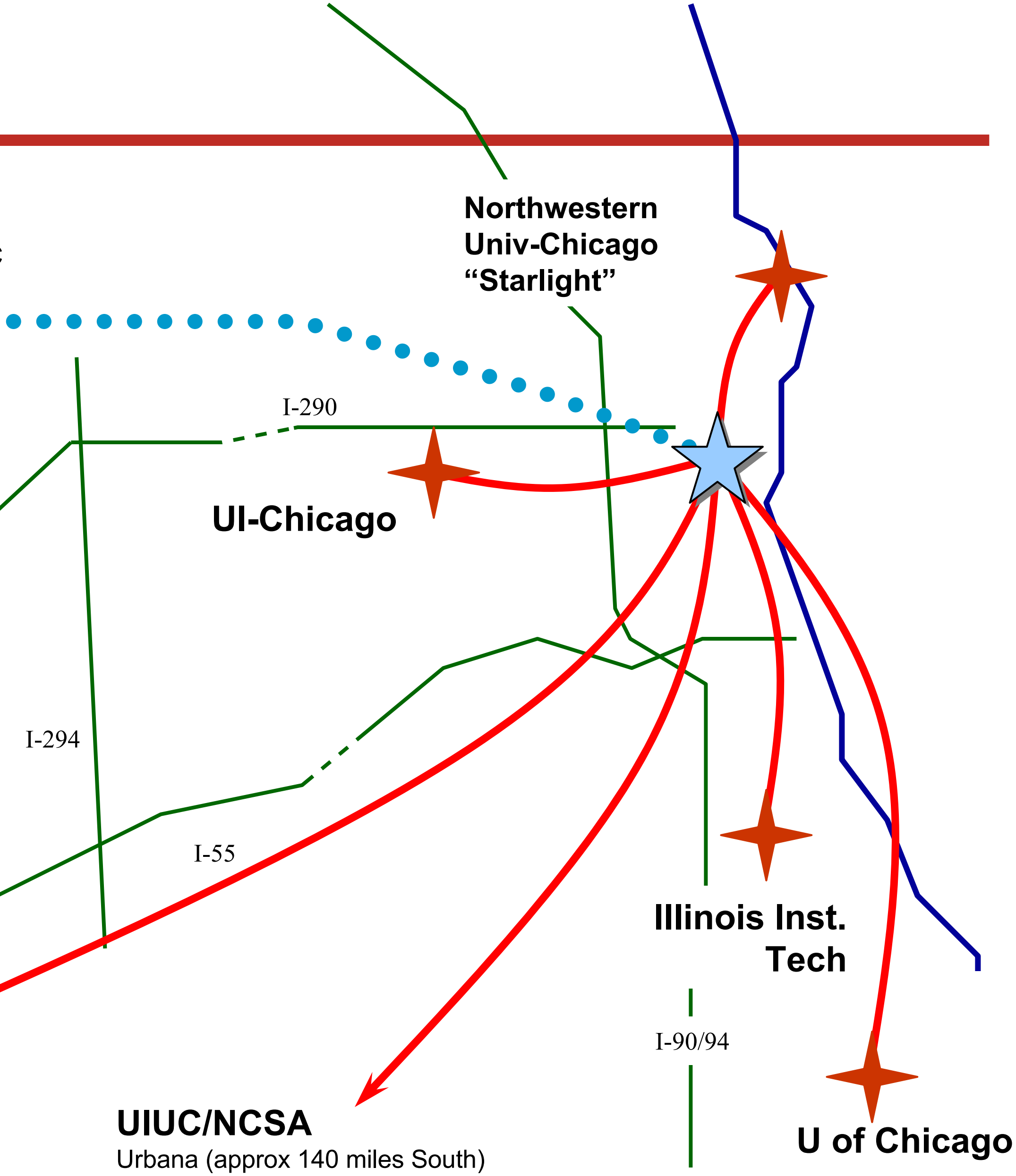
(Illinois Wired/Wireless Infrastructure for Research and Education)

- **State Funded Dark Fiber Optical Infrastructure**
  - $7.5M Total Funding 1998-2003
  - Application Driven
    - *Access Grid: Telepresence & Media*
    - *TeraGrid: Computational and Data Grids*
  - New Technologies Proving Ground
    - *Optical Network Architecture*
    - *Dense Wave Division Multiplexing*
    - *Optical network control and management*
    - *Advanced middleware infrastructure*



I-WIRE Geography

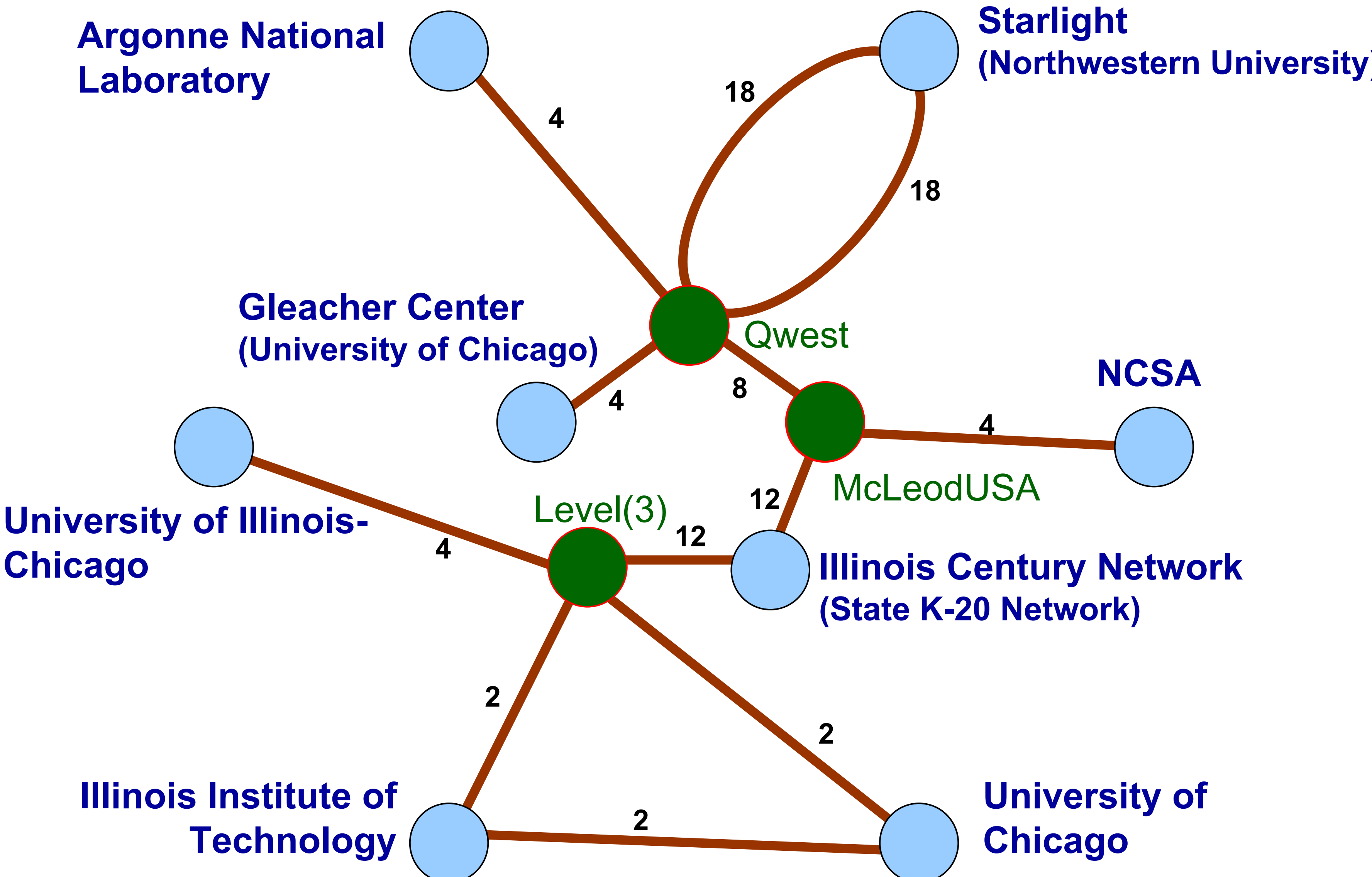




- Investigating extensions to Northwestern Evanston, Fermilab, Northern Illinois University (DeKalb), Chicago State University, Milikin University (Decatur), Bradley University (Peoria)

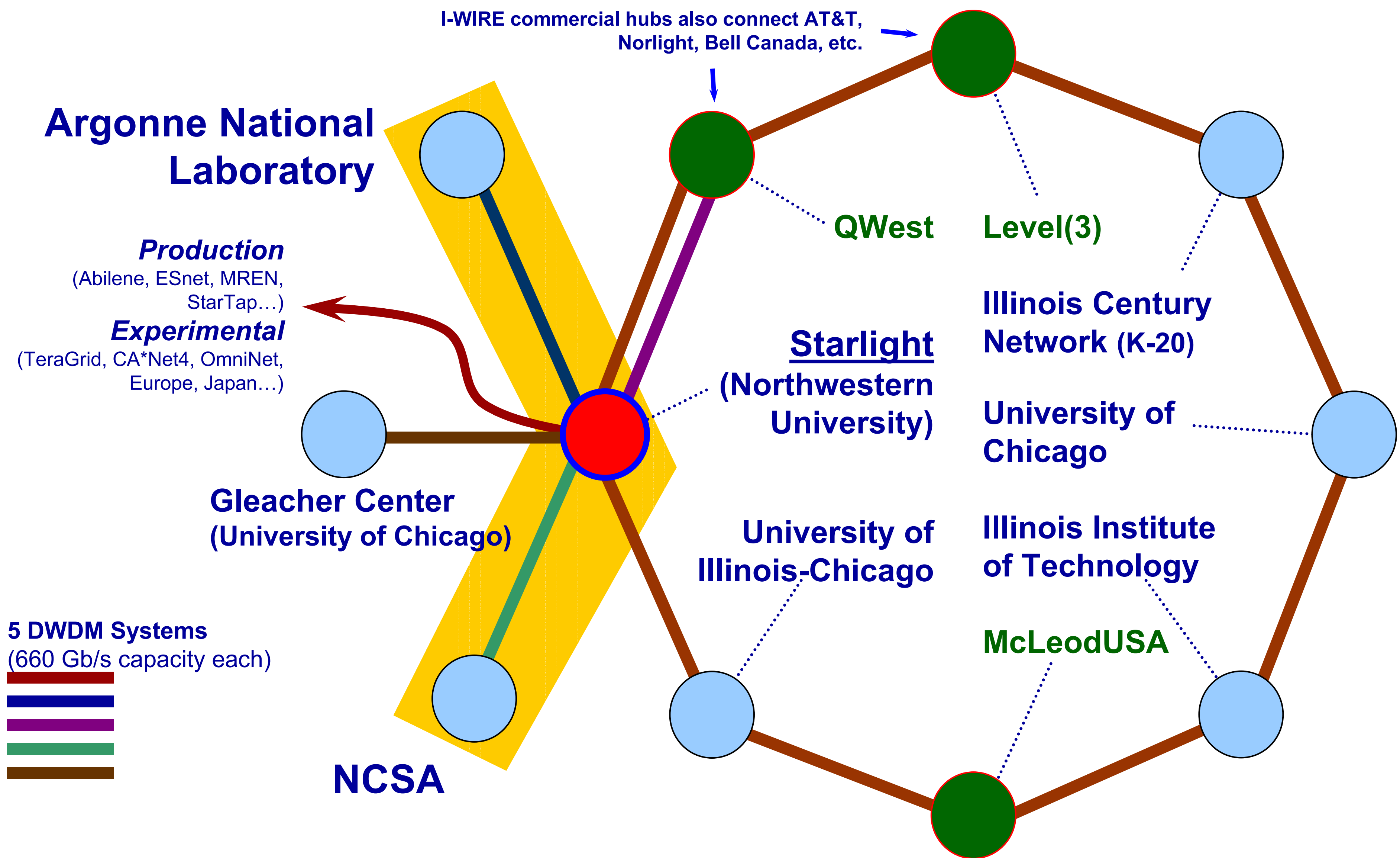
Charlie Catlett (catlett@mcs.anl.gov)

I-WIRE Fiber Topology



Charlie Catlett (catlett@mcs.anl.gov)

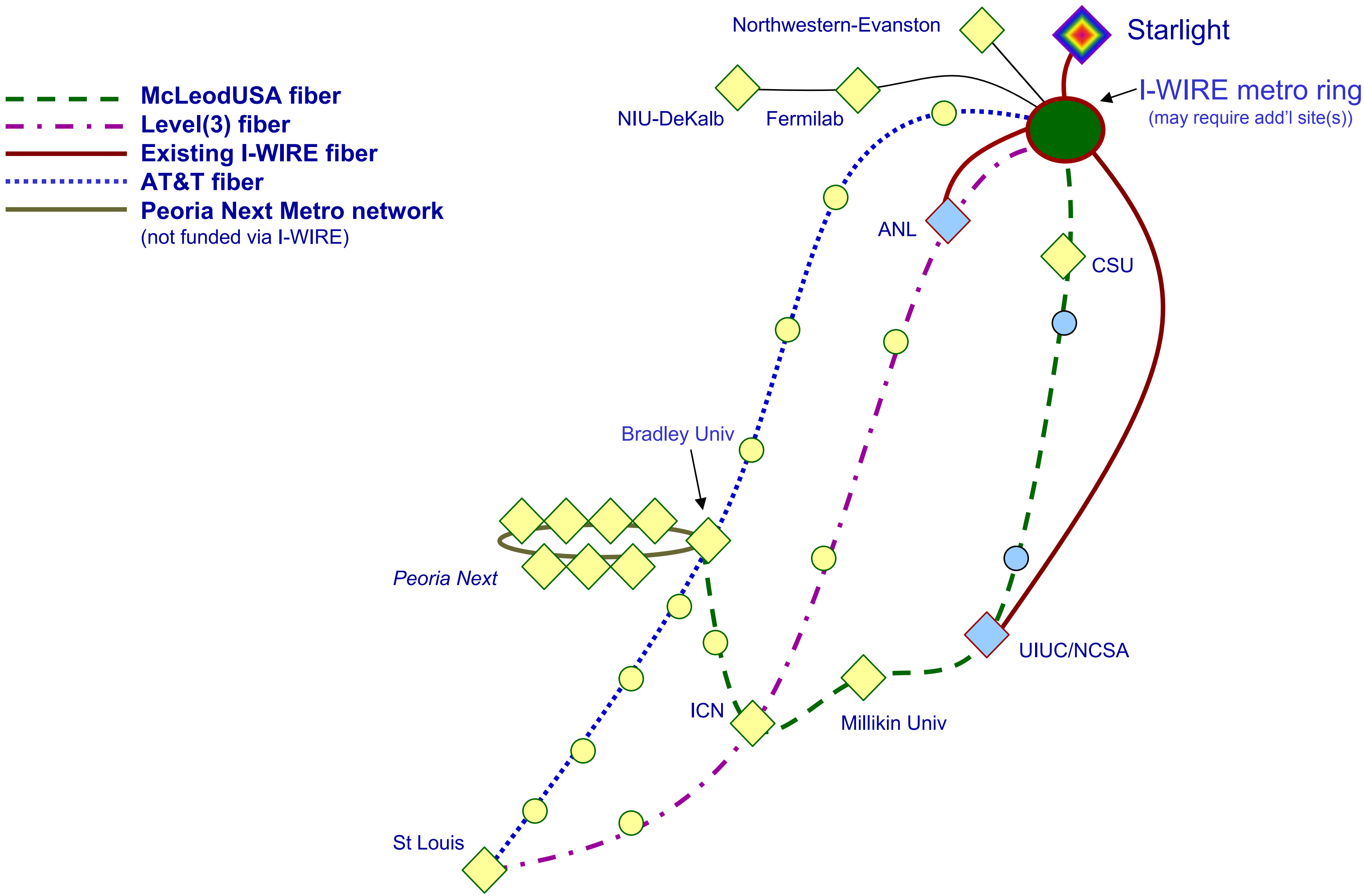
Initial I-WIRE DWDM Systems



I-WIRE Economics

- **Initial Costs**
  - $4M for fiber
    - *20-year IRU Fiber (existing fiber) ($700 to $5,000/strand-mile)*
    - *New construction (for 'last mile') ($30 to $100/foot ($160-530k/mile))*
  - $2.5M for equipment
- **Annual costs**
  - $100k for fiber maintenance, equipment space/power
  - $250k for equipment maintenance, engineering staff
- **Example of Potential Savings to the State of Illinois**
  - NCSA: 622 Mb/s Urbana to Chicago, $50,000 per month
    - *Replace with 2.5 Gb/s channel (4x capacity) using I-WIRE*
      - $35k equipment, ~$k/mo maintenance
- **Benefits to the State of Illinois (so far)**
  - Chicago area institutions saving $3-400,000/year in Internet costs
  - Leveraged to win $88M TeraGrid project (joint with NCSA and 3 other labs).
  - Creates a unique environment for attracting research funding, world-class researchers, and companies to the Chicago area and to Urbana.

Proposed 2003 I-WIRE Expansion



Charlie Catlett (catlett@mcs.anl.gov)

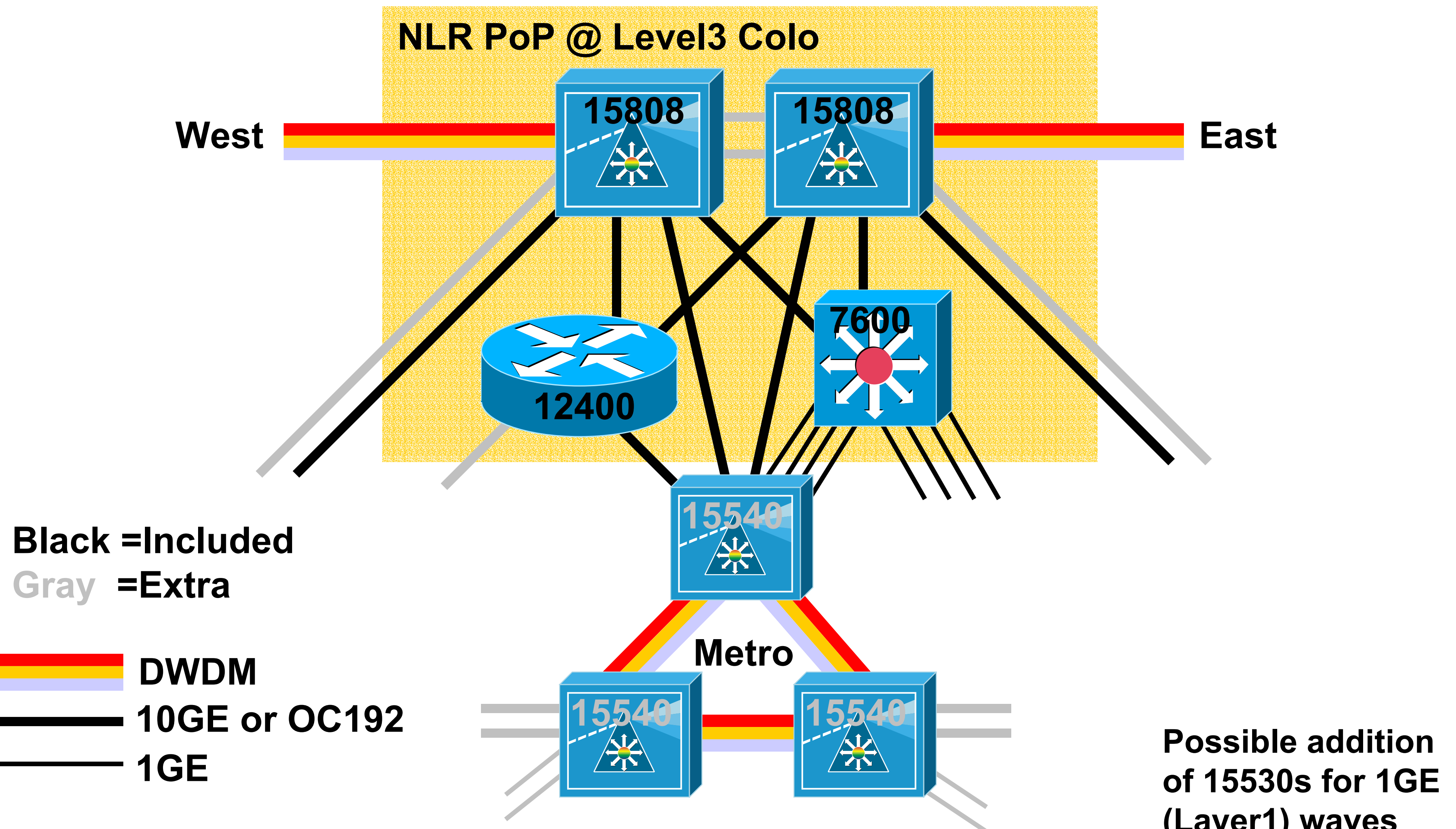
Optical Networks

- “High Performance Nets”
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - TeraGrid
  - I-WIRE
  - National Light Rail
- Architectures and Futures

Case Study: National Light Rail (in progress)

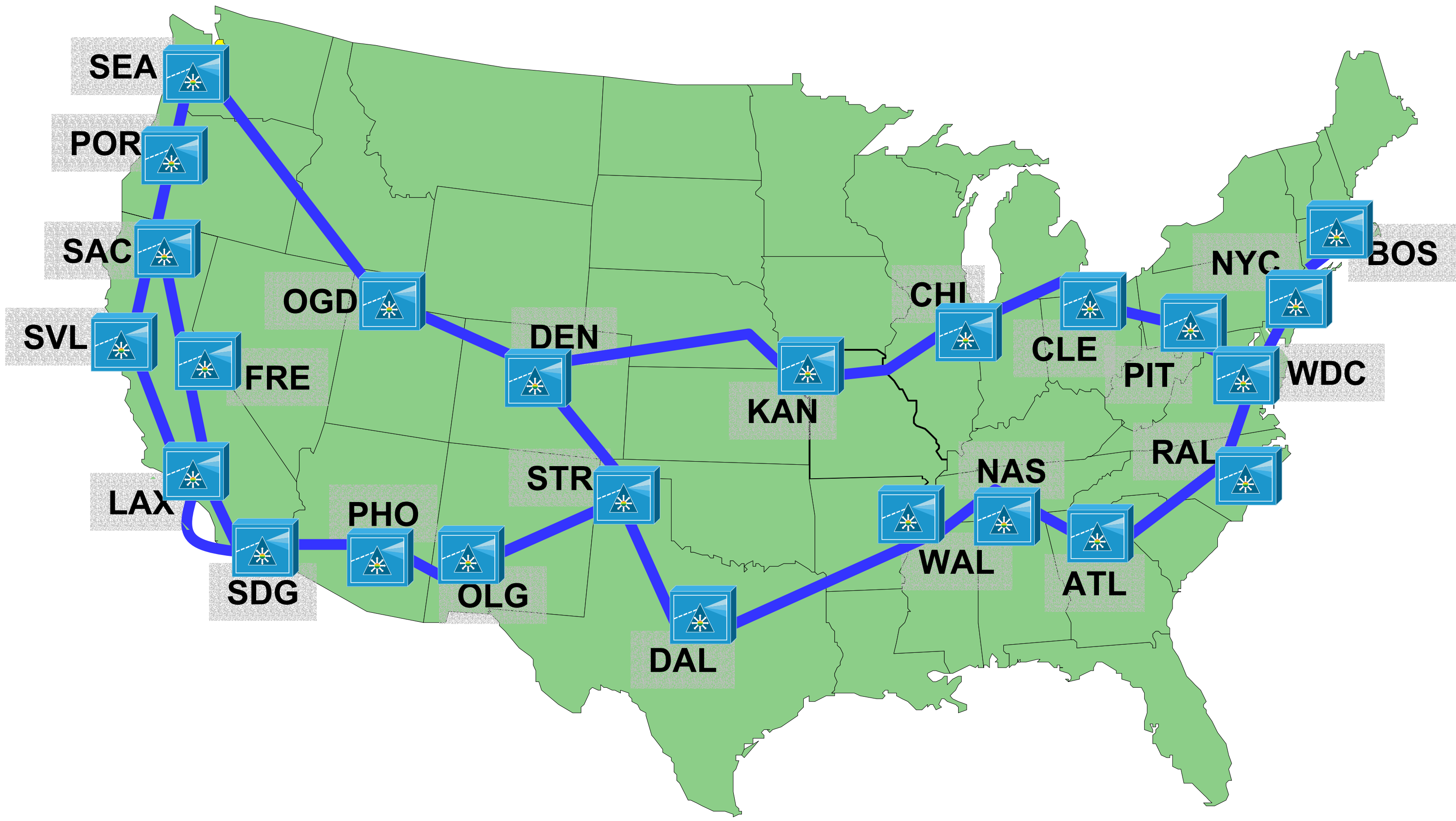
- **7000 mile national fiber footprint**
- **Four initial lambdas**
- **Three "networks"**
  - IP routed network (1 lambda)
  - 10 Gb/s experimental wavelengths (2 lambdas)
  - 1 Gb/s experimental GbE service (1 lambda)
- **Total Cost $800-100M depending on final topology**
  - 5 year cost, includes $2-3M/yr operations
- **Status**
  - Timeframe: 2003-4
  - Funding: Final negotiations with partners, enough momentum (i.e. partners with money committed) to build at least 1/2 of proposed system

NLR PoP Architecture



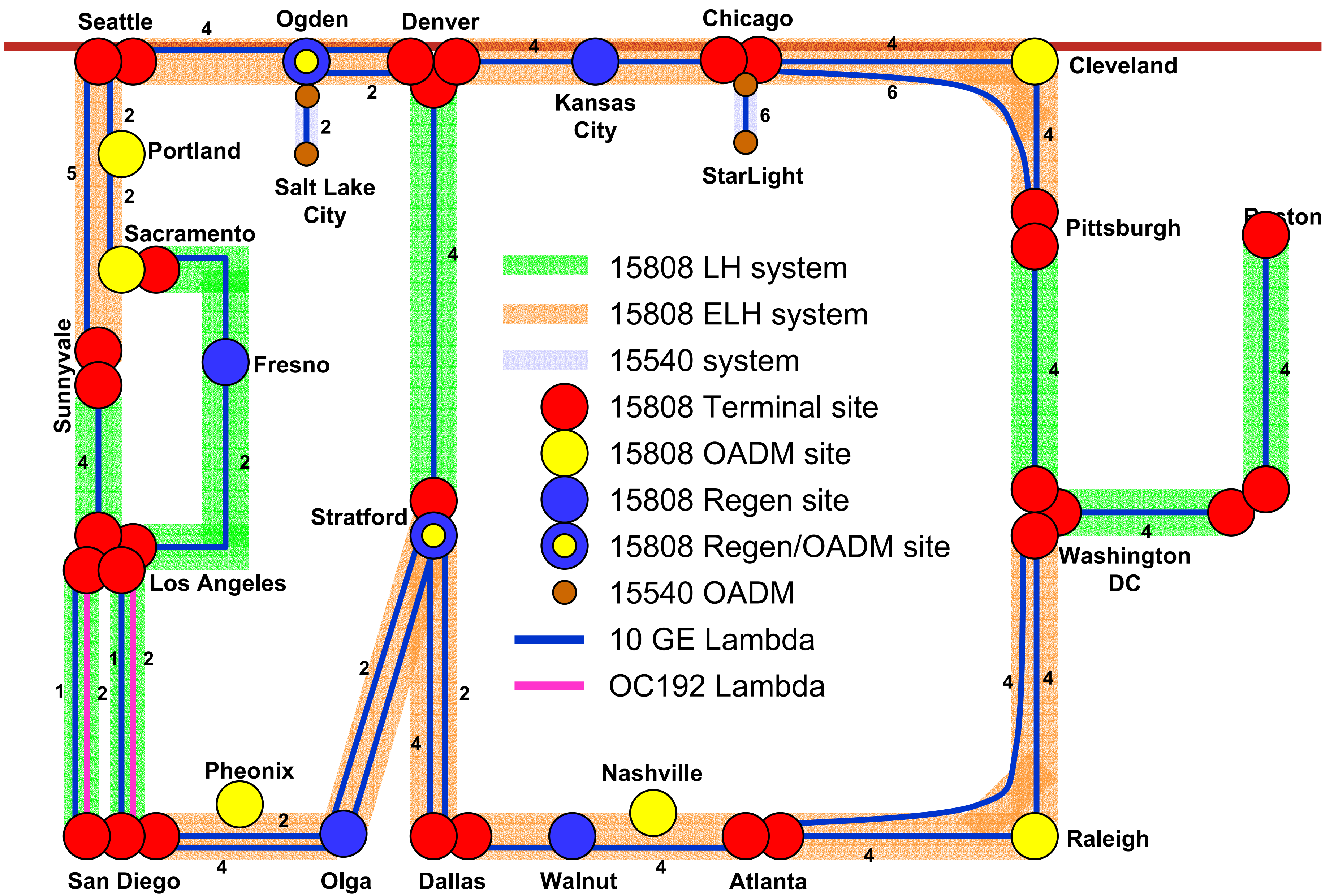
Source: NLR, Cisco

NLR Footprint and Layer 1 (optical) Topology



Source: NLR, Cisco

NLR 15808/15540 Wave Map



Source: NLR, Cisco

Evaluation Strategies and Factors

- **Cost (Optical portion only)**
  - Metric: Cost of lambda per mile per year
  - Examples:
    - *TeraGrid Qwest Partnership: Under $500*
    - *Recent Vendor Quotes: $6-800*
    - *NLR: $6-800 for initial waves, lower for additional waves*
- **Notes**
  - Lambdas are incremental relative to base infrastructure
  - More lambdas, lower per-lambda cost
- **Options**
  - Lease from Vendors (see recent vendor quotes above)
  - Build Dedicated
    - *Low cost for many (>16) lambdas, high initial cost*
  - Build with Consortium
    - *Share high initial cost with others*
  - Build with Vendor
    - *Leverage vendor infrastructure to lower initial cost*

Selecting Fiber Sources

- Footprint Routes and Coverage
  - **AT&T goes North, others cut across middle (skipping Northern Plains). Inter-city route distances vary.**
- Fiber types
  - **Level(3) is homogeneous LEAF fiber. AT&T is multiple types. Any impact? (maybe for ULH)**
- Repeater hut spacing
  - **AT&T spacing is closer. More amplifiers. But at higher speeds will closer spacing be critical?**
- Space and Power costs, access policies
  - **250+ locations on a 16,000 mile footprint**
- Metro space locations
  - **relative to regional/local fiber, other vendors (carrier neutral colo or single vendor controlled?)**
- Fiber Costs and Terms
  - **IRU, $/strand-mile, willing to "strand"?, $/mile maintenance (strand-mile or route mile?)**
  - **Legal issues- indemnification, terms for "forced relocation," bankruptcy contingency!**

Context: Fiber build costs range from $30-100/ft (up to $500k/mile).
If you only want 2 strands… cost ranges from $1000-5000 per mile.

Optical Networks

- “High Performance Nets”
- Optical Network Building Blocks
- Optical Transport
- Case Studies:
  - •TeraGrid
  - •I-WIRE
  - •National Light Rail
- Architectures and Futures

End-to-End Architecture Issues

- **Today's High Performance *Production* Networks**
  - Packet Switched: IP Routers
  - Fixed topology between routers
    - *Optical networks used for fixed point-to-point transport*
    - *No ultra-long haul in production, anywhere*
    - *DWDM systems use fixed-wavelength optics*
    - *Provisioning takes hours to days*
  - Maximum 10 Gb/s Individual Channel speeds
    - *Typically 2.5 Gb/s backbones*
    - *3-5 10 Gb/s backbones in service as of late 2002*
    - *1 multi-10 Gb/s backbone, TeraGrid at 40 Gb/s*
- **Testbeds and Proven in the Laboratory**
  - Lambda Switching/Routing
  - Tunable lasers
  - Optical Burst Switching

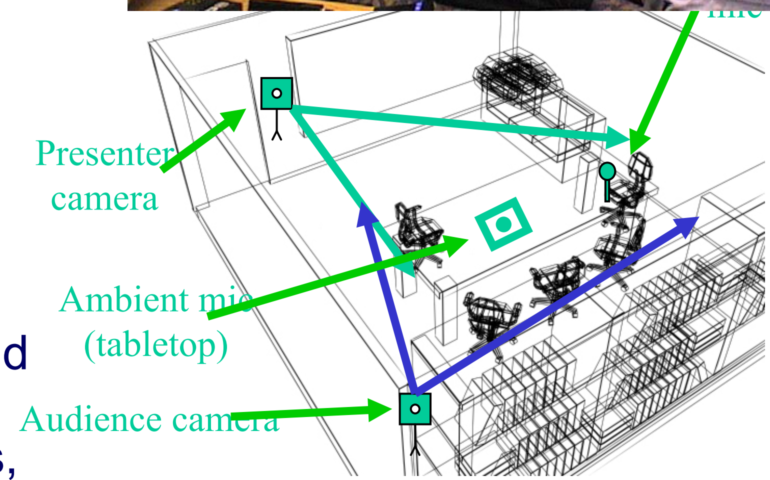
Optical Bandwidth on Demand

- **Digital Theater**
  - Each movie theater in a large area (SF, New York, Houston) requests 1 hour of bandwidth a week (OC192)
  - All movies transferred during this time
  - Efficient use of expensive but necessary fat pipe
- **Sporting Events**
  - E.g. Football Stadiums - need high-bandwidth 8-10 times/year
  - Today's answer: drive a truck to the event with a satellite dish
    - *OK for today's model*
  - Experiments with multiple HDTV streams
    - *Too much for satellite*
    - *Experiments in stadiums with user-selectable camera view*

Source: George Porter, Tal Lavian UC-Berkeley

Many-to-Many Interactive Collaboration

- **Access Grid**
  - Enable collaborative work at dozens of sites worldwide, with strong sense of shared presence
  - Combination of commodity audio/video tech + Grid technologies for security, discovery, etc.
  - 50+ sites worldwide, number rising rapidly
- **Exploration of Game Consoles**
  - PS2 or Xbox as Access Grid node
- **Network Flow Engine**
  - Heavy Dependence on Multicast, moving toward SSM
  - Need to break large streams into multiple formats (variable quality), bind some streams, present participants (sites) with information about streams, etc.



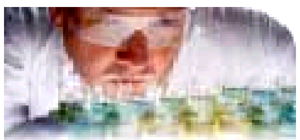
Source: Rick Stevens, ANL

“ Cause Computing”

Mathematical Research

Push the limits of our theoretical understanding of the abstract mathematics tools that civilization is built upon and explore these new frontiers with your computer.

Fighting Diseases

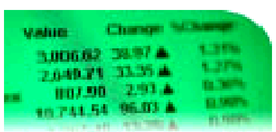


Striving to improve quality of life and eliminate suffering, Entropia members can support disease research projects.

Entertainment

Entropia is building new technologies to help accelerate wonderful new digital entertainment productions. Your computer can be a key part of bringing this exciting new technology to life!

Economics Research



Long term stability of the world economies has become crucial to growing prosperity. Entropia members can help illuminate the meanings of global economic behavior through rigorous research models.

Scientific Research

Science is a foundation on which many of the greatest human achievements rest. The Entropia community can participate in some of the most intense scientific research underway today.

Environmental Research

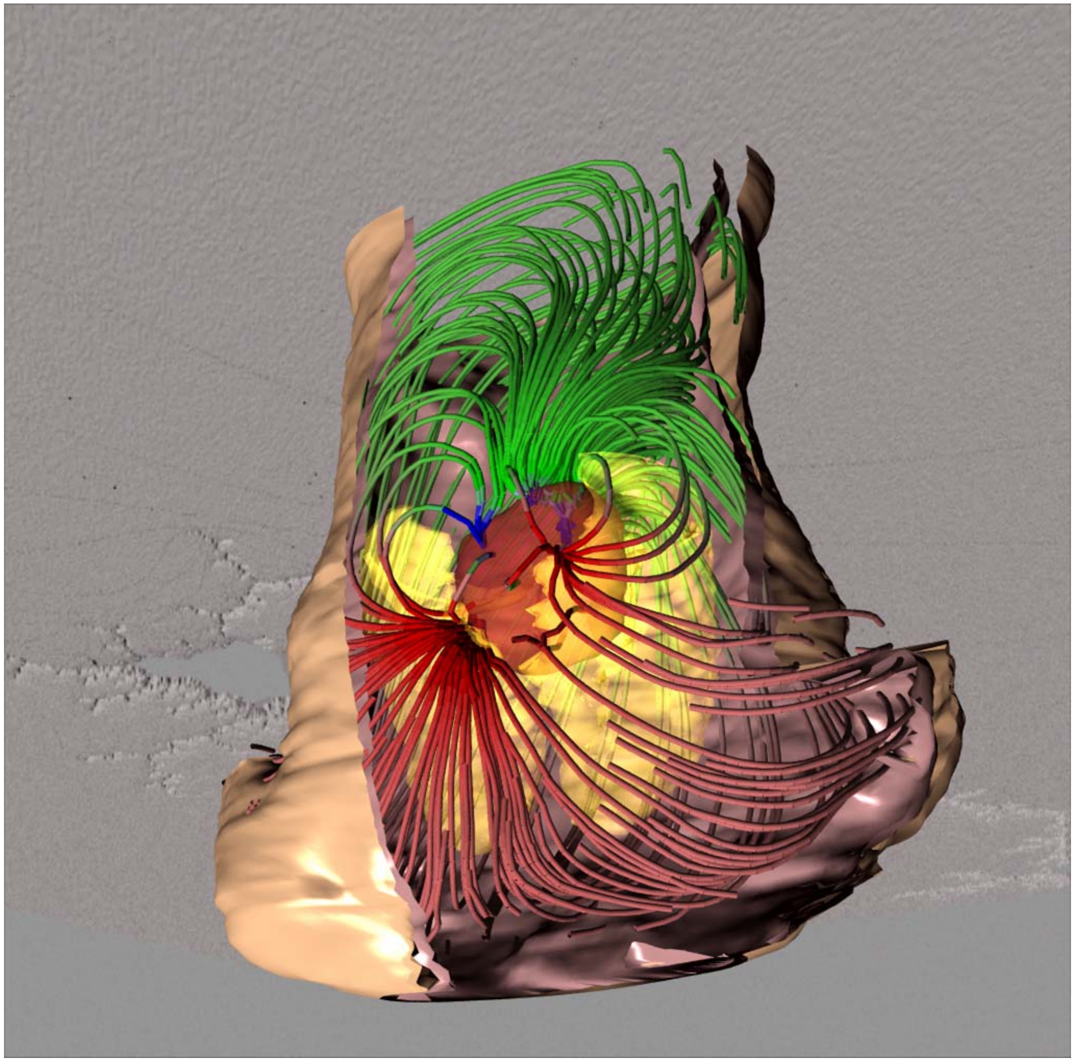
Responsible management and preservation of Earth's environment requires a deep understanding of the complex effects of many factors. Entropia members can help researchers determine the most important aspects of planet stewardship for future generations.

Product Design

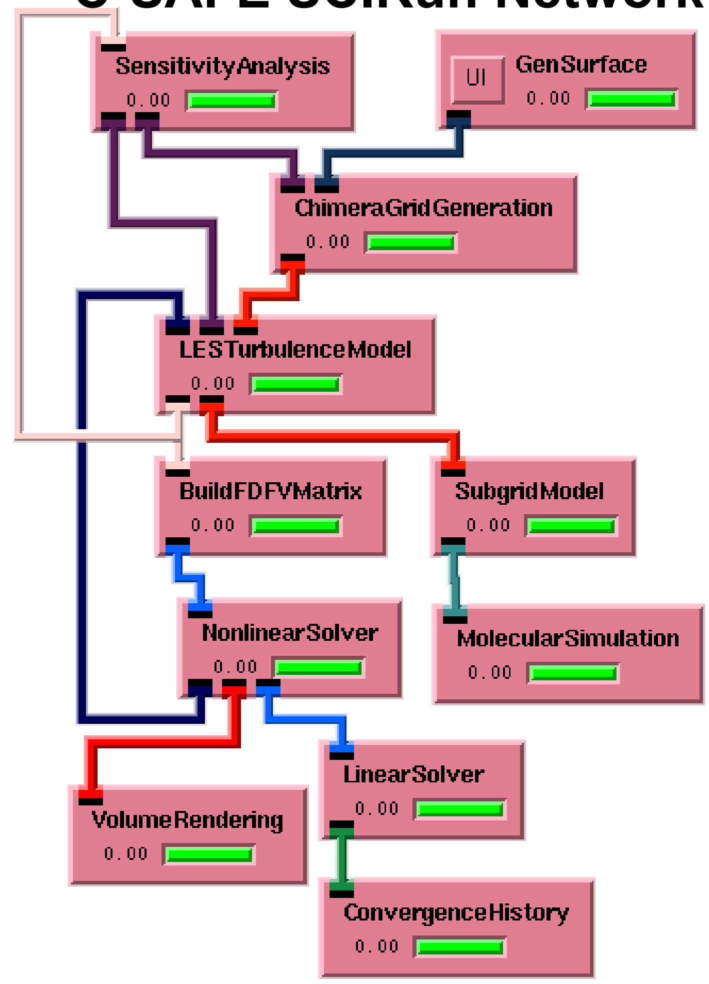
Researching safe product designs quickly and effectively requires an ever increasing amount of computing power to test and refine them before manufacturing even begins. Your computer can play a crucial role in making safer medicines, transportation, appliances, clothing, toys and more!

Source: Entropia

Medical Simulation: Remote Visual Steering



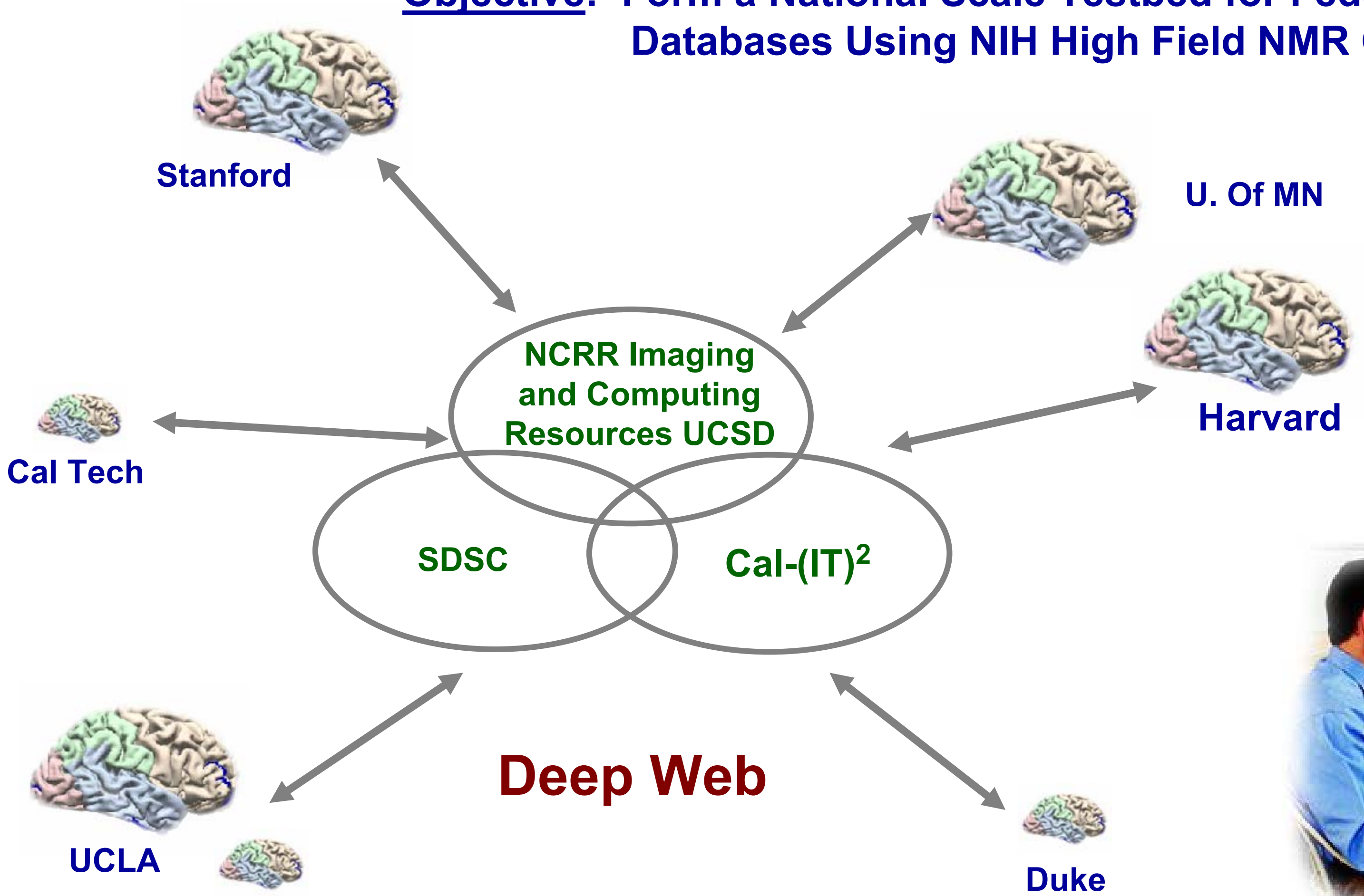
Example Combustion C-SAFE SCIRun Network



Source: Chris Johnson, Univ. of Utah

The Brain Data Grid

Objective: Form a National Scale Testbed for Federating Large Databases Using NIH High Field NMR Centers



Surface Web

Wireless "Pad" Web Interface

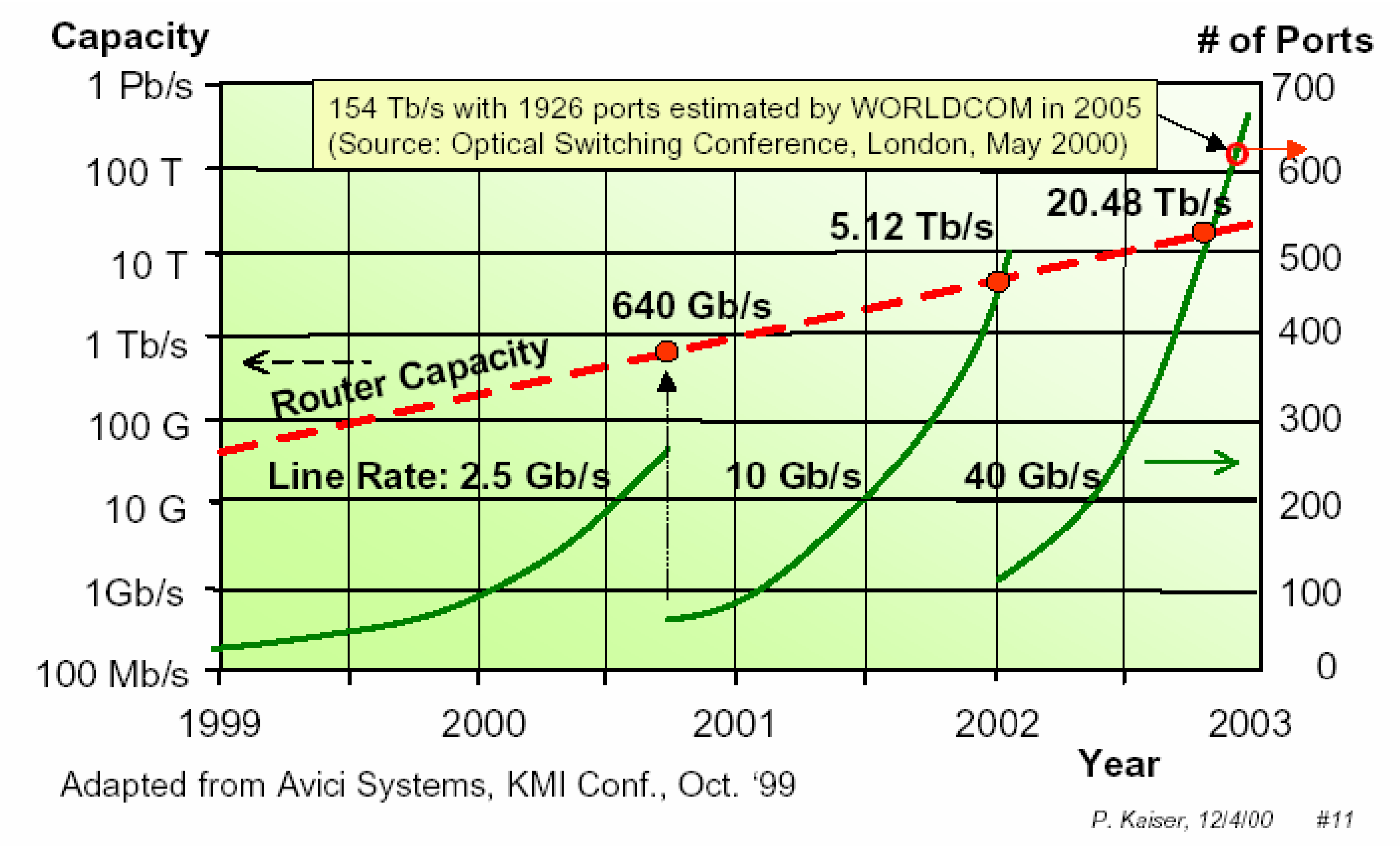
Cyber Infrastructure Linking Tele-instrumentation, Data Intensive Computing, and Multi-scale Brain Databases.

Source: Mark Ellisman and Larry Smarr, UCSD

Provisioning and Guaranteed Bandwidth

- **MPLS**
  - Multiprotocol Label Switching (MPLS)
  - Routers set up Label Switch Paths through the network
    - *Collections of routes are fast-tracked through these paths without the router having to do a route lookup.*
  - Assumes large shared pipes, MPLS paths of some smaller amount of bandwidth
- **Example- TeraGrid**
  - Application running on 100 cluster nodes at SDSC and 200 cluster nodes at NCSA
  - Set up MPLS path for all traffic between these nodes, e.g. at 25 Gb/s. Leave remaining bandwidth for general traffic
- **Future: G-MPLS**
  - Generalized MPLS (aka MPλS)
  - Routers set up optical paths through the network
  - Assumes large number of available optical paths

Evolution of IP Routers

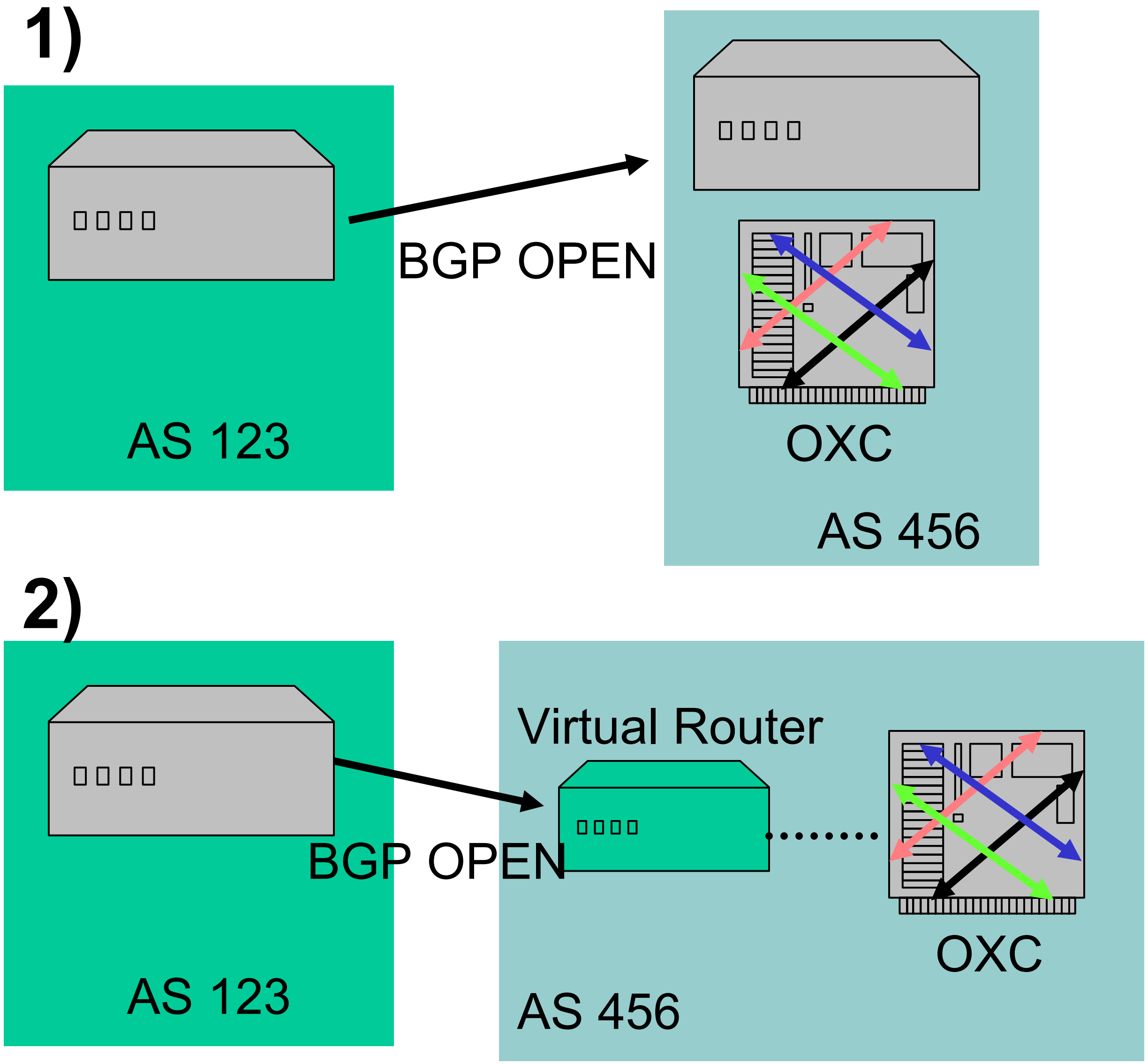


Source: P. Kaiser

Canarie's approach

- **OBGP (Optical BGP)**
- **Routers spawn "virtual BGP" processes that peers can connect to**
- **By modifying BGP messages, lightpath information can be traded between ASes**

Canarie Virtual Routers and OBGP



- BGP OPEN message sent to router with information about optical capabilities
- A virtual BGP process is spawned
- A BGP session is initiated independently with new BGP process
- The virtual process (running on the router) configures the OXC to switch the proper optical wavelengths

Optical BGP Networks

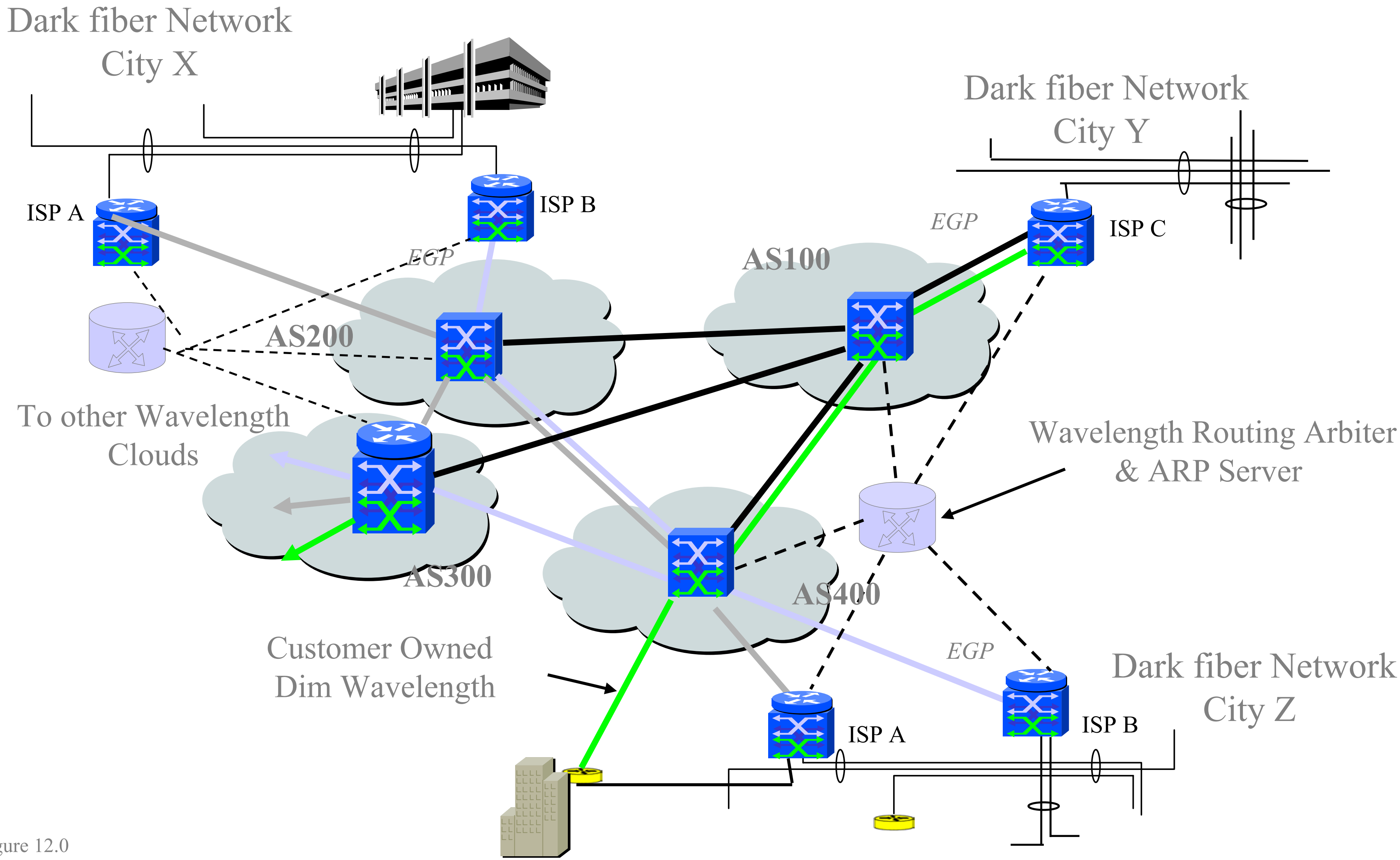
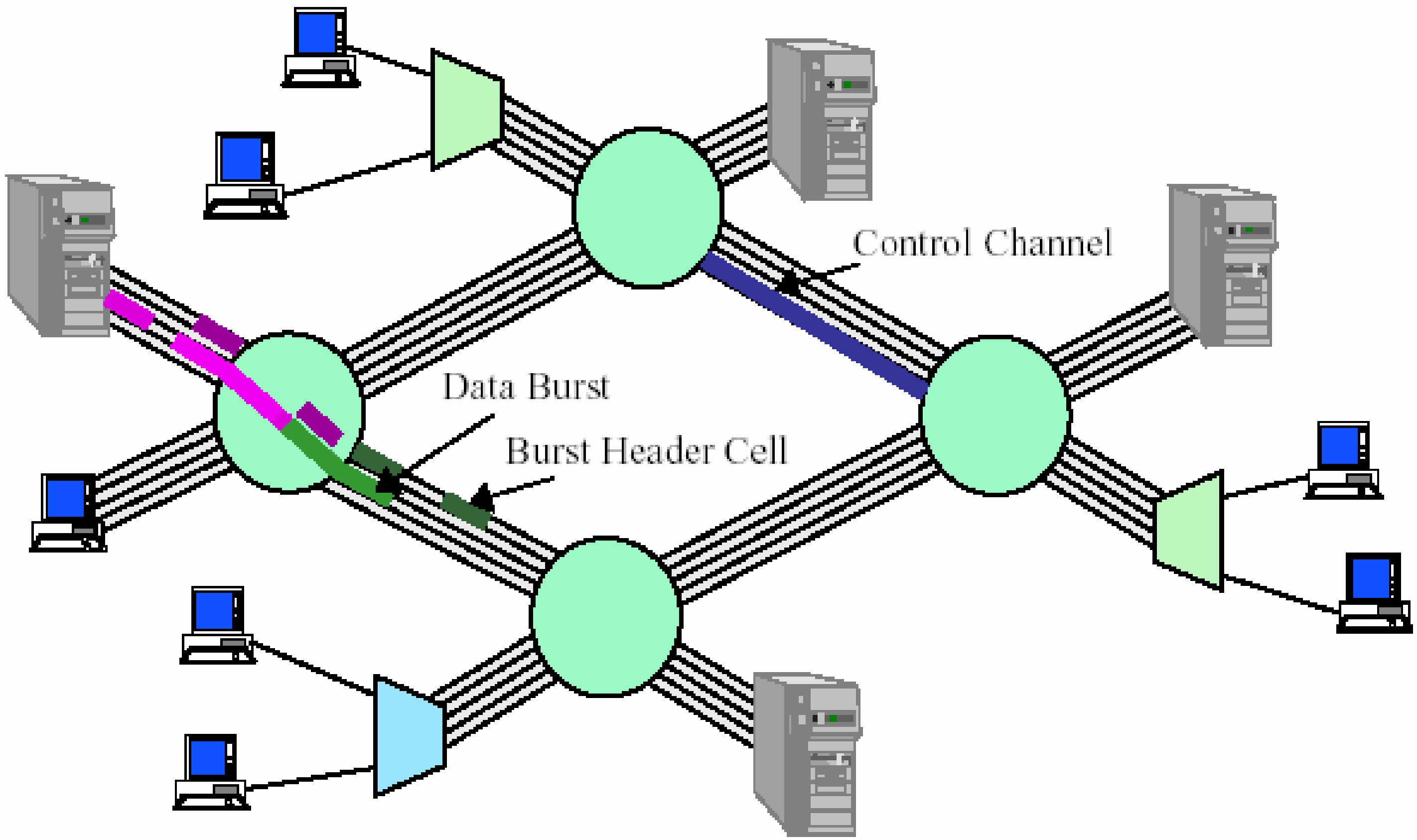


Figure 12.0

Source: CANARIE

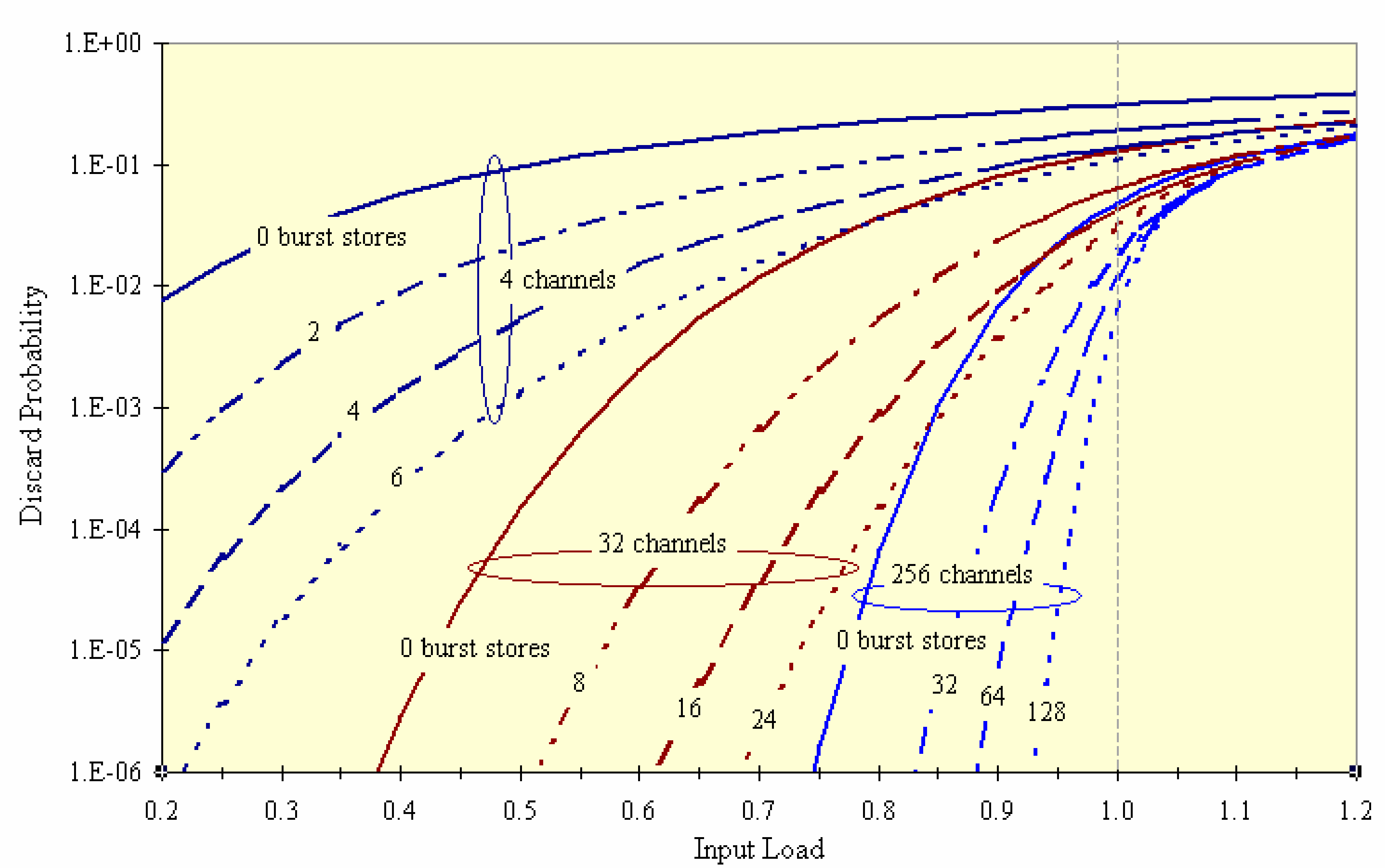
Burst Switching



- Separate control and data channels
- Burst Header Cel sent to set up next-hop path. BHC contains offset (time from first bit of BHC to first bit of Burst) and length (of Burst)
- Without optical buffering, critical design point is to avoid blocking.

Source: Jonathan Turner, Washington University St. Louis

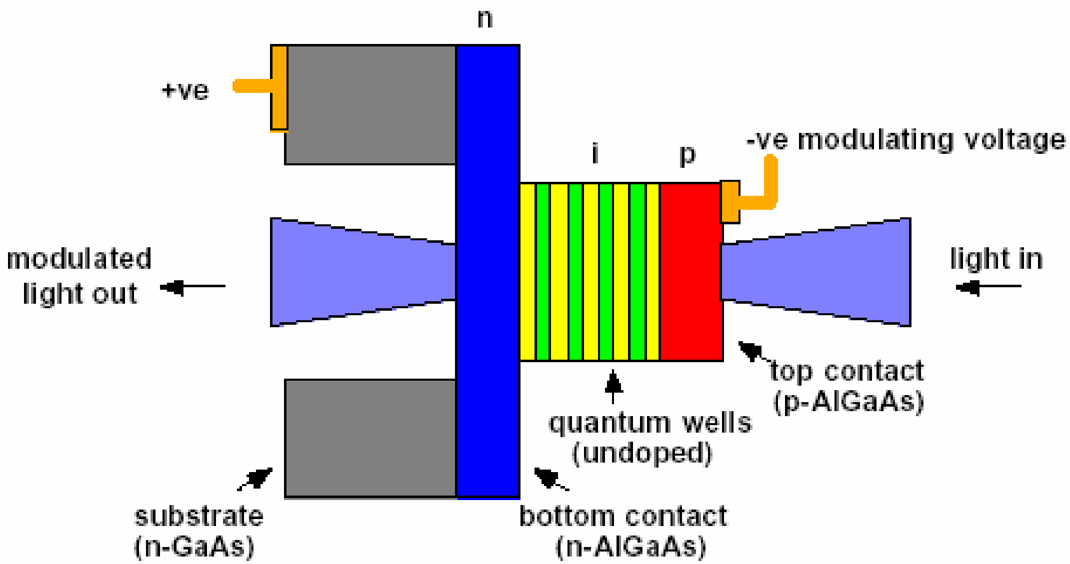
At Issue for *Optical* Burst Switching: Blocking



Source: Jonathan Turner, Washington University St. Louis

Futures- Optical Chip Interconnects

Example Optoelectronic Device for Chip Interconnects - Quantum Well Electroabsorption Modulator



essentially no internal speed limitations

no "threshold"

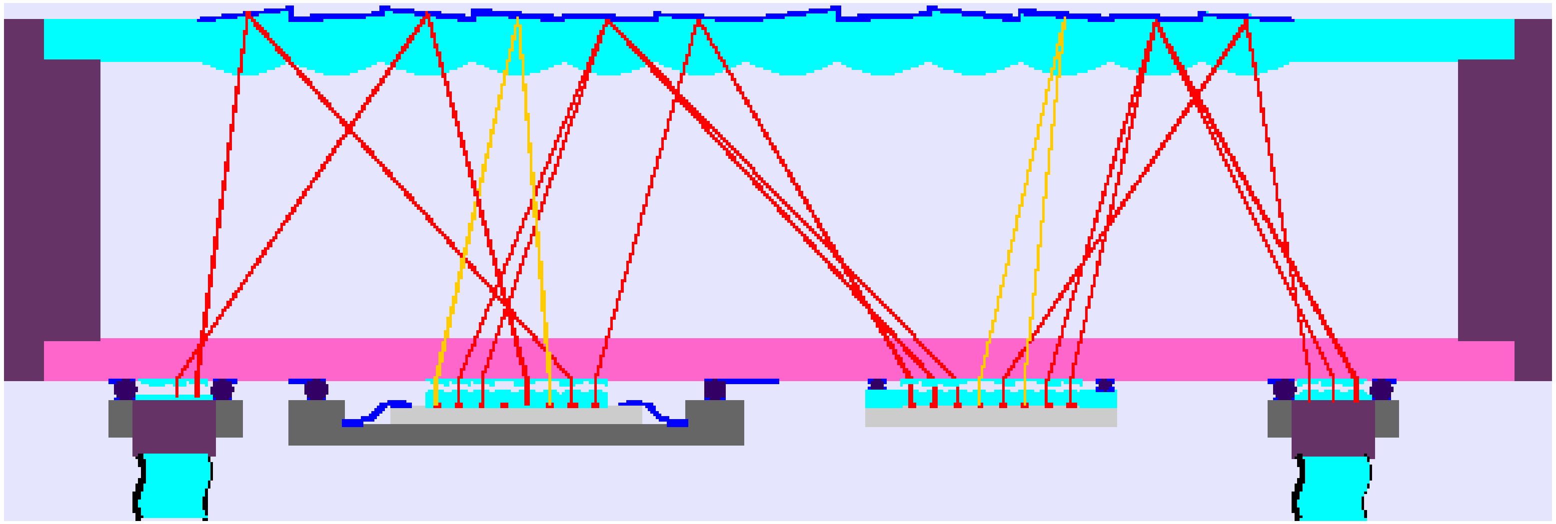
successfully fabricated and bonded in large arrays (e.g., 4000)

use in transmission or, with internal mirror, reflection

same mechanism as high-performance modulators on externally modulated lasers (EMLs)

Source: David Miller, Stanford

Optical Module of the Future



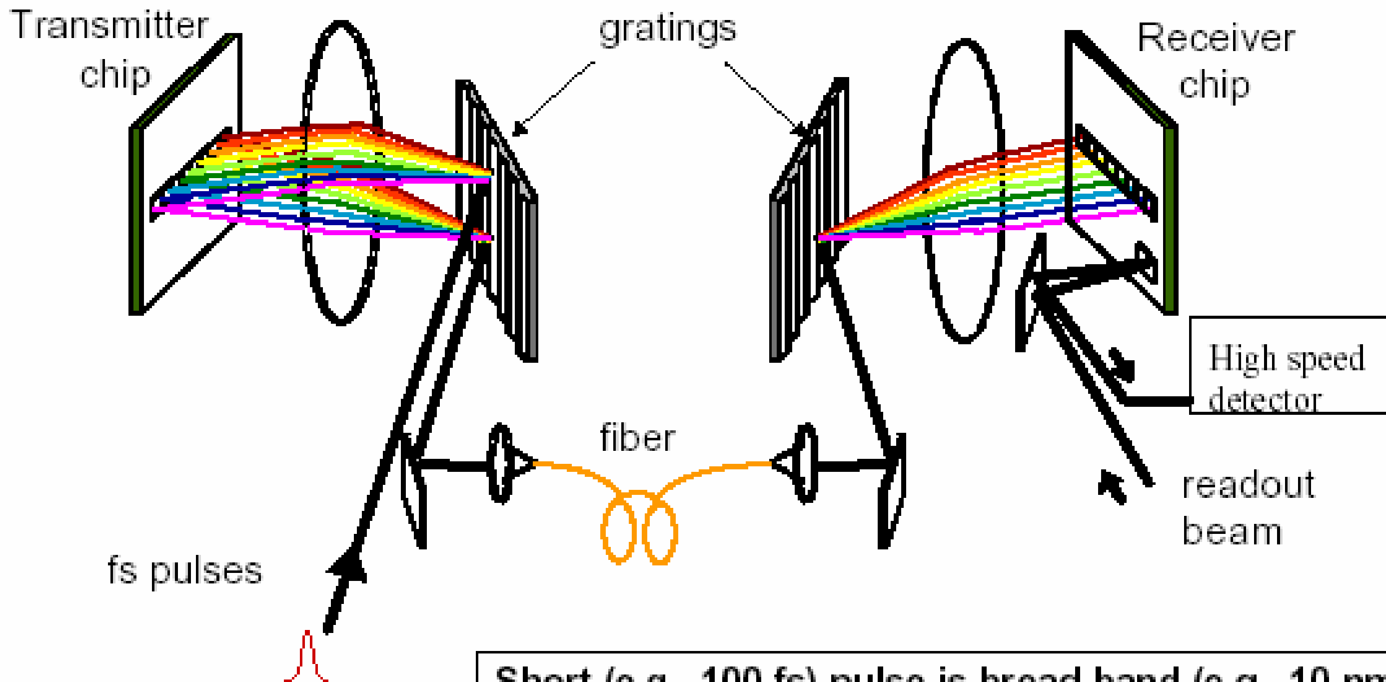
Hypothetical module with

- silicon CMOS chips with optoelectronic devices and lenslet arrays
- electrical wiring and waveguide layer
- massively parallel free space optics within and between chips
- flexible fiber connections off the module

T. J. Drabik and D. A. B. Miller, Stanford

Source: David Miller, Stanford

WDM to the Chip



Short (e.g., 100 fs) pulse is broad band (e.g., 10 nm wavelength range) source

- **spread wavelengths over reflective modulator array**
  - **send reflected signals over single fiber to receiver array**

Multiple channel interconnect with single fiber and single laser

D. Agarwal, G. A. Keeler, B. E. Nelson, and D. A. B. Miller, LEOS Annual Meeting, San Francisco, CA (November 8-11, 1999). Paper ThT4.

Source: David Miller, Stanford

More Information (from SURA OpCook)

- **SURA Optical Network Cookbook – www.opcook.sc.edu**
- **Books**
  - · "Fiber-Optic Communications Technology" by Lowell L. Scheiner & Djafar K. Mynbaev, Prentice Hall; ISBN: 0139620699
  - · "Understanding Fiber Optics" by Jeff Hecht, Prentice Hall; ISBN: 0130278289
  - · "Optical Networking Crash Course" by Steven Shepard, McGraw-Hill Professional; ISBN: 0071372083
- **Web**
  - Light Reading Beginner's Guides
    - *www.lightreading.com/*
  - Corning Cable Systems Basic Principles of Fiber Optics
    - *www.corning.com/ and www.corningcablesystems.com*
  - Illustrated Fiber Optics Dictionary
    - *www.fiber-optics.info/glossary-a.htm*
  - How Fiber Optics Works
    - *www.howstuffworks.com/fiber-optic.htm*
  - Introduction to Fiber Optics
    - *www.commspecial.com/fiberguide.htm*
  - Optics for kids
    - *www.opticalres.com/kidoptx.html*