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

Interim Status Report

J. Patrick Gary

Network Projects Leader

Earth and Space Data Computing Division

NASA Goddard Space Flight Center

pat.gary@nasa.gov

301-286-9539

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

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

Summary of Purpose

- "…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."

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

Statement of Objectives

1. Introducing elements of the NSF cyberinfrastructure and associated middleware to the research community at Goddard

2. Demonstrating how the lambda net and accompanying middleware can help to build "virtual communities" for science research

3. Developing and nurturing collaborations with the UCSD university consortium, particularly with SIO, one of the key consortium members, which will be the focus of the prototype

4. Extending the capabilities of the Earth System Modeling Framework (ESMF) such that interoperability between models, model components, geospatial services, and other functions can exist across distributed computing platforms

5. Understanding and resolving technical issues associated with the lambda network and demonstrate its capabilities to the Goddard community

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

Initial Year Planned Accomplishments/Milestones

1. Formalize a working relationship with key members of Code 900 for this pilot demonstration. Quantify requirements for the L-Net and acquire the system components and make the NLR partnership arrangements to support an end-to-end science scenario test. The scenario will make use of ESMF software tools and methods to enable rapid integration of the ensembles of climate data predictions on the fly across the country.

2. Establish requirements for a working network system internal to GSFC. Also establish a design for prototyping the interfaces between components of the network system. The design should also accommodate a fully functional system.

3. Meet with key members of the SIO/UCSD and GMAO/ESMF teams to discuss technical issues. Address requirements for new methods and/or messages that would be desirable for supporting GSFC models and data assimilation. Also establish a working relationship with the network vendors and define requirements for utilization of the L-Net.

4. Execute the prototype L-Net, demonstrating the functionality of the APIs supporting both the SIO end and the GSFC end.

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

Technical Approach

- Transcontinental, Regional, and GSFC Local Networking
  - » Become a member of the NLR in collaboration with the NREN Project, or directly through a Mid-Atlantic Terascale Partnership membership arrangement
  - » Connect GSFC with the NLR at McLean via fibers or lambdas from Level3 and/or DRAGON
  - » Use NLR's initial lambdas for science/application experiments between GSFC and UCSD/SIO
  - » Enable a connection from GSFC's *Thunderhead* cluster in building 28 across the NLR to other clusters such as those at SIO or others on UCSD's OptIPuter network first at 1-GE, then 10-GE

- 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

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

Overall Schedule Given 2Mar04 "effective start" (I.e., date of permission to use awarded funds)

- By 31May04 demonstrate initial use of 10-GE and multi-wavelength network technologies within GSFC in test of L-Net designs

- By 30Aug04 with DRAGON project demonstrate use of optical network technology switching among three 2.4 Gbps wavelengths among GSFC, UMCP, and USC/ISI-E in Arlington, VA

- By 30Nov04 connect GSFC at 10 Gbps with NLR at McLean, VA

- By 28Feb05 demonstrate science/applications between GSFC and UCSD/SIO across the NLR

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

Significant Accomplishments To Date
Acknowledgements

- Most contributions to this IRAD's accomplishments and plans were achieved through the Subgroups of GSFC's Information Technology Pathfinder Working Group (ITPWG), chaired by Milton Halem (900 Emeritus)

    » Applications Subgroup, chaired by Mike Seablom (586)

    » Middleware Subgroup, chaired by Walt Truszkowski (588)

    » Network Infrastructure Subgroup, chaired by Pat Gary (930)

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

Significant Accomplishments To Date
Application Development

- Characterizing new science/applications to be enabled by the "Lambda Network" (L-Net), e.g.:

  » With GEON PI's from UCSD and UTEP
  - Videoconference   9Sep03
  - GSFC visit   16Oct03

  » With SIO senior scientists and UCSD-based OptIPuter PI
  - Videoconference   24Mar04
  - Summary write-up of CEOP application   11Apr04

  » With GISS' Jim Hansen (940) and Gavin Schmidt (940)
  - GISS visit by Seablom and Tom Clune (930)   2Apr04
  - Planned ensemble forecasting using lambda network / grid computing
  - Received matrix of experiments that will be executed by Greg Toth (586)

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

Significant Accomplishments To Date

Application Development (continued)

- Proposal for "UMBC Research Support for the GSFC Cyberinfrastructure Initiative" received   13May04

  - » Period of performance: 15Jun04 - 14Jun05

  - » PI: Dr. Milton Halem

  - » Provides UMBC faculty and graduate student support to assist Goddard in demonstrating the usefulness of the L-Net technology

    - – Kevin Fisher under the guidance of Dr. Padma Mundur

    - – Marcella Wilson under the guidance of Dr. Yelena Yesha

    - – TBD under the guidance of Dr. Y. J. (Ray) Chen

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

Significant Accomplishments To Date

Detailing L-Net Links: Three Major Parts

- Transcontinental Network Between San Diego, CA and McLean, VA

- Regional Network Between McLean and GSFC

- GSFC Local Network

Considerations for Transcontinental Backbone Network
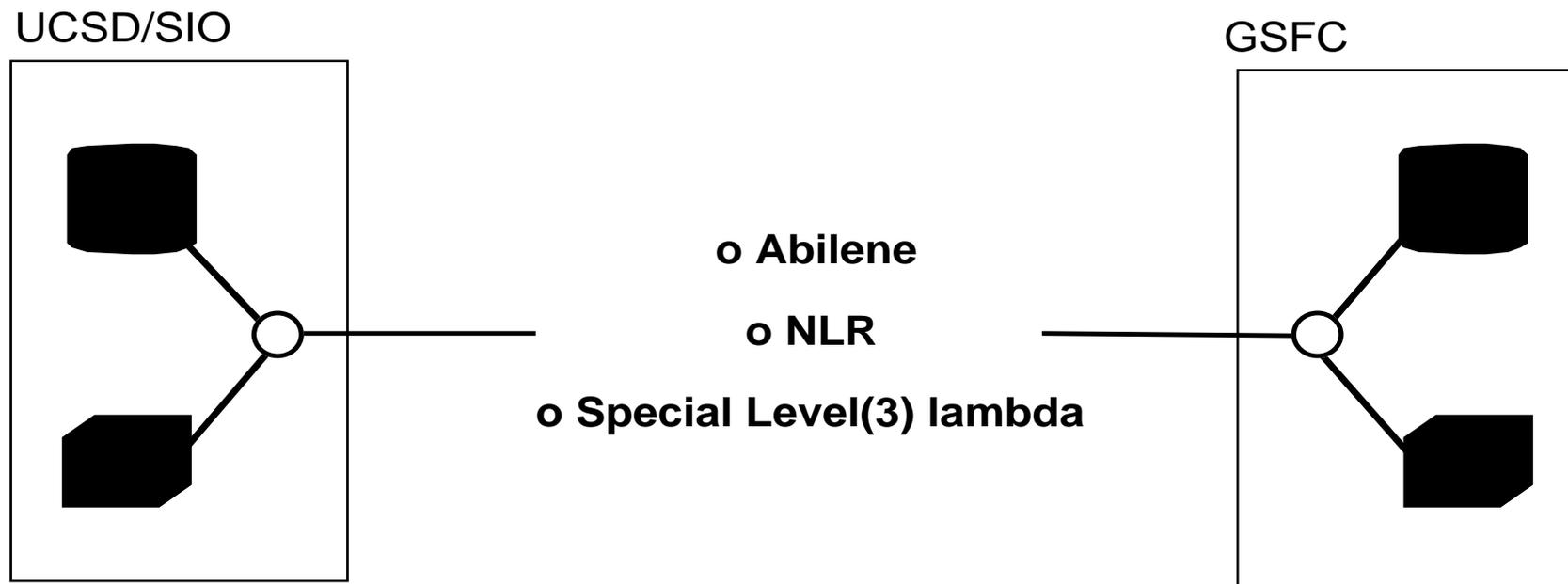
UCSD/SIO

GSFC

o **Abilene**

o **NLR**

o **Special Level(3) lambda**

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

Significant Accomplishments To Date

Detailing L-Net Links

- Transcontinental Network Between San Diego, CA and McLean, VA

  » Started establishing baseline throughput performance using Internet2's Abilene network

  » Obtained free use of 2 dedicated 1 Gbps VLAN's among GSFC-UCSD-UIC from Level3 Communications, starting in Jul04 for at least two months

  » Confirmed access to the National LambdaRail's (NLR's) 10 Gbps IP Backbone and National Switched Ethernet "lambdas" with only yearly funding of 5 year membership in the Mid-Atlantic Terascale Partnership, starting in Sep04

  » Initiated discussions to participate in Internet2's Hybrid Optical and Packet Infrastructure experiments on the NLR's third 10 Gbps lambda

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

Abilene Network Operations Center
Indiana University
www.abilene.iu.edu
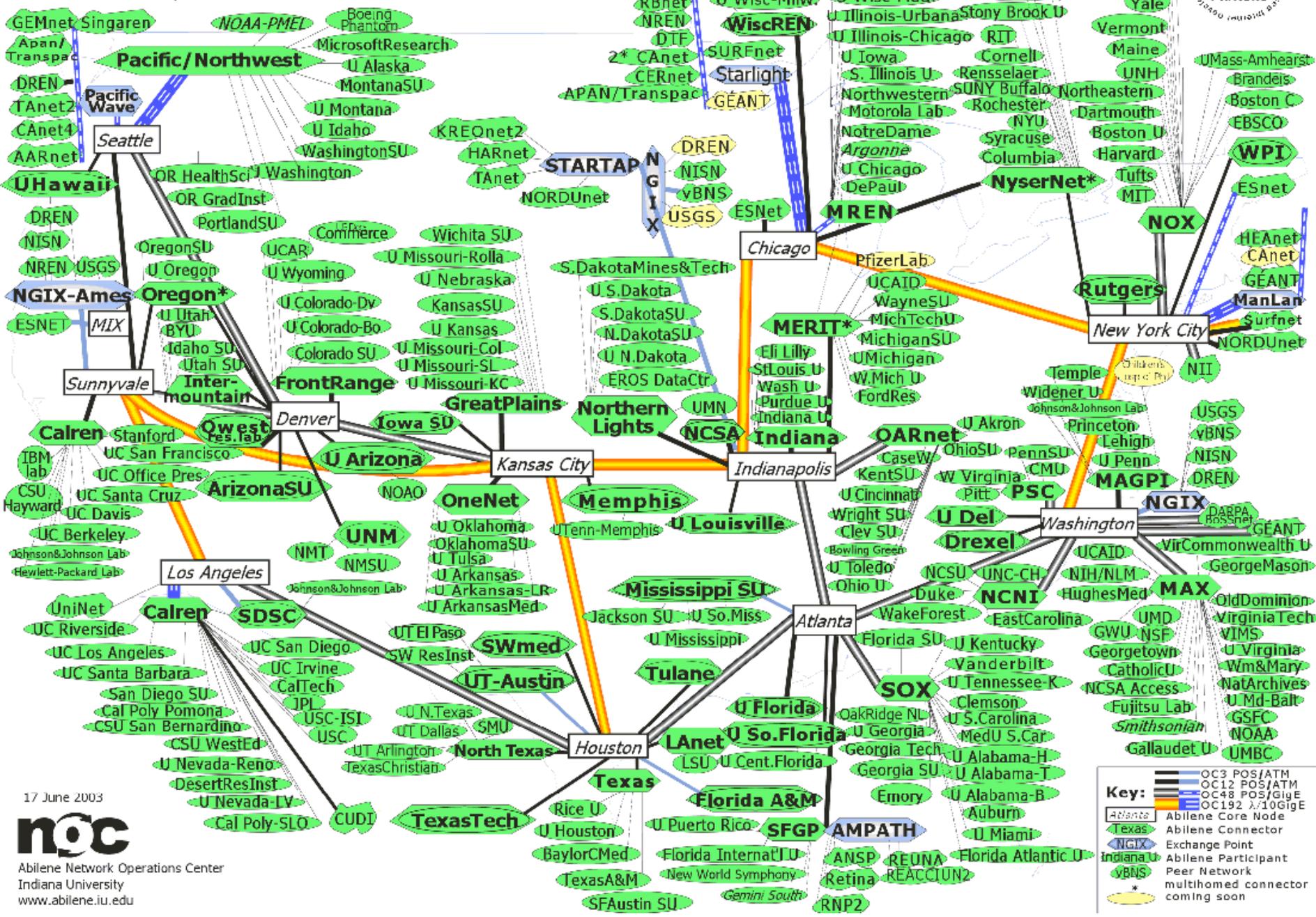
Key:
- OC3 POS/ATM
- OC12 POS/ATM
- OC48 POS/GigE
- OC192 λ/10GigE
- Abilene Core Node
- Abilene Connector
- Exchange Point
- Abilene Participant
- Peer Network
- multihomed connector
- coming soon

Network-based Limitations of Abilene Removed with NLR

- Applications traffic must be IP-based

- 1 GE present limits at access POP's

- Shared 10 GE backbone

- Typically 13 store-and-forward router hops between GSFC and SIO; ~75 msec RTT

- Private addresses of UCSD's OptIPuter not advertised via Abilene

# 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)

J. P. Gary
05/21/04
16

CALit2 Donation Design

Level (3) COMMUNICATIONS

it2

Los Angeles Level 3 Gateway	Chicago Level 3 Gateway	Washington DC Level 3 Gateway

GbE   GbE   GbE   GbE   GbE   GbE

San Diego Super Computer	Starlight 710 North Lake Shore Dr	NASA/Goddard

Key Points:
1) Redundant 3Flex GbE ports in each City
2) Each Port should be directly connected to **separate** switches
3) Each VC should be mapped to highest available customer queues
4) Each VC should be set to full CDRs
5) Local Loops provided by CALit2

Legend

VLAN 1	← →
VLAN 2	← →
VLAN 3	← →
CPA	← →

National LambdaRail (http://www.nationallambdarail.org/)
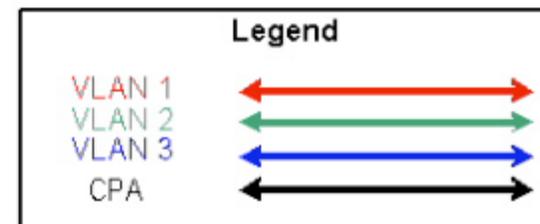
- Provide an enabling network infrastructure for new forms and methods for research in science, engineering, health care, and education as well as for research and development of new Internet technologies, protocols, applications and services.

- Provide the research community with direct control over a nationwide optical fiber infrastructure, enabling a wide range of facilities, capabilities and services in support of both application level and networking level experiments and serving diverse communities of computational scientists, distributed systems researchers and networking researchers.

J. P. Gary
05/21/04
18

Considerations for Regional Access Network for GSFC
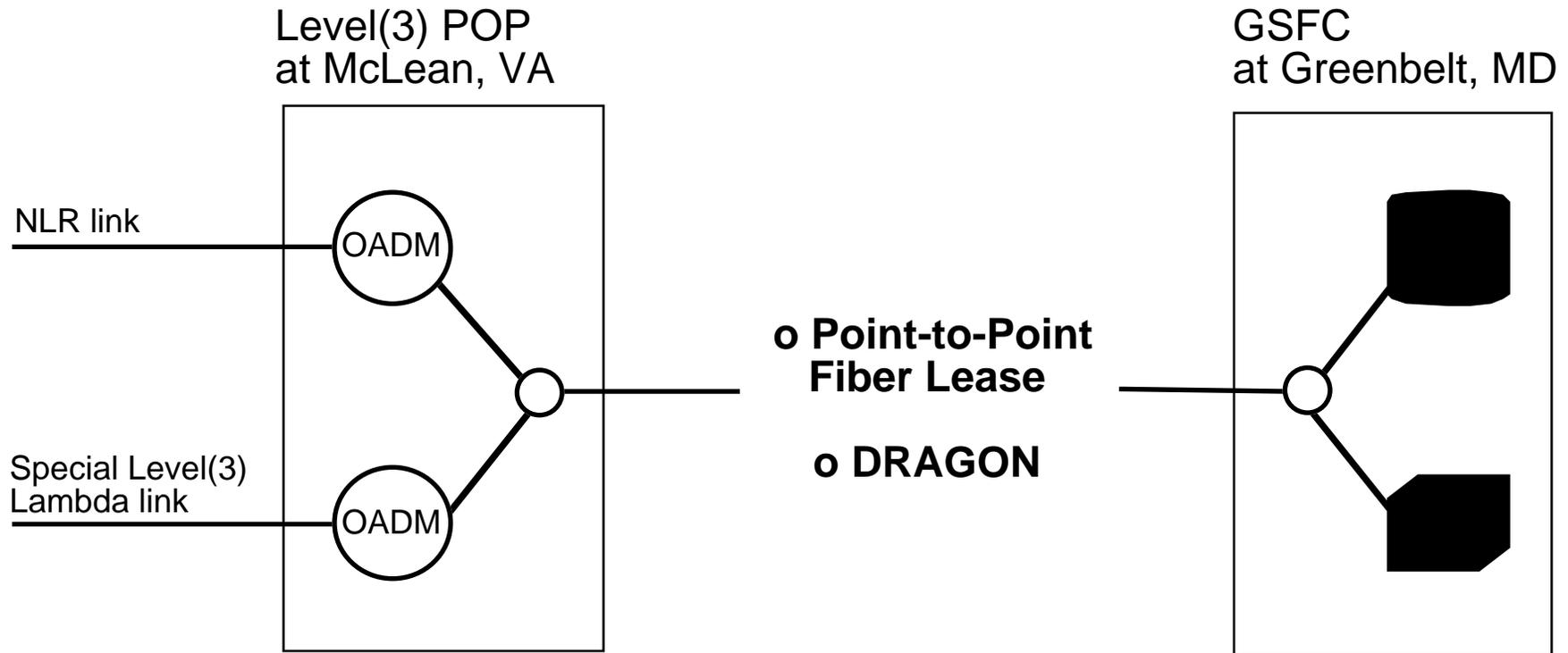
Level(3) POP
at McLean, VA

GSFC
at Greenbelt, MD

NLR link

OADM

Special Level(3)
Lambda link

OADM

o **Point-to-Point
Fiber Lease**

o **DRAGON**

JPG 03/08/04

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

Significant Accomplishments To Date

Detailing L-Net Links (continued)

- Regional Network Between McLean and GSFC

  » Obtained fiber lease quotes from Level3 and DRAGON, respectively, for alternate plans to close existing "air gaps"

  » Identified network equipment and finalizing CLIN's to enable one 10-GE and two 1-GE "lambdas"matched to the transcontinental links identified above

  » Plan to finish PR submissions by Jun04, and implement links within one month of equipment delivery

  » Plan to demonstrate initial DRAGON optical network with three 2.4 Gbps lambdas among GSFC, UMCP, and USC/ISI-East in Arlington, starting in Jun04

Fibers Considered in Point-to-Point Fiber Lease Approach

- 9Sep03 quotes obtained from Level (3) Communications by GSFC's Jeff Smith

  - » Only addressed provisioning of a lateral between GSFC and existing Level (3) fibers, not end-to-end lease/IRU between GSFC and McLean

  - » Required extensive construction, making it unattactive in both cost and schedule

- 10May04 quotes obtained from Level (3) Communications by ARC's Kevin Jones and GSFC's Pat Gary

  - » Still under consideration

Level(3) 10May04 Proposed Goddard to McLean Fiber Solution
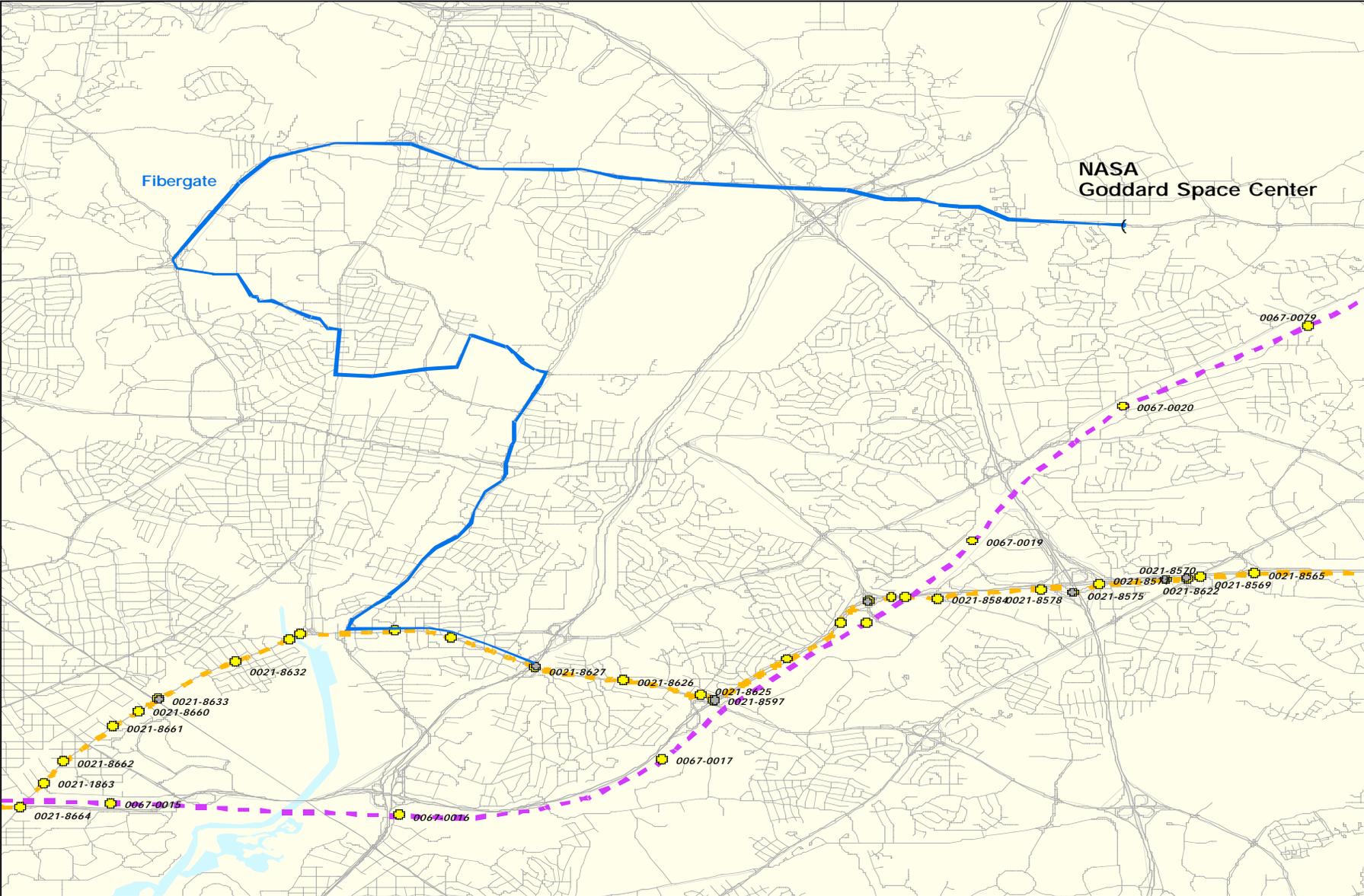
Fibers Considered in DRAGON-based Approach

- Fibers already "deployed" for all needed DRAGON links; and usage contracts in various states of progress

  - » "Link 1" GSFC -- UMD at College Park (UMCP): Order placed with FiberGate; expect available use in ~1 month

  - » "Link 2" UMCP -- ISI/E at Arlington & GWU in DC: In contract negotiations with Qwest; expect available use in ~2 months

  - » "Link 3" ISI/E (& GWU) -- Level(3) POP in McLean: Have quote for 20 year Indefeasible Right to Use (IRU)

DRAGON - Complete Network by Year 3
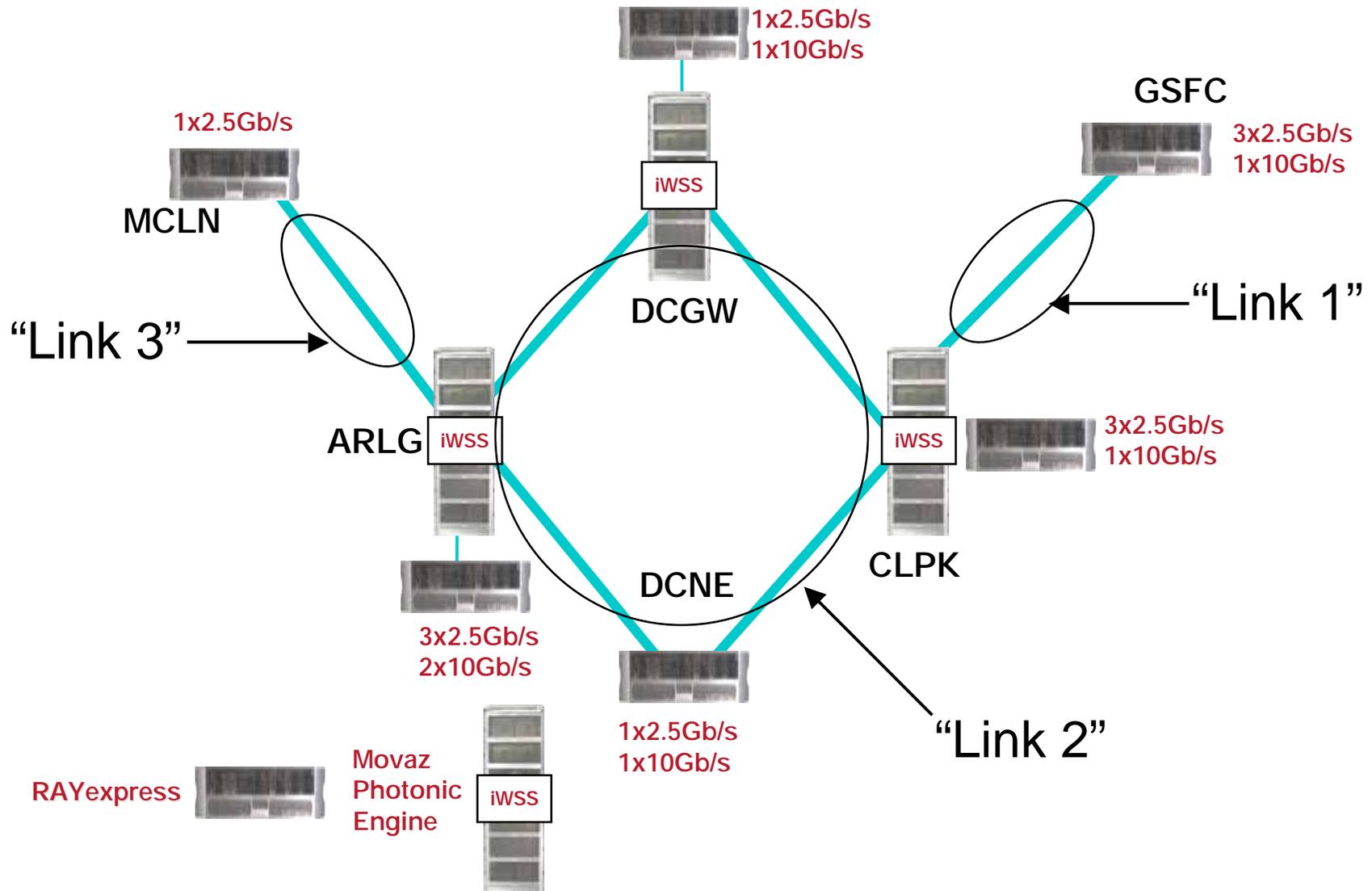
1x2.5Gb/s
1x10Gb/s

GSFC

1x2.5Gb/s

3x2.5Gb/s
1x10Gb/s

MCLN

iWSS

"Link 1"

DCGW

"Link 3"

ARLG   iWSS

iWSS

3x2.5Gb/s
1x10Gb/s

CLPK

DCNE

3x2.5Gb/s
2x10Gb/s

"Link 2"

1x2.5Gb/s
1x10Gb/s

RAYexpress

Movaz
Photonic
Engine

iWSS

J. P. Gary

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

Significant Accomplishments To Date

Detailing L-Net Links (continued)

- GSFC Local Network

    » Validated strategy of 10x1-GE to/from 1x10-GE up/down link designs between clusters, via GE and Myrinet throughput performance baselining completed between 16 cpu-pairs in the Thunderhead cluster

    » Validated feasibility of Scientific and Engineering Network (SEN) inter-building 10-GE upgrade plans, via full line rate loopback tests completed between a Force10 E300 10-GE switch in GSFC building 28 and a loaner Force10 E300 10-GE switch temporarily deployed in GSFC building 33

    » Plan to test a temporary full 10-GE network link between separate clusters in GSFC buildings 28 and 33 by Jul04

    » Plan to test UMBC/Ray Chen's optical switch in SEN's 10-GE environment, starting in Aug04

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 E300

← 10GigE →

Force10 E300

GigE

HECN GigE Switch

GigE

S1   S2

S1   S2

G4   G4

G4   G4
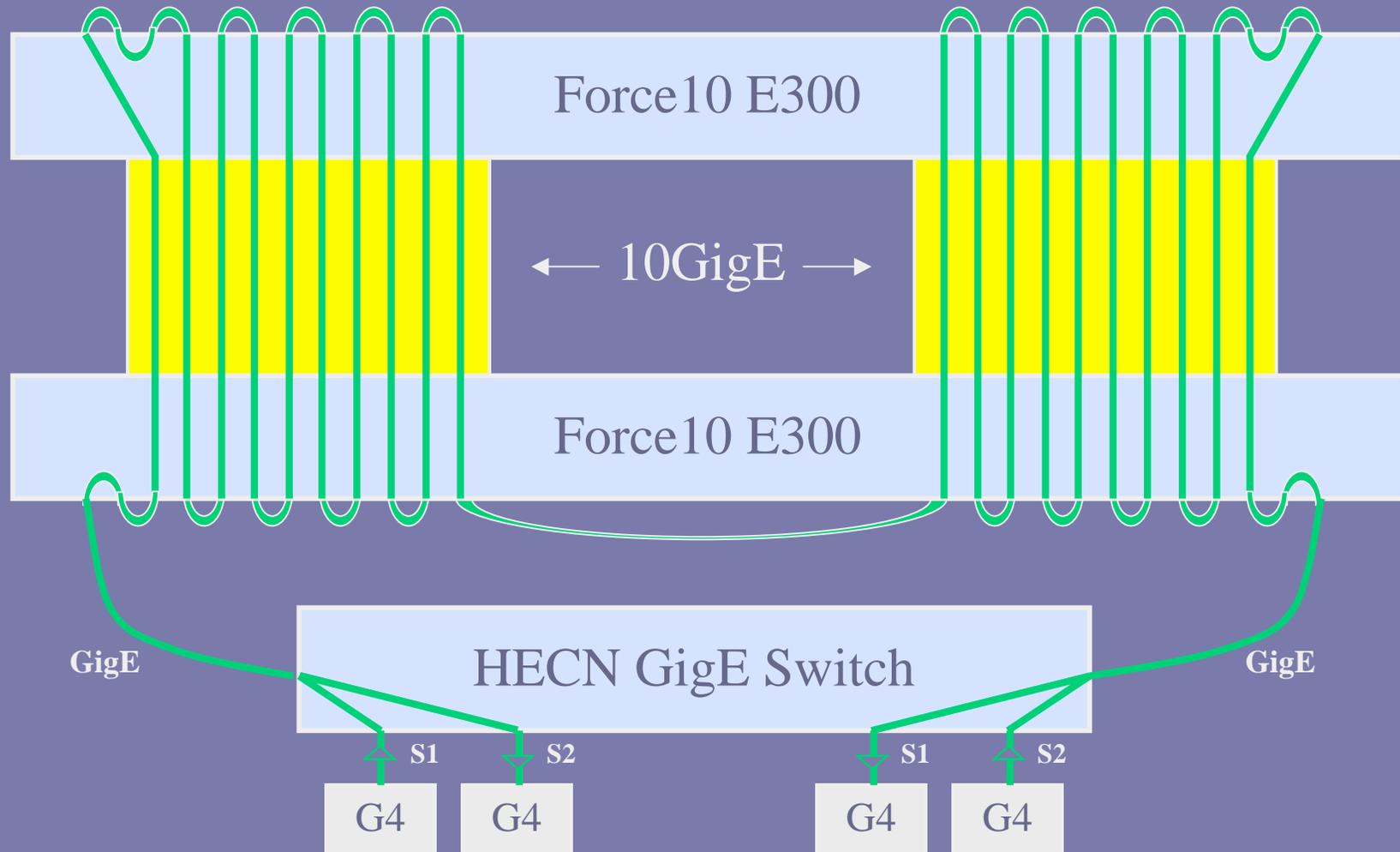
Test data generated with nuttcp and VLAN looping

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
```

$(30687801.2734 \text{ MB} + 30695954.5781 \text{ MB})*10*2/1024/1024/1024 = 1.14336 \text{ PB}$

* Over 0.25 PB transferred bidirectionally across each 10GigE link

Force10 Petabyte Challenge
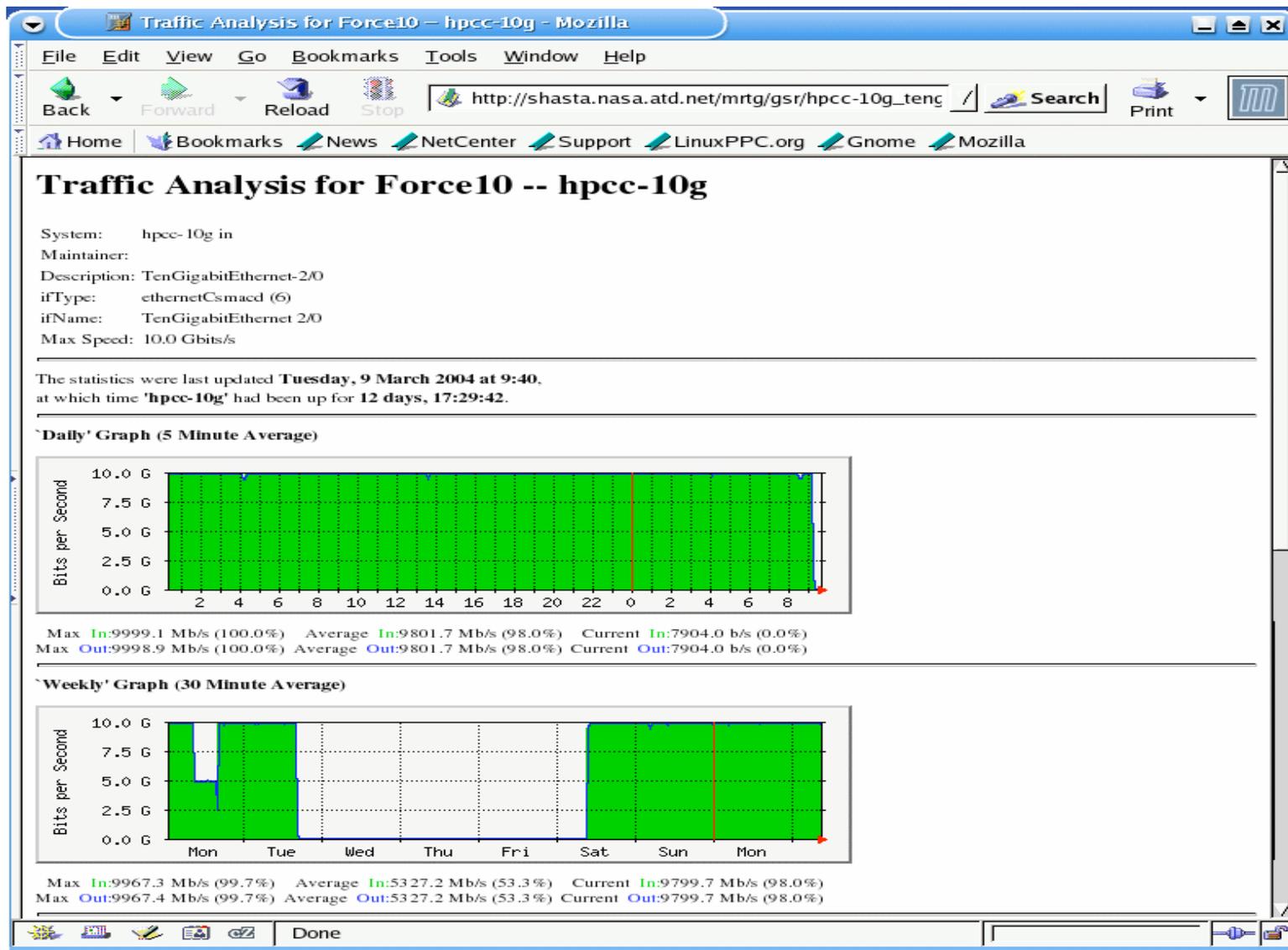
Graph for TenGigabitEthernet-3/0 basically identical

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

Significant Accomplishments To Date

Detailing L-Net Links (continued)

- Other

  » George Uhl (SWALE) attended Optiputer All Hands Meeting 1/14/04 - 1/16/04

    – Strategy meeting to discuss GSFC/SIO lambda service implementation

  » Established dual 2.4 Gbps lambda service using Marconi optical add/drop multiplexers between the High End Computer Network (HECN) and EOSDIS Network Prototyping Lab (ENPL) network labs in B28 and B32, respectively

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

<u>Backup Slides</u>

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

## _Examples of Initial Primary Users/Applications (1 of 2)_

- Drs. Paul Houser and Mike Bosilovich of Code 970 are collaborating with Dr. John Roads of SIO on the Coordinated Earth Observing Program under GEWEX

- Dr. Roads with Dr. Max Suarez of Code 900.3, Mike Seablom of Code 560, and a UMBC graduate student working with Dr. Milton Halem, GSFC Emeritus, plan to run interactive distributed regional model forecasts using boundary forcing conditions from the Global Modeling and Assimilation Office (GMAO) global climate model

- Dr. Yoram Kaufman of Code 910 is collaborating with Dr. Ramanathan of SIO on an Aerosol project

- Dr. J. Herman of Code 910 is the Co–I with Dr. Francisco Valero of SIO who is the PI on the Triana mission

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

## _Examples of Initial Primary Users/Applications (2 of 2)_

- Dr. Michelle Rienecker of Code 900.3 is collaborating with Dr. Tim Barnett of SIO on the assimilation of global sea height data from TOPEX and GRACE

- SIO's Prof. Richard Sommerville has one of his modelers remotely providing computational science support to the NCCS of Code 930

- Code 920 has collocated one of its geophysical scientists at SIO

- UCSD's Geosciences Network PI Dr. Dogan Seber has identified some of GSFC's solid earth research data sets and models for developing collaborative research efforts with Dr. Weijia Kuang and others from Code 920

Daily Median Hops to UCSD: 1-Nov-2003 to 8-Mar-2004

GSFC-ICESAT   GSFC-MAX   LATIS

J. P. Gary
05/21/04
34

Daily Median RTT to UCSD: 1-Nov-2003 to 8-Mar-2004

GSFC-ICESAT   GSFC-MAX   LATIS

J. P. Gary
05/21/04
35
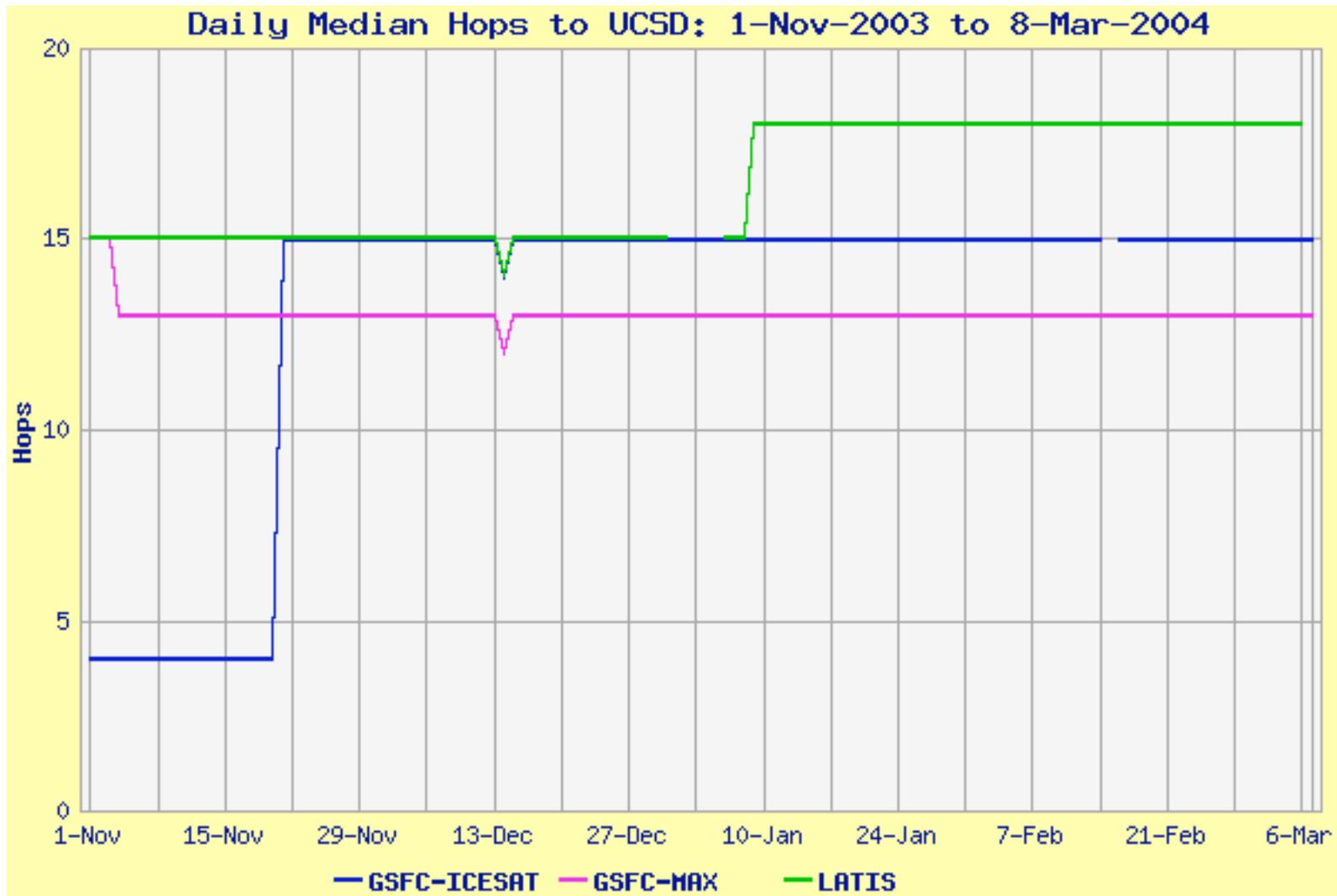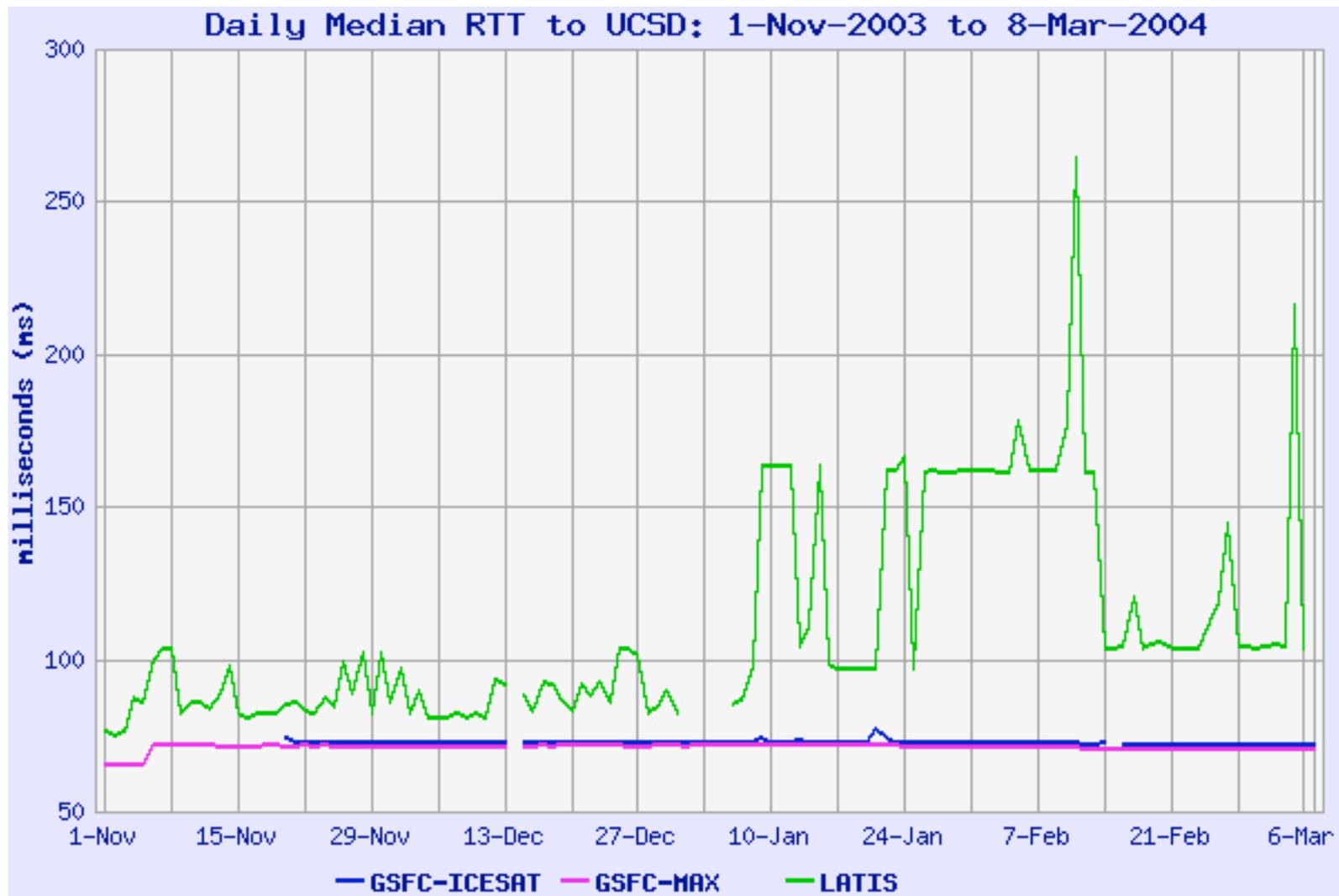
NLR – Optical Infrastructure - Phase 1

Seattle
Portland
Boise
Ogden
Denver
KC
Chicago
Clev
Pitts
Wash DC
Raleigh
Atlanta
Jacksonville
LA
San Diego

NLR Route
NLR MetaPOP
NLR Regen or OADM

NLR Phase 1 - Installation Schedule
Will Complete Aug 2004

Seattle
Portland
Complete
May
Sunnyvale
Los Angeles
San Diego
Boise
August
Ogden
Denver
June
Kansas
Chicago
StarLight
Complete
Cleveland
Pitts
Complete
Wash DC
Raleigh
Complete
Atlanta
July
Jacksonville
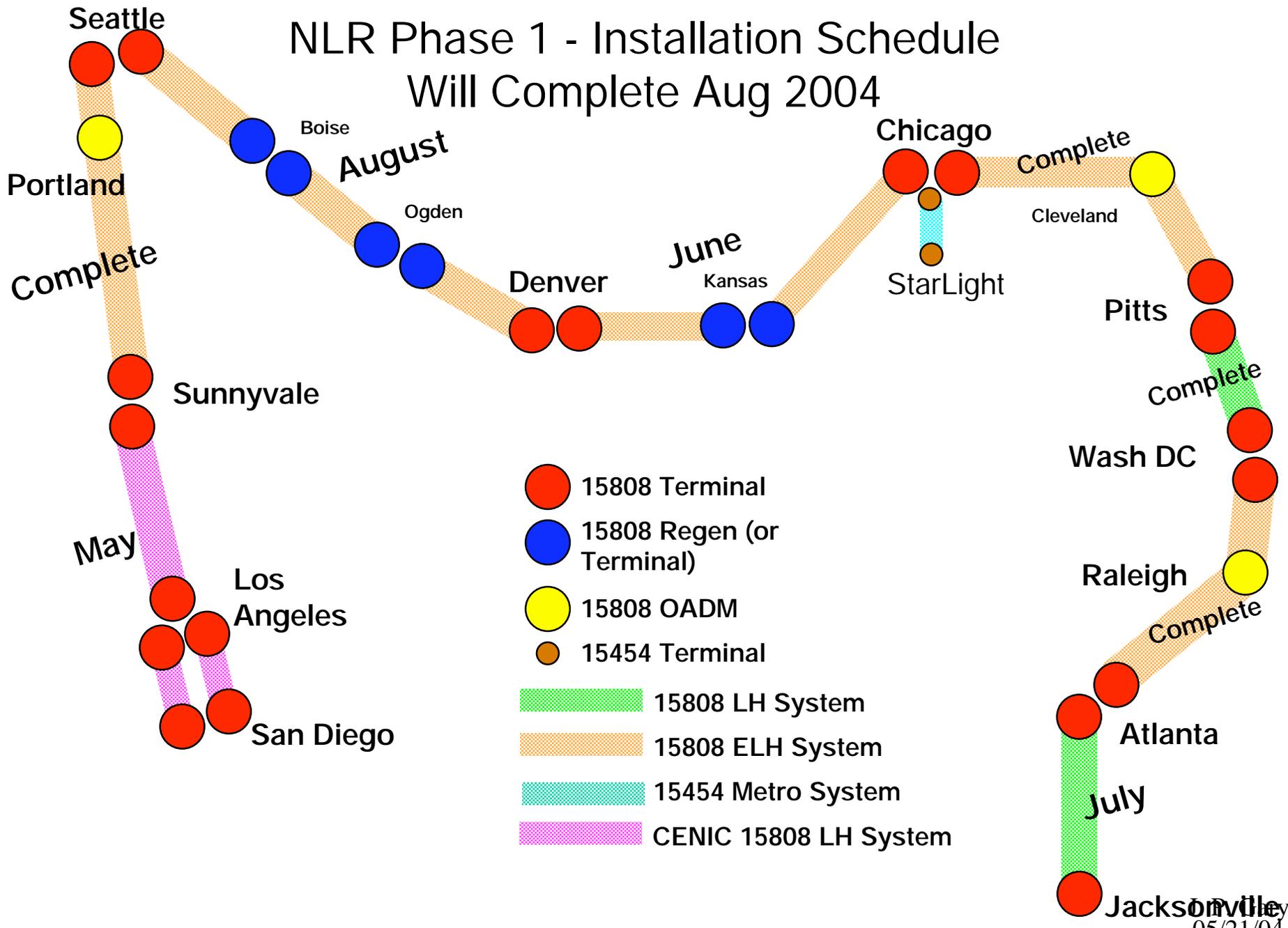
- 15808 Terminal
- 15808 Regen (or Terminal)
- 15808 OADM
- 15454 Terminal
- 15808 LH System
- 15808 ELH System
- 15454 Metro System
- CENIC 15808 LH System

05/21/04
37

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

Seattle
Portland
Boise
Chicago
Clev
Pitts
New York
Denver
Ogden/
Salt Lake
KC
Wash DC
LA
San Diego
Raleigh
Atlanta
Jacksonville

— NLR Route
▪▪▪ ORNL Waves
••• Waves
● NLR MetaPOP
● NLR Regen or OADM

J. P. Gary
05/21/04
38

NLR – Proposed Southern Route
Wiltel Fiber in the West -- AT&T Fiber in the Southeast
Qwest Fiber North-South
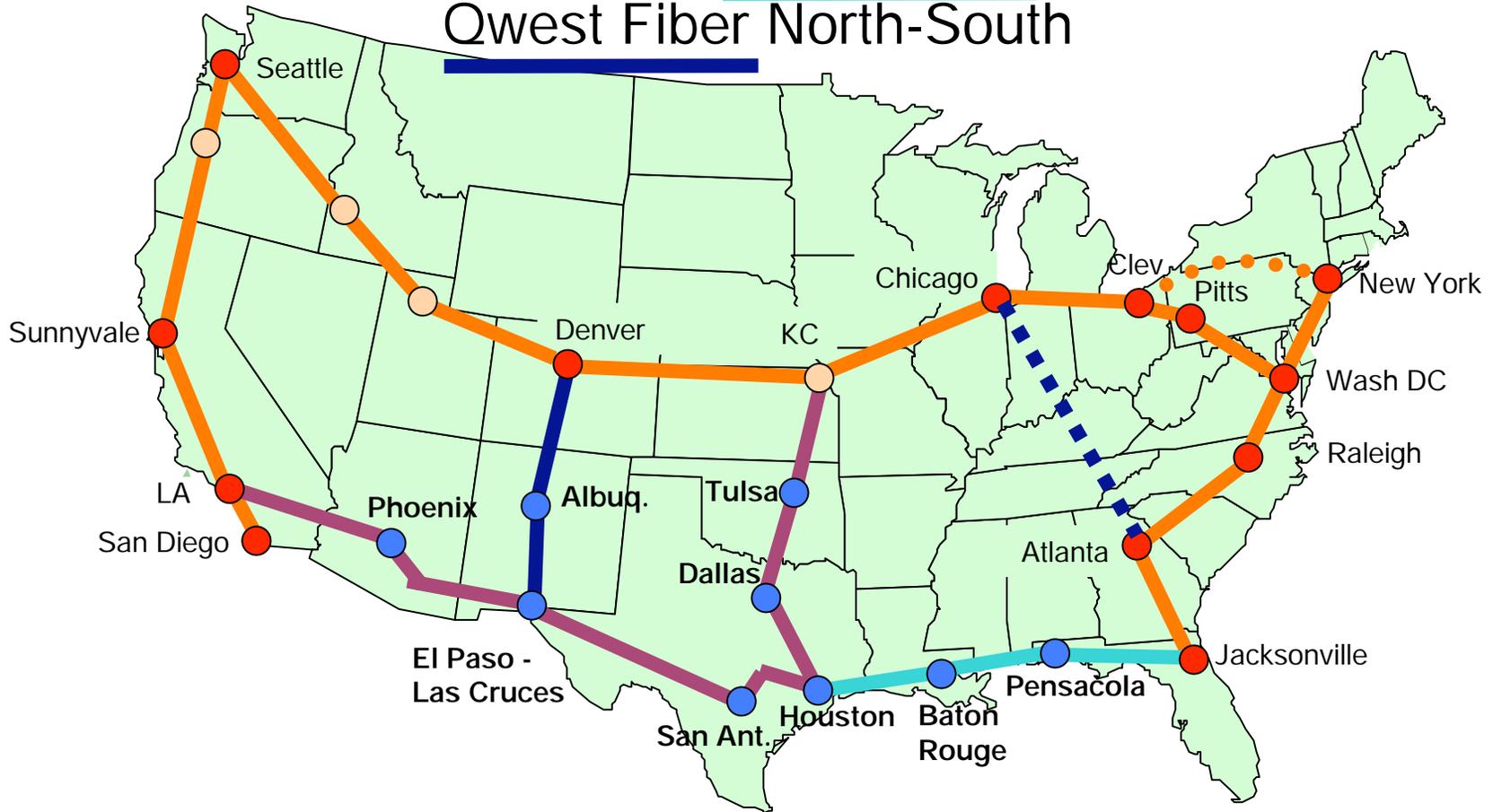
Using Cisco 15454 DWDM gear

Layer 2 Offering

- Nationwide GigE service

- 1 GigE Ports at all NLR "MetaPOPs" (where members are located)

- Cisco switches - 6509/7609 - connected by multiple 10 GigE lambdas (backbone)

- Ethernet edge / Ethernet core

Layer 2 Services Expectations (Continued)

Primary Purposes for NLR Layer 2 network:

- Lower Speed Transport (i.e. 1 Gbps):
  - » Provide a 1Gbps+ nationwide transport service over the NLR Layer 1 optical backbone for users that require **reliable connectivity, but not** necessarily the degree of dedicated transport provided by **lambda based** services.

- **Provide a flexible infrastructure** and tools which allow users to **make, or signal for, service level changes** to their individual Layer 2 paths, dramatically reducing provisioning delays while making more efficient use of the network.

DRAGON

Dynamic Resource Allocation via GMPLS Optical Networks

Jerry Sobieski
University of Maryland (UMD)
Mid-Atlantic Crossroads (MAX)

Tom Lehman
University of Southern California
Information Sciences Institute (USC ISI)

Bijan Jabbari
George Mason University (GMU)

National Science Foundation

J. P. Gary
05/21/04
42

The DRAGON Project
Key Features/Objectives

- Dynamic provisioning of guaranteed end-to-end resources
- Application specific network topologies
- Reserve resources and topology in advance, instantiate when needed
- Protocol, format, framing agnostic
- Direct transmission of HDTV, ethernet, sonet, fibreChannel, or any optical signal

The DRAGON Project
Key Features/Objectives

- Uses all optical transport in the metro core
  - Edge to edge Wavelength switching (2R OEO only for signal integrity)
  - Push OEO demarc to the edge, and increasingly out towards end user
- Standardized GMPLS protocols to dynamically provision intra-domain connections
  - GMPLS-OSPF-TE and GMPLS-RSVP-TE
- Develop the inter-domain protocol platform to
  - Distribute Transport Layer Capability Sets (TLCS) across multiple domains
  - Perform E2E path computation
  - Resource authorization, scheduling, and accounting
- Develop the "Virtual LSR"
  - Abstracts non-GMPLS network resources into a GMPLS "virtual LSR".
- Simplified API
  - Application Specific Topology definition and instantiation
  - Resource resolution, proxy registration and signaling