NGI Program at DARPA

PITAC Review
Jan 14, 2000
Mari Maeda
DARPA
mmaeda@darpa.mil
Outline

• Introduction and Goals
• NGI Research
• SuperNet Testbed Status
• New Applications
Number of flows

- 10 MB
- 10 K
- 20 K

Time

- 0.01 sec
- 0.1 sec
- 1 sec
- 10 sec
- 1 min
- 22 hours

Mean: 10 kilobyte
Scaling the Internet

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

mean hop distance = 16
Program Time Line

BAA 98-02
- Proposal eval.
- Contract negotiations

BAA 99-06
- Today

1998 1999 2000 2001

$5M to HPCC

Funding level

- $20M Quorum
- $20M
- $35M reduced from $40M
- $10M 'pork'
- $35M reduced from $40M
- $1M 'pork'?

NGI BAA open
NGI Program Components

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

- Simplified protocol layering - IP over dynamic Optical Network.
- End-to-end performance
- Testbed

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

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

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

Diversity and number of end devices
A range of projects addressing a range of near-term and long-term problems

- ultra-fast all-optical logic
- 1-10 Gbps NIC/SIC
- Optical Access
- optical burst switch
- multi-Gbps/multi-wavelength on demand
- inter-planetary Internet

Environment is very dynamic; industry is moving very rapidly. Goals and deliverables need to be adjusted accordingly. (Surprise us.)
Gbps++ End to End Delivery

- end-host architecture
- local/metro/regional network architecture (ONRAMP, HELIOS)
- protocol research (gbps tcpip, flow switch to bypass routers..)
- optical burst switch, label switch, packet switch
Wavelength Division Multiplexing

- Savings in regenerator equipment and new fiber build costs
- Stretch capacity per fiber (#, rate per channel) and unregenerated transmission span

- Fiber
- Gain-flattened amplifier
- Sources
- Modulation techniques
- Multiplexing techniques
- Nonlinearities
- Dispersion compensation
Optical Networking

- Dynamically configure new lightpaths to optically switch long sessions

Router initiated optical flow
courtesy of Adel Saleh
BIT Program’93-’99: Optical Networking

- Optical path set up by NC&M system
- Optical layer restoration
  (path vs line switched rings; mesh restoration)

becoming commercialized, largely by startups
Optical Flow Switching

• End-Host triggered Optical Flow Switching
  (1 - 10 Gbps end-to-end)

courtesy of N. Froberg, MIT LL
Advanced Optical Networking
(bypassing/offloading electronics)

Reconfigurable Optical Networking > minutes

Optical Flow Switching > 100 msec
(router or end-host triggered)

Optical Burst Switching > 10 \( \mu \text{sec} \sim 1 \text{ msec} \)

Optical Packet Switching > \( \mu \text{sec} \)

All-Optical Switching > nsec
Feeder & Distribution Network Architecture

To backbone network

- 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
Today's High Speed Access Solution

- individual fiber to end-customer sites requesting large bw
- multiplexing / switching in upstream office

A better solution needed that
- doesn't waste so much fibers
- enable very fast provisioning or very large bw on demand
Optical Access

- Need fiber infrastructure to be put in place - but use passive and active optical nodes in the distribution network.

- ECV
  - mini CO
  - manhole
  - bldg tel. room
  - pedestal

- star, tree, bus, ring
Distribution of Routed Wavelengths

Passive Components in Distribution Network

- Wavelength Add/drop
- High/Low Band Splitter
- Customer

Access Node

Distribution Fiber Pairs

Add/Drop Fiber Pairs

Fork Point

Customer
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
Multi-In Multi-Out (MIMO)

- Multiple antennas all transmitting in the same band
- Leverage heavy multipath environment
- Receiver signal processing:
  - Treat each sub-channel as “desired” signal, rest as “interferers” and use Adaptive Antenna Array-like technique to detect each (i.e. linear combinatorial nulling)
- Theoretically N-fold enhancement in spectral efficiency
- Acquisition time / training sequence overhead tradeoff

*indoor conditions demonstrated: 20-40 bps/Hz over 30 kHz
*investigate higher bit-rates and outdoor (urban, suburbs, open terrain) system
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
• Everlog 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

transit traffic
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
Internet Tomography

hop count histogram

temporal behavior
Scaling the number of Flows

Flow-Size Histogram
(March 1999)

OC3 Link
60,000~100,000 flows over
5 minute period
timeout after 1 minute

Fine-Grained Networking

SuperNet Technology

web-centric traffic
Name Lookup Today

- local lookup
  “http://www.nato.int/kosovo/video.htm”
- if not cached, go to root name servers & get remote DNS IP address
- go to remote server and get server IP address
- send to 152.152.96.6
Next-generation networking and service environment

- Network with devices & sensors, plus computers
- Devices, users (computers) and services may be mobile
- Services may be composed of groups of nodes
- Problems: configuration, routing, discovery, adaptation, security

App should be able to conveniently (i) specify a resource and (ii) send messages to it
iNAT Project (MIT)

- Intelligent naming
  - Intentional Naming System (INS)
  - Resource discovery in future networks
- Adaptive transmission
  - Congestion Manager (CM)
  - End-system congestion management and adaptation in the future Internet
InterPlanetary Networking

- Time-dependent, high latency, lossy paths in deep space
- Between planetary gateways, internets, platforms
- Layer 2, 3, 4 protocols
- IP address space and naming in space (domain name server)
> 100 Gb/s All-Optical Logic and Time-Division Multi-Access Network

- ultra-fast all-optical logic
- instead of demultiplexing hierarchically into lower rates, enable users to seize 100Gbps+ stream
- implementation of network nodes and TDM-based LAN
- performer: MIT Lincoln Laboratory
All-Optical Logic Gates

- nonlinear Kerr effect induces a phase shift in presence of control pulse

\[ I_{\text{out}} \sim \cos^2\left(\frac{\phi_b + \Delta\phi_{\text{nl}}}{2}\right) \]

\[ \Delta\phi_{\text{nl}} = \left(\frac{2\pi}{\lambda}\right) n_2 L I_c \]

- nonlinear element: fiber or SOA (semiconductor optical amplifier)
- different configurations - TOAD, NOLM, UNI
All-Optical Switch Implementation

*Ultrafast Nonlinear Interferometer (UNI)*

- Ultra short pulse train sequence generated via split/delay/combine technique
- Ultrafast refractive index effects yield differential phase shift between two orthogonal signal components.
- Differential phase shifts translate into polarization changes of the temporally realigned signal components (pi shift translates to polarization rotation)

**AND** Operation: UNI biased OFF, control pulses turn signal ON

**INVERT** Operation: UNI biased ON, control pulses turn signal OFF
Networking Architecture and Protocol

Folded Unidirectional Bus

1) Head end generates empty slots and distributes credits
2) Nodes with data to send and a credit may access any empty slot on the network
3) Nodes receive packets from the bus, read the headers and process data intended for them

User Interface Nodes

RECEIVER

2 psec pulses
100 Gbps (10 psec bit period)
Today’s Electro-Optic Regeneration

- **rcvr**
- **txm**

electro-optic conversion
re-timing, re-shaping,
amplification

All Optical Regeneration

Input: **Signal in**

- **isoler**
- **BRF**

Output: **Switch out**

- **50/50 splitter**
- **SOA**
- **Bandpass Filter**
- **polarization rotator**
- **polarizer**
SUPERNET TESTBED (www.ngi-supernet.org)

NTON II
4 λ's
@ 10 Gb/s per λ

multi-vendor (opt, gigE, layer2, layer3)
multi-service provider
Sites Supporting 1 Gigabit+ Connectivity

• Connected Today
• Planned or under discussions for 2000
  • {connected in at lower bw}

• MIT Campus
• MIT Lincoln Laboratory
• Cabletron
• AT&T/TCG
• Harvard

• Drexel
• U of Penn
• Sarnoff Laboratory
  • {Johns Hopkins Medical School}
  • {UMDNJ}

• ISI -East/DARPA
• NRL
• NASA
• NSA-U of MD
• DISA
• DIA
• CNRI
• North Carolina -MCNC, UNC
  • {Walter Reed Army Hospital}
  • {NIST}
  • {NIH/NLM}
  • {NIMA}
  • {Naval Surface Warfare}
  • {Holocaust Museum}
  • {Office of Naval Intelligence}

• U of Washington
• Microsoft
• CMU
• U Pitt Medical Center
• Pittsburgh SuperComputing
• Colorado State U
• NYC

• JPL
• CalTech
• LLNL
• Sandia NL
• SLAC
• NASA Ames
• USC - almost
• Boeing
• Tektronix
• SDSC
• Spawar
• Network Elements Inc.
  • {Silicon Valley Test Track - Sprint, sun, sgi, xerox park ..}
ATDNET-MONET TESTBED

• 2.5 - 10 Gbps / wavelength
• limited ‘transparency’
1. > 10 Gb/s capacity on 4 λ’s
   - transparent to format
   E.g. ATM/Sonet, SCM...

2. 1/2 reserved for “7x24” Sonet
   (next business day repairs)
BOSSNET Testbed

• Four fibers along inland/coast rail routes between Washington DC and Boston
• 29 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)
Recent SuperNet Experiments and Demonstrations

- 5x270 Mbps HDTV/POS transmission over 300 km
- 1.2 Gbps TCPIP between desktops; POS 300 km
- 1.5 Gbps HDTV/ATM Transmission over 500 km
- 600 km gbE over MONET/ATDNET
- 10 Gbps dynamic path set up over MONET/ATDNET
- Optical mid-span meet: multi-vendor protection switching demonstration
- Automatic optical layer topology discovery
100’s of Gigabit Desktops connected over SuperNet

• Currently targeting 5 campuses (MIT, USC, CMU, UWash, Berkeley)

• Discussions underway with a number of equipment vendors (desktop machines, NIC cards, gE switches, routers)

• Designing a qualification test

• Goal: approximately 50 desktops per each of 5 campuses, 2Q ‘00
Networking Radars

• Distributed radar control
• Remote data viewing and processing

• New operational paradigm
• Bring down cost/time of research (e.g. design of next generation aircraft one radar)
• Training
CSU-CHILL Radar for Remote Sensing and Meteorological Analysis

EM transparent dome, trailer, radar -- mobile

240 MBps-2.88 GBps

to the network

polarimetric, Doppler radar

300-600 km radius ($\Delta r=30m$)

rain, hail, ice crystals, turbulence...

Returned Power, Differential Reflectivity, Differential Propagation Phase, Radial Velocity, for each plane
MATISSE

Networking of sites for testing, characterization, fabrication, users

Example of characterization setup: Computer Microvision Workstation
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
Video Blanket
... see, remember and understand everything ...

• Enhanced visualization
  – Summarize live video from thousands of cameras into a few integrated displays.
  – Video from multiple cameras overlaid in real time over 3D site models to provide scene context.
    • Fly in for best perspective for objects of interest
  – Track events/ people across cameras.

• Active video surveillance
  – Close up views co-registered with wide angle views.
  – Virtual walk around of static/ mobile objects of interest
  – Detect/ recognize people (or vehicle) by comparing face/ iris (or license plate) to a database.

• Analysis of events
  – Motion tracks, who was where, when and went where.
Video Tracking using a panning camera: Dynamic Mosaic Videos

Original Video
Tracking Suspect Along Street

Dynamic Mosaic Video

Synopsis Mosaic

prc_c.avi

prdyn.avi
Visualizing disjoint views

Input data

3D Model

Re-projection to new view-point after alignment of image frames to 3D model

Merging projected frames 1 and frames 2 to create a composite frame from a new view-point

Frame 1

Frame 2
HUBS Telemedicine Application (UPMC, Johns Hopkins)

Beyond text based electronic patient records and proprietary picture archiving/communication systems

• users: radiologists, clinicians, researchers, educators
• 1 image 1 - 10 MegaBytes
• one study = 2 to 100 images or 20-100 MB/study
• Hopkins and UPMC 700,000 studies/year
• 35 terabytes/yr or 15 terabytes/yr compressed

• 3000 new studies/day requires
• 50 Mbps, x 2 old studies/new studies, 3x clinical transaction = 300 Mbps

• Use next-generation VPN service to tie together multiple archiving sites together with users and demonstrate technical feasibility as well as user acceptance
**NGI Multicast Applications and Architecture**

**New Ideas**
- **Digital Amphitheater** – use dynamic video merge servers in a large-scale multicast architecture to meaningfully conference on the order of a thousand participants, with wide range of participant-controls for the display.
- **Multicasting HDTV** – use RTP and multicast to transmit digital television signals and related data objects on NGI.

**Impact**
- Push beyond commercial limits of network video.
- Prototype high definition multimedia while retaining economic benefits of commodity computers.
- In collaboration with Corporation for Public Broadcasting, bring about technology transfer between broadcast DTV and broadband networking.

**Schedule**

<table>
<thead>
<tr>
<th>Project Start</th>
<th>1999 Q3</th>
<th>2000 Q1</th>
<th>2000 Q2</th>
<th>2000 Q3</th>
<th>2000 Q4</th>
<th>2001 Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDTV over RTP Specs</td>
<td>Video Merge Server Alpha</td>
<td>HDTV Multicast on Supernet</td>
<td>Largest-scale Digital Amphitheater Meeting</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

USC/Information Sciences Institute: Allison Mankin
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.
Infrastructure: .geo domain

- Use DNS to encode latitude/longitude for any element in a hierarchical scheme.
- minutes.degrees.tendegrees.geo
- e.g. 37e47n.1e5n.10e20n.geo