NGI Program at DARPA

PITAC Review

Oct. 6, 1999

Mari Maeda

DARPA

mmaeda@darpa.mil

Outline

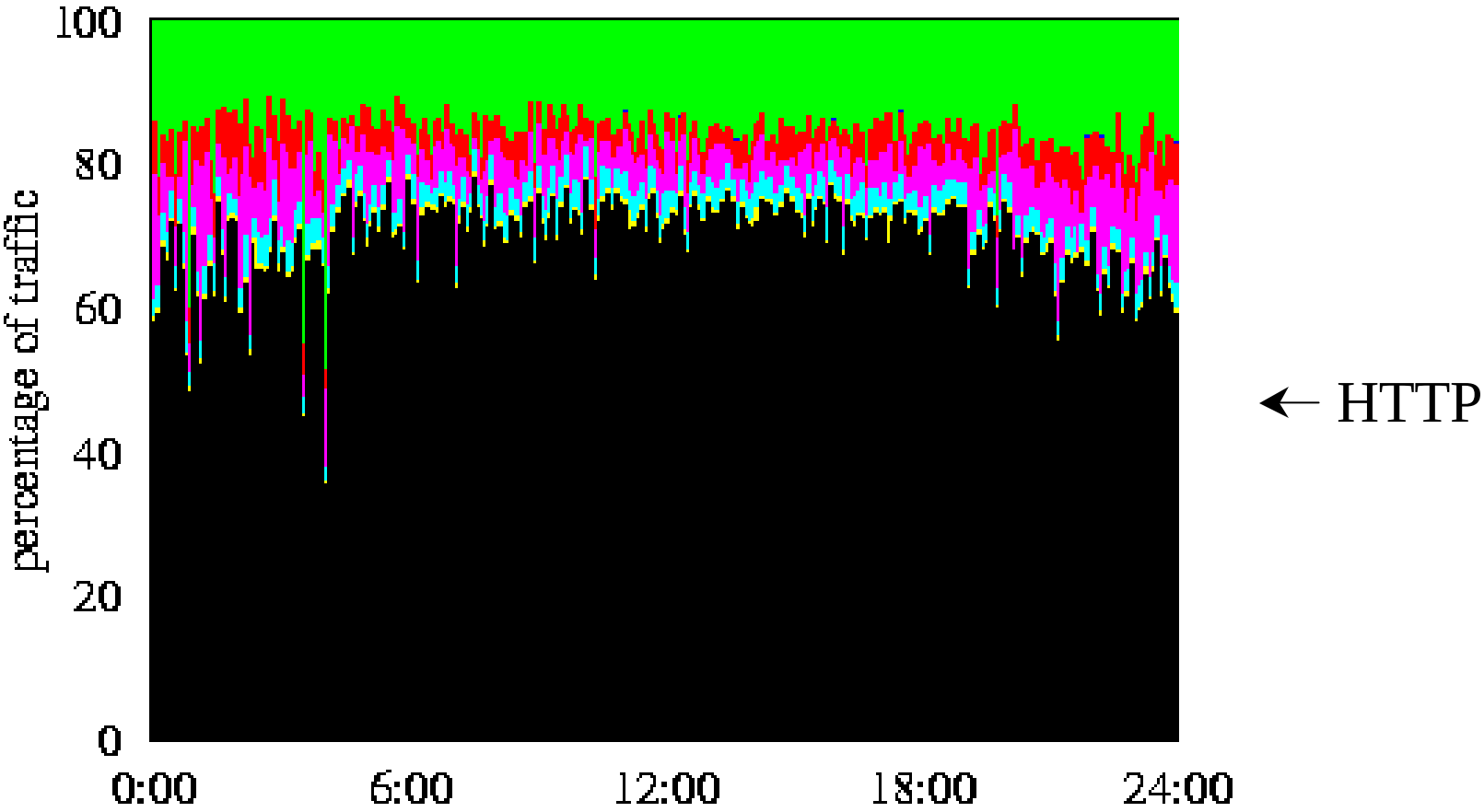
- Program Objectives
- SuperNet Technology
- Network Engineering
- SuperNet Testbed Deployment
- New Applications Development
- Tech Transfer

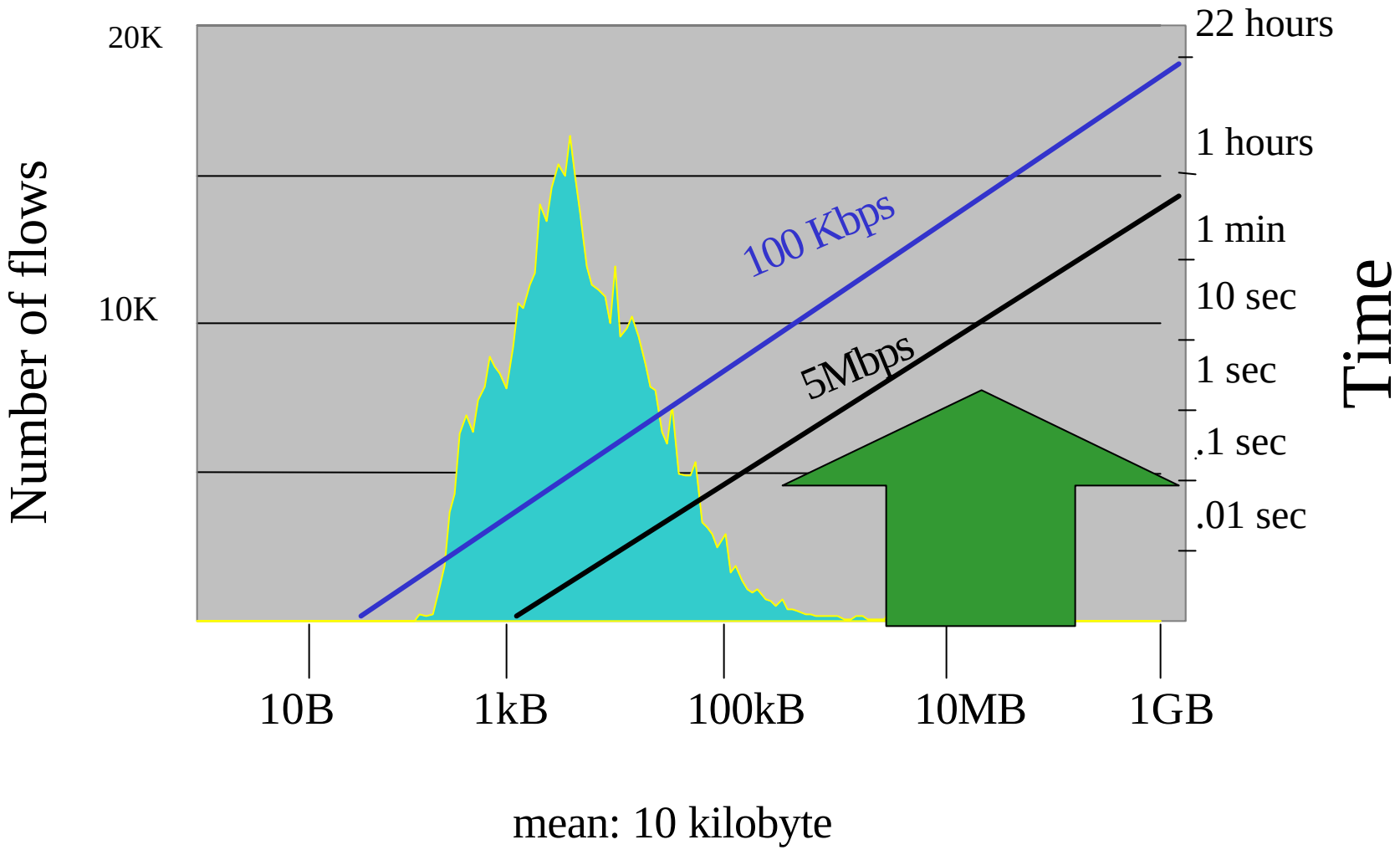
Information Technology Programs

Related programs (with networking component) supported by DARPA

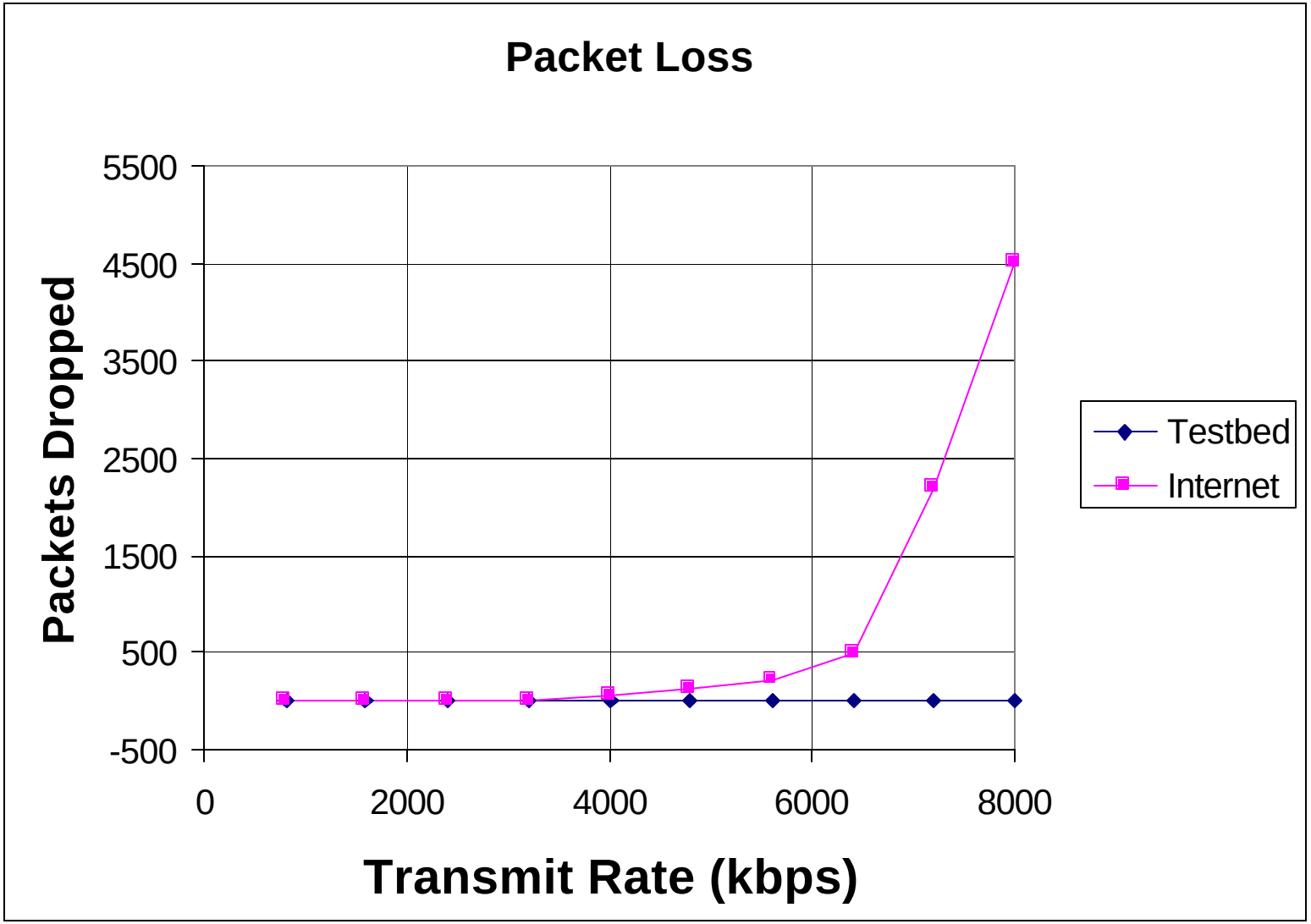
- Tolerant Networks
- Inherent Survivability
- Simulations and Modeling
- SensIT
- Quorum
- Active Networks

Today's Internet Traffic Makeup



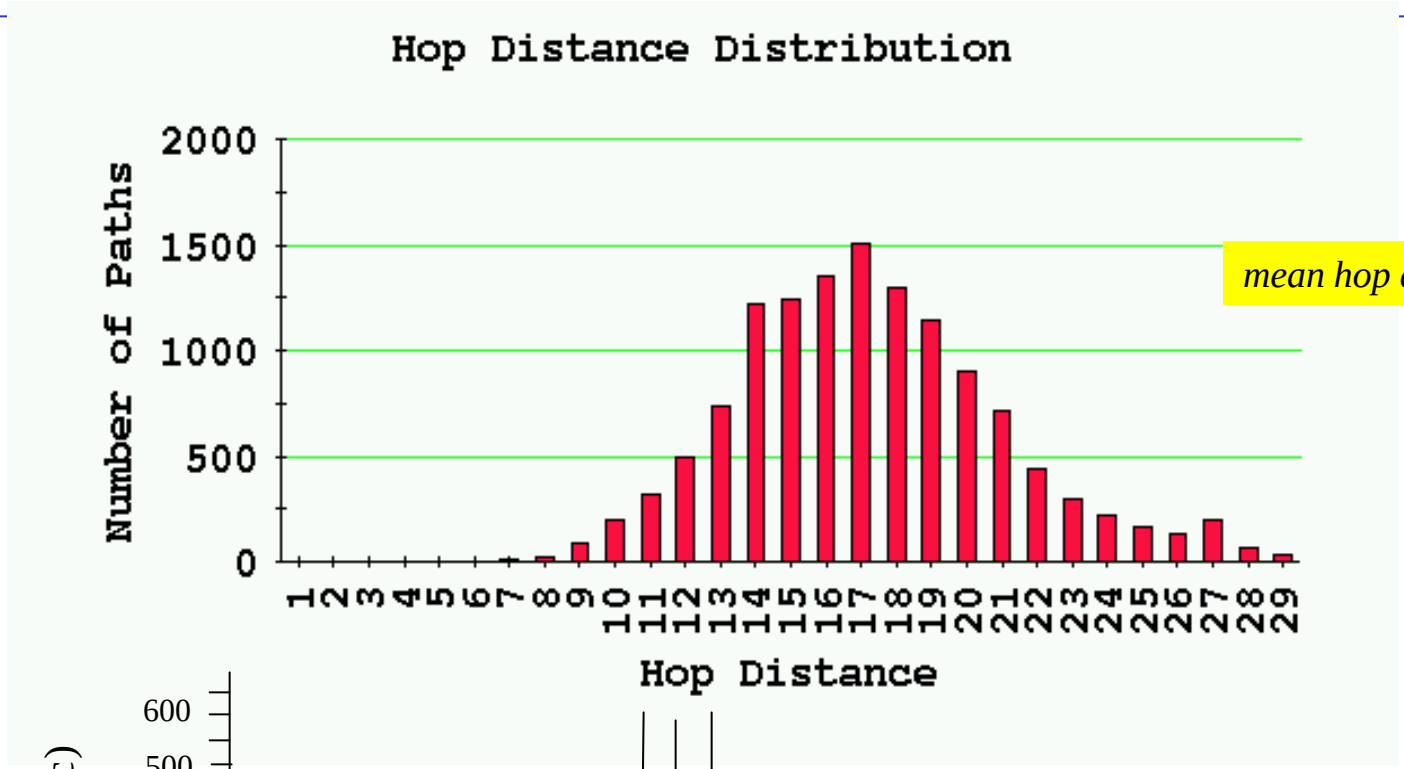

20K
10K
Number of flows
22 hours
1 hours
1 min
10 sec
1 sec
.1 sec
.01 sec
Time
100 Kbps
5Mbps
10B
1kB
100kB
10MB
1GB
mean: 10 kilobyte

Today's Internet Packet Loss vs Transmission Rates

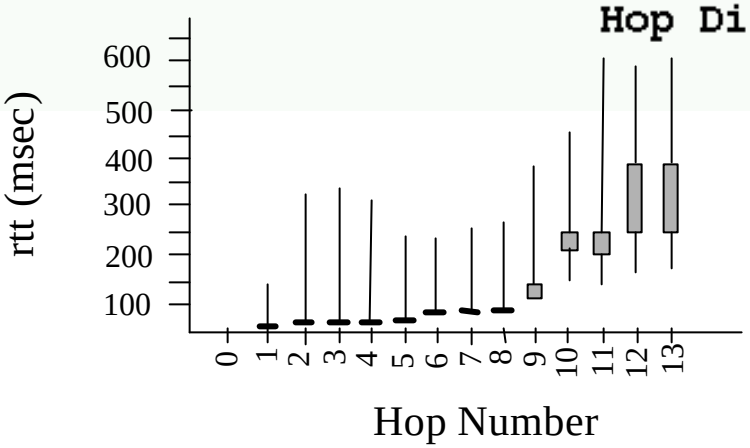


Cambridge to L.A.

Scaling the Internet



mean hop distance = 16



increased:
- loss probab.
- delay
- delay variation
- security compromise

Goals

To enable ultra-high bandwidth on demand over national networks guaranteed over the shared infrastructure

SuperNet

- *Simplified protocol layering - IP over dynamic Optical Network.*
- *End-to-end perf : regional & local access network.*

Create tools that greatly automate planning and management functions enablingnetworks to grow while limiting the cost and complexity of network management and control

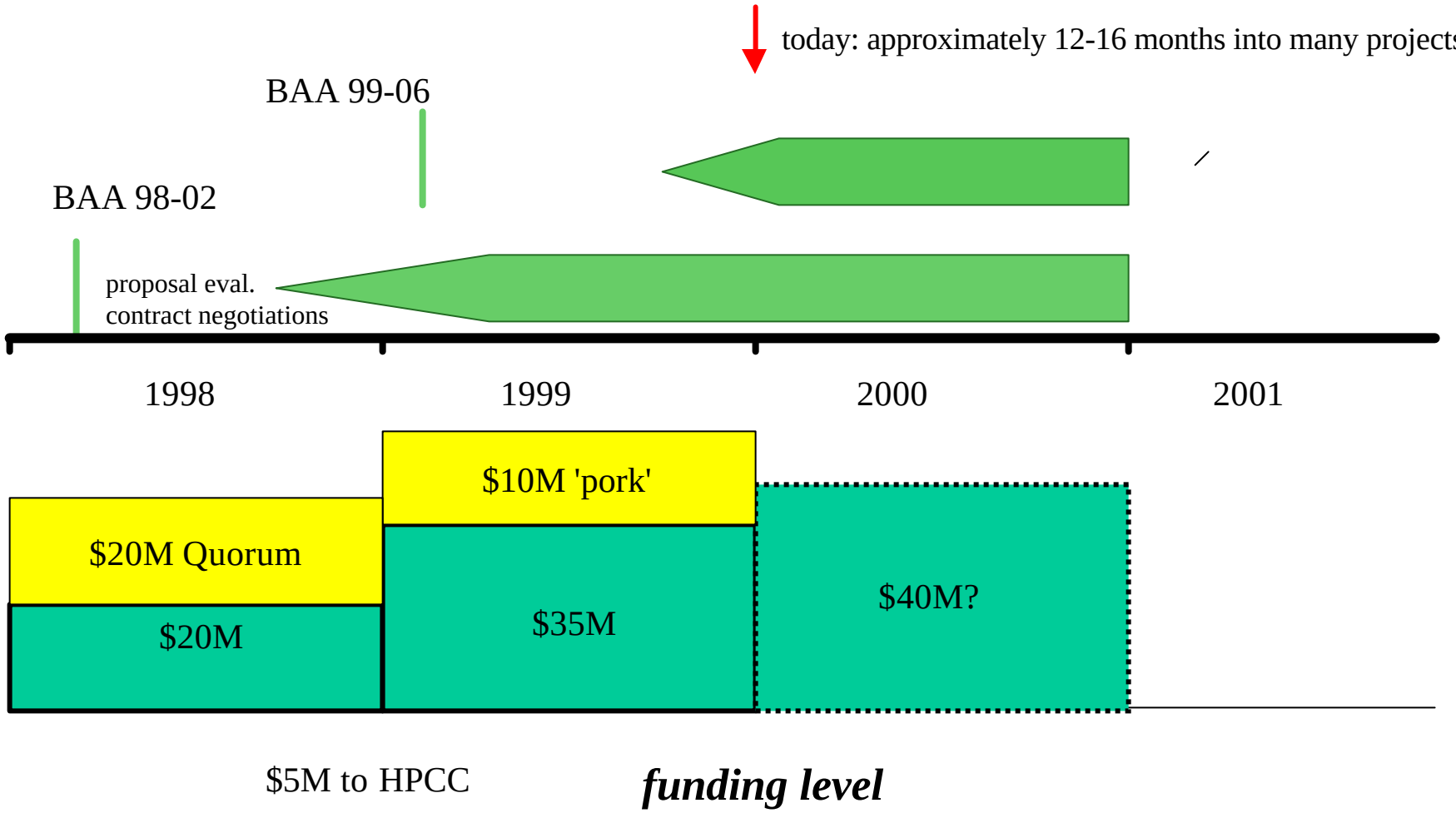
Network Engineering

- *Adaptive network management and control software*
- *Large-scale network monitoring/analysis/visualization tools*

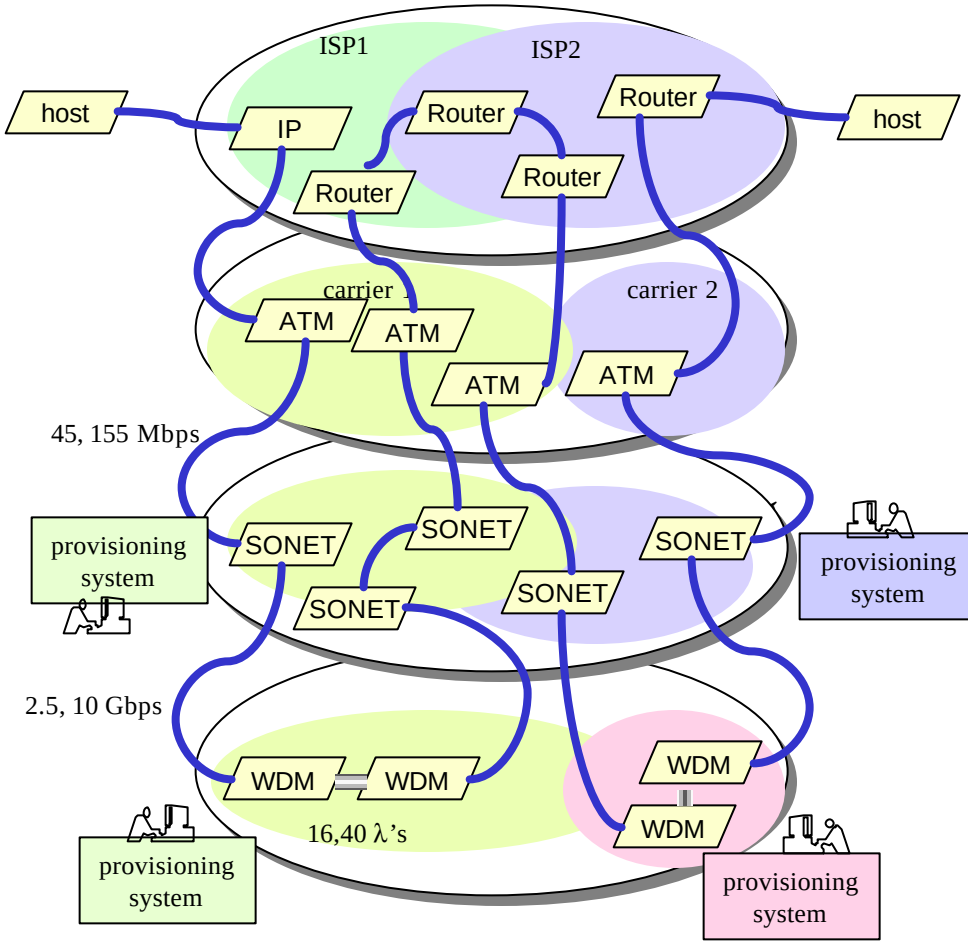
Develop, test, deploy applications requiring gigabit end to end throughput

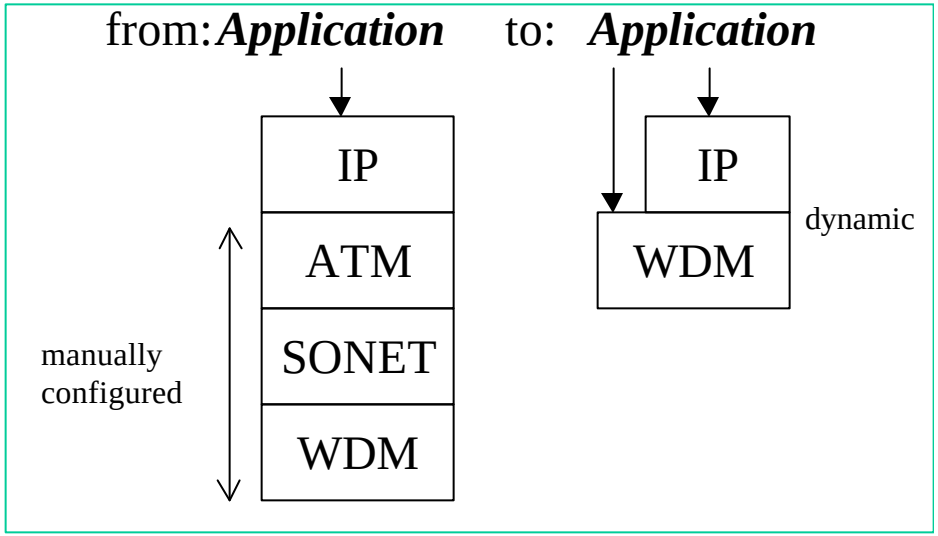
Apps

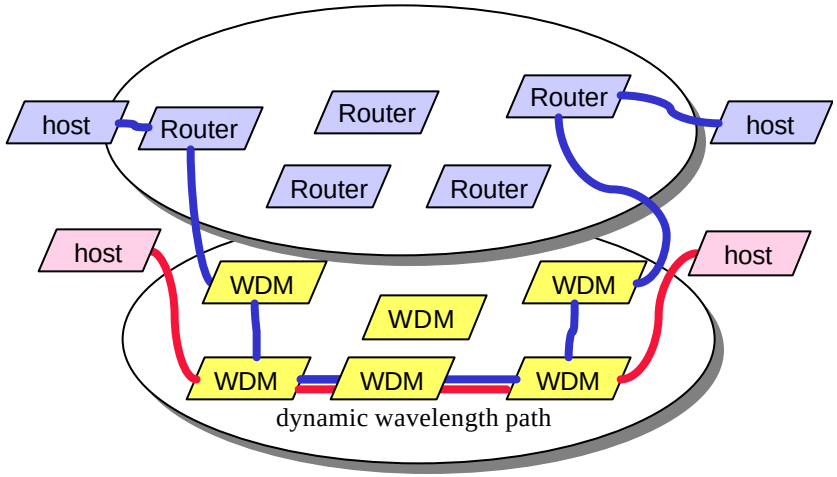
Program Time Line



SuperNet: Simplifying Protocol Stacks







Optical Network

“Optical Networking” trends

- 1996 1Tbps transmission demonstration using WDM
- 1999 3 Tbps transmission over limited span but 1 Tbps transmission over 10,000 km unregenerated span
- deployed ‘optical networks’ uses high capacity WDM links on a point to point spans in core networks
- emergence of startups based on ‘reconfigurable’ (provisionable) optical networking nodes supporting optical layer network restoration

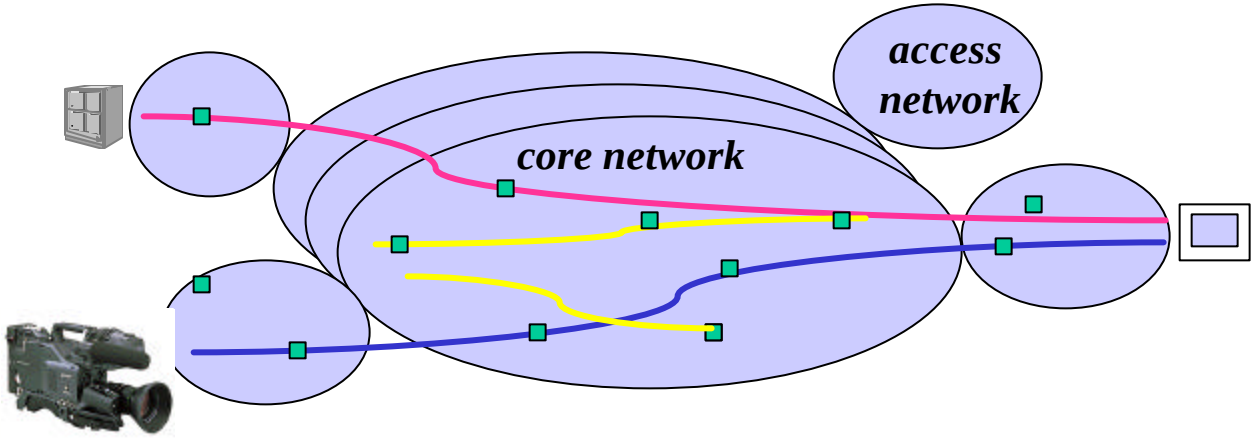
SuperNet Goals

To enable ultra-high bandwidth on demand over national networks, guaranteed over the shared infrastructure

Target: Multi-Gbps end to end

Approach:

- *Streamlined networking protocol stacks*
- *Stress end-to-end architecture and performance*
- *Technologies for regional, metropolitan, local area network.*
- *Take advantage of dynamically controlled optical layer*



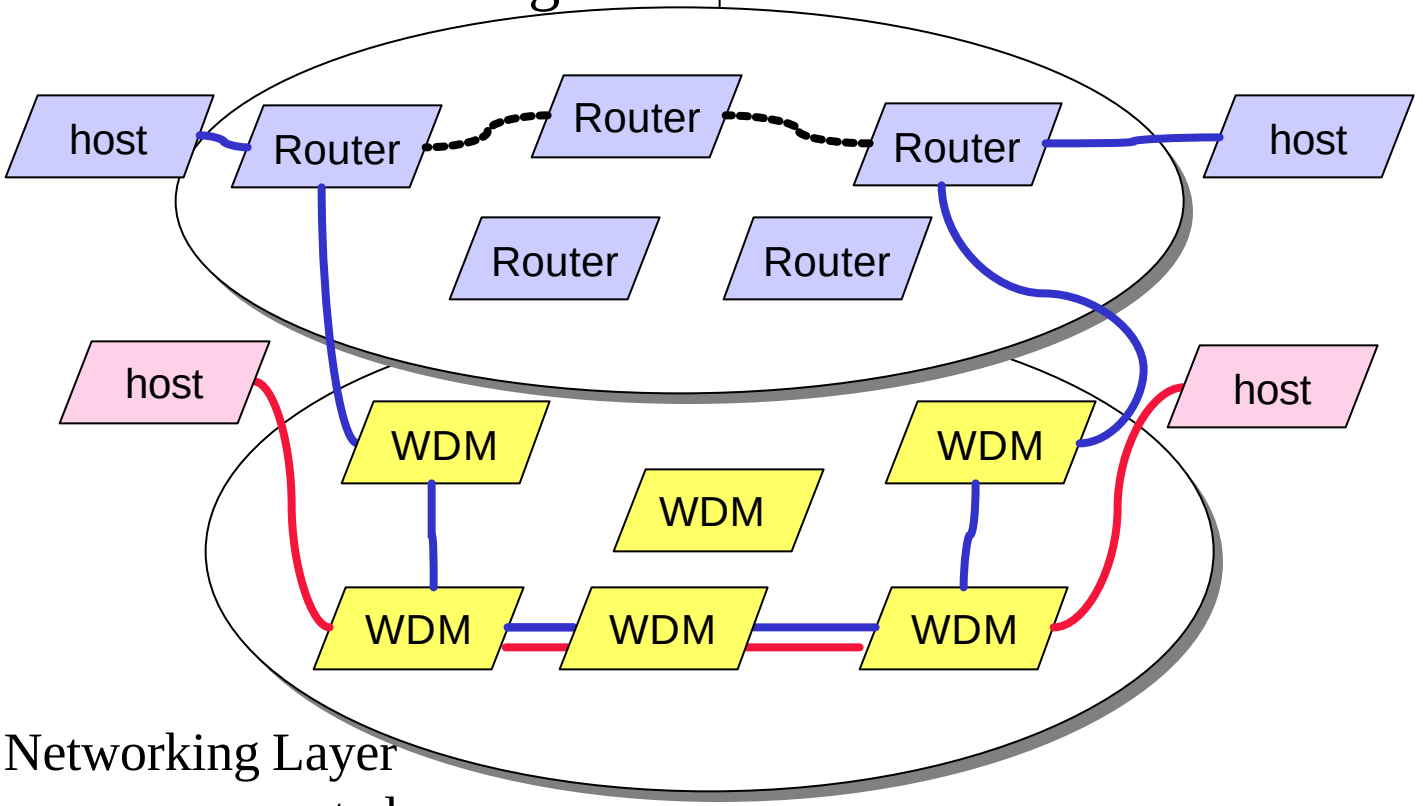
End-node performance

- Raw speed is not the issue:
  - TCP can break a gigabit/second.
    - Over Myrinet (1.2Gb/s), TCP goes 1.17 Gb/s.
    - Over Onramp technology, TCP goes .97 Gb/s for 150 km.
- Issues:
  - Dynamics of TCP window adaptation.
    - Congestion control at Gb/s may not work well with current schemes.
    - How to get started.
    - Need to change speed fast. Flow switching.
  - Asymmetric capacity in flow switching.
- Goal: write the next chapter on TCP implementation.

Optical Networking

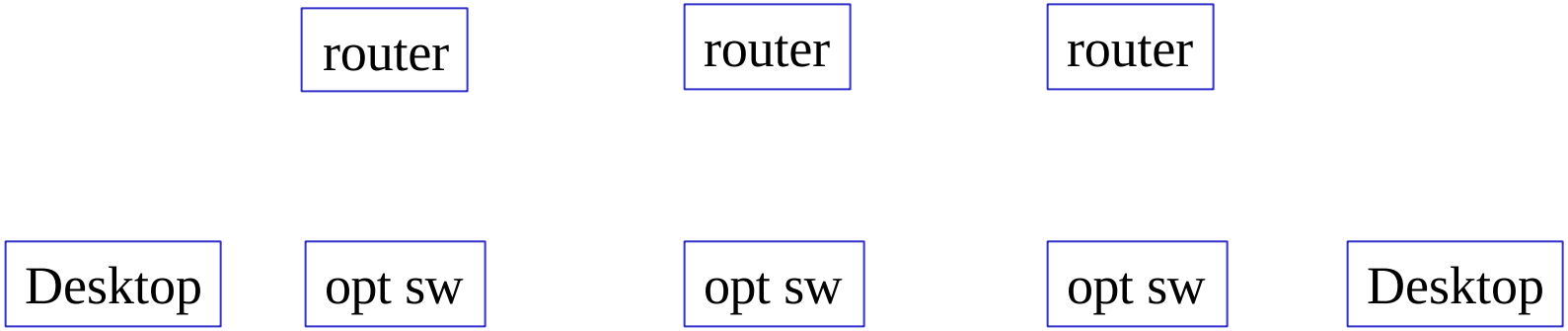
- •WDM based router bypass
- •Optical Flow Switching -- based on aggregate traffic change
- •Host-triggered path setup -- holding times ~>seconds
- •Optical burst switch -- holding times ~>microseconds

end-to-end perf, flexibility



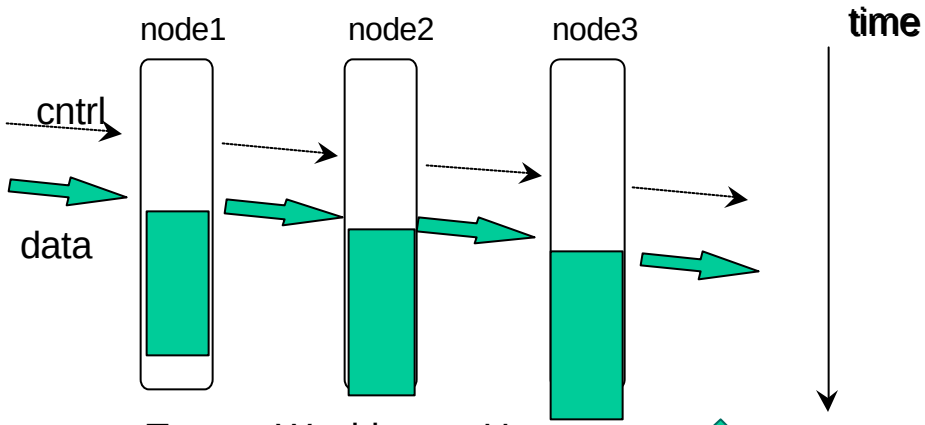
Dynamic Optical Networking Layer
transparenet, opaque, or regenerated

Optical Flow Switch (Optical Bypass)



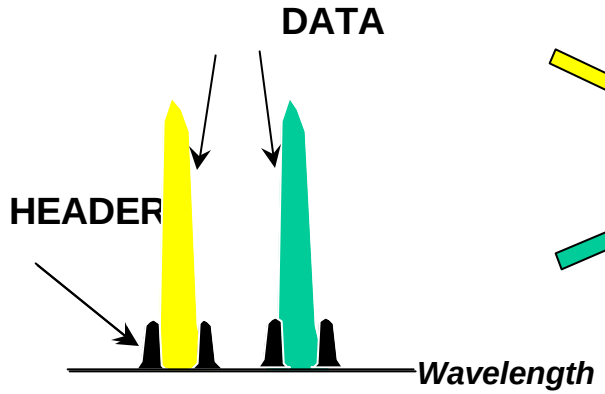
Optical Label Switching

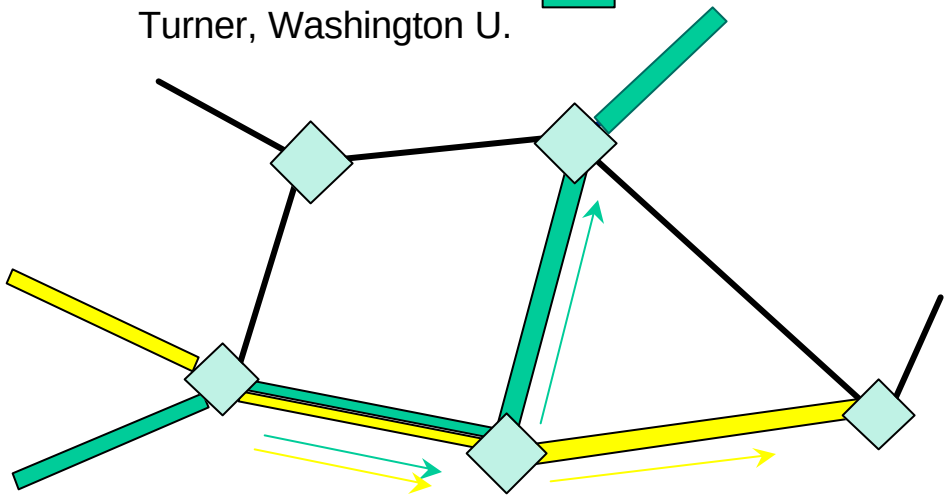
Optical Burst Switch



Turner, Washington U.

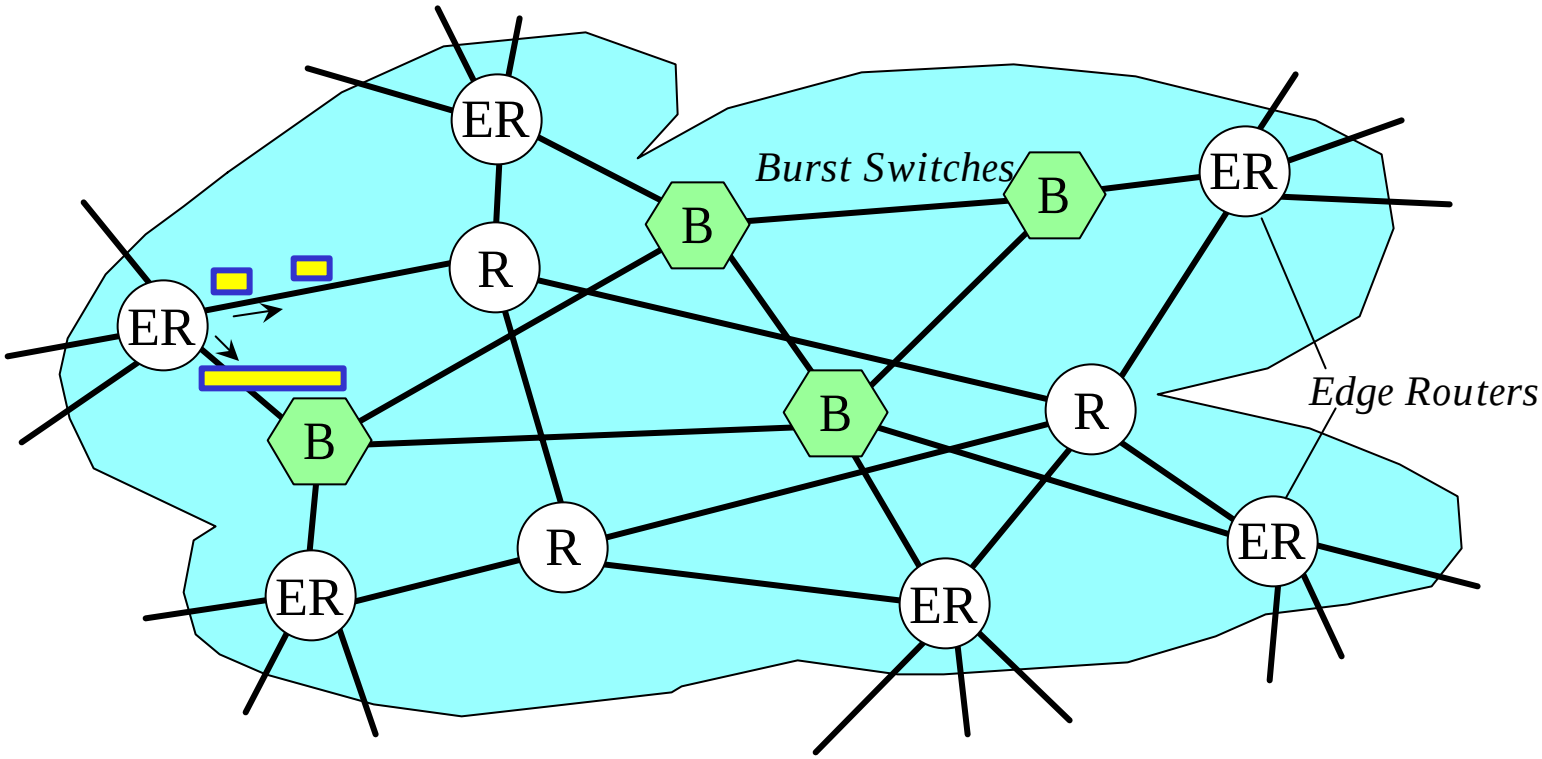
Optical Label Switching





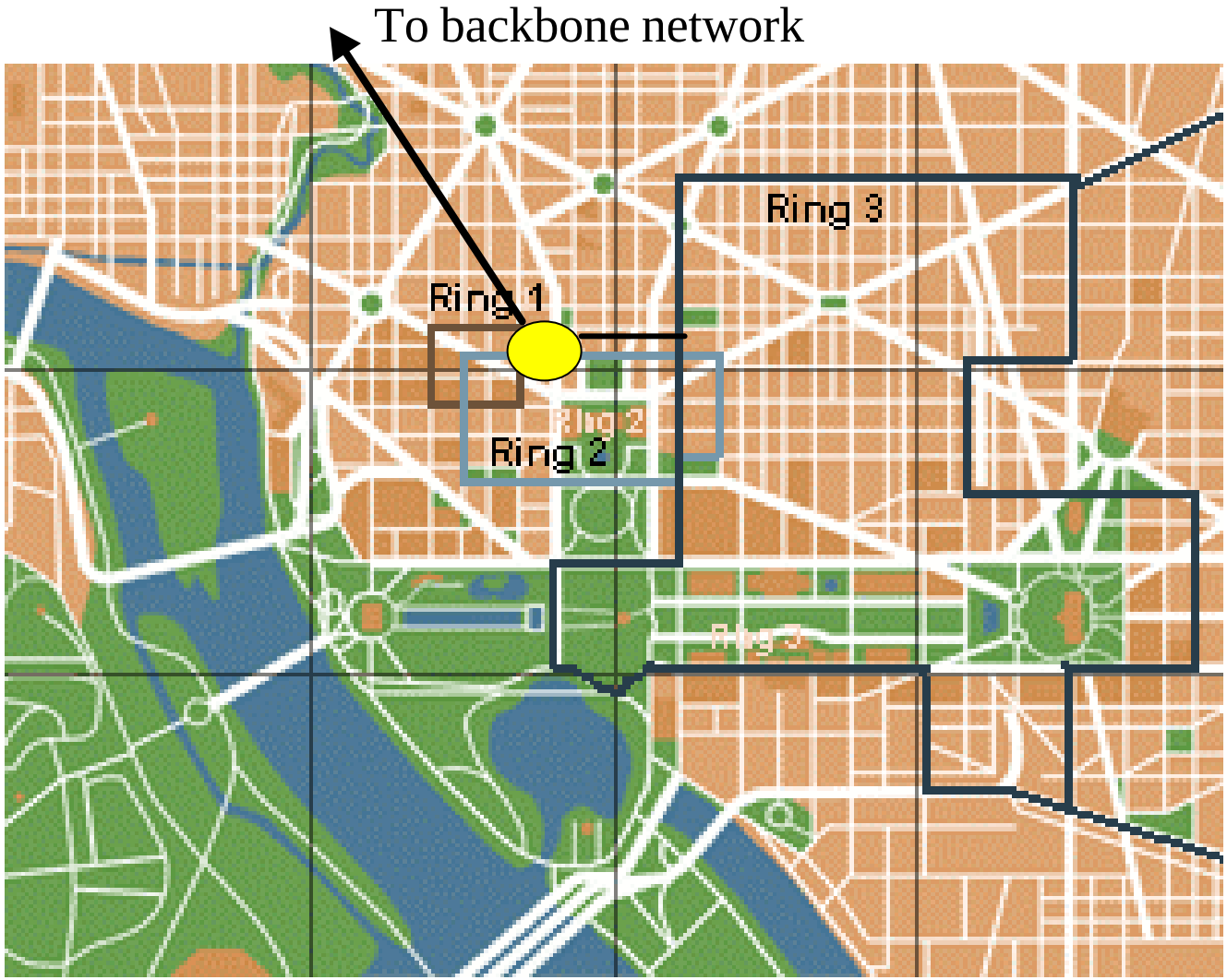
Telcordia.

Inserting Burst Switches into Data Network



- Use packet classification in *Edge Routers* to separate packets by length.
  - long packets (ftp, http replies) to *Burst Switches*, short packets (acks, http requests, DNS lookups) to routers.
- Long packets usually part of longer data transfers.
  - allows burst switch interfaces to assemble larger bursts for efficient transmission
- Study burst formation using trace-based simulation & prototype if promising.

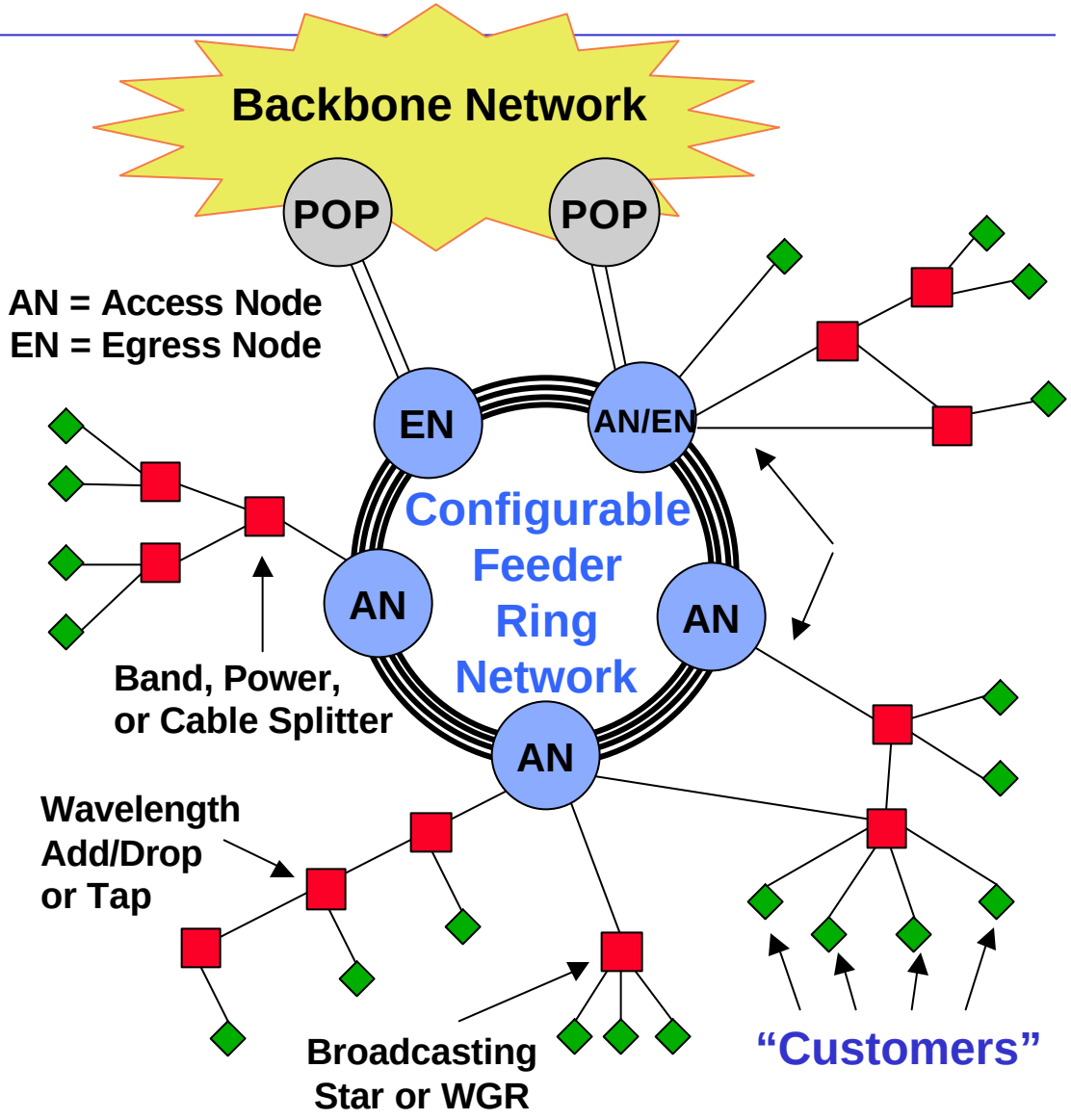
Access & Distribution Network Architecture



- service flexibility
- cost-effective architecture

ONRAMP
MIT, LL,AT&T,
JDS,Bay,Cabletron

ONRAMP Testbed



- Regional Access Network Architecture (10-1000 sq miles)

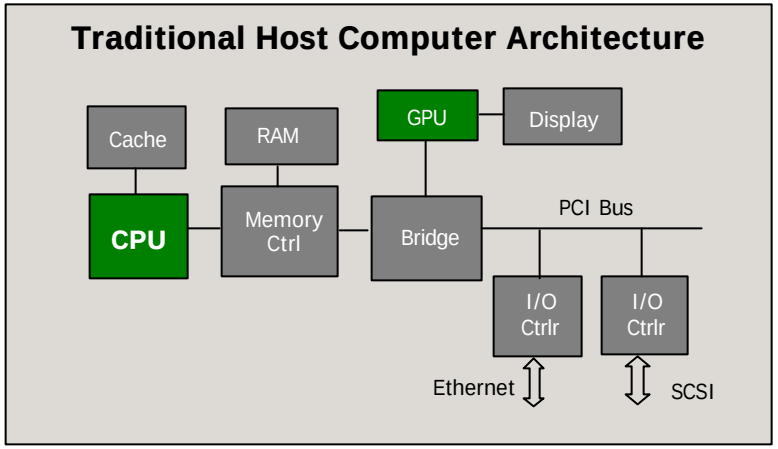
- Feeder Ring Network
  - -multi-fiber WDM ring
  - -reconfigurable Access Nodes
  - -full optical restoration

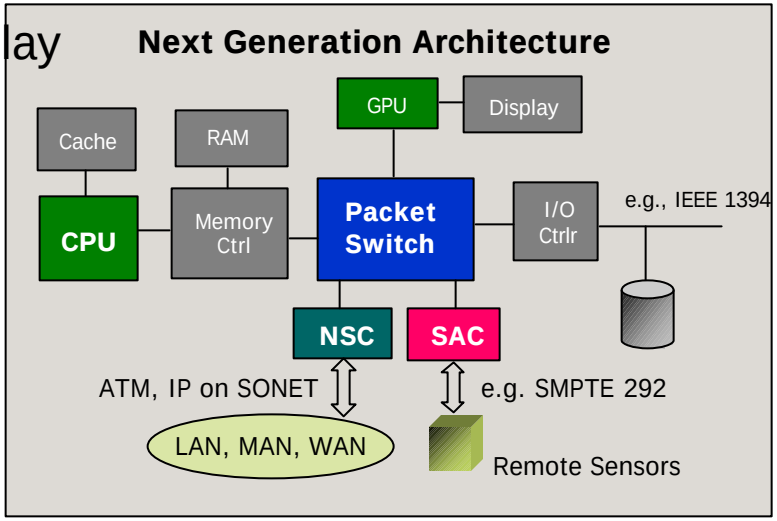
- Distribution Network
  - -cost sensitivity
  - -passive, transparent WDM
  - -tree/bus/ring topology

- BW squandering to mitigate complexity?
- wavelength density in feeder vs distribution network?
- shared or routed wavelengths?
- optical bypass, MAC protocol
- push end-node performance

Gigabit/second Host Platform

- Enable gigabits (up to 10 Gbps) to the end user
- Cell/Packet switch replaces traditional bus with its bottleneck
- Two new adapter cards being designed to plug into host switching backplane
- Network Service Card
  - -offload many higher layer functions from host CPU, ATM & IP,
  - -bursty & streams,
- Sensor Adaptor Card
  - - multigigabit (bitrate agile)real-time stream from remote sensors to host for processing, storage, display





Universal Network Access Module

Electronic IO modules at the core and the periphery of the network that can

-recognize and lock to the bit rate (bit-rate adaptability)

-recognize and handle different protocols (protocol agility)

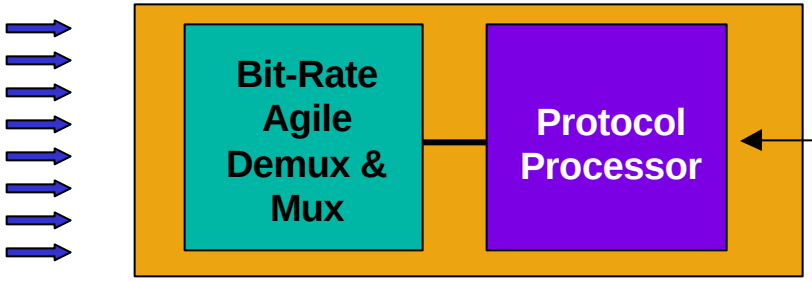
•Protocol/bit-rate transparent IO for dynamically reconfigurable or burst switched networks

•Automated network upgrades without replacing hw (lock-on or sw downloads)

- *•rapid deploymet*
- *•adapt to new types of sensors, CPE's*
- *•minmum inventory*

•Development & testing of new protocols

- •OC3/12/48c ATM / SONET
- •OC3/12/48c IP/SONET
- •Gigabit ethernet
- •SMPTE 25/292
- •IEEE 1394 (firewire)
- •G-Link
- •FDDI
- •Fibre Channel
- •"ngi protocol" e.g. IP/WDM



how economic?

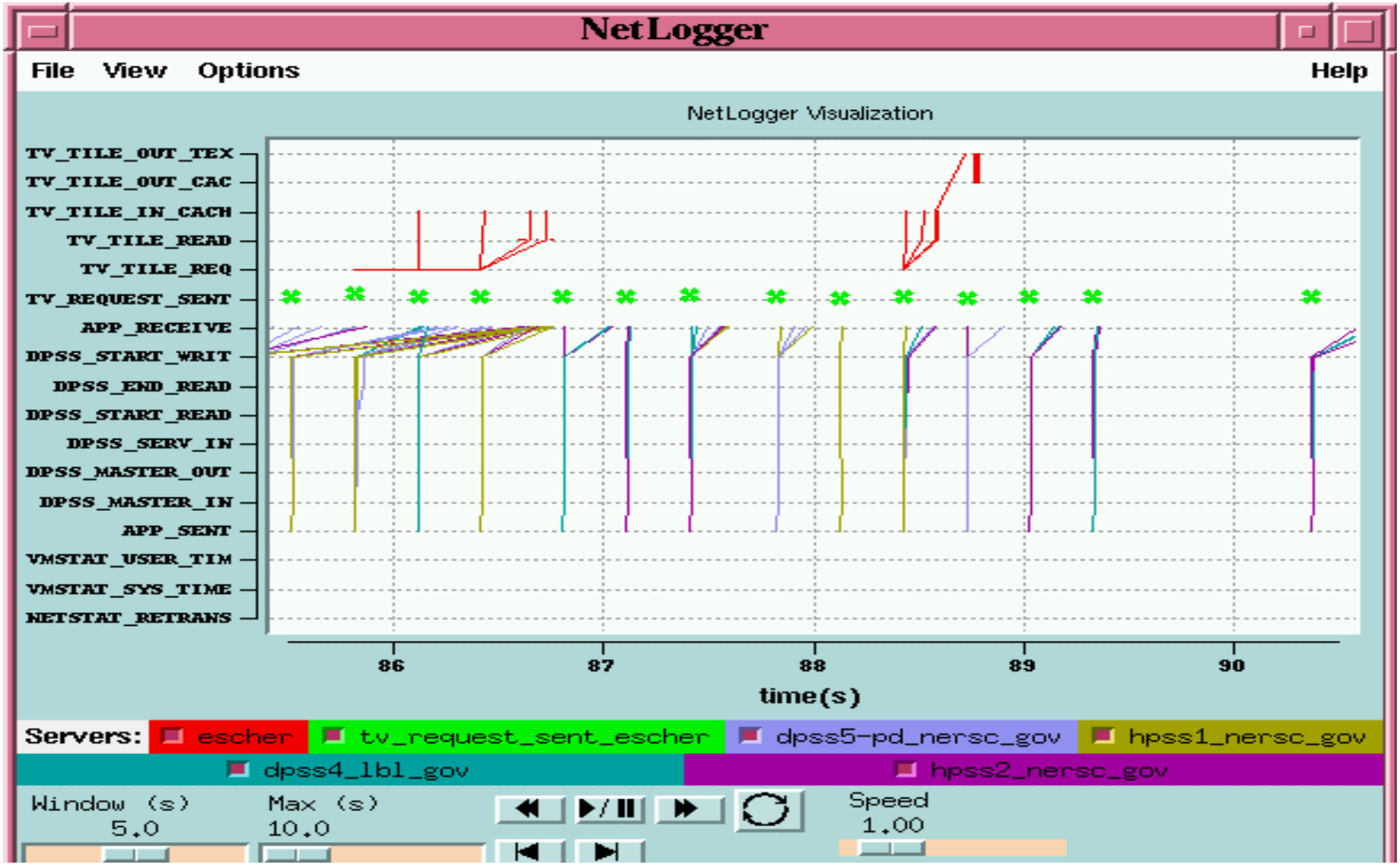
Tektronix: UNAS Project
Bellcore: VBRI Project
(see postdeadline paper)

Networked Applications Performance Analysis: NetLogger Toolkit

application, host, network

- Application to application performance analysis tool
- Identifies bottlenecks in path of data flow: application , operating system, network level (e.g. CPU load, interrupt rate, TCP retransmission, window size...)
- Post-hoc and real-time analysis
- Eveng Log Generation , Analysis and Visualization Tools (depict event points, load-line, lifeline)

NetLogger/NLV analysis of a TerraVision with DPSS



Network Engineering:
Network Monitoring, Analysis and Visualization

- Monitor and automate the discovery of the topology and traffic behavior of the Internet and future networks on a global scale.
- What makes this hard:
  - –no central authority
  - –scale (span and speed)
  - –capturing dynamic behavior
  - –visualization

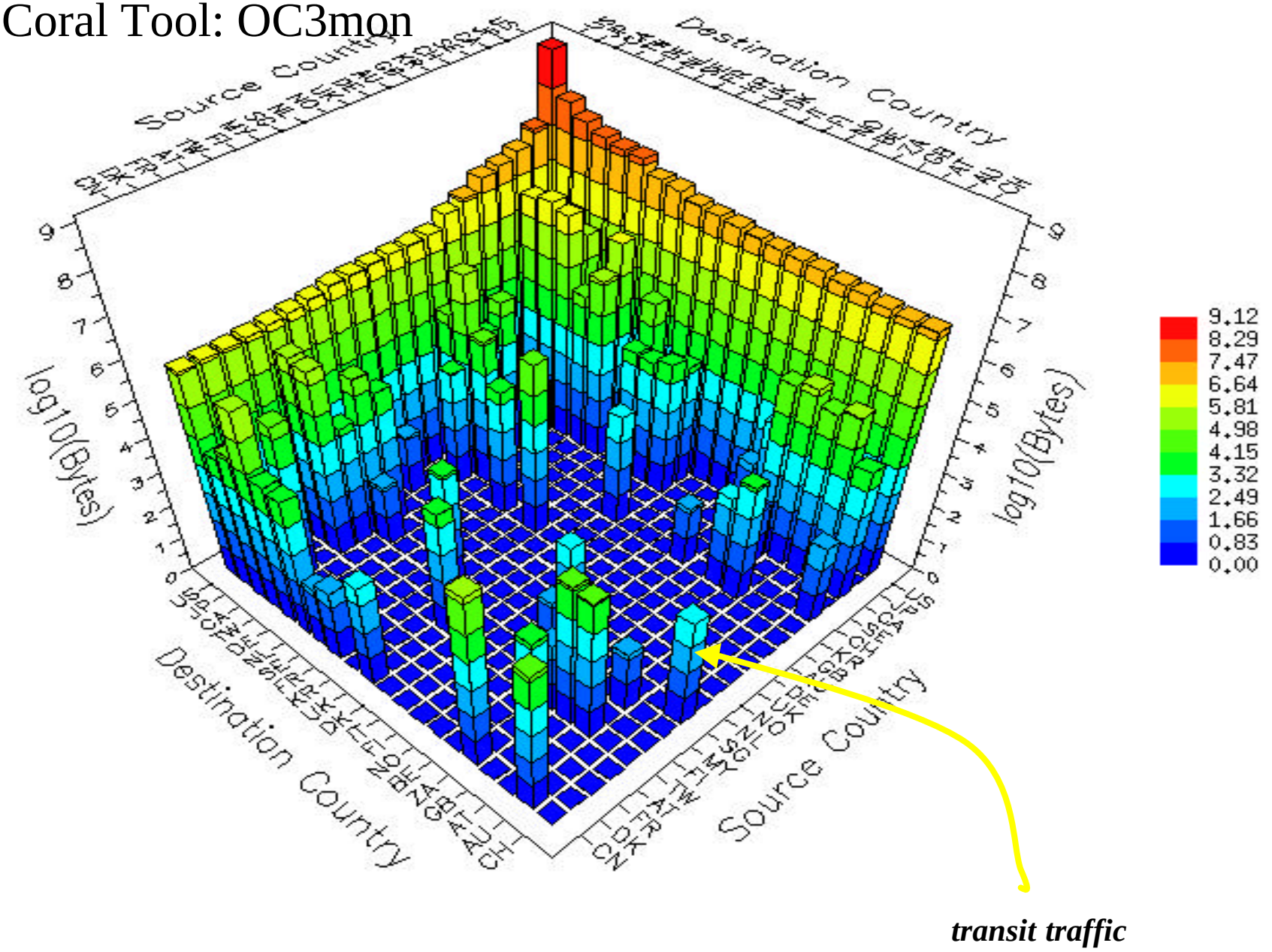
Tools :

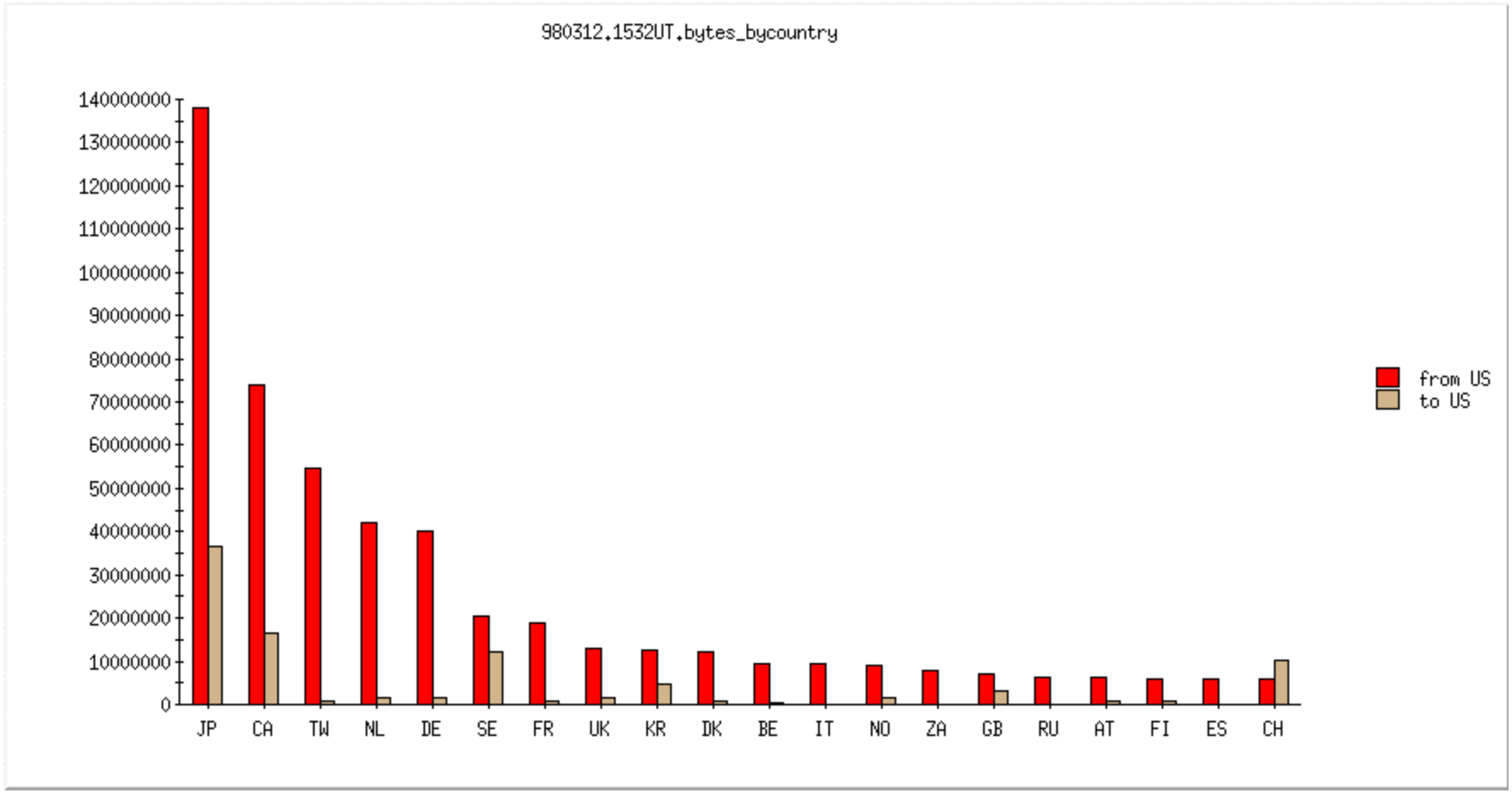
"*skitter*" (active measurements: performance, topology)

"*coral*" monitors (passive measurements over high speed links)

UCSD/CAIDA
(Cooperative Association for Internet Data Analysis)

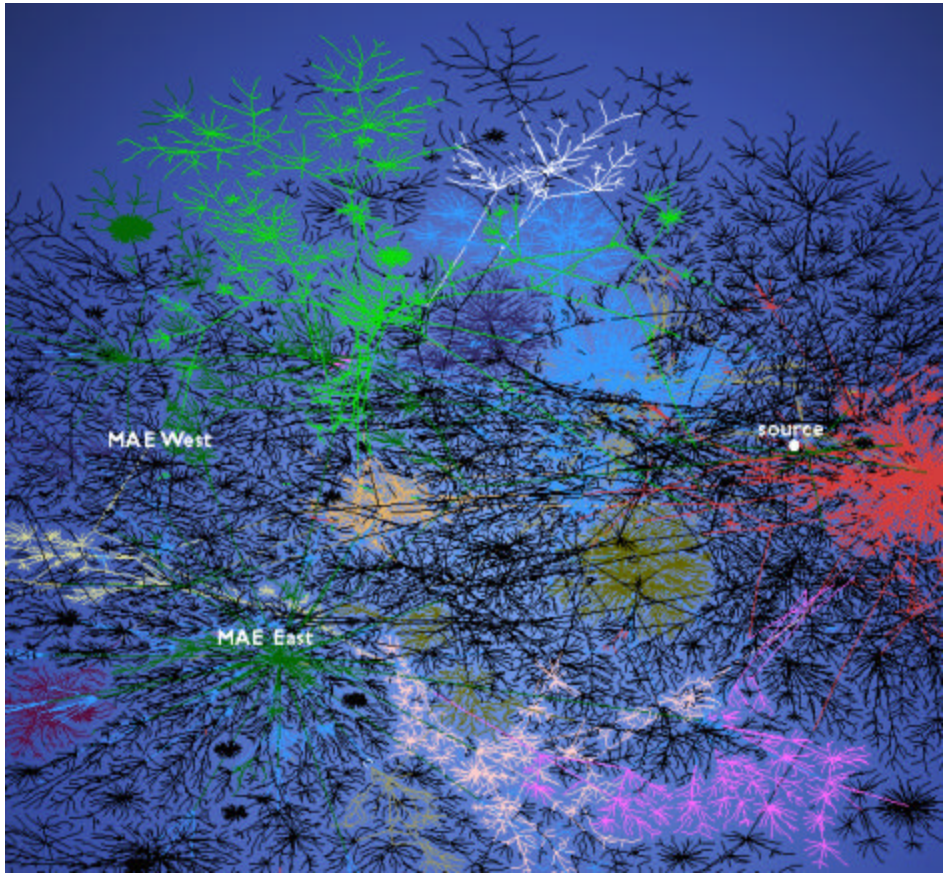
Coral Tool: OC3mon




980312.1532UT.bytes_bycountry
140000000
130000000
120000000
110000000
100000000
90000000
80000000
70000000
60000000
50000000
40000000
30000000
20000000
10000000
0
JP
CA
TW
NL
DE
SE
FR
UK
KR
DK
BE
IT
NO
ZA
GB
RU
AT
FI
ES
CH
from US
to US

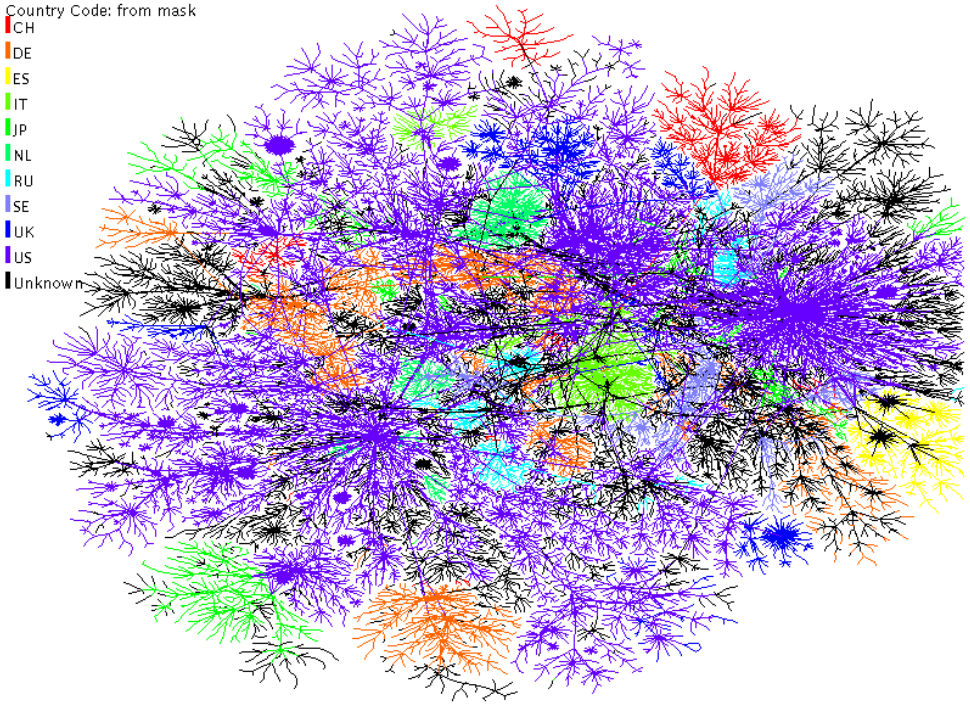
Network Tomography

- Network “Radar” :Global connectivity information
- Measure IP paths (“hops”) from source to MANY (~10^4) destinations
- Use 52 byte ICMP echo requests (every 30 min.) as probes
- Challenges:
  - pervasive measurement with minimal load on infrastructure
  - visualization



Visualizations of connectivity / topology

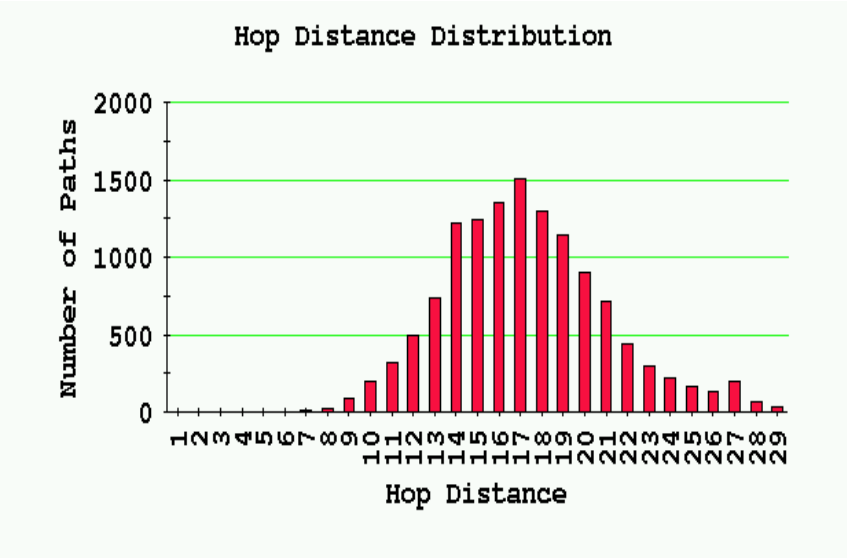
colored by country-level domain



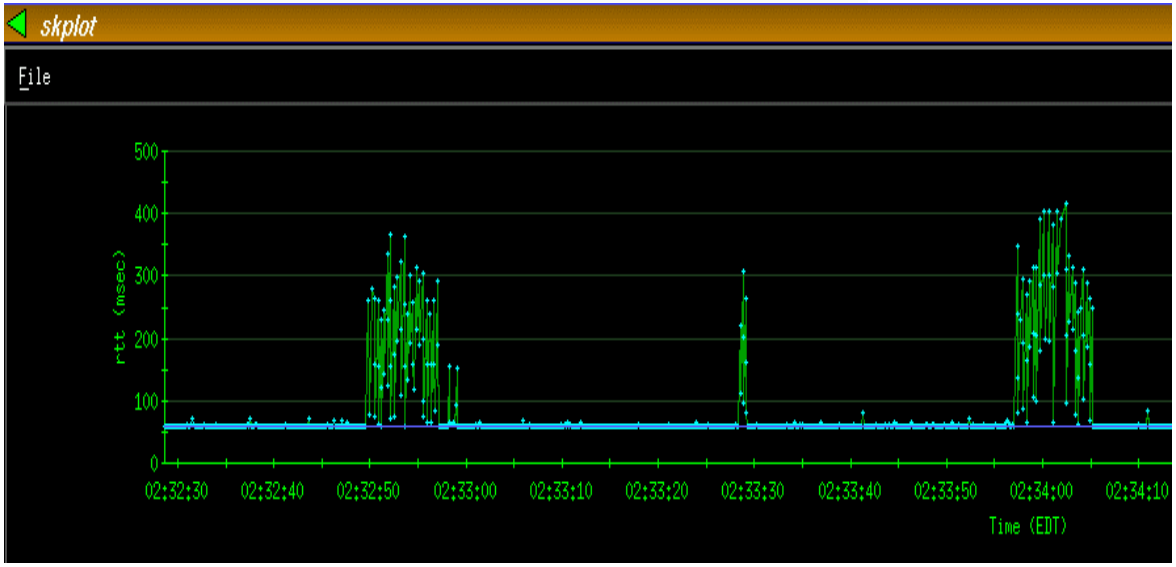
alternative visualization

Internet Tomography

hop count histogram

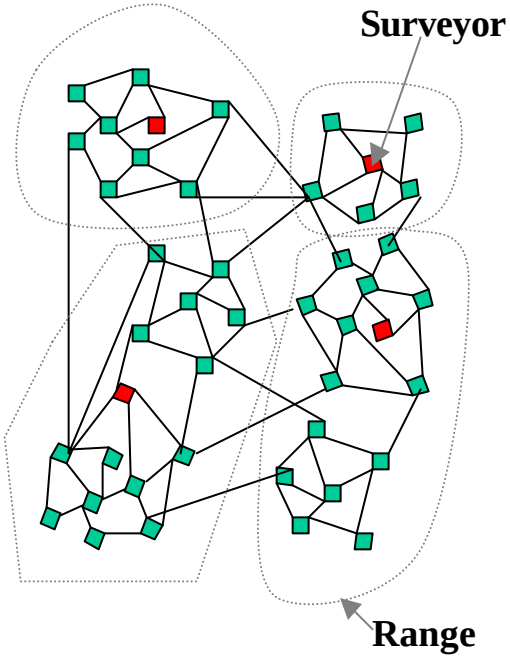


temporal behavior



Network Engineering:
Adaptive Network Management Project

Large-scale network fault isolation



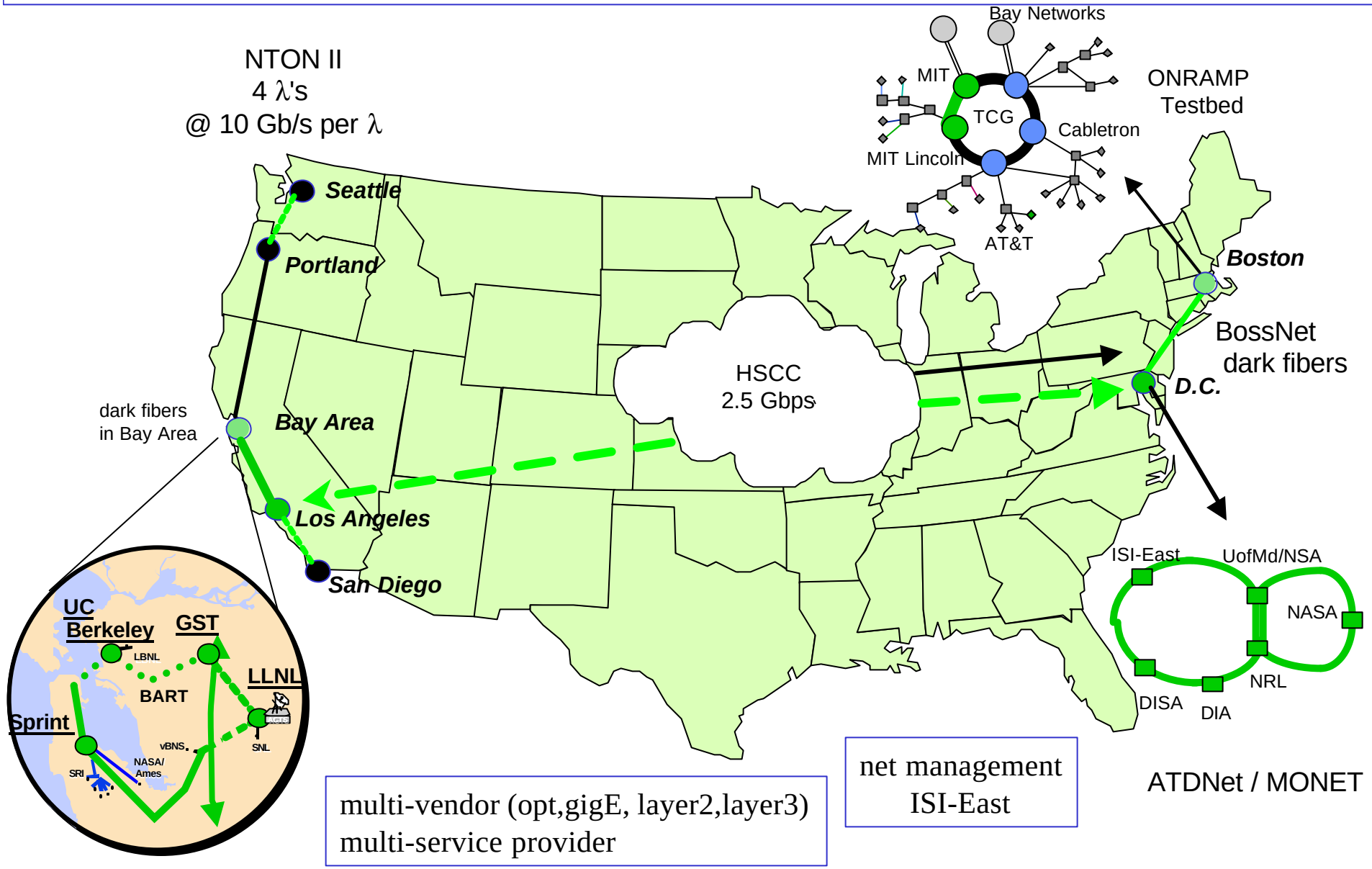
Self-configuring network monitors

- Surveyors map neighborhood
- They coordinate with other surveyors to adjust their ranges
- Careful multicast based self-organization
  - continuous range expansion
  - range description exchange
  - back off
- …eventually adapts to surveyor failure, network partitions

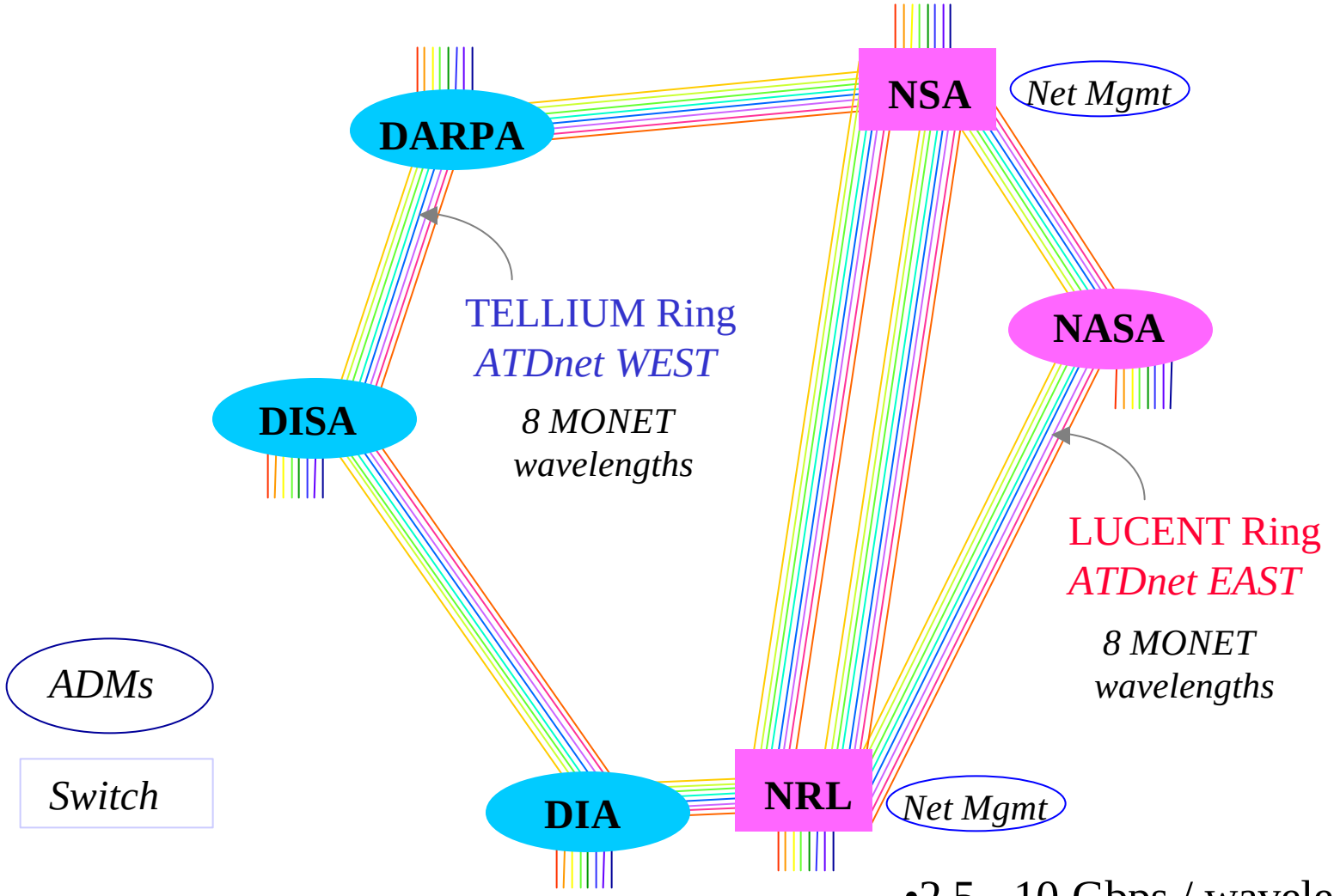
Adapts to network traffic & fault (link cut, node failure, congestion, network partition).

SCAN: UCSC & ISI

SUPERNET TESTBED (www.ngi-supernet.org)



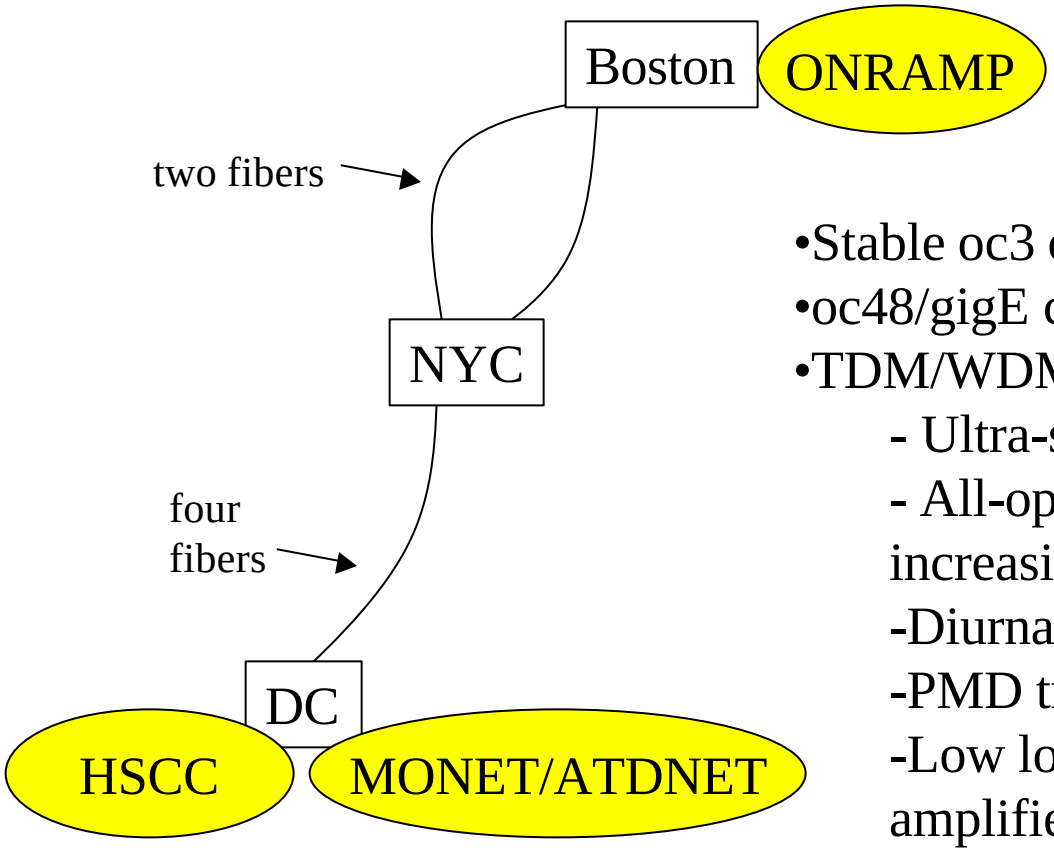
ATDNET-MONET TESTBED



- 2.5 - 10 Gbps / wavelength
- limited 'transparency'

BOSSNET Testbed

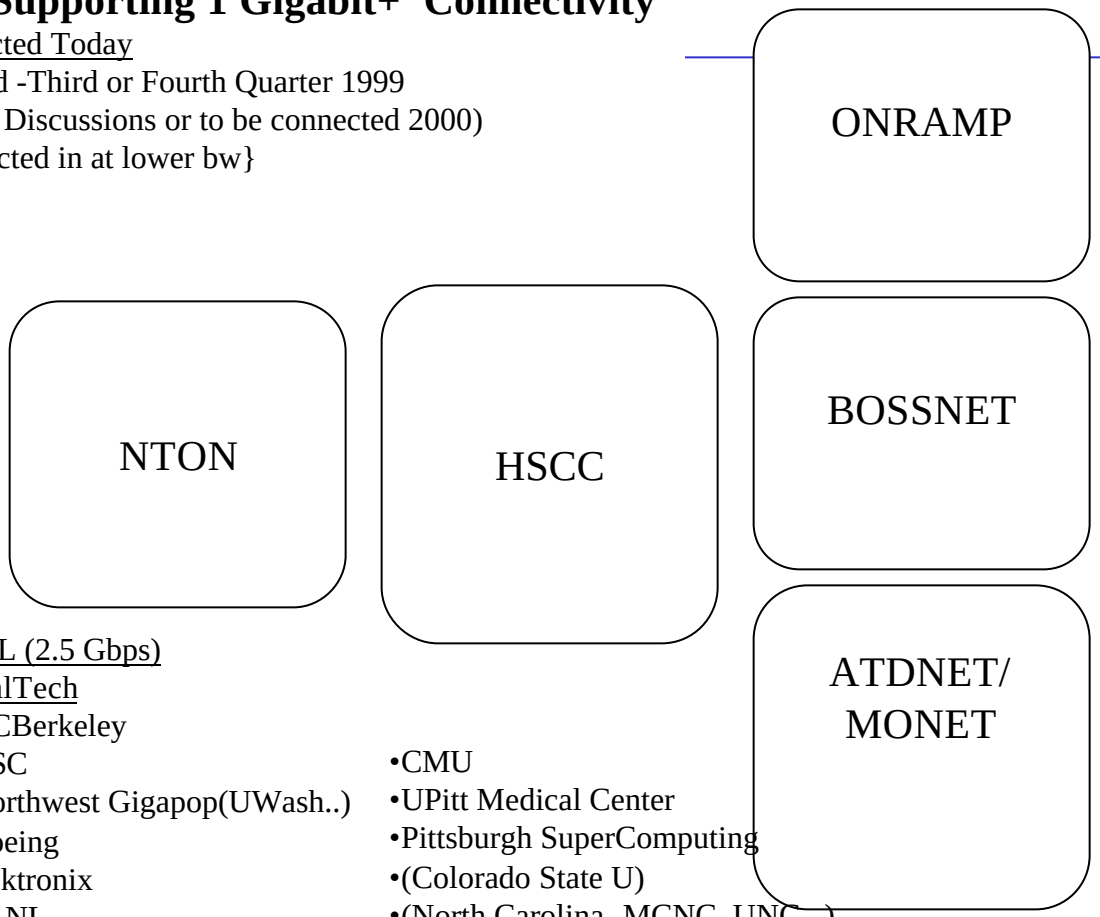
•four fibers along inland/coast rail routes between Washington DC and Boston
• multiple huts being populated with custom equipment (span length 40-100km)
•connection between HSCC, MONET/ATDNet, ONRAMP networks



•Stable oc3 channel
•oc48/gigE channel ⟹ for apps
•TDM/WDM Experiments over installed fiber:
- Ultra-short-pulse 100 Gbps transmission
- All-optical R^2 regeneration over increasing spans (incl. loopbacks)
-Diurnal clock recovery investigations
-PMD tracking/mitigation techniques
-Low loss window extensions (Raman amplifiers)

Sites Supporting 1 Gigabit+ Connectivity

•Connected Today
•Planned -Third or Fourth Quarter 1999
•(Under Discussions or to be connected 2000)
•{connected in at lower bw}



•MIT Campus
•MIT Lincoln Laboratory
•Bay Networks
•(Cabletron)
•(AT&T/TCG)
•(Harvard)

•(Drexel)
•(U of Penn)
•(Johns Hopkins MedicalSchool)
•(UMDNJ)

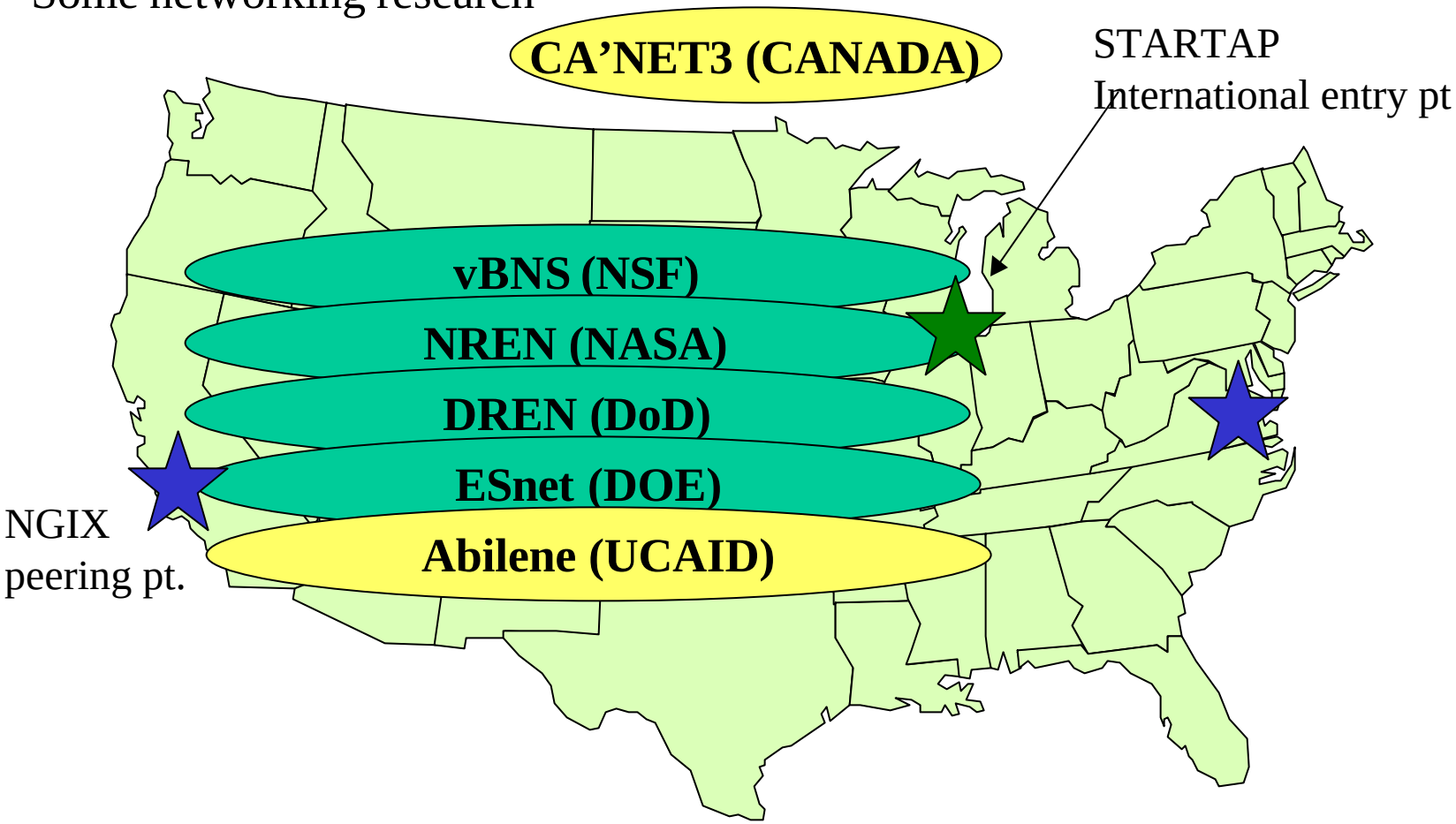
•JPL (2.5 Gbps)
•CalTech
•UCBerkeley
•USC
•Northwest Gigapop(UWash..)
•Boeing
•Tektronix
•LLNL
•Sandia NL
•SLAC
•SDSC
•Spawar
•(Microsoft)
•{Silicon Valley Test Track - Sprint, sun, sgi, xerox park ..}

•CMU
•UPitt Medical Center
•Pittsburgh SuperComputing
•(Colorado State U)
•(North Carolina -MCNC, UNC ..)
•(NYC)

•ISI -East/DARPA (8λ)
•NRL
•NASA
•NSA-Uof MD
•DISA
•DIA
•(CNRI)
•{NIST}
•{NIH}

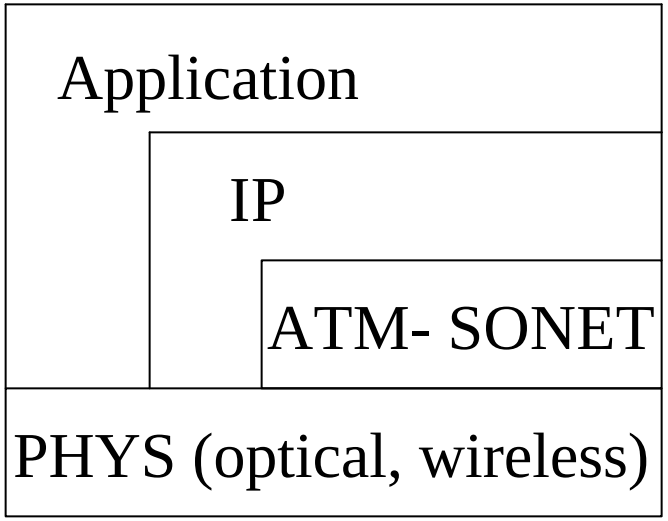
Other NGI Testbeds

- Stable applications focused testbed connecting primarily for universities, govt labs.
- Most built on carrier service, POS or IP/ATM ,backbone speeds OC3-OC48
- Some networking research

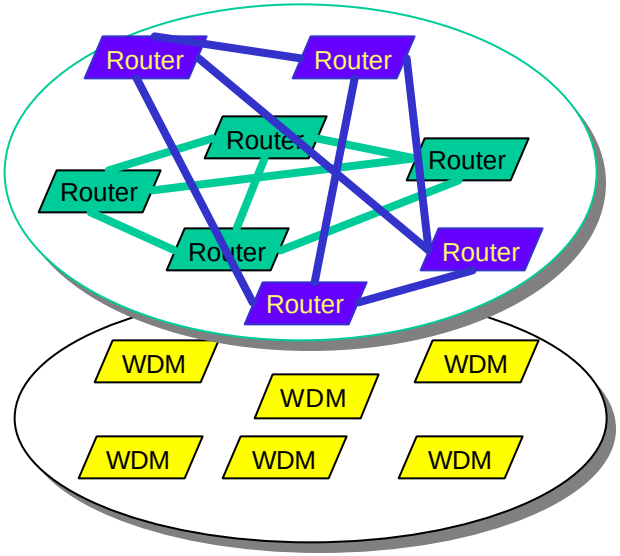


SuperNet Research Testbed

•networking research AND applications research network
•research in different networking layers (incl. phys) over the shared resource
•trick: use overlays (e.g. atm pvc's or wavelength defined subnet)
•default network state == working, NOT DOWN!!
•network THAT CAN BE BROKEN but carefully coordinated for service-affecting experiments



overlayed layer 3 testbeds e.g.
•*CAIRN*
•*ABONE*



e.g. optical or ATM layer

CAIRN: Layer 3+ research

- Collaborative Inter-Agency Research Network
- Stable layer 2
- Researchers control programmable routers over wide-area network
- Examples of research: multicast routing, IPsec, DNS security, video conferencing
- Currently 31 sites (industry, government, university) over commercial ATM service

Gigabit Desktop

Original Implementation Plan: 10’s of end nodes at gbps speeds

Planning:

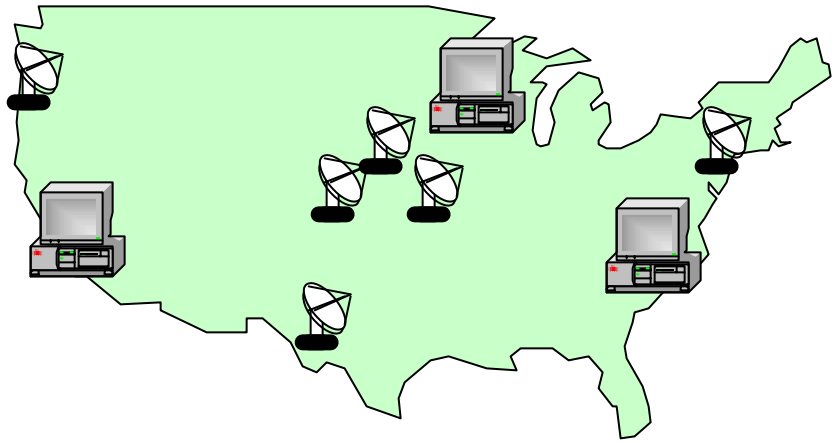
Multiple tens of end sites, each with a few gbps end-systems

Gigabit desktop technology development:
-Gbps host/NIC/Service Cards (won’t be available for another 14 months)
-Wavelength to the desktop (user or app triggered flow/burst switch)
-TCP+

But what can we immediately with COTS equipment (namely gigE) now?
Can we connect 100’s of gbps desktops over SuperNet?

Real-Time Testing and Experimentation Applications

- Network unique or costly facilities/resources with users & one another.
- Distributed radar control
  - virtual radar console for real time experiments by remote users
    (experienting with design of aircraft cone radar, surveillance satellite radars..)
  - new operational paradigm
- Device fabrication and testing
  - network labs for MEMS device fabrication, testing/characterization, designers, users.



CSU-CHILL Radar for Remote Sensing and Meterological Analysis

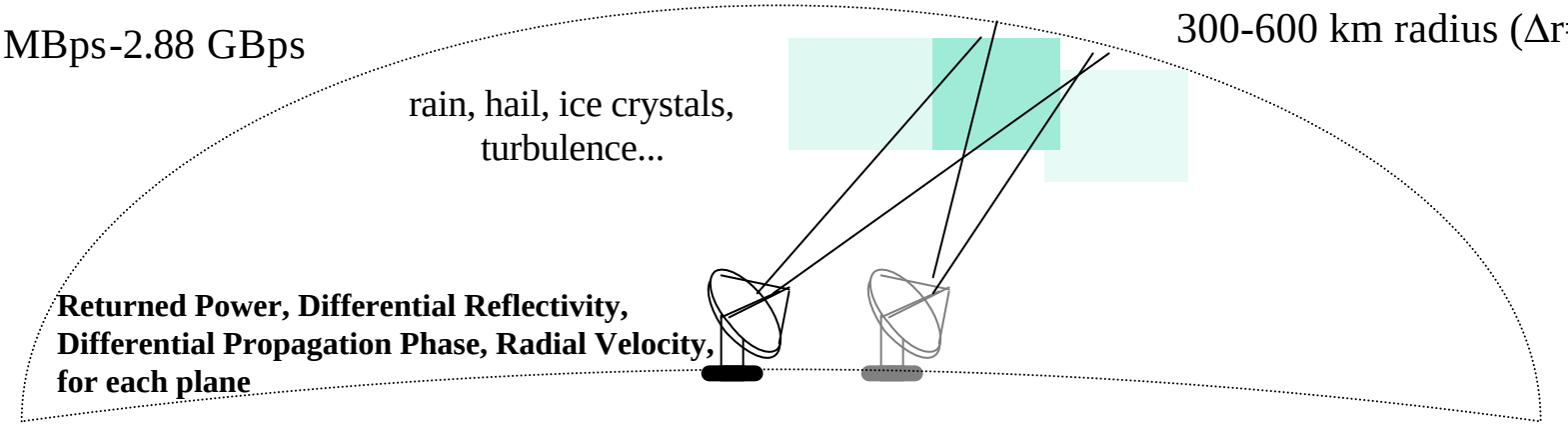
EM transparent dome, trailer, radar -- mobile

to the network

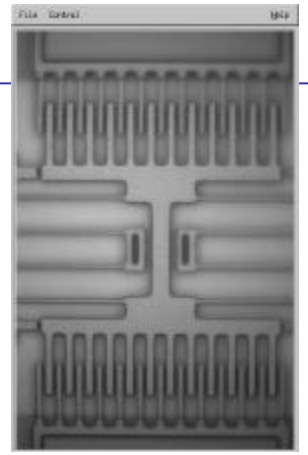
polarimetric, Doppler radar

240 MBps-2.88 GBps

300-600 km radius (Δr=30m)



MATISSE:Computer Microvision Workstations



Characterize MEMS devices by applying cw signal (variable amplitude/freq.)
Optically monitor device response over varying focal planes

Acoustic/vibrational isolation chamber

Waveform Generator:
- 12-bit waveform generation
- MHz frequencies with mHz resolution
- flexible stroboscopic control

Scientific Microscope:
- ultra-high resolution motor control
- stroboscopic LED illumination

CCD camera system:
- Megapixel camera & frame grabber

typical dataset 10 Gbytes

MIT, CMU, Berkeley, LBL

Digital Earth

Open, distributed, scalable multi-resolution 3-D representation of the earth into which massive quantities of geo-referenced information can be embedded.

•Use Domain Name System to develop a hierarchy of servers responsible for geographic cells of earth.
•Enhance today's text-indexing with geographic indexing web to geographically indexed.
• With Virtual Reality Modeling Language (VRML), so with standard browser with plug-in & ~ 50 Mbps, navigate the 3-D model.
•Collaboration between SRI, Planet9 Studios, Sprint.

Others

- digital amphitheater (conferencing 1000 people)
- multicasting HDTV (collab. with PBS)
- medical app
- crisis management

Inter-Agency Collaboration

Joint Reviews
- -Proposal review
- -Project reviews
- -PI meeting

Funding of development of tools used on NGI testbed

Cofunding of projects & workshops (CAIDA, CAIRN, NTON, Dec. Internet Economics Workshop)

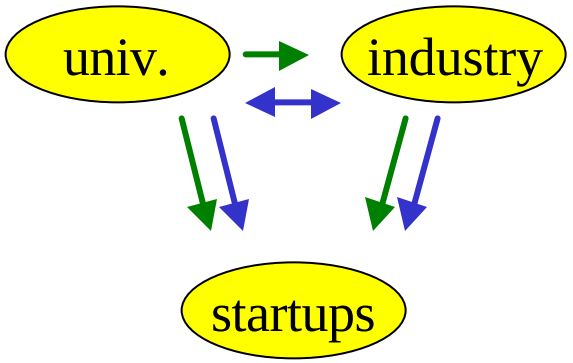
Experiment, demonstration, field trial coordination using testbeds (NRT meetings, recent Terravision demo)

Participation in research (e.g. NASA, NIST, researchers, also NRL, DISA, NSA..)

Technology Transfer

Who are the researchers?

- academic institutions
- industry (basic/applied research depts)
- startups (new, spawned from programs)
- consortia (NTON, ONRAMP, HSCC..)
  - -mix of univ,eqpt vendor,
  - service providers,heavy cost-sharing
- others FFRDC, govt labs, non-profit (CAIDA..)



Productization, New Service Definition, New architectures

People Transfer (e.g. students, new startups...)

IP Licensing

Sponsorship, membership (consciousness raising ...)