GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Summary Status of
GSFC 10 Gbps "Lambda Network" (L-Net)
February 10, 2005

J. Patrick Gary

Network Projects Leader

Computational & Information Sciences and Technology Office/606

NASA Goddard Space Flight Center

pat.gary@nasa.gov

301-286-9539

Outline for End of Year Review

- **Motivation**
  - » **Advances in Networking Technology**
  - » **Enabling New NASA Science Needs**
- **Goals**
- **Key Challenges and Solution Designs**
- **Implementation Status**
- **Next Steps**

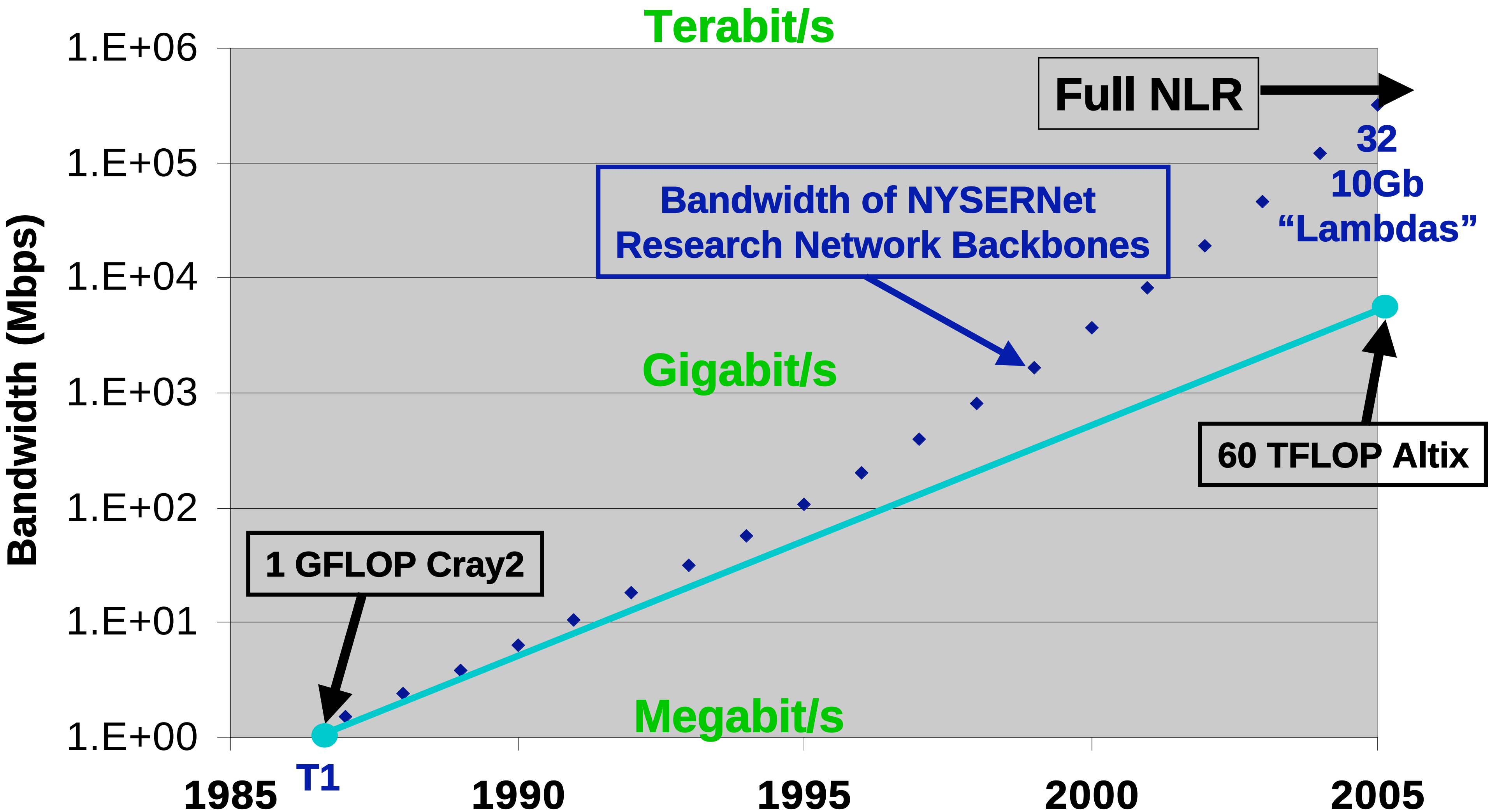
GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Motivation Outline

- **Advances in Networking Technology**
  - » **Bandwidth growth rate greater than Tflops growth rate**
  - » **National LambaRail (NLR) implementation**
  - » **Global Lambda Integrated Facility (GLIF) cooperation**
  - » **Latest Internet2 IPv4 Land Speed Record**
  - » **Personal Computer Interface advances**

- **New NASA Science Needing Gigabit per Second (Gbps) Networks**
  - » **Coordinated Earth Observing Program (CEOP)**
  - » **Hurricane Predictions**
  - » **Global Aerosols**
  - » **Remote viewing & Manipulation of Large Earth Science Data Sets**
  - » **Integration of Laser and Radar Topographic Data with Land Cover Data**
  - » **Large-Scale Geodynamics Ensemble Simulations**

Optical WAN Research Bandwidth Has Grown Much Faster than Supercomputer Speed!



Source: Timothy Lance, President, NYSERNet

NLR Will Provide an Experimental Network Infrastructure for U.S. Scientists & Researchers

NLR
light the future

"National LambdaRail" Partnership

Serves Very High-End Experimental and Research Applications
4 x 10Gb Wavelengths Initially
Capable of 40 x 10Gb wavelengths at Buildout

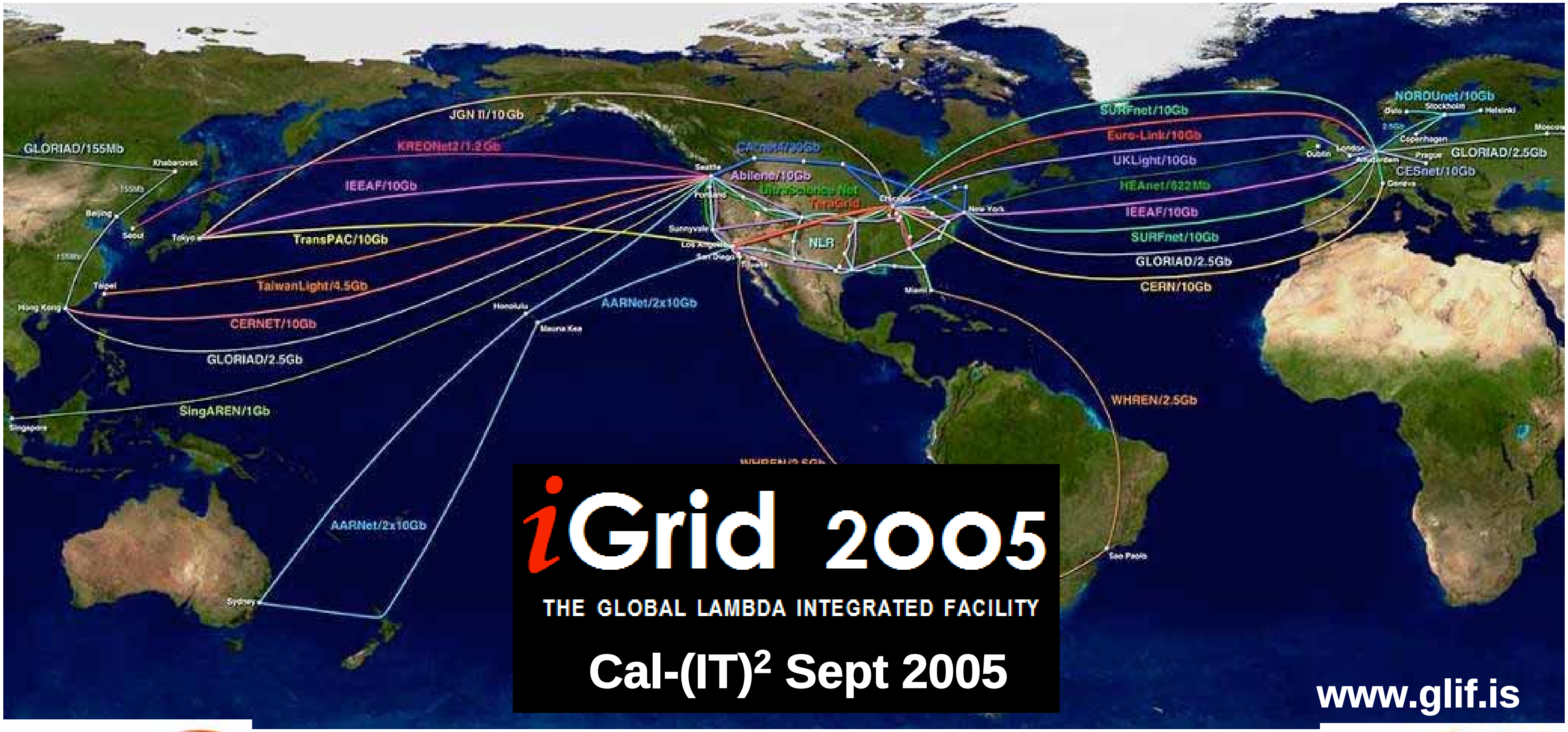
Links Two Dozen State and Regional Optical Networks

First Light September 2004

For more information regarding NLR see http://www.nlr.net or contact info@nlr.net

Global Lambda Integrated Facility: Coupled 1-10 Gb/s Research Lambdas

Predicted Bandwidth, to be Made Available for Scheduled Application and Middleware Research Experiments by December 2004



Visualization courtesy of Bob Patterson, NCSA

Internet2 Land Speed Record

(Rules and current records: http://lsr.internet2.edu/)

Latest IPv4 Single Stream Record (http://data-reservoir.adm.s.u-tokyo.ac.jp/lsr)

- **7.21 Gbps** (TCP payload), standard frame, 148.850 Petabit meter / second
- 20,645 km connection between SC2004 booth and CERN through Tokyo, Latency 433 ms RTT

Network used in the experiment

Personal Computer Interface (PCI) Advances

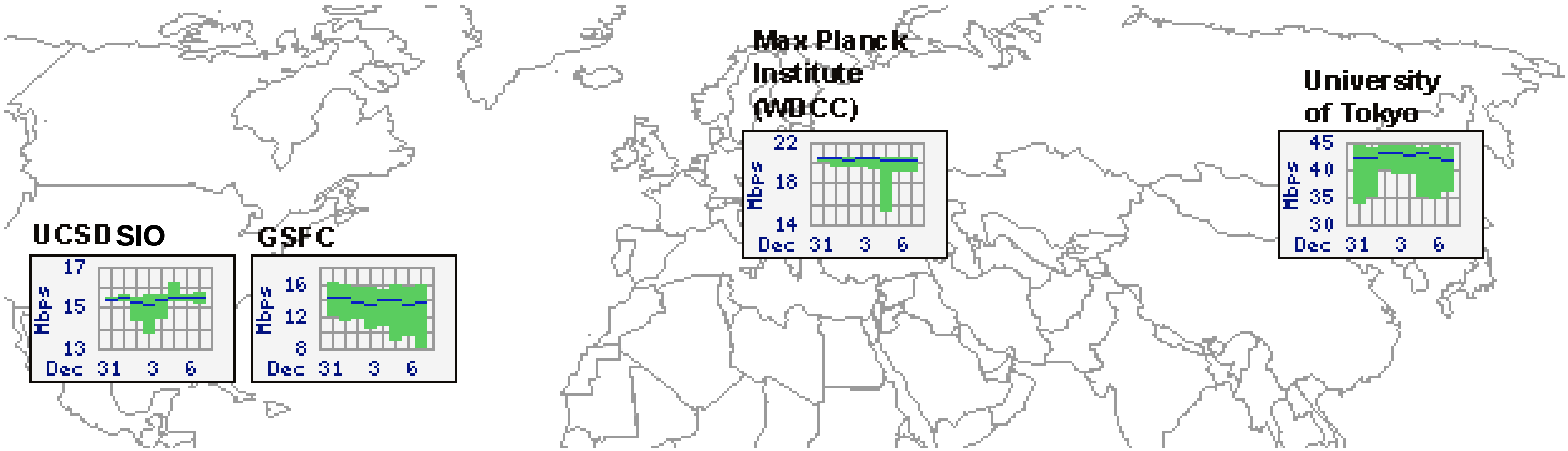
- **Shared Parallel Bus**
  - » **PCI 1.0 (32-bit, 33 MHz): 1.056 Gbps (1 direction at a time)**
  - » **PCI 2.3 (64-bit, 66 MHz): 4.224 Gbps (1 direction at a time)**
  - » **PCI-X 1.0 (64-bit, 133 MHz): 8.448 Gbps (1 direction at a time)**
  - » **PCI-X 2.0 (64-bit, 266 MHz): 16.896 Gbps (1 direction at a time)**

- **Dedicated Serial Interface (4 wires per "lane")**
  - » **PCI Express:**
    - – **2.5 Gbps (raw) per lane each direction**
    - – **2.0 Gbps (without encoding overhead) per lane each direction (maximally 4.0 Gbps bi-directional)**
    - – **Up to 32 lanes**

Next Step: OptIPuter, NLR, and Starlight Enabling Coordinated Earth Observing Program (CEOP)

Source: Milt Halem, NASA GSFC

Accessing 300TB's of Observational Data in Tokyo and 100TB's of Model Assimilation Data in MPI in Hamburg -- Analyzing Remote Data Using GRaD-DODS at These Sites Using OptIPuter Technology Over the NLR and Starlight



Note Current Throughput 15-45 Mbps: OptIPuter 2005 Goal is ~10 Gbps!

http://ensight.eos.nasa.gov/Organizations/ceop/index.shtm

OptIPuter and NLR will Enable Daily Land Information System Assimilations

- The Challenge:
  - » More Than Dozen Parameters at ~ 50 GB per Parameter , Produced Six Times A Day, Need to be Analyzed
- The LambdaGrid Solution:
  - » Sending this Amount of Data to NASA Goddard from Project Columbia at NASA Ames for Human Analysis Would Require < 15 Minutes/Day Over NLR
- The Science Result:
  - » Making Feasible Running This Land Assimilation System Remotely in Real Time

Source: Milt Halem, NASA GSFC

NLR/GSFC Applications: Hurricane Prediction

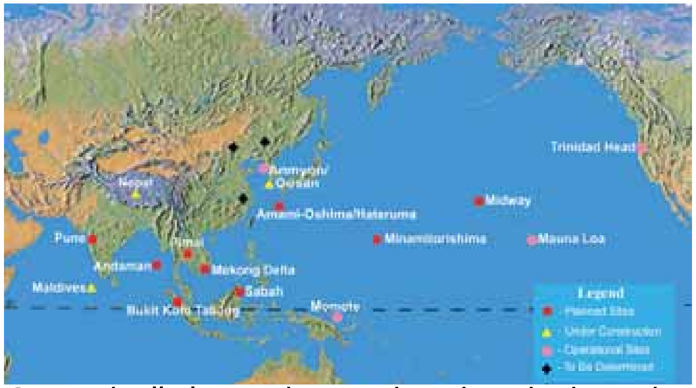
- The NASA Finite-Volume General Circulation Model (fvGCM) has been producing real-time, high-resolution (~25 km) weather forecasts focused on improving hurricane track and intensity forecasts.
- During the active 2004 Atlantic hurricane season, the fvGCM provided landfall forecasts with an accuracy of ~100 km up to 5 days in advance.
- The 50–100 Mbps throughput available between fvGCM users at GSFC and the Columbia supercomputer at ARC greatly hindered carrying out time-critical simulations of the hurricanes that devastated Florida.
- The 10 Gbps NLR access will enable remote, 3D visualization analysis as soon as forecast variables become available.
- Key Contacts: Ricky Rood, Bob Atlas, Horace Mitchell, GSFC; Chris Henze, ARC.

In an fvGCM forecast, Hurricane Frances makes landfall on the Gulf Coast of Florida while Hurricane Ivan intensifies in the tropical Atlantic. Visualization by J. Williams, GST.

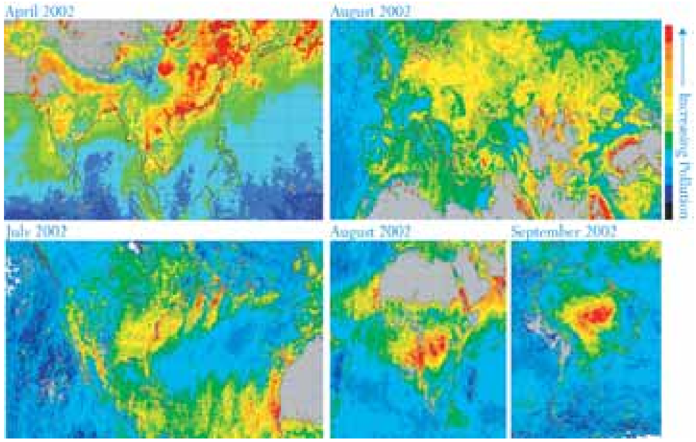
http://fvnwp.gsfc.nasa.gov

NLR/GSFC Applications: Global Aerosols

- Project Atmospheric Brown Clouds (ABC) is an international effort to discover and analyze areas of brown colored atmosphere to learn how dust and pollution particles are transported and what impacts they have on the environment, climate, agricultural cycles, and quality of life.
- GSFC and the Scripps Institution of Oceanography (SIO) are planning a collaboration to predict the flow of aerosols from Asia across the Pacific to the U.S. on timescales of days to a week.
- GSFC will provide an aerosol chemical tracer model (GOCAR) embedded in a high-resolution regional model (MM5) that can assimilate data from Indo-Asian and Pacific ground stations, satellites, and aircraft.
- Remote computing and analysis tools running over the NLR will enable acquisition and assimilation of the Project ABC data.
- Key Contacts: Yoram Kaufman, William Lau, GSFC; V. Ramanathan, Chul Chung, SIO.



Strategically located ground stations in the Indo-Asian and Pacific regions monitor atmospheric pollution.

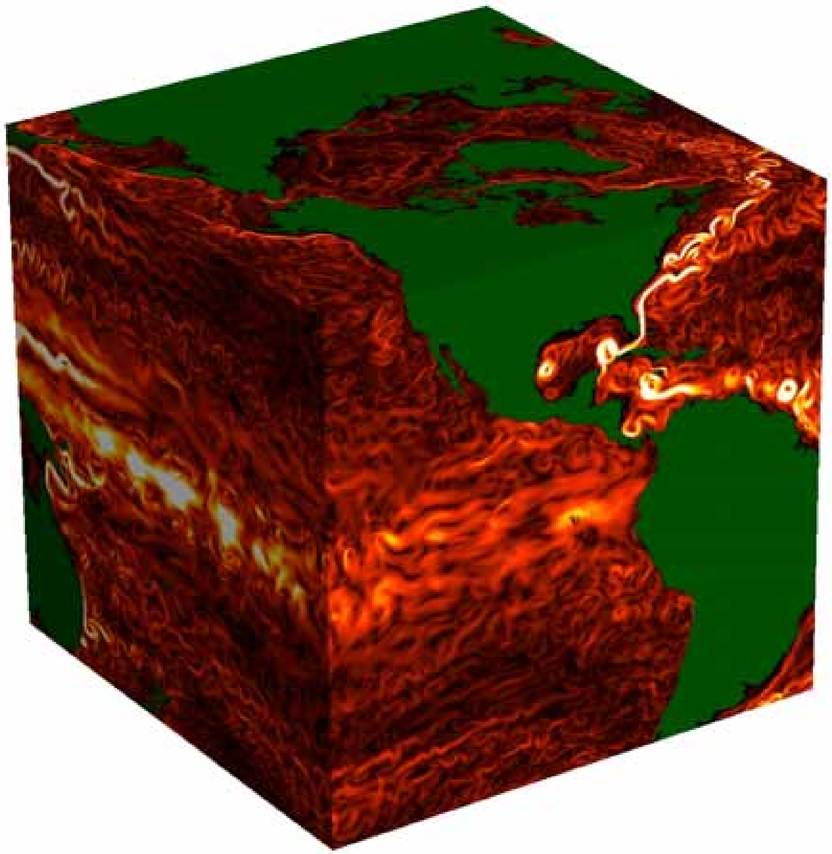


The global nature of brown clouds is apparent in analysis of NASA MODIS Data. Research by V. Ramanathan, C. Corrigan, and M. Ramana, SIO.

http://www-abc-asia.ucsd.edu

NLR/GSFC Applications: Remote Viewing and Manipulation of Large Earth Science Data Sets

- Remote viewing and manipulation of data sets at GSFC and JPL is needed to support EOSDIS and Earth system modeling.
- GSFC's EOSDIS Clearing House (ECHO) and JPL's GENESIS prototype science analysis system (iEarth) will become connected over the NLR. The link will enable comparison of hundreds of terabytes of data, generating large, multi-year climate records.
- Initial work will focus on the Estimating the Circulation and Climate of the Ocean (ECCO) modeling team. Besides ready access to the NLR, the team will need versatile subsetting and other data manipulation functions to reduce compute and bandwidth requirements as well as a set of Grid-accessible statistical analysis and modeling operators to refine and validate the ECCO models.
- Key Contacts: ECHO metadata gateway team, GSFC; GENESIS team, led by Tom Yunck, JPL.

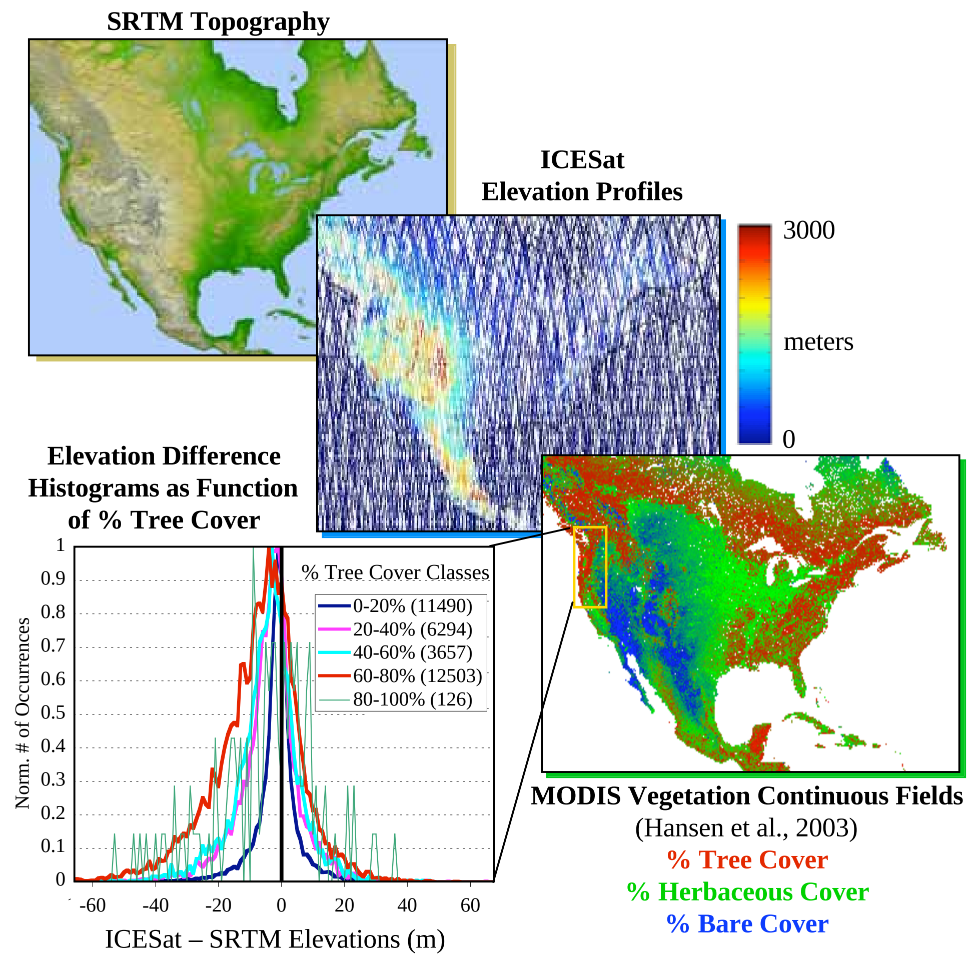


Near-surface (15-m) ocean current speed from an eddy-permitting integration of the cubed-sphere ECCO ocean circulation model. Research by JPL and MIT. Visualization by C. Henze, Ames.

http://www.ecco-group.org

NLR/GSFC Applications: Integration of Laser and Radar Topographic Data with Land Cover Data

- NASA has executed two advanced missions to create an accurate high-resolution topographic model of the Earth: the Shuttle Radar Topography Mission (SRTM) and ICESat, with its Geoscience Laser Altimeter System (GLAS).
- The agency now has the opportunity to merge the two data sets, using SRTM to achieve good coverage and GLAS to generate calibrated profiles. Proper interpretation requires extracting land cover information from Landsat, MODIS, ASTER, and other data archived in multiple DAACs.
- Use of the NLR and local data mining and subsetting tools will permit systematic fusion of global data sets, which are not possible with current bandwidth.
- Key Contacts: Bernard Minster, SIO; Tom Yunck, JPL; Dave Harding, Claudia Carabajal, GSFC.



http://icesat.gsfc.nasa.gov
http://www2.jpl.nasa.gov/srtm
http://glcf.umiacs.umd.edu/data/modis/vcf

High speed networking and Grid computing for large-scale simulation in geodynamics

W. Kuang[1], W. Jiang[2], S. Zhou[3], P. Gary[1], M. Seablom[1], W. Truszkowski[1], J. Odubiyi[4], D. Liu[2], J. Palencia[5], G. Gardner[6]

[1] NASA Goddard Space Flight Center, [2] JCET, UMBC, [3] Northrop Grumman IT/TASC, [4] Bowie State University, [5] Raytheon ITSS, [6] INDUSCORP

Introduction

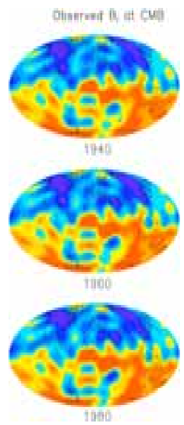
Now large-scale simulation has been wide-spread in many disciplines of solid Earth science research. A typical numerical test in the simulation can easily reach 10^{13} flops and beyond.

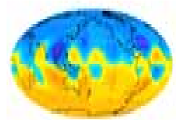
One such research problem that we are working on now is to establish a framework on predicting geomagnetic secular variation on decadal and longer time scales, utilizing surface geomagnetic/paleomagnetic records and our MoSST core dynamics model (Figure 1). In this approach, model forecast results and observations are weighted to provide initial state for assimilation (Figure 2). Typically 30 independent numerical tests are necessary for a reasonable ensemble size. This could easily require a computing cycle on orders of petaflops and larger.

A single super-computing facility for such studies is not an optimal choice, due to many limitations, in particular those on user management and administration. But it is relatively easy for users (researchers) to manage because of a unified system environment.

Grid computing can be a much better choice so that independent numerical tests can be carried out independently on different systems. However, researchers (users) have to deal with heterogeneous systems and other problems, such as those on network communication.

In this poster, we discuss our activities in GSFC on application of grid computation to geodynamics modeling.





Numerical results from our MoSST core dynamics model

Figure 1. Radial component of the magnetic field at the CMB inverted from surface geomagnetic observation (left panel) and from numerical modeling (top).

Geomagnetic data assimilation

$$x^a = x^f + K\left(x^o - Hx^f\right)$$

X^a: Assimilation solution
X^f: Forecast solution
X^o: Observation data

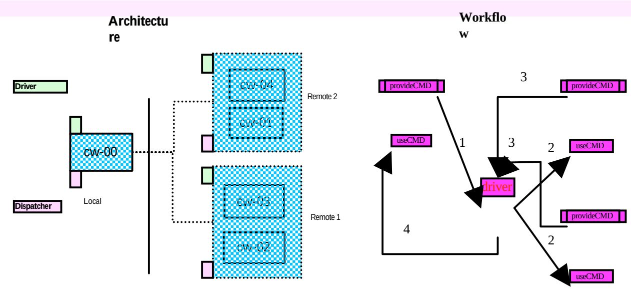
Figure 2. Mathematical foundation of data assimilation. The common gain **K** depends on knowledge of error statistics of observations and of models. If ensemble Kalman-filter approach is applied. An ensemble size of at least 30 (i.e. independent tests) is required.

Prototype on MoSST simulation with independent systems

The objective of this prototype work is to test operability of executing our MoSST core dynamics model on independent computing systems. Individual computing units are slated out from selected components of our beowulf system to mimic independent computing environment. The prototype program for grid computing is built upon xcat3 framework (based on java/python). See Figure 3 for conceptual layout of our prototype experiment.

The sample script and the execution process are shown in Figure 4.

Our prototype experiment is very successful. With this experiment, we can proceed further our test on real remote systems. Also with this experiment, we can identify the needs from the user's considerations on supporting environment and other middleware that makes grid computing "friendly".



System configuration

OS: Fedora core 2; MPICH-1.2.5.2; Intel Fortran Compiler; Java 2
PE: Dual Intel Xeon, 2.4 Ghz, 1 GB, 1 GigEthernet

Figure 3. Prototype layout

```
import sys
import cca

from jarray import zeros

from java.lang import System
from java.lang import String, Object
# get the absolute location for XCAT
.
# create the TypeMap for the user component
# create component wrappers
geo1Provides = cca.createComponentWrapper("geo1Provider",
geo1ProviderMap)
geo2Provides = cca.createComponentWrapper("geo2Provider",
geo2ProviderMap)
#uses = cca.createComponentWrapper("user", userMap)
# assign a machine name
cca.setMachineName(uses, "cw-00")
cca.setMachineName(provides, "cw-00")
cca.setMachineName(geo1Provides, "cw-03")
cca.setMachineName(geo2Provides, "cw-04")

# set a creation mechanism to in-process
cca.setCreationMechanism(uses, "local")
cca.setCreationMechanism(provides, "local")
cca.setCreationMechanism(geo1Provides, "ssh")
cca.setCreationMechanism(geo2Provides, "ssh")
.
# connect their ports
cca.connectPorts(uses, "dispatchUsesPort", provides,
"dispatchProvidesPort")
cca.connectPorts(uses, "geo1UsesPort", geo1Provides,
"geo1ProvidesPort")
cca.connectPorts(uses, "geo2UsesPort", geo2Provides,
"geo2ProvidesPort")
# invoke the method
cca.invokeMethodOnComponent(uses,
                                              portClassName,
                                              portType,
                                              providesPortName,
                                              methodName,
                                              methodParams)
```

```
wjiang@corewulf:~/XCAT_HOME[wjiang@corewulf XCAT_HOME]$ xcat.jython.RunScript
src/jython/samples/usesdispatchgeo2ssh.py
. . . . . .
Executing jython script : src/jython/samples/usesdispatchgeo2ssh.py
[ 20:00:11.004 main: xcat.ccacore.XCATServicesImpl.java:118 <init> ] called with name: CompositionTool
[ 20:00:12.531 main: xcat.jython.CompositionTool.java:353 createTypeMap ] called
[ 20:00:12.554 main: xcat.jython.ComponentWrapper.java:338 getInstanceName ] called
. . . . . . .
[ 20:00:12.555 main: xcat.jython.CompositionTool.java:233 instantiate ] called for component: user
. . . . . .
[ 20:00:12.591 main: xcat.ccacore.XCATComponentIDServerImpl.java:110 <init> ] exporting ComponentID as
a Grid service
. . . . . .
[ 20:00:12.597 main: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:62 setServices ] adding
dispatchUsesPort
[ 20:00:19.228 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:152 startExecuting ]
Command /home/wjiang/research/app/test1/run_main_mpi is sent to cluster cw-03
[ 20:00:19.250 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:154 startExecuting ]
Command /home/wjiang/research/app/test1/run_main_mpi is received!
[ 20:00:19.358 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:158 startExecuting ] Post
info: Application is running on Cluster cw-03. Result will be stored in ~/data/test1
[ 20:00:19.359 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:161 startExecuting ]
Command /home/wjiang/research/app/test1/run_main_mpi is sent to cluster cw-04
[ 20:00:19.380 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:163 startExecuting ]
Command /home/wjiang/research/app/test1/run_main_mpi is received!
[ 20:00:19.491 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:166 startExecuting ] Post
info: Application is running on Cluster cw-04. Result will be stored in ~/data/test2
. . . . . . .
[ 20:00:19.491 Thread-0: samples.esmfcca2.src.dispatch.CMDPortDispatch.java:95 useCMD ]
CMDPortDispatch.java: useCMD(...)!
CMDPortDispatch: useCMD(), geoCMD is Application is running on Cluster cw-04. Result is stored in
~/data/test2
[ 20:00:19.493 Thread-0: xcat.ports.BasicPortImpl.java:95 destroyImpl ] called; going down in 5
seconds
. . . . . .
[ 20:00:19.579 Thread-0: samples.esmfcca2.src.usesgeo2.UsesDispatchGeo2.java:189 startExecuting ] Your
simulation is successful!!*!!!
```

Figure 4. Prototype Operation Script (left) and Screen Caption (right)

Discussions

1. Our research on geomagnetic data assimilation can greatly benefit from grid computing.
2. Our prototype experiment is successful and can be readily expanded to systems with identical settings and SSH communication protocol.
3. Our prototype experiment is limited in many areas, such as handling network communication between independent systems (e.g. instant feedback of remote systems to host systems), heterogeneous environment (e.g. prior knowledge on participating systems is necessary), authentication (e.g. prototype cannot handle high level access security requirement). Therefore, further experiment is needed to improve our work, such as integrating our work with other (developed and developing) middleware handling the problems.

Related work at GSFC

There are parallel, but related research going on in GSFC on networking and software development. These research activities are updated in http://esdcd.gsfc.nasa.gov/LNetImplement.html. Recent overview of GSFC research activities is given by Dr. M. Halem and can be found in http://esdcd.gsfc.nasa.gov/L-Netpdfs/ESSAAC_MHpres9904.pdf. Some of the activities listed in the report are shown in Figures, 5 and 6. These activities work towards establishing 21st century cyber infrastructure for large-scale scientific teamwork based on fast network.

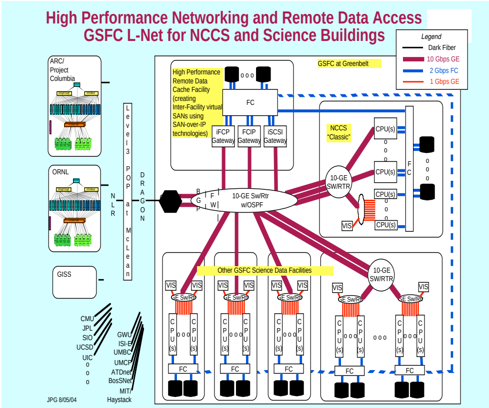
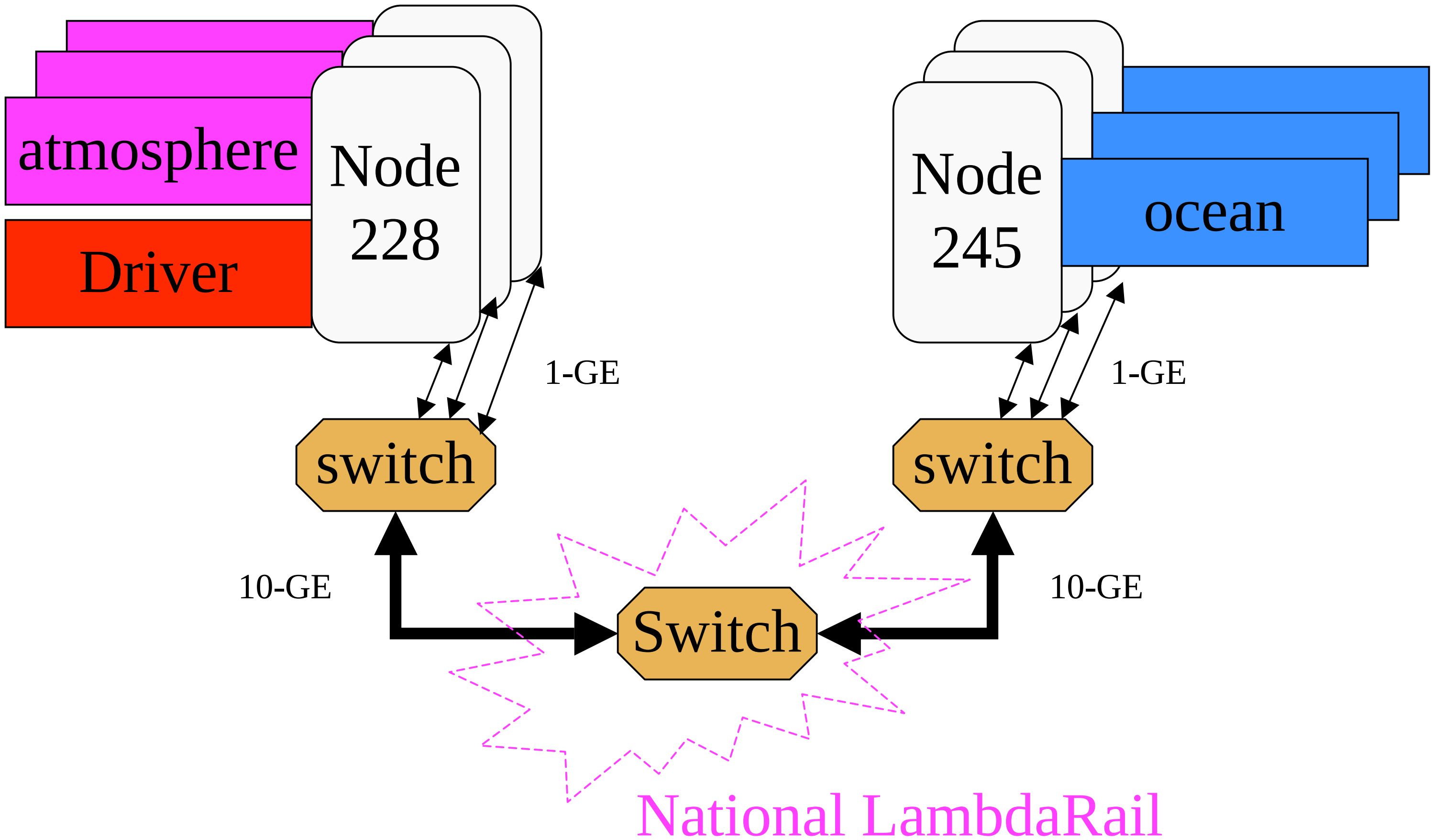


Figure 5. NASA GSFC IRAD work on regional fast network

Figure 6.

APPLICATIONS -
Future GRID on 10-GE Network



Dr. Zhou is working on applying Grid Computing and High-Speed Network to large-scale distributed computing in Earth and Space Science. More details can be found at http://esto.nasa.gov/conferences/estc2004/papers/a4p1.pdf.

GSFC FY04 IRAD Proposal "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure”

Project Goals

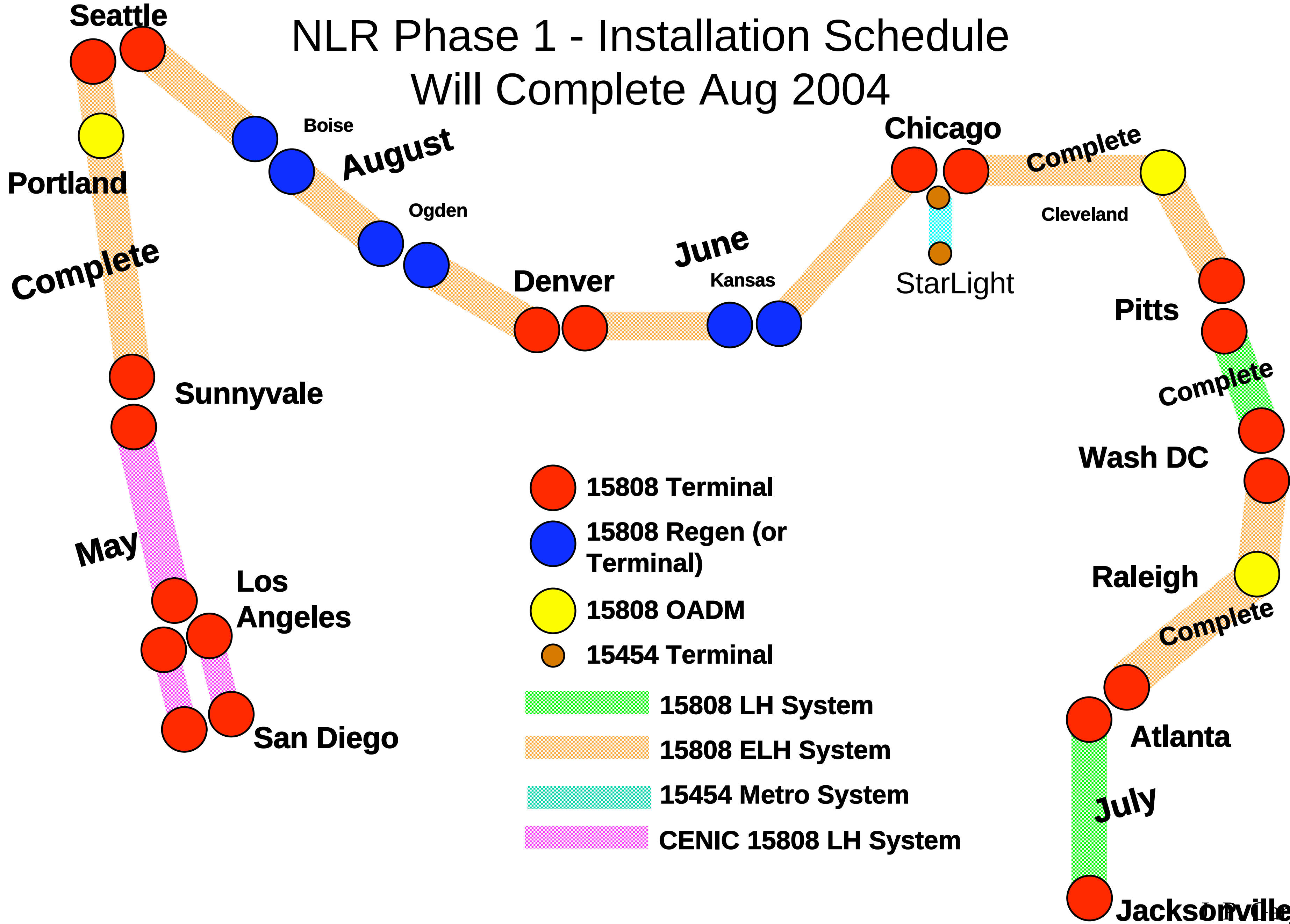
- **“…establish a “Lambda Network” (in this case using optical wavelength technology and 10 Gbps Ethernet per wavelength) from GSFC’s Earth science Greenbelt facility in MD to the Scripps Institute of Oceanography (SIO) through the University of California, San Diego (UCSD) facility over the National Lambda Rail (NLR), a new national dark optical fiber infrastructure.”**
- **“…make data residing on Goddard’s high speed computer disks available to SIO with access speeds as if the data were on their own desktop servers or PC’s.”**
- **“…enable scientists at both institutions to share and use compute intensive community models, complex data base mining and multi-dimensional streaming visualization over this highly distributed, virtual working environment.”**

Key Challenges and Solution Designs Outline (1 of 2)

- **Implementing 10-Gbps Computer Networks End-to-End (ISO Layers 1-3)**
  - » **Transcontinental Network Part**
    - – **NLR Phase 1/Year 1**
  - » **Regional Network Part**
    - – **DRAGON Phase 1/Year 1**
  - » **Local Area Network Part**
    - – **10-GE upgrade to GSFC's Scientific and Engineering Network**

NLR – Optical Infrastructure - Phase 1

NLR Phase 1 - Installation Schedule
Will Complete Aug 2004

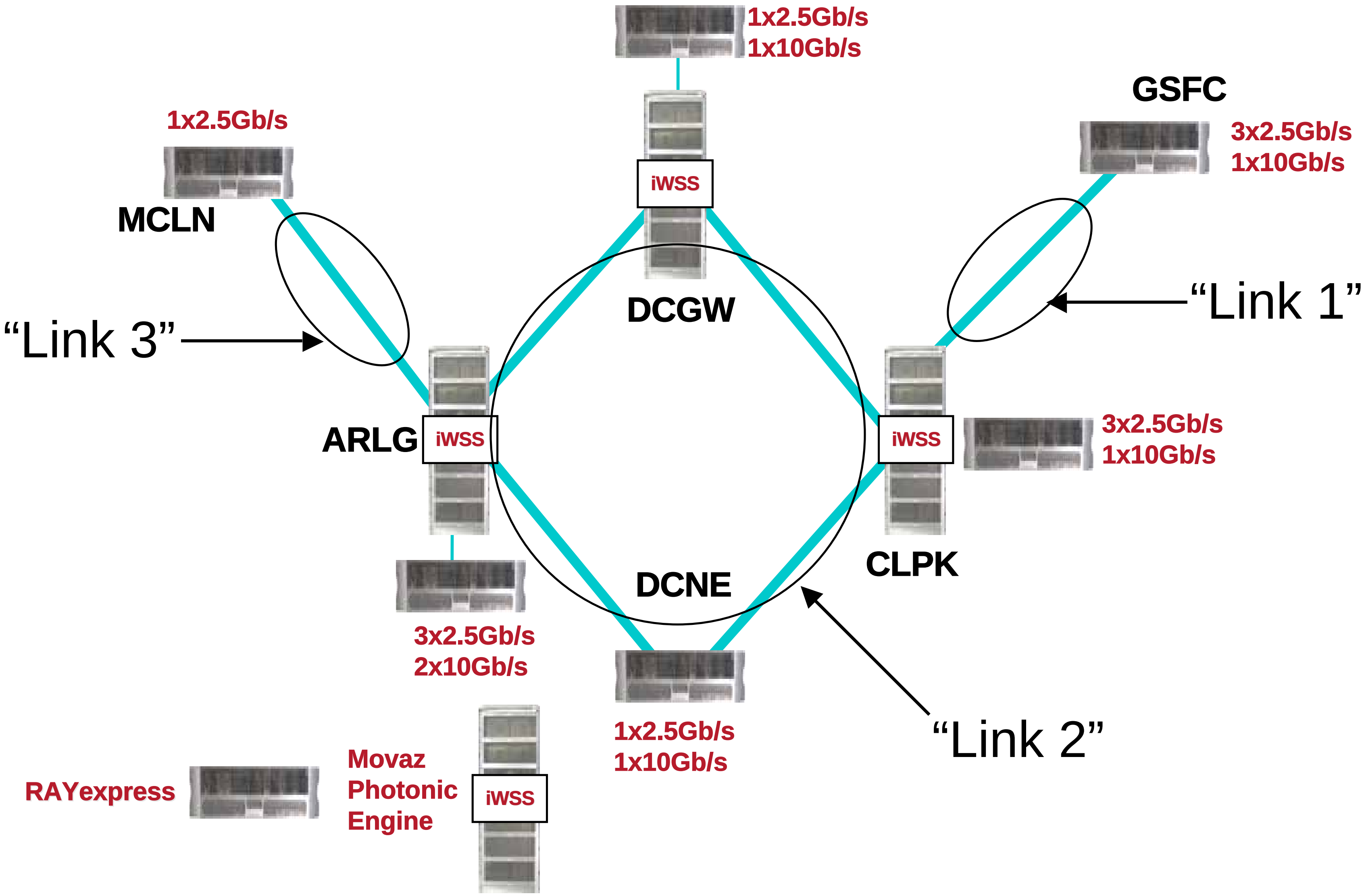


NLR Wavelengths

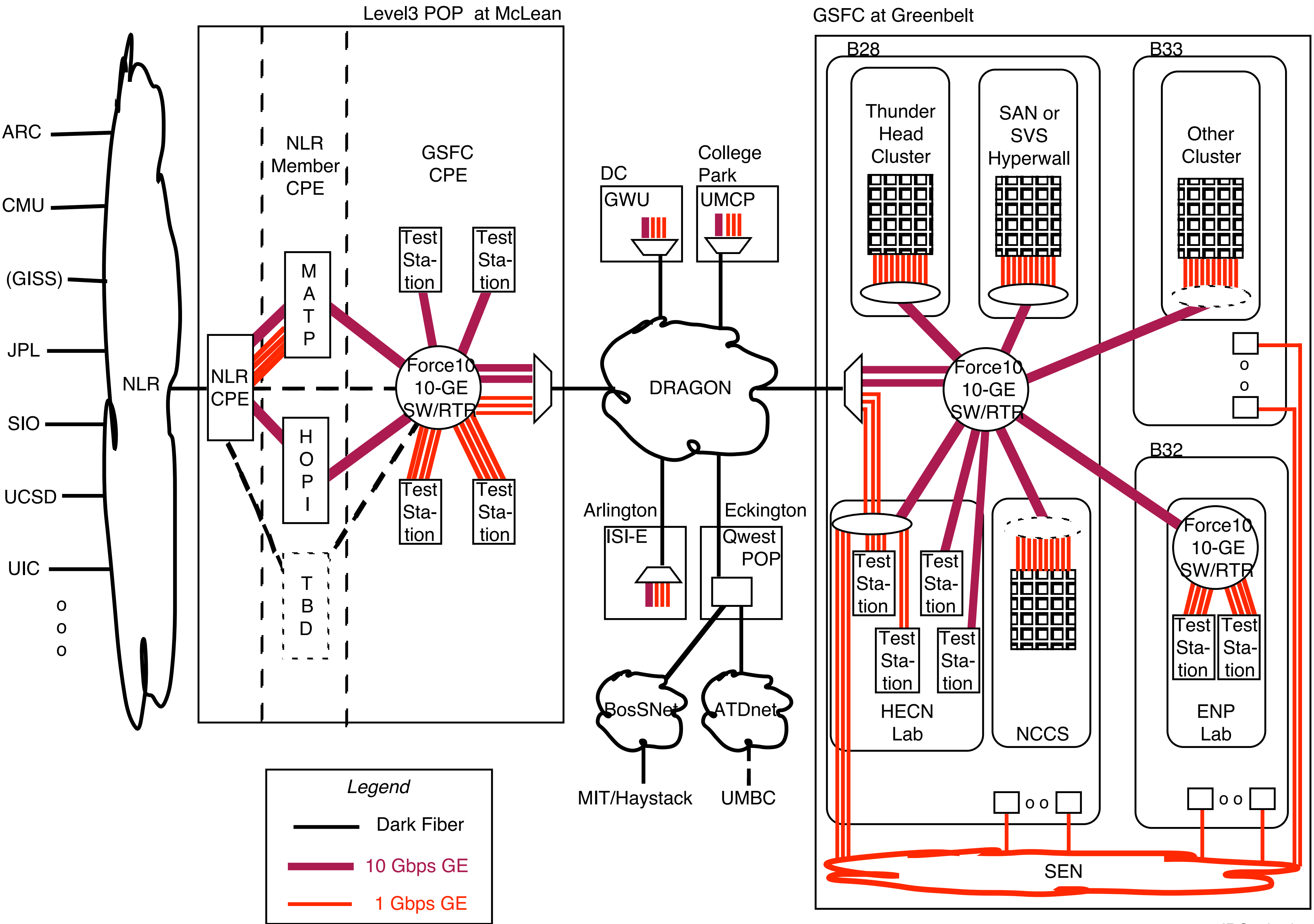
- **Initial complement of 4 λs installed and available at outset**
  - One λ for national switched Ethernet experimental network
  - Another λ for national 10 Gbps IP network to support internetworking and end-to-end transport protocol experiments
    - Similar to Internet2's Abilene except routers will be available for measurement and experimentation
  - Third λ will serve as a quick start facility for new research projects
  - Fourth λ will be used by Internet2's HOPI testbed

More λs will be activated as needed to support the research and operational objectives of the community

Dynamic Resource Allocation with GMPLS on Optical Networks (DRAGON) Configuration



GSFC L-Net Configurations at McLean and Greenbelt



JPG 6/02/04

GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Key Challenges and Solution Designs Outline (2 of 2)

- **Tuning Applications for High Performance Networks Use (ISO Layers 4-7)**
  - » **Large round-trip-time latencies for packet acknowledgements**
    - – **TCP Alternates or Enhancements**
  - » **Slow disk access times**
    - – **Pre-fetch caching to RAM**
  - » **Interactive data steaming to 100 mega-pixel displays**
    - – **Multiple GE interfaces to visualization clusters**
  - » **GrADS/DODS**
    - – **Porting to OptIPuter connected hosts**

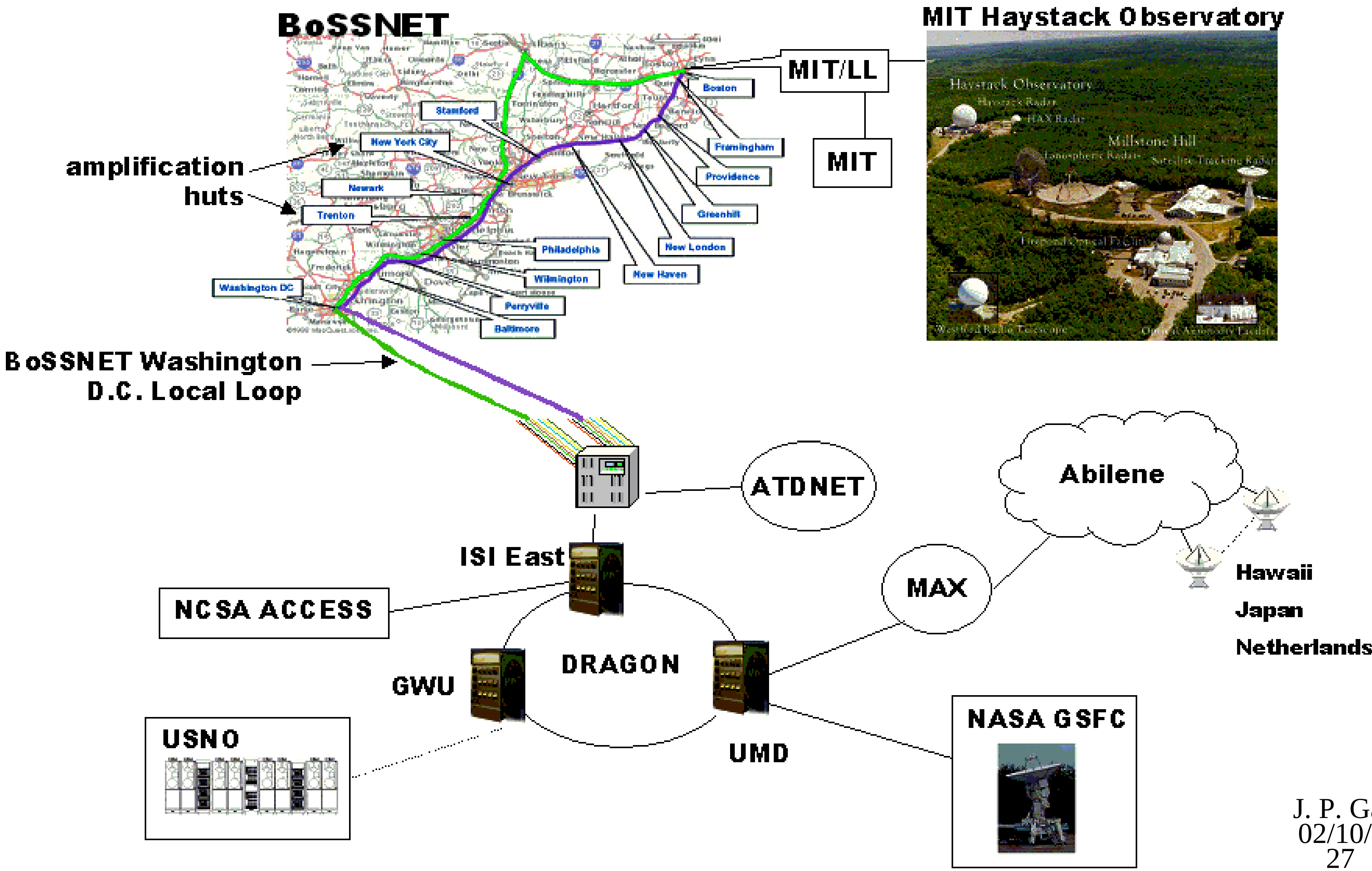
Outline for End of Year Review

- **Motivation**
- **Goals**
- **Key Challenges and Solution Designs**
- **Implementation Status**
- **Next Steps**

Key Accomplishments for the Year

- Partner with NSF-funded OptIPuter Project - national leaders in optical WAN networking, distributed cluster computing, and mega-pixel visualization display research
  - » Early 10-GE connection with NLR/CAVEwave lambda at SC2004
  - » Free use of 10-Gbps WASH-STAR lambda
  - » OptIPuter networking with Scripps Institute of Oceanography
- Partner with NSF-funded DRAGON Project - national leaders in optical MAN networking research
  - » Two 10-Gbps and three 2.4-Gbps lambdas initially, of 40 possible
- Access to Other 10-Gbps NLR lambdas: Shared IP, GE VLANs, HOPI
- First 10-Gbps network within GSFC
- Leading NASA's way in NLR use for ARC's Project Columbia

DRAGON eVLBI Experiment Configuration



J. P. Gary
02/10/05

NASA GSFC Among First 10 Users of the NLR

• **GSFC's initial 10-Gbps connection to the NLR was enabled via cooperation with the National Science Foundation (NSF)-funded OptIPuter Project (**http://www.optiputer.net**) and the NLR (**http://www.nlr.net)

•**GSFC's initial 10-Gbps NLR connection was used to transmit Earth science data sets in real time to an OptIPuter 15-screen tiled display at the SC2004 conference in Pittsburgh, PA.**

• **"The involvement of NASA Goddard demonstrated the capabilities of NLR and showed just how researchers in 'big science' will need this kind of capacity to make new discoveries about aspects of our world and to help transfer this knowledge to practical uses by others in carrying out important tasks that improve our lives."**

  •**Tom West, President and CEO of the NLR**

NASA GSFC in the NLR booth with the OptIPuter-provided 15-screen tiled display cluster during SC2004

• **Earth science data sets created by GSFC's Scientific Visualization Studio were retrieved across the NLR in real time and displayed at the SC2004 in Pittsburgh**

• **Animated Earth (**http://aes.gsfc.nasa.gov/**) data sets were retrieved from OptIPuter servers in Chicago and San Diego and from GSFC servers in McLean, VA**

• **Land Information System (**http://lis.gsfc.nasa.gov/**) data sets were retrieved from OptIPuter servers in Chicago, San Diego, & Amsterdam**

NLR booth at SC2004 with OptIPuter-provided 15-screen tiled display cluster. Photo Source: Randall Jones, NASA GSFC

L-Net SC2004 Photo Gallery
http://esdcd.gsfc.nasa.gov/LNetphoto3.html

Interactive Retrieval and Hyperwall Display of Earth Sciences Images on a National Scale

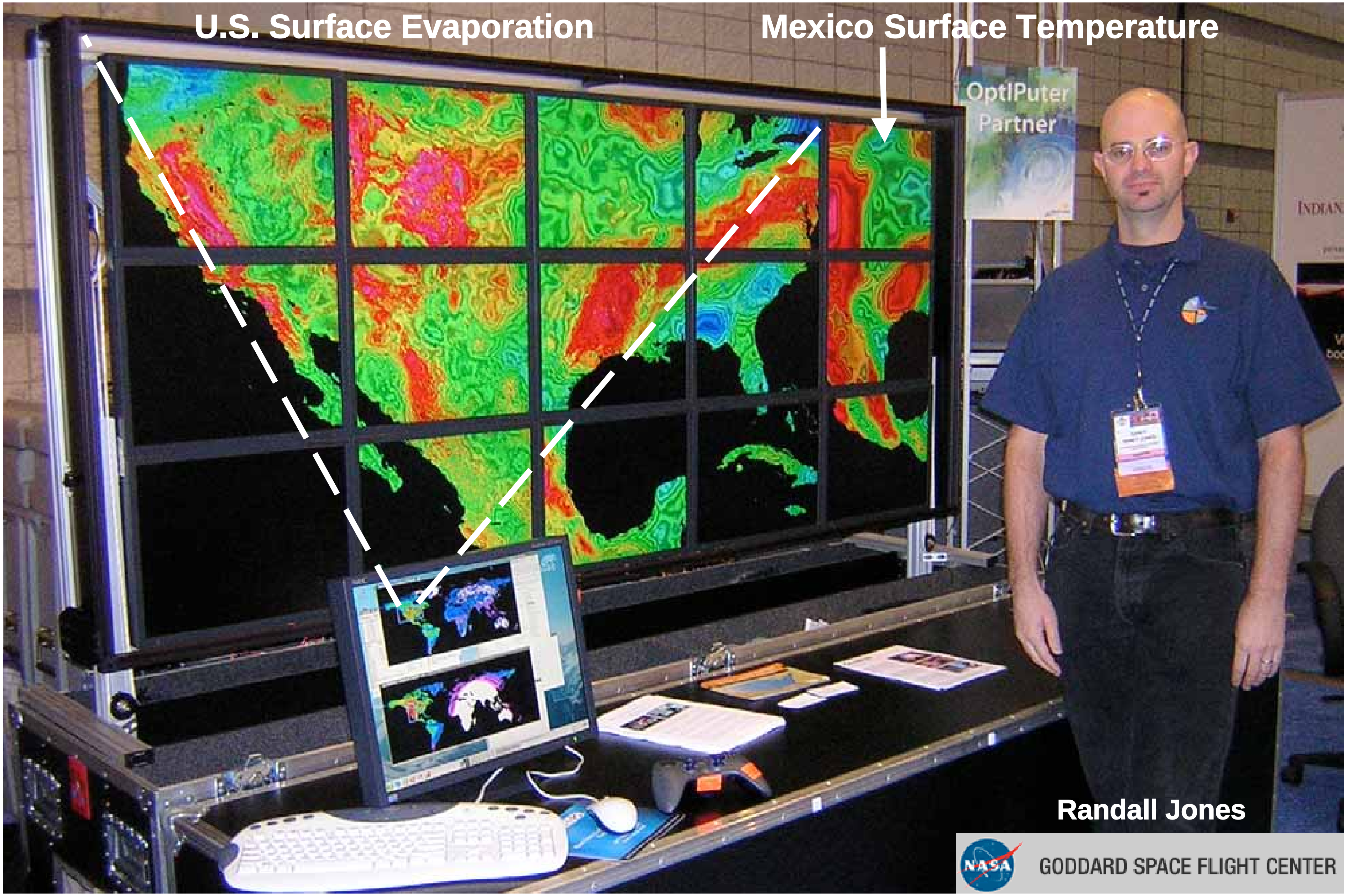
Enables Scientists To Perform Coordinated Studies Of Multiple Remote-Sensing Or Simulation Datasets

Source: Milt Halem & Randall Jones, NASA GSFC & Maxine Brown, UIC EVL

Earth science data sets created by GSFC's Scientific Visualization Studio were retrieved across the NLR in real time from OptIPuter servers in Chicago and San Diego and from GSFC servers in McLean, VA, and displayed at the SC2004 in Pittsburgh

OptIPuter

http://esdcd.gsfc.nasa.gov/LNetphoto3.html



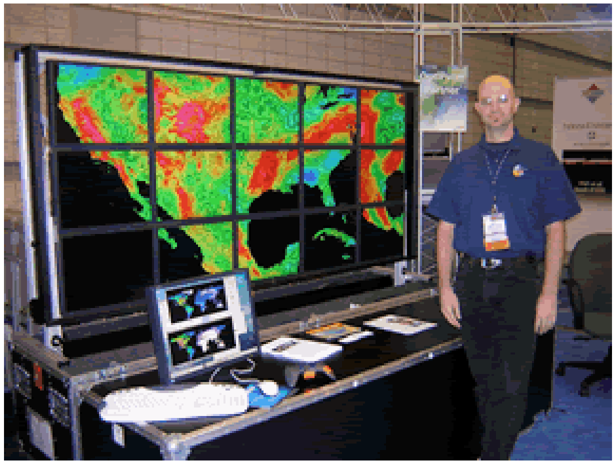
Global 1 km x 1 km Assimilated Surface Observations Analysis
Remotely Viewing ~ 50 GB per Parameter

NASA GSFC in the NLR booth with the OptIPuter-provided 15-screen tiled display cluster during SC2004

NLR booth at SC2004 with OptIPuter-provided 15-screen tiled display cluster.

Eric Sokolowsky (GST, Inc.) of GSFC's SVS interactively views model and observation data (set 1) from NASA's Animated Earth project with hyperwall paradigm.

Eric Sokolowsky (GST, Inc.) of GSFC's SVS with model and observation data (set 2) from NASA's Animated Earth project in hyperwall paradigm.



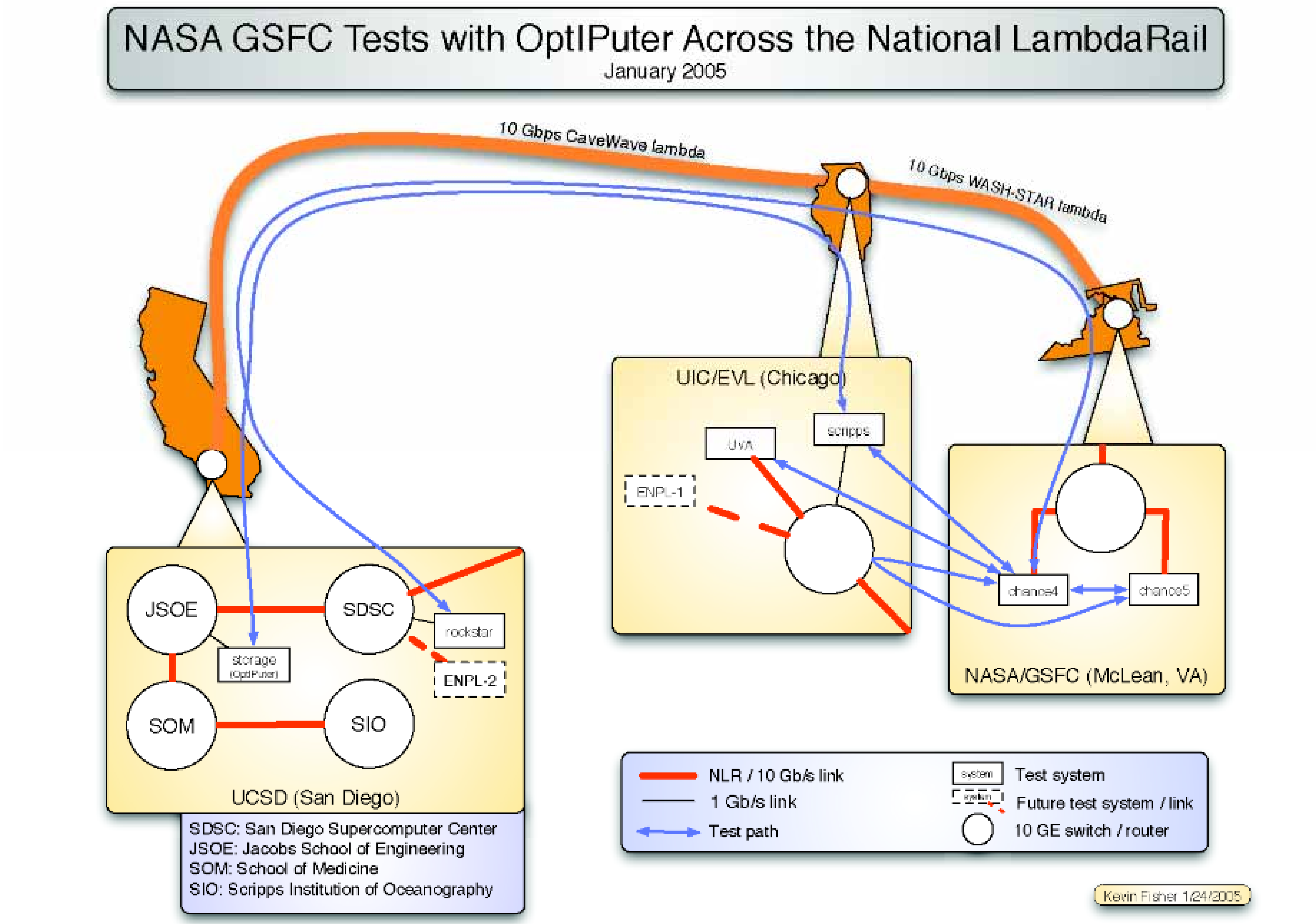
Randall Jones (GST, Inc.) of GSFC's SVS with model data from NASA's Land Information System in OptIPuter's display paradigm.

Various visitors to the NLR booth being briefed by Tom West, president and CEO of the NLR.

Rear view of the OptIPuter-provided 15-screen tiled display cluster.

L-Net SC2004 Photo Gallery: http://esdcd.gsfc.nasa.gov/LNetphoto3.html

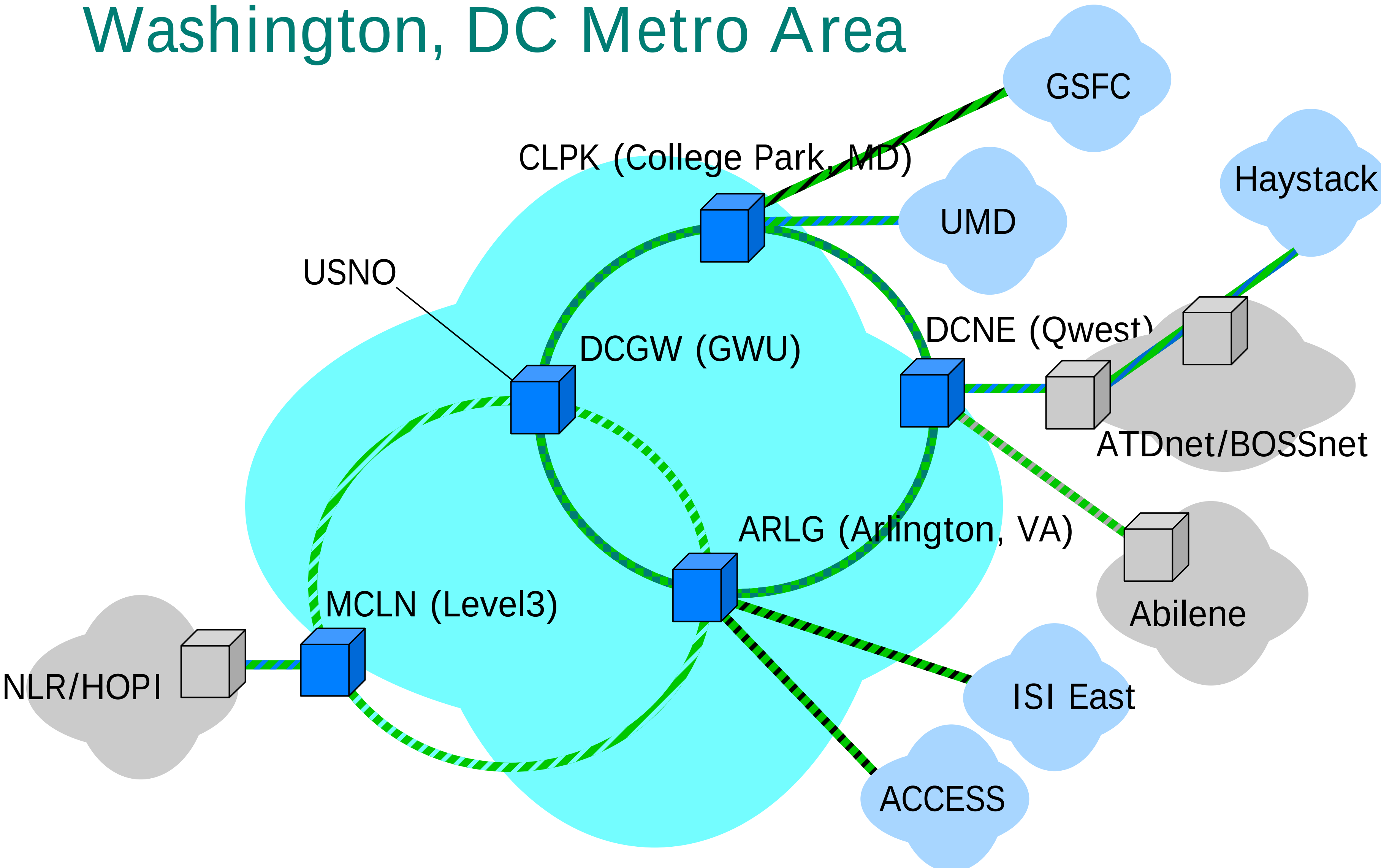
Photo Sources: Randall Jones, NASA GSFC


NASA GSFC Tests with OptIPuter Across the National LambdaRail
January 2005
10 Gbps CaveWave lambda
10 Gbps WASH-STAR lambda
UIC/EVL (Chicago)
UvA
scripps
ENPL-1
chance4
chance5
NASA/GSFC (McLean, VA)
JSOE
SDSC
rockstar
storage (OptIPuter)
ENPL-2
SOM
SIO
UCSD (San Diego)
SDSC: San Diego Supercomputer Center
JSOE: Jacobs School of Engineering
SOM: School of Medicine
SIO: Scripps Institution of Oceanography
NLR / 10 Gb/s link
1 Gb/s link
Test path
Test system
Future test system / link
10 GE switch / router
Kevin Fisher 1/24/2005

NASA GSFC Among First 10 Users of the NLR

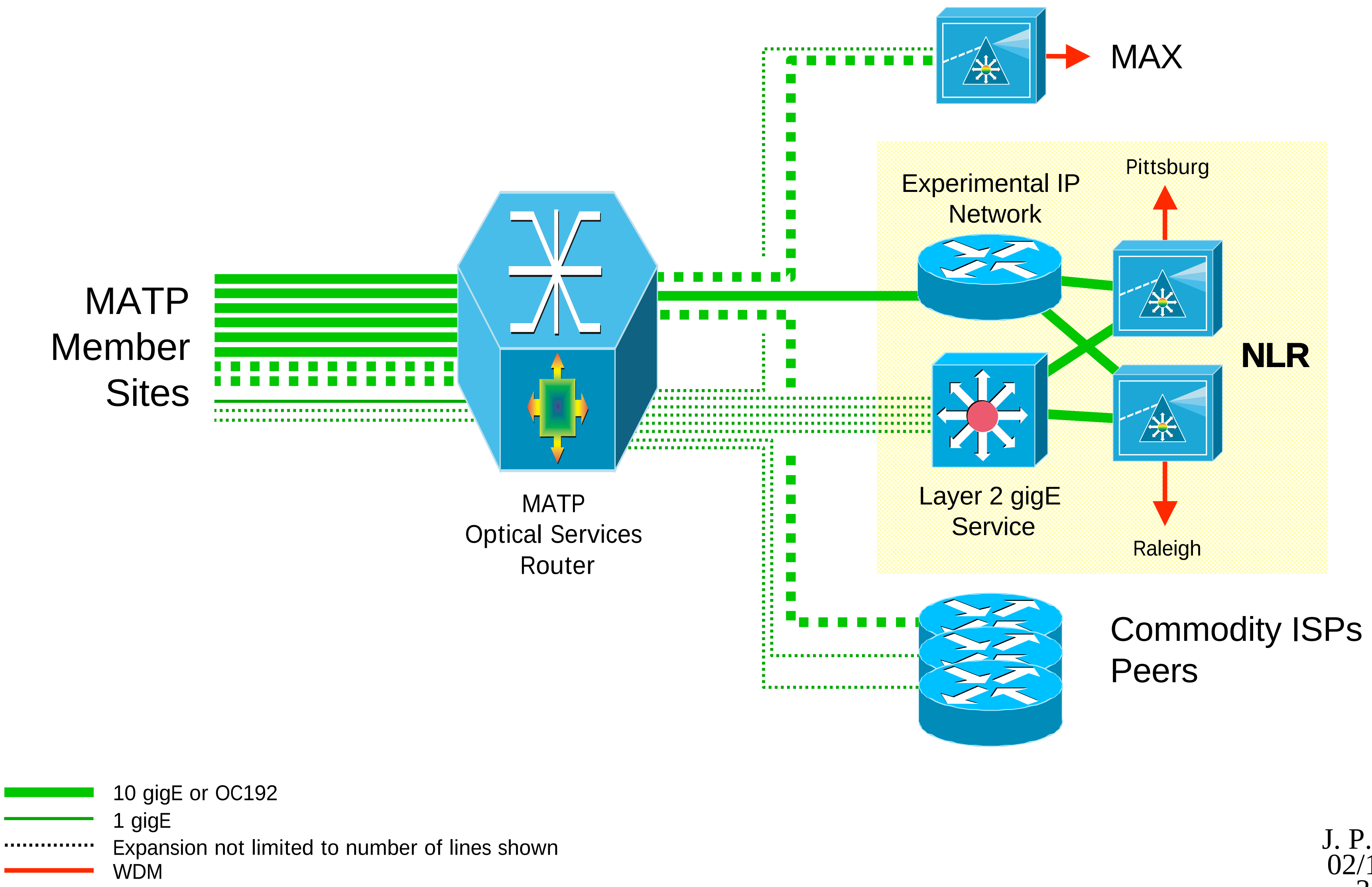
- Presently GSFC's computers connected to the NLR are located in the NLR suite at the Level3 Communications' optical fiber "carrier hotel" facility in McLean, VA
- In early March of 2005, two 10-Gbps connections will be enabled across the NSF-funded multi-wavelength Dynamic Resource Allocation via GMPLS Optical Network (DRAGON) research network (http://dragon.east.isi.edu)
- These DRAGON-based connections will link NLR/McLean with several high-performance computers at GSFC's main site in Greenbelt, MD, as well as with computers at other sites on the Washington, DC-area DRAGON
- Access to other 10-Gbps NLR lambdas is planned via membership in Mid-Atlantic Terascale Partnership (for the Shared IP and GE VLAN lambdas) and participation in Internet2's Hybrid Optical and Packet Infrastructure

DRAGON Network
Washington, DC Metro Area

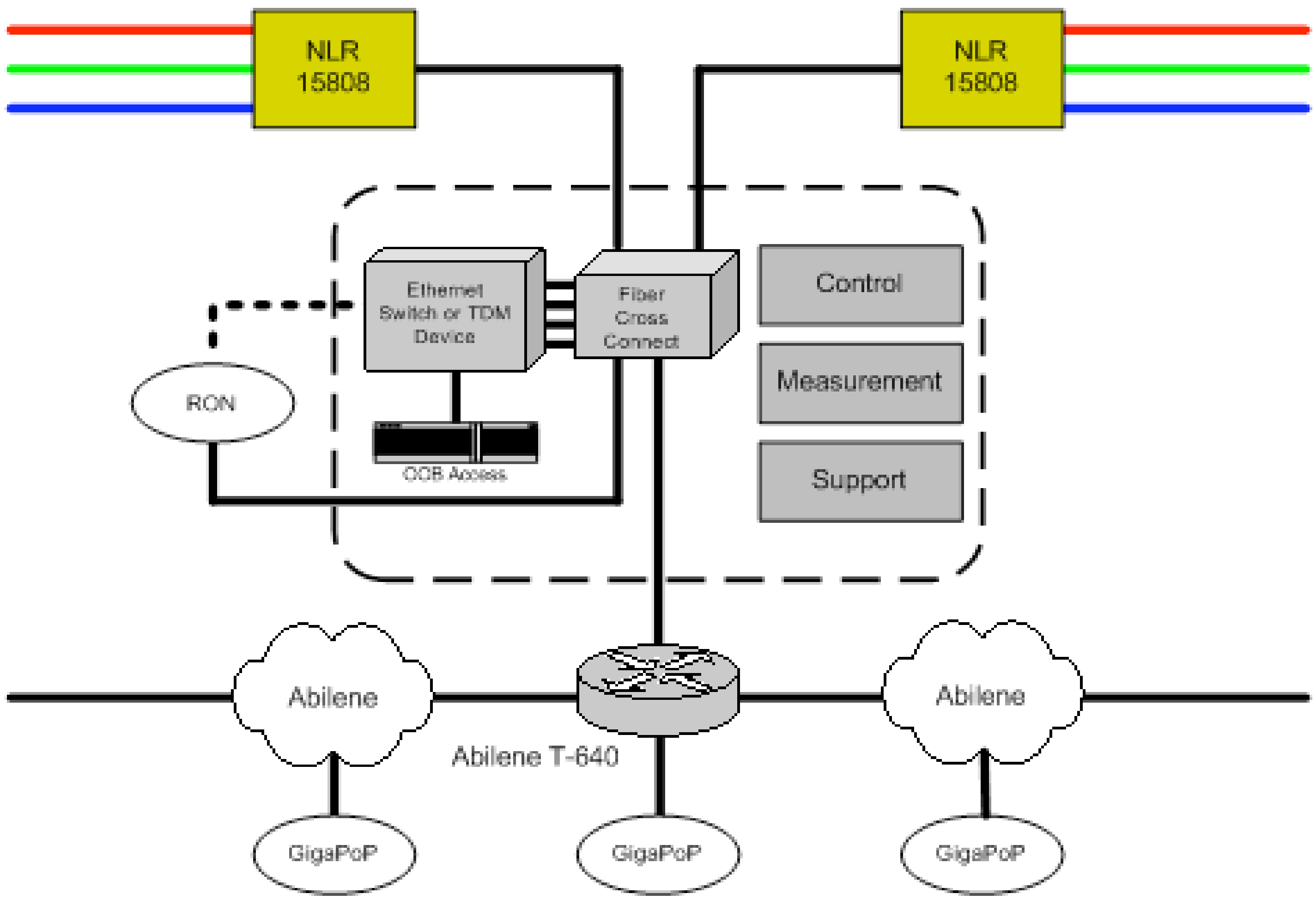


MATP Aggregation Facility Architecture

DRAFT



HOPI Node


NLR
15808
NLR
15808
Ethernet
Switch or TDM
Device
Fiber
Cross
Connect
Control
Measurement
Support
RON
OOB Access
Abilene
Abilene
Abilene T-640
GigaPoP
GigaPoP
GigaPoP

Outline for End of Year Review

- **Motivation**
- **Goals**
- **Key Challenges and Solution Designs**
- **Implementation Status**
- **Next Steps**

Future Work

- **MAP Core Integration LambdaGrid Infrastructure**
  - » **New science drivers and evaluators of NLR interconnection among USCD/SIO, UIC, GSFC, JPL, ARC**
    - – **Coordinated Earth Observing Program**
    - – **Hurricane Predictions**
    - – **Global Aerosols**
    - – **Remote viewing & Manipulation of Large Earth Science Data Sets**
    - – **Integration of Laser and Radar Topographic Data with Land Cover Data**
  - » **Collaboration among PI Larry Smarr (UCSD/Cal-(IT)2), Co-I's John Orcutt (UCSD/SIO), Tom DeFanti (UIC), Milt Halem (UMBC), and several scientists at GSFC, JPL, & ARC**

- **High-Speed Networking, Grid Computing, and Large-Scale Ensemble Simulations in Geodynamics, Weijia Kuang (GSFC), Shujia Zhou (GSFC) et al**

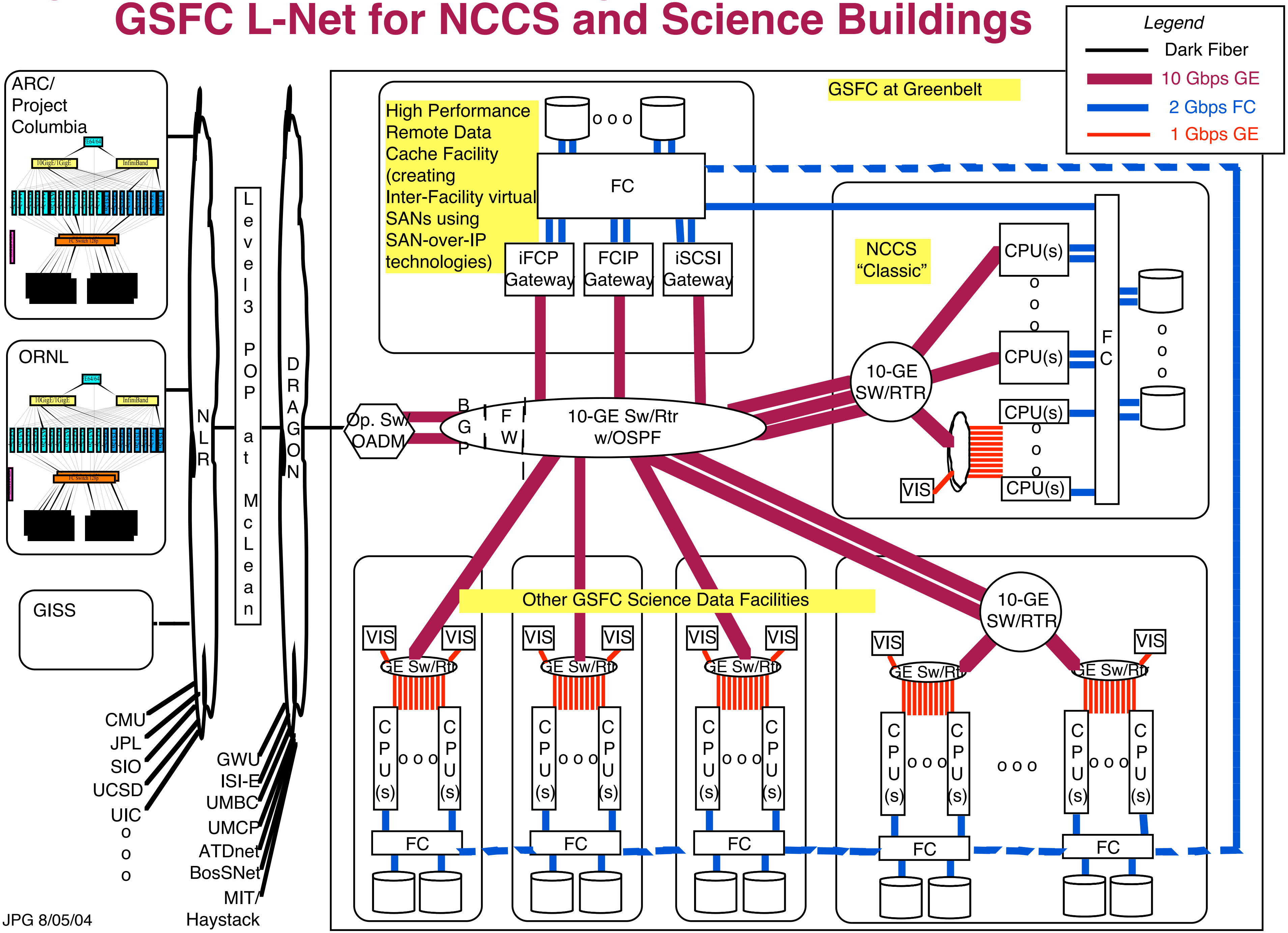
- **Expanding 10-GE L-Net**
  - » **More science buildings/clusters within GSFC; More NLR dedicated lambdas, e.g: ARC, ORNL, GISS; Wide Area SAN for NCCS; Optical switching within GSFC**

GSFC L-Net Enabling New NASA Science Needs

- New science drivers and evaluators of NLR interconnection among USCD/SIO, UIC, GSFC, JPL, ARC
  - » Coordinated Earth Observing Program
  - » Hurricane Predictions
  - » Global Aerosols
  - » Remote viewing & Manipulation of Large Earth Science Data Sets
  - » Integration of Laser and Radar Topographic Data with Land Cover Data

  Reference: "MAP Core Integration LambdaGrid Infrastructure" proposal, January 14, 2005
    - – PI: Larry Smarr (UCSD/Cal-(IT)2)
    - – Co-I's: John Orcutt (UCSD/SIO), Tom DeFanti (UIC), Milt Halem (UMBC)
- W. Kuang et al., "High Speed Networking and Large-Scale Simulation in Geodynamics", abstract/poster, Fall AGU 2004
- S. Zhou et al., "High-Speed Network and Grid Computing for High-End Computation: Application in Geodynamics Ensemble Simulations", submitted for 13th Annual Mardi Gras Conference, February 2005

High Performance Networking and Remote Data Access
GSFC L-Net for NCCS and Science Buildings



JPG 8/05/04

GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Special Acknowledgements

GSFC Internal

- IT Pathfinder Working Group
  - Chair: Dr. Milton Halem/Emeritus & UMBC
  - Applications Lead: Mike Seablom/610.3
  - Middleware Lead: Walt Truszkowski/588
  - Network Lead: Pat Gary/606.1
- High End Computer Network Team
  - Bill Fink/606.1
  - Kevin Kranacs/585
  - Paul Lang/ADNET/606.1
  - Aruna Muppalla/ADNET/606.1
  - Jeff Martz/CSC/606.2
  - Mike Stefanelli/CSC/606.2
  - George Uhl/SWALES/423
  - Steve Booth/SWALES/423
  - Kevin Fisher/586/UMBC coop

GSFC External

- National LambdaRail
  - CEO: Tom West
  - Net Eng Lead: Debbie Montano
- OptIPuter Project (NSF-funded)
  - PI: Dr. Larry Smarr/UCSD
  - Co-PI: Dr. Tom DeFanti/UIC
  - PM: Maxine Brown/UIC
  - UCSD Net Eng: Greg Hidley, Arron Chin, Phil Papodopolos
  - UCIC Net Eng: Alan Verlo, Linda Winkler
- DRAGON Project (NSF-funded)
  - PI: Jerry Sobieski/UMCP
  - Co-I: Tom Lehman/USC-ISI/E
  - Net Eng: Chris Tracy
- NASA Research and Education Network
  - DPM: Kevin Jones/ARC

GSFC FY04 IRAD Proposal "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

For More Info Related to the GSFC L-Net

- **DRAGON - Dynamic Resource Allocation via GMPLS Optical Networks**
  - » **http://dragon.east.isi.edu/**
  - » **http://www.itrd.gov/iwg/lsn/jet/conferences/20040413/jetroadmapworkshop03.pdf**
  - » **http://duster.nren.nasa.gov/workshop7/pps/04.Sobieski.DRAGON.ppt**

- **HOPI - Hybrid Optical and Packet Infrastructure**
  - » **http://networks.internet2.edu/hopi/**
  - » **http://www.itrd.gov/iwg/lsn/jet/conferences/20040413/jetroadmapworkshop10.pdf**
  - » **http://duster.nren.nasa.gov/workshop7/pps/11.Winkler.HOPI.ppt**

- **Level3 - Level 3 Communications, Inc**
  - » **http://www.level3.com/**
  - » **http://www.itrd.gov/iwg/lsn/jet/conferences/20040413/jetroadmapworkshop31.pdf**

For More Info Related to the GSFC L-Net (continued)

- **MATP - Mid-Atlantic Terascale Partnership**
  - » **http://www.midatlantic-terascale.org/**
- **Movaz Networks**
  - » **http://www.movaz.com/**
  - » **http://www.itrd.gov/iwg/lsn/jet/conferences/20040413/jetroadmap workshop14.pdf**
- **NLR - National Lambda Rail**
  - » **http://www.nlr.net/**
  - » **http://www.itrd.gov/iwg/lsn/jet/conferences/20040413/jetroadmap workshop43.pdf**
  - » **http://duster.nren.nasa.gov/workshop7/pps/10.Farber.LambdaRail .ppt**

For More Info Related to the GSFC L-Net (continued)

- **OptIPuter**
  - **http://www.optiputer.net/**
  - **Smarr, L., Chien, A., DeFanti, T., Leigh, J., and Papadopoulos, P., "The OptIPuter", Communications of the ACM, November 2003, Vol. 46, No. 11, pp. 59-67; http://delivery.acm.org/10.1145/950000/948410/p58-smarr.pdf?key1=948410&key2=4972772901&coll=ACM&dl=ACM&CFID=25879183&CFTOKEN=50598646**
  - **http://duster.nren.nasa.gov/workshop7/pps/07.Smarr.OptIPuter.ppt**

GSFC FY04 IRAD Proposal "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Principal Investigator & Co-Investigators

- **Name:** **Pat Gary (930) & Jeff Smith (585) Co-PI's & GSFC's Information Technology Pathfinder Working Group (ITPWG) as Co-I's**
- **Organizations:** **Code 420, Code 580, Code 920, & Code 930**
- **Telephone:** **301-286-9539 & 301-614-5038 for Co-PI's**
- **E-mail:** **Pat.Gary@nasa.gov, Jeff.Smith@nasa.gov for Co-PI's**

Project Website

- **http://esdcd.gsfc.nasa.gov/IRAD_Lambda.html**

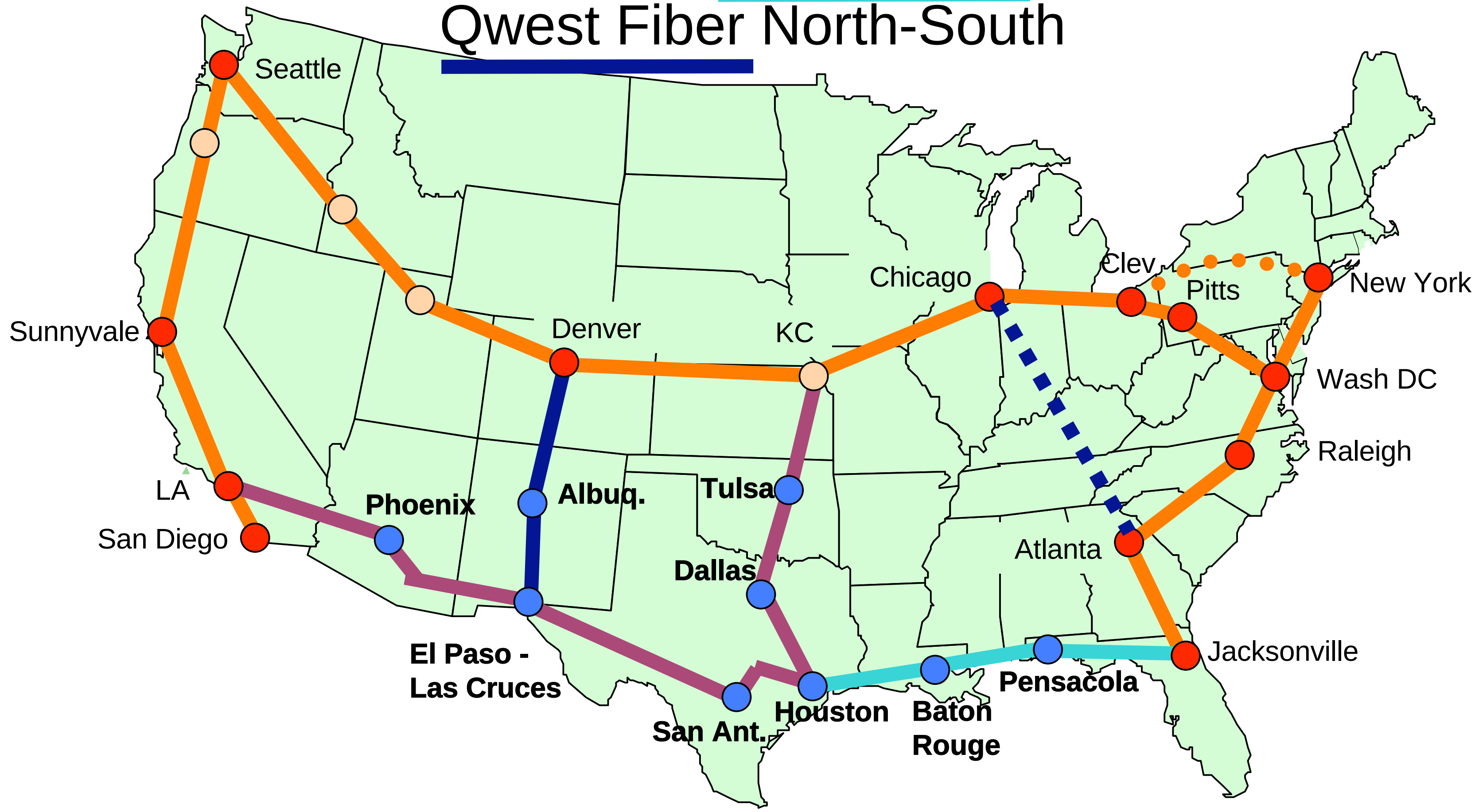
Backup Slides

NLR – Phase 1 +
Route to NY & ORNL Waves & Salt Lake City

NLR – Proposed Southern Route

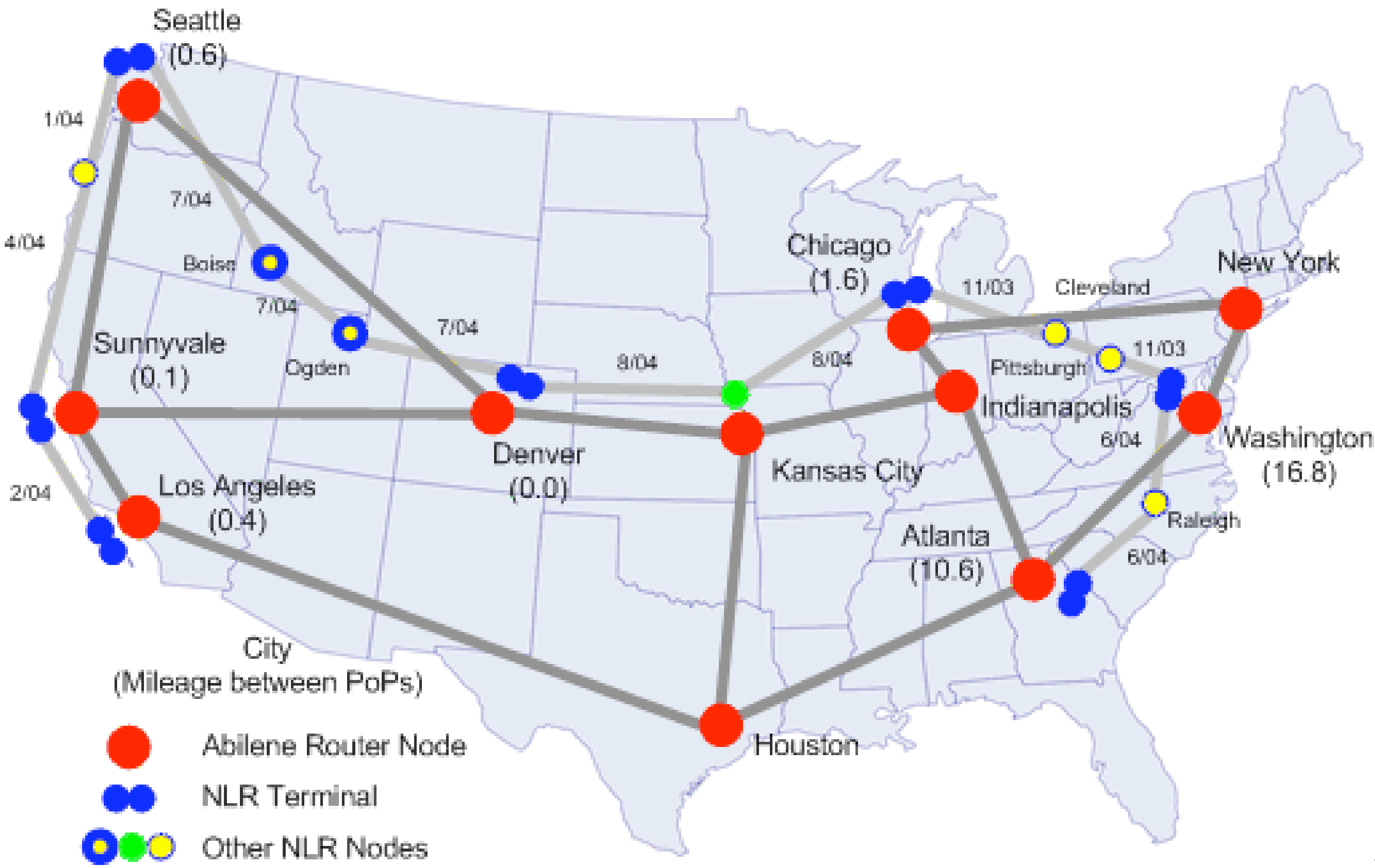
Wiltel Fiber in the West -- AT&T Fiber in the Southeast

Qwest Fiber North-South



Using Cisco 15454 DWDM gear

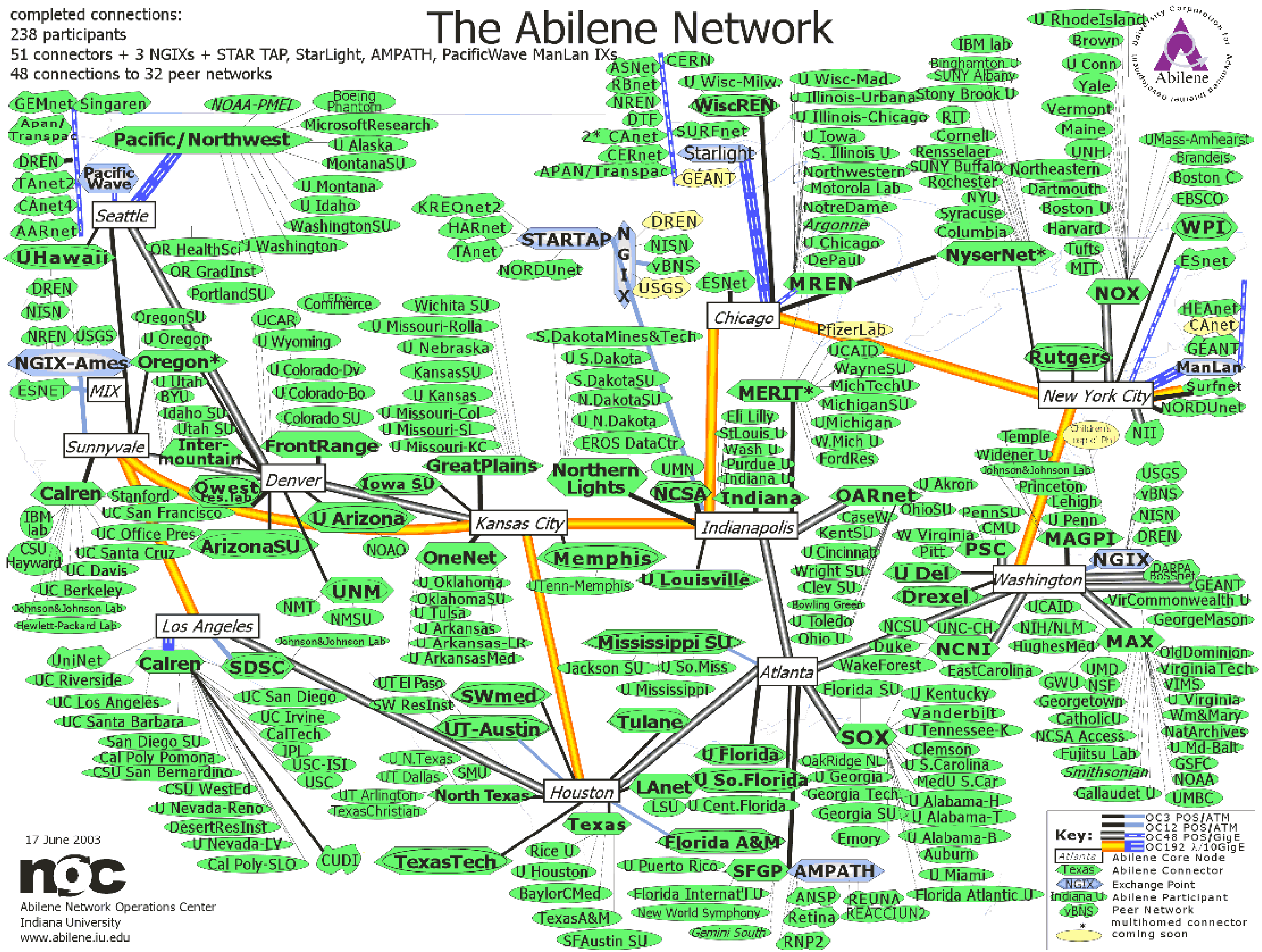
Abilene/NLR Map



J. P. Gary
02/10/05

The Abilene Network

completed connections:
238 participants
51 connectors + 3 NGIXs + STAR TAP, StarLight, AMPATH, PacificWave ManLan IXs
48 connections to 32 peer networks



17 June 2003

noc
Abilene Network Operations Center
Indiana University
www.abilene.iu.edu

NLR Members as of 19Jul04

- CENIC
- Pacific Northwest GigaPOP
- Pittsburgh Supercomp. Center
- Duke (coalition of NC univers.)
- Mid-Atlantic Terascale Partnership
- Cisco Systems
- Internet2
- Florida LambdaRail
- Georgia Institute of Technology
- Committee on Institutional Cooperation (CIC)
- Texas / LEARN
- Cornell
- Louisiana Board of Regents
- University of New Mexico
- Oklahoma State Regents
- UCAR/FRGP

Plus Agreements with:

- SURA (AT&T fiber donation)
- Oak Ridge National Lab (ORNL)

(Courtesy of Dr. Larry Smarr/UCSD)
Goddard Space Flight Center IRAD

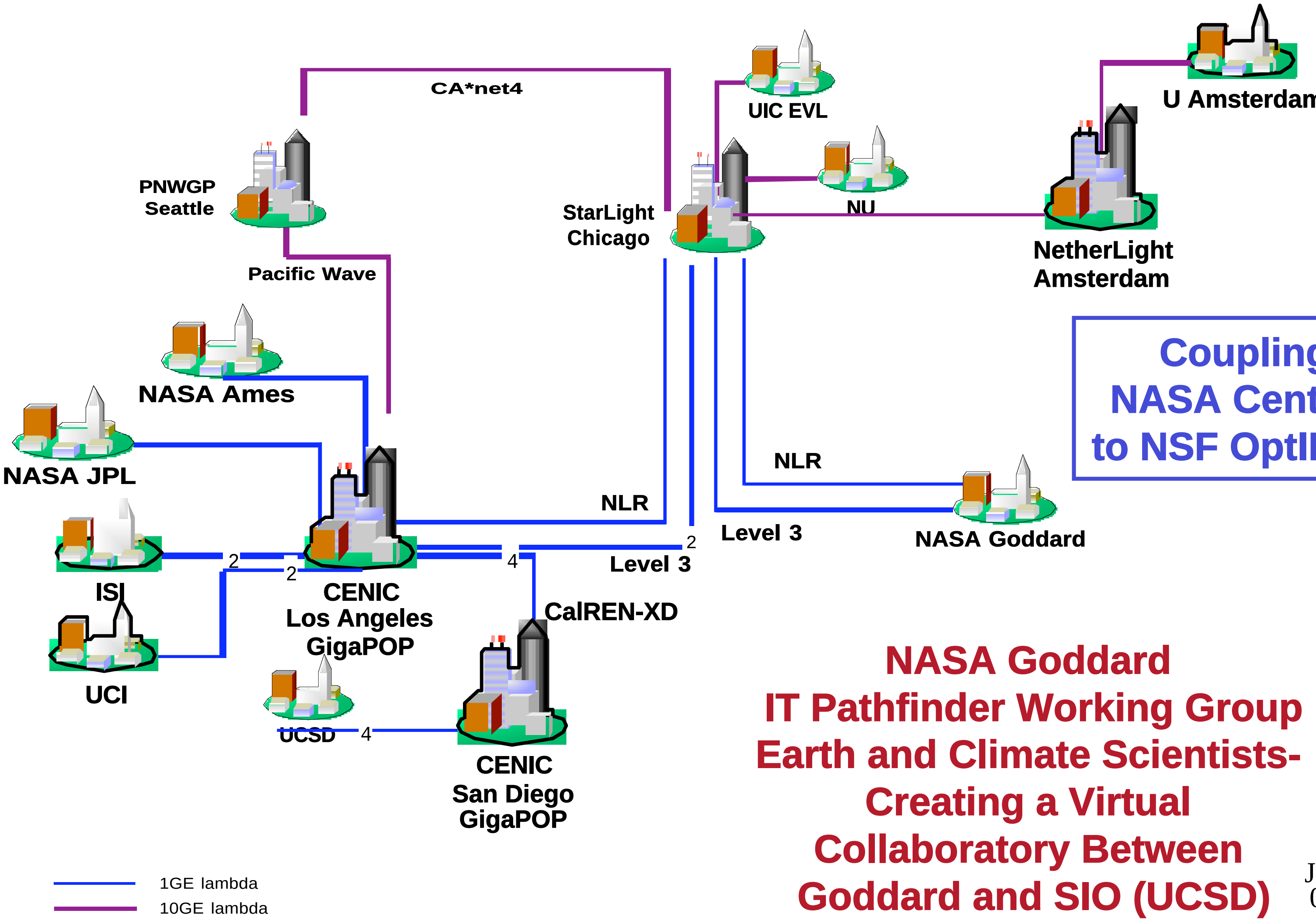
- Establish a “Lambda Network”
  - » Use optical wavelength technology (10Gbps Ethernet per wavelength)
  - » Connect GSFC’s Earth Science Greenbelt facility in Maryland to Scripps Institution of Oceanography in through the University of California at San Diego (UCSD)
  - » Ride the National LambdaRail
- Synergies
  - » Falls in with reinvigorating the nation’s cyber infrastructure
  - » Takes advantage of next generation networking technologies (Lambda-Nets)
  - » Makes use of NSF funded compute, storage and visualization resources being implemented at UCSD (OptIPuter and GEON)
- Benefits
  - » Develop real-time interactive collaborations with leading Earth and space science academic institutions
  - » Enable scientists at both institutions to share and use compute intensive community models, complex data base mining and multi-dimensional streaming visualization
  - » Creates a virtual laboratory and SIO wing within GSFC’s Building 33
  - » Several NASA missions benefit from the high bandwidth connection, eg:
    - – HPC between ARC/Project Columbia and GSFC/NCCS
    - – CEOP data analyses between Rhodes (SIO) and Bosilovich (GSFC)

(Courtesy of Dr. Larry Smarr/UCSD)

GSFC FY04 IRAD Proposal

- "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure”
- Transcontinental, Regional, and Local Networking
  - » Became a member of the NLR, with NREN Project assistance, through a Mid-Atlantic Terascale Partnership membership arrangement
  - » Deploying GMPLS-managed Movaz Optical Switches and Add Drop Muxs at GSFC and in DRAGON regional multiwavelength network
  - » Utilize NLR lambdas between GSFC and UCSD/SIO
  - » Interconnect GSFC’s Thunderhead and other clusters to UCSD/SIO’s OptIPuter network at 10GigE via this Lambda Network
- Application Development
  - » Integrate Earth System Modeling Framework software with GRID middleware by constructing prototype interfaces between the components
  - » Identify requirements for new methods and/or messages that would be desirable for supporting GSFC models and data assimilation

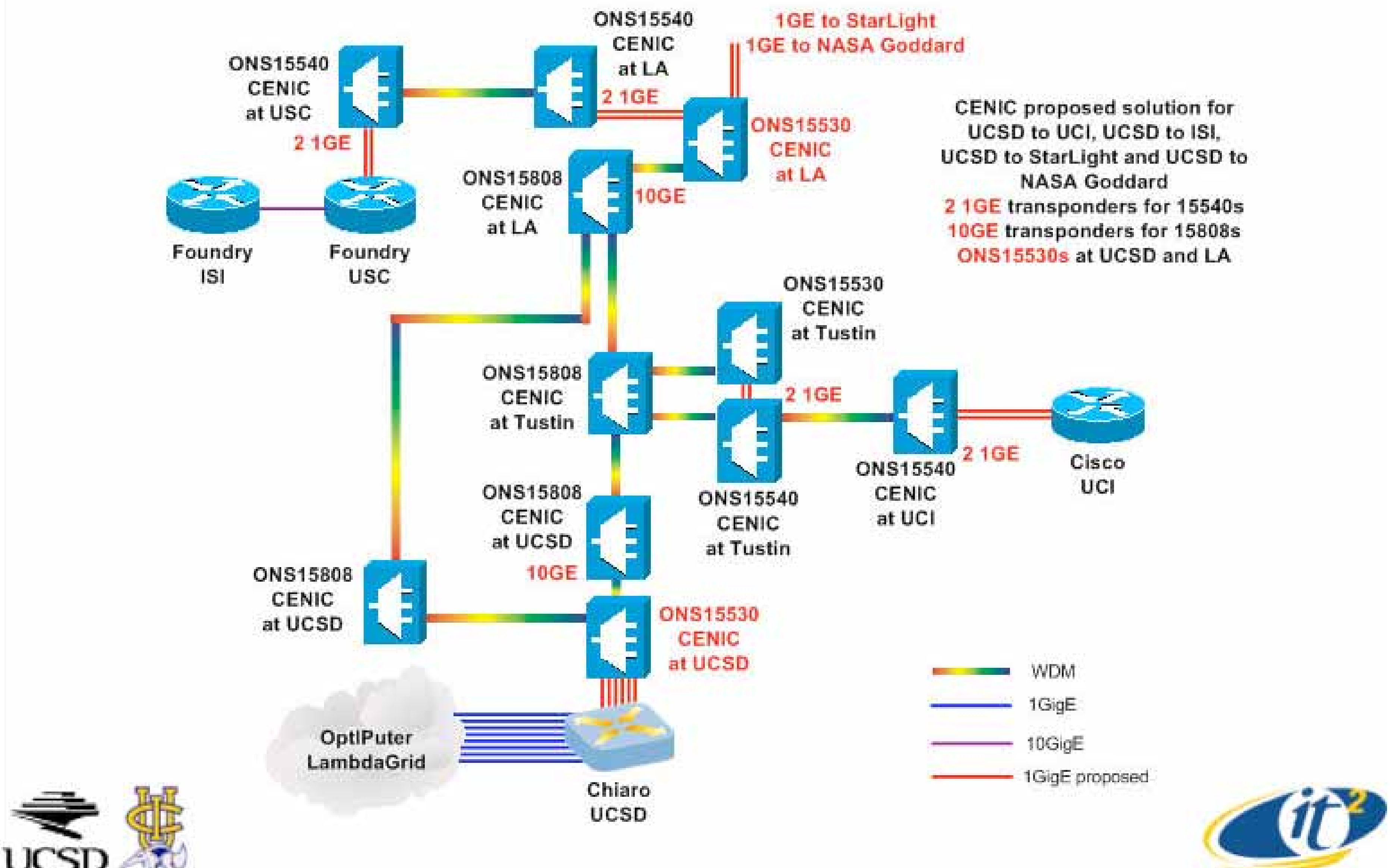
R&D Test: Move to Internet Protocol Over Dedicated Optical Lightpaths



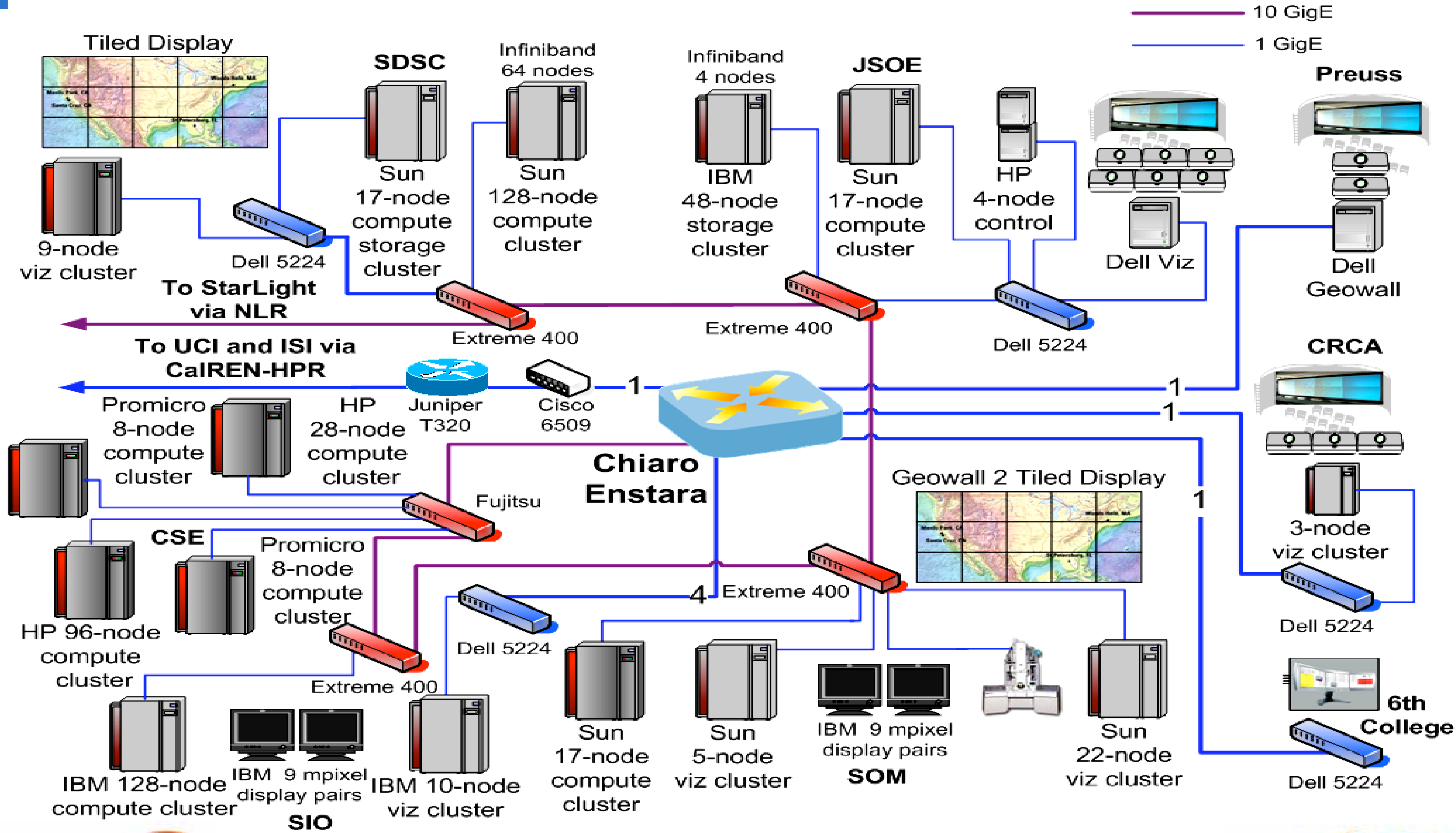
Coupling NASA Centers to NSF OptIPuter

NASA Goddard IT Pathfinder Working Group Earth and Climate Scientists- Creating a Virtual Collaboratory Between Goddard and SIO (UCSD)

Southern California CalREN-XD Build Out



UCSD Packet Test Bed
OptIPuter Year 2



NASA GSFC in the NLR booth with the OptIPuter-provided 15-screen tiled display cluster during SC2004

Goddard Space Flight Center's connection to the National LambdaRail sent Earth science data sets in real time to a 15-screen tiled display at the SC2004 conference in Pittsburgh, PA. Above, Eric Sokolowsky (GST, Inc.) of the Scientific Visualization Studio interactively views model and observation data from NASA's Animated Earth project.

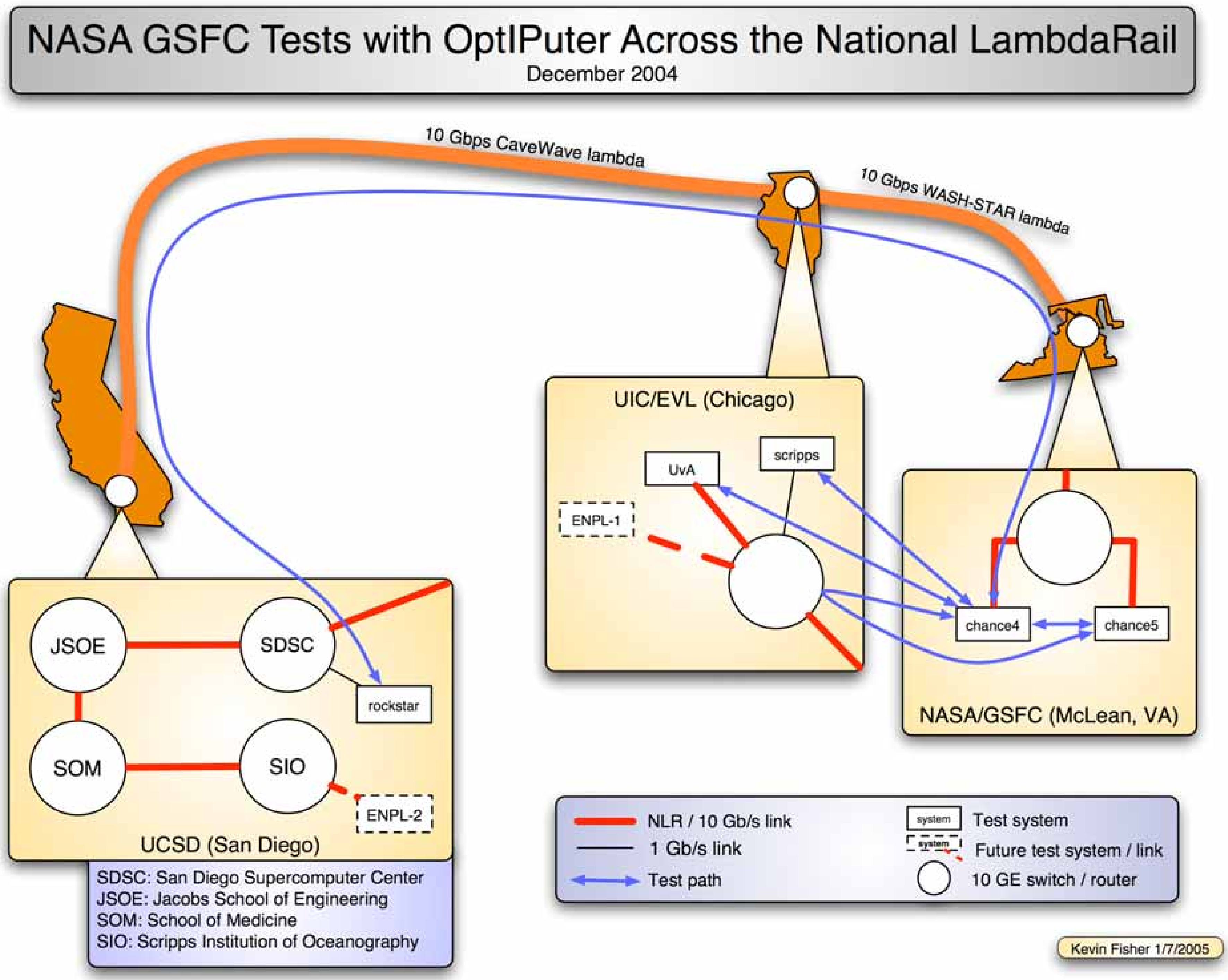
http://esdcd.gsfc.nasa.gov/L-Netpdfs/L-Net_NLR_TechBrief.pdf

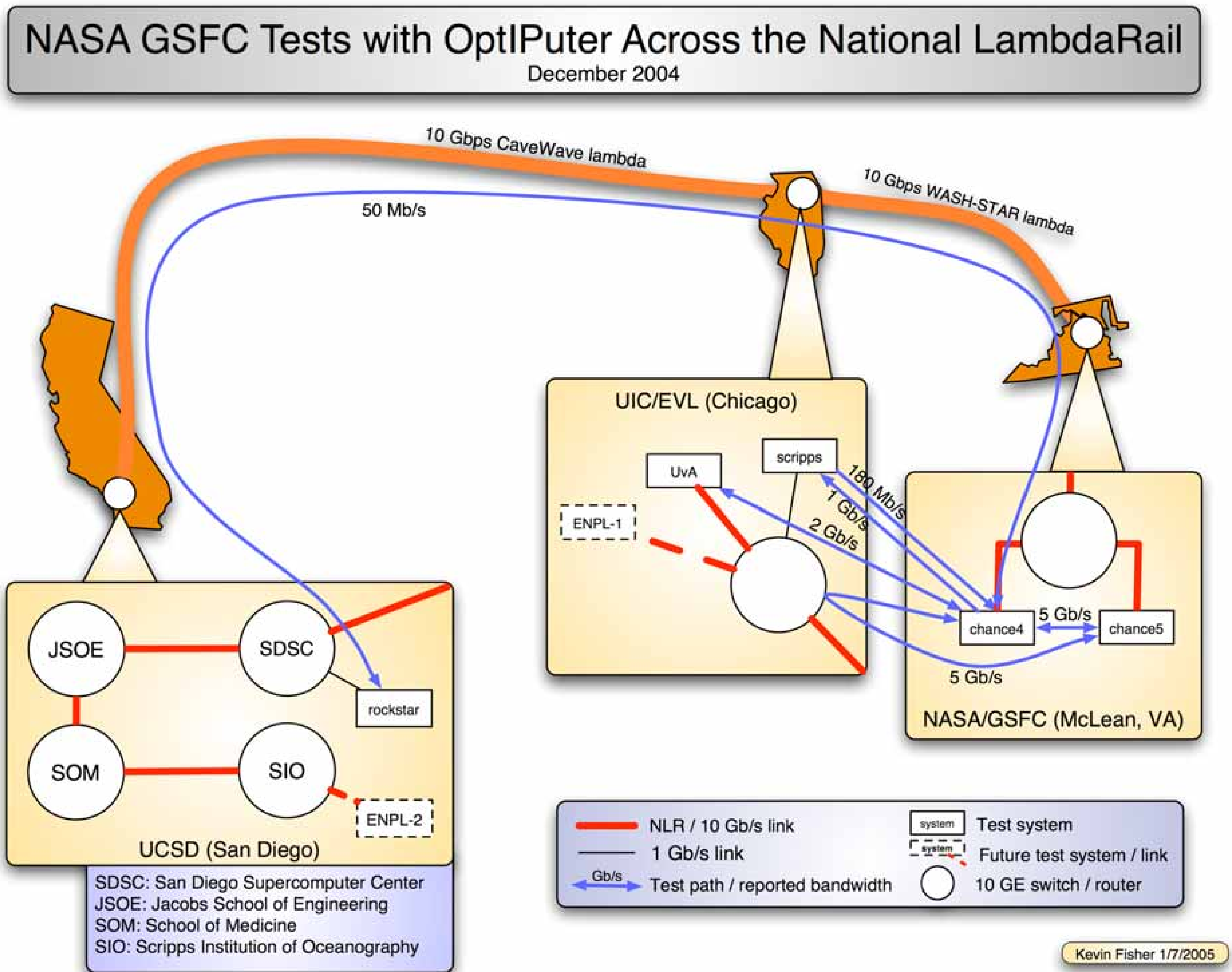
NASA GSFC Among First 10 Users of the NLR

• On 2Dec04 NLR's Layer 1 Network Operations Center activated the 10-Gbps WASH-STAR lambda directly connecting the McLean-based GSFC 10-GE switch/router and servers with the University of Illinois Chicago (UIC)/Starlight facility

• OptIPuter's personnel provided static routes on their previously existing NLR/CaveWave lambda and GSFC network test user accounts on their NLR /CaveWave-connected servers, respectively at UIC/StarLight and University of California San Diego (UCSD)

• Using the above and GSFC's two 10-GE NLR-connected servers located in McLean, GSFC's Bill Fink installed his nuttcp network throughput performance testing software (ftp://ftp.lcp.nrl.navy.mil/pub/nuttcp/), and conducted initial memory-to-memory data flow tests between various pairs of these servers


NASA GSFC Tests with OptIPuter Across the National LambdaRail
December 2004
10 Gbps CaveWave lambda
10 Gbps WASH-STAR lambda
UIC/EVL (Chicago)
UvA
scripps
ENPL-1
chance4
chance5
NASA/GSFC (McLean, VA)
JSOE
SDSC
rockstar
SOM
SIO
ENPL-2
UCSD (San Diego)
SDSC: San Diego Supercomputer Center
JSOE: Jacobs School of Engineering
SOM: School of Medicine
SIO: Scripps Institution of Oceanography
NLR / 10 Gb/s link
1 Gb/s link
Test path
system
Test system
system
Future test system / link
10 GE switch / router
Kevin Fisher 1/7/2005



NASA GSFC Tests with OptIPuter Across the National LambdaRail
December 2004
10 Gbps CaveWave lambda
50 Mb/s
10 Gbps WASH-STAR lambda
UIC/EVL (Chicago)
UvA
scripps
ENPL-1
180 Mb/s
1 Gb/s
2 Gb/s
chance4
5 Gb/s
chance5
5 Gb/s
NASA/GSFC (McLean, VA)
JSOE
SDSC
rockstar
SOM
SIO
ENPL-2
UCSD (San Diego)
SDSC: San Diego Supercomputer Center
JSOE: Jacobs School of Engineering
SOM: School of Medicine
SIO: Scripps Institution of Oceanography
NLR / 10 Gb/s link
1 Gb/s link
Gb/s
Test path / reported bandwidth
system
Test system
system
Future test system / link
10 GE switch / router
Kevin Fisher 1/7/2005

Fibers Used in DRAGON-based Approach

- **"Link 1" GSFC -- UMD at College Park (UMCP): Fibers leased from FiberGate became available to be characterized on 10Jun04, were extended to GSFC building 28 on 22Jun04, were characterized on 13Jul04, and initially used in DRAGON on 1Sep04**

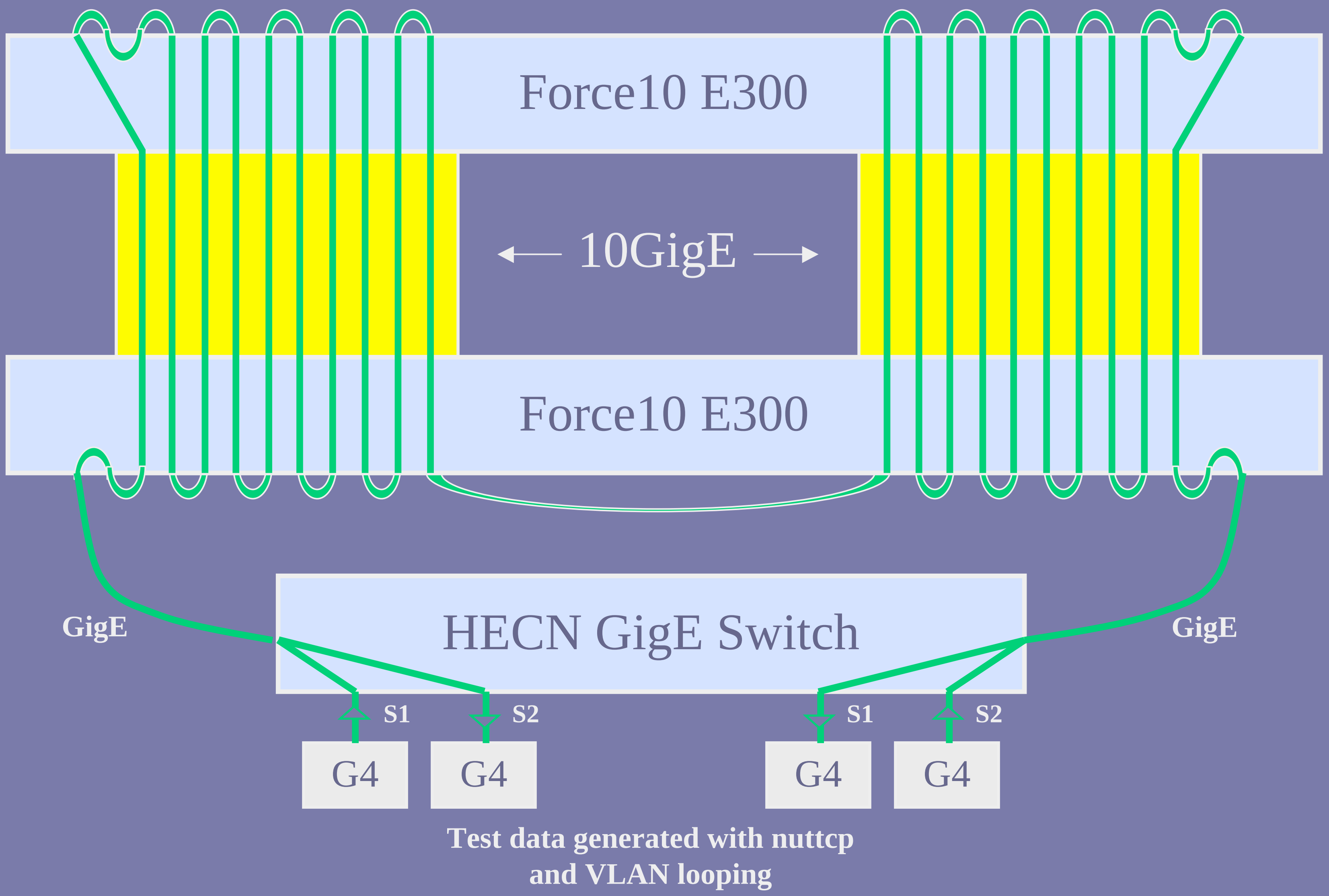
- **"Link 2" UMCP -- GWU in DC -- ISI/E at Arlington --Qwest POP at Eckington in DC -- UMCP: Fibers leased from Qwest became available to be characterized on 28Jun04, were characterized during the week of 12Jul04, and initially used in DRAGON during the week of 5Sep04**

- **"Link 3" Level(3) POP in McLean -- GWU -- ISI/E -- Level(3) POP in McLean: Fibers leased from Level3. Level(3) POP -- GWU fibers are expected to become available to be characterized during the week of 1Mar05, characterized the week of 1Mar05, and initially used in DRAGON the week of 1Mar05. GWU -- ISI/E -- Level(3) POP fibers are expected during the week of 1May05**

Testing of Force10 E300 10-GE Switch Capability Completed by Bill Fink (930) and Paul Lang (ADNET)

- **Throughput testing**
- **Link Aggregation (two 10GigE links, similar physical and VLAN as above)**
- **Initial QoS testing (port based)**
- **VLAN Stacking**

Force10 10GigE Testing



Force10 10GigE nuttcp Stress Test
(Over 1 Petabyte in 3 days*)

S1: porthos (2x800 MHz PowerMac G4) -> clifford (867 MHz PowerMac G4)
S2: underdog (867 MHz PowerMac G4) -> bigdog (2x1 GHz G4 XServe)

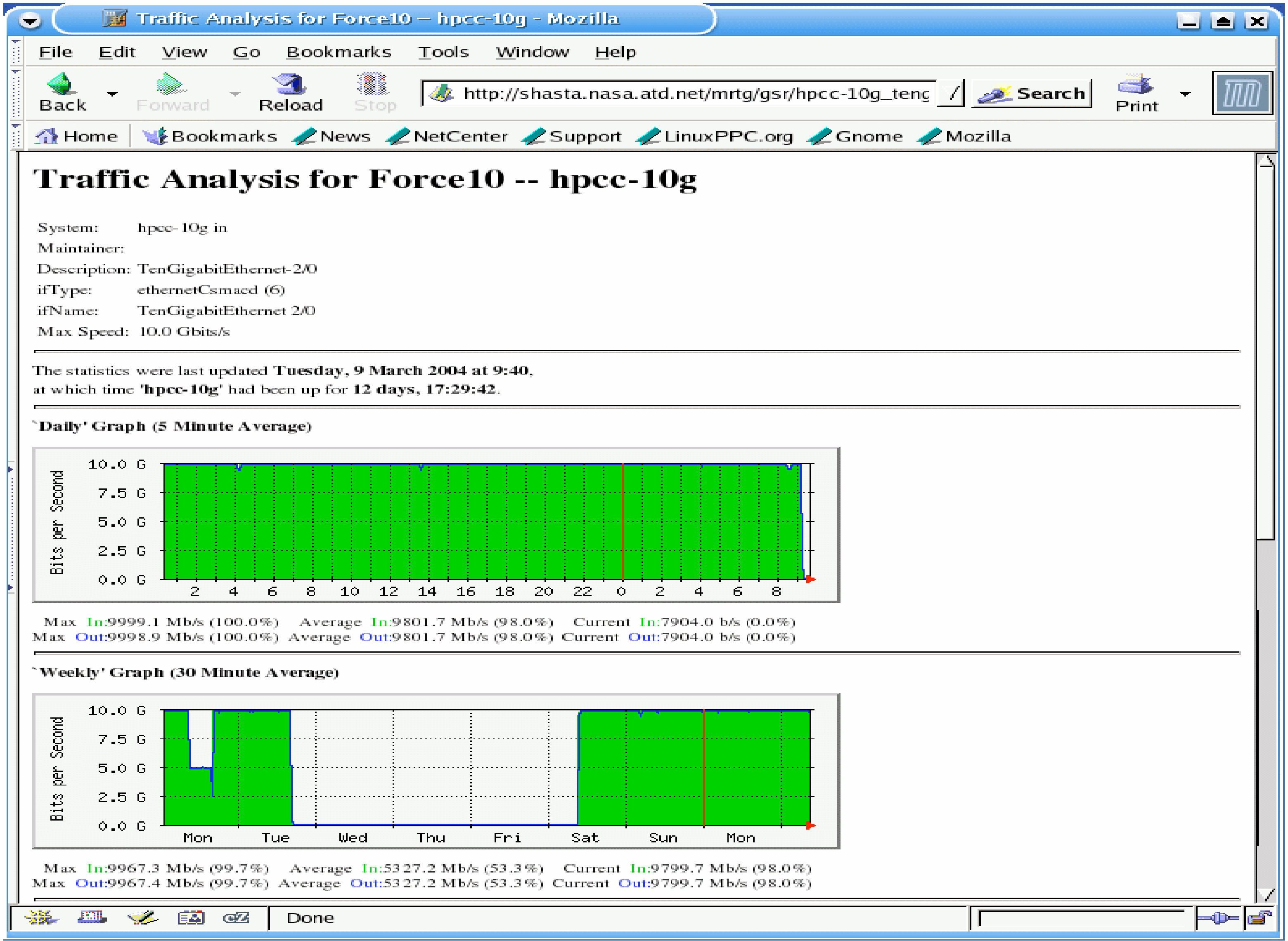
All systems running YellowDog Linux with Intel Pro/1000 Server Adapter or NetGear GA620T NIC, and using a 9K Jumbo Frame MTU

```
clifford% nuttcp -u -r -T72h -w2048 -Iporthos2clifford porthos & \
          nuttcp -u -T72h -w2048 -Iunderdog2bigdog underdog bigdog
underdog2bigdog:  30687801.2734 MB / 259772.86 sec =  990.9732 Mbps 74 %TX 29 %RX
                  1950222 / 3929988785 drop/pkt 0.05 %loss
porthos2clifford: 30695954.5781 MB / 259773.58 sec =  991.2337 Mbps 65 %TX 17 %RX
                  6056720340 / 9985802526 drop/pkt 60.65 %loss
```

(30687801.2734 MB + 30695954.5781 MB)*10*2/1024/1024/1024 = 1.14336 PB

* Over 0.25 PB transferred bidirectionally across each 10GigE link

Force10 Petabyte Challenge



Graph for TenGigabitEthernet-3/0 basically identical

GSFC FY04 IRAD Proposal "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Key Initial Features of GSFC L-Net Design

- **GSFC Local Network Part**
  - **10-GE (Gigabit Ethernet) connections for:**
    - **Thunderhead cluster in B28**
    - **SVS Hyperwall and/or GSFC SAN Pilot interface in B28**
    - **Network test stations in HECN lab in B28 and ENP lab in B32**
    - **Optical switch from UMBC/Ray Chen**
  - **10-GE ports for future connections:**
    - **NCCS in B28**
    - **One other cluster such as Houser's in B33**
    - **Others can be easily planned**
  - **Four 1-GE connections with GSFC's Science and Engineering Network**

GSFC FY04 IRAD Proposal "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

Key Features of GSFC L-Net Design (continued)

- **Regional Network Part**
  - » **Two 10-GE connections with DRAGON at GSFC in Greenbelt**
  - » **Two 10-GE connections with DRAGON at Level3 POP in McLean**
  - » **Two 10-GE and multiple 1-GE connections for network test stations at Level3 POP in McLean**
- **Transcontinental Network Part**
  - » **10-GE connection with NLR/MATP's IP Backbone and Switched Ethernet lambdas**
  - » **10-GE connection with NLR/Internet2's HOPI lambda**

GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure"

<u>Major Significance (1 of 2)</u>

- **Partner with NSF-funded OptIPuter Project**
  - » **Collaboration with national leaders in optical WAN networking, distributed cluster computing, and mega-pixel visualization display research**
  - » **Early 10-GE connection with NLR/CAVEwave lambda for SC2004**
  - » **Free use of 10-Gbps WASH-STAR lambda**
  - » **OptIPuter networking with Scripps Institute of Oceanography**
- **Partner with NSF-funded DRAGON Project**
  - » **Collaboration with national leaders in optical MAN networking research**
  - » **Two 10-Gbps and three 2.4-Gbps lambdas initially, of 40 possible**
- **Access to Other 10-Gbps NLR lambdas**
  - » **Shared IP and GE VLANs via membership in Mid-Atlantic Terascale Partnership**
  - » **Internet2's Hybrid Optical and Packet Infrastructure**

GSFC FY04 IRAD Project "Preparing Goddard for Large Scale Team Science in the 21st Century: Enabling an All Optical Goddard Network Cyberinfrastructure”

Major Significance (2 of 2)

- **First 10-Gbps network within GSFC: inter- and intra-buildings connecting with science user compute/storage/visualization clusters**
- **Enabling new NASA science needs**
  - » **Coordinated Earth Observing Program**
  - » **Hurricane Predictions**
  - » **Global Aerosols**
  - » **Remote viewing & Manipulation of Large Earth Science Data Sets**
  - » **Integration of Laser and Radar Topographic Data with Land Cover Data**
  - » **Large-Scale Geodynamics Ensemble Simulations**
- **Leading NASA’s way in NLR use for ARC’s Project Columbia**