"The Jump to Light Speed – Data Intensive Earth Sciences are Leading the Way to the International LambdaGrid"

Keynote the 15th Federation of Earth Science Information Partners Assembly Meeting: Linking Data and Information to Decision Makers

San Diego, CA

June 14, 2005

Dr. Larry Smarr

Director, California Institute for Telecommunications and Information Technology

Harry E. Gruber Professor,

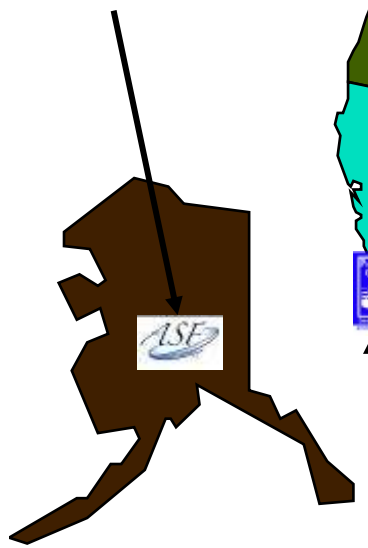
Dept. of Computer Science and Engineering

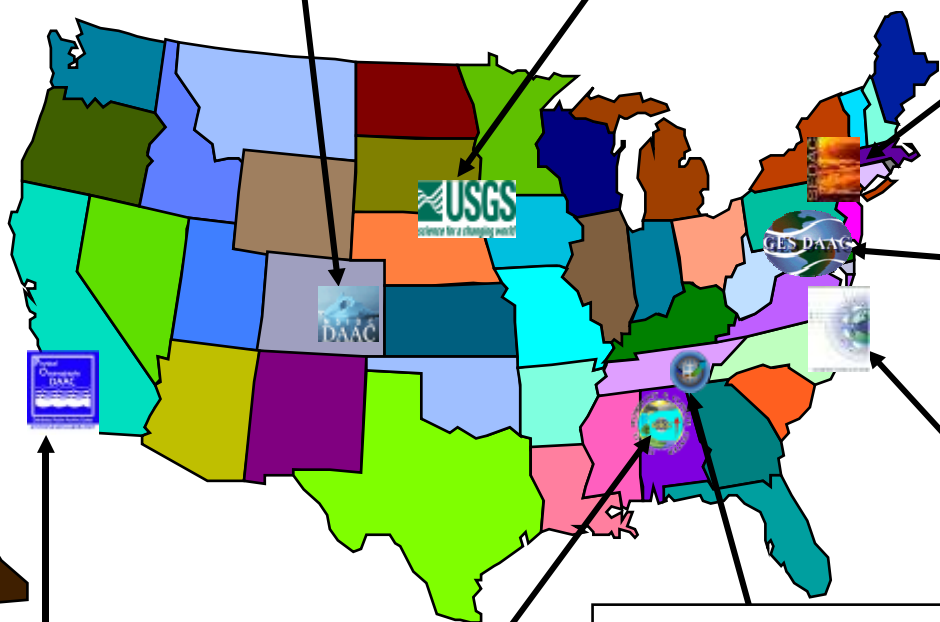
Jacobs School of Engineering, UCSD

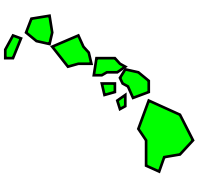
OptIPuter

it²

Earth System Enterprise-Data Lives in Distributed Active Archive Centers (DAAC)







NSIDC (67 TB)
Cryosphere
Polar Processes

LPDAAC-EDC (1143 TB)
Land Processes
& Features

ASF (256 TB)
SAR Products
Sea Ice
Polar Processes

SEDAC (0.1 TB)
Human Interactions in
Global Change

GES DAAC-GSFC (1334 TB)
Upper Atmosphere
Atmospheric Dynamics, Ocean Color, Global Biosphere, Hydrology, Radiance Data

ASDC-LaRC (340 TB)
Radiation Budget,Clouds
Aerosols, Tropospheric
Chemistry

ORNL (1 TB)
Biogeochemical
Dynamics
EOS Land Validation

GHRC (4TB)
Global
Hydrology

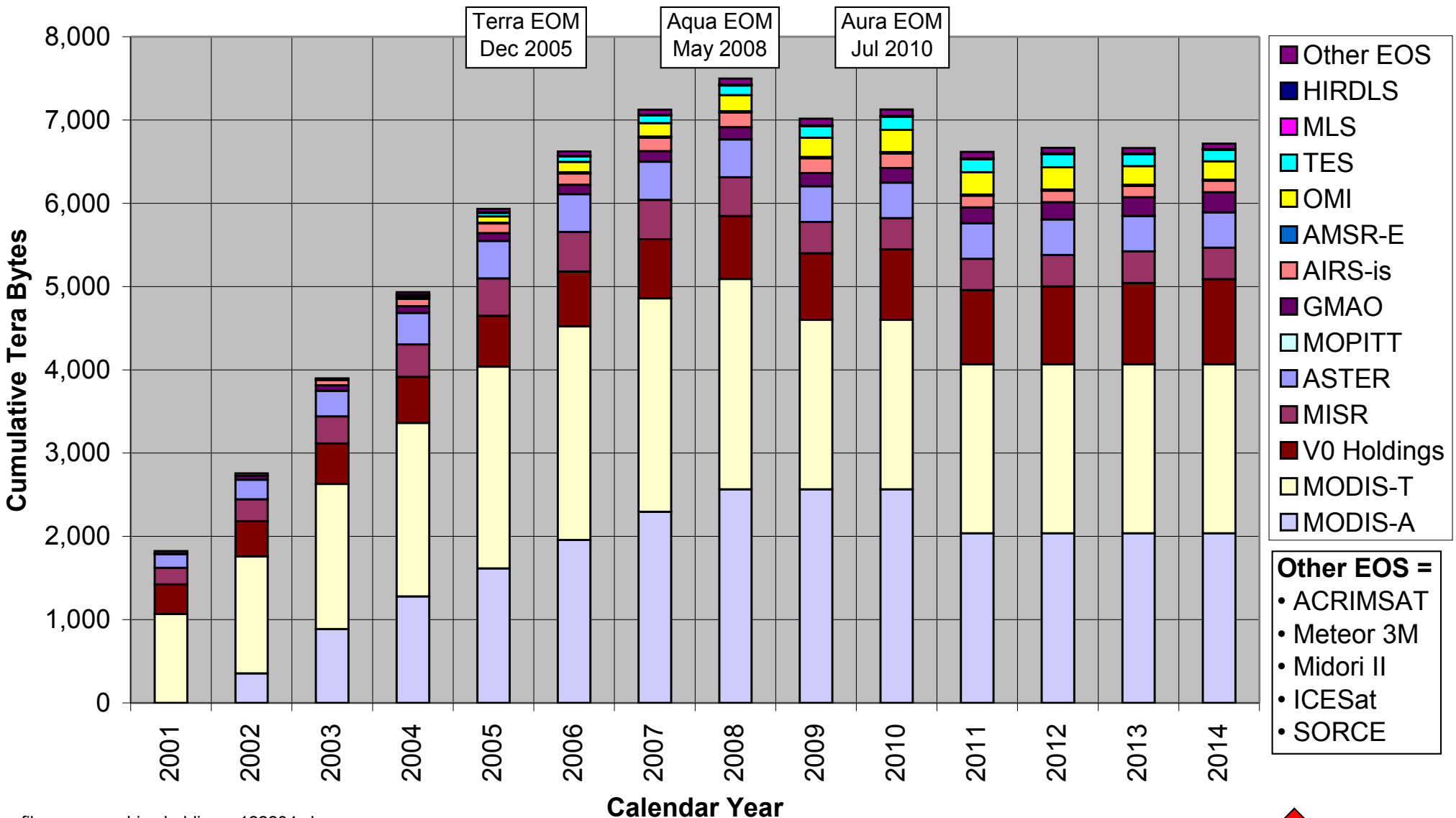
PODAAC-JPL (6 TB)
Ocean Circulation
Air-Sea Interactions

Challenge: How to Get Data Interactively to End Users Using New Technologies

OptIPuter

it2

Cumulative EOSDIS Archive Holdings--Adding Several TBs per Day



file name: archive holdings_122204.xls
tab: all instr bar

NOTE: Data remains in the archive pending transition to LTA

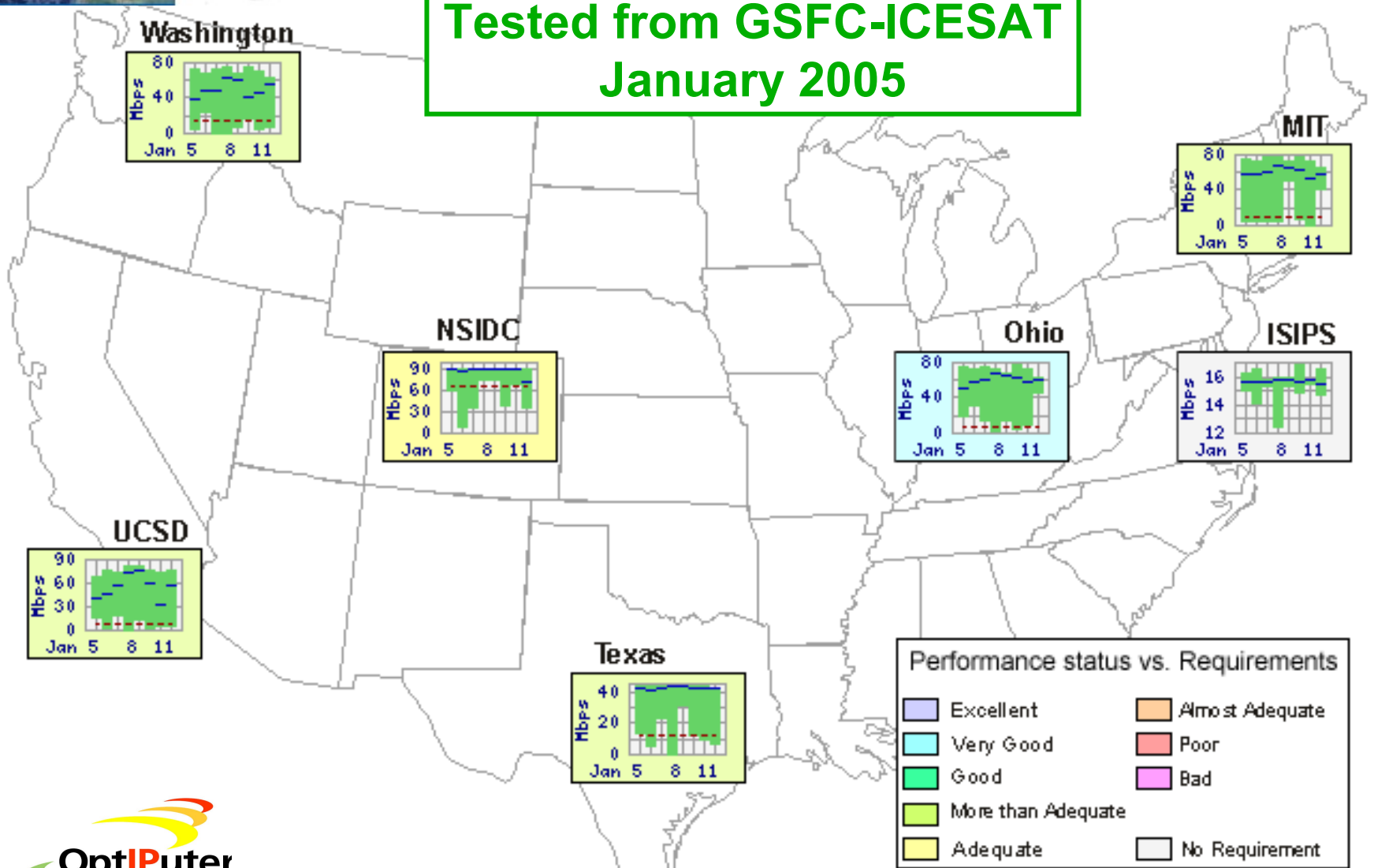
Source: Glenn Iona, EOSDIS Element Evolution Technical Working Group January 6-7, 2005

Barrier: Average Throughput of NASA Data Products to End User is Only < 50 Megabits/s



ENSIGHT Active Testing ICESat Destinations

Tested from GSFC-ICESAT January 2005



http://ensight.eos.nasa.gov/Missions/icesat/index.shtml

OptIPuter

it2

High Resolution Aerial Photography Generates Images With 10,000 Times More Data than Landsat7

Landsat7 Imagery
100 Foot Resolution
Draped on elevation data





New USGS Aerial Imagery
At 1-Foot Resolution

~10x10 square miles of 350 US Cities
2.5 Billion Pixel Images Per City!

OptIPuter

it2

Multi-Gigapixel Images are Available from Film Scanners Today

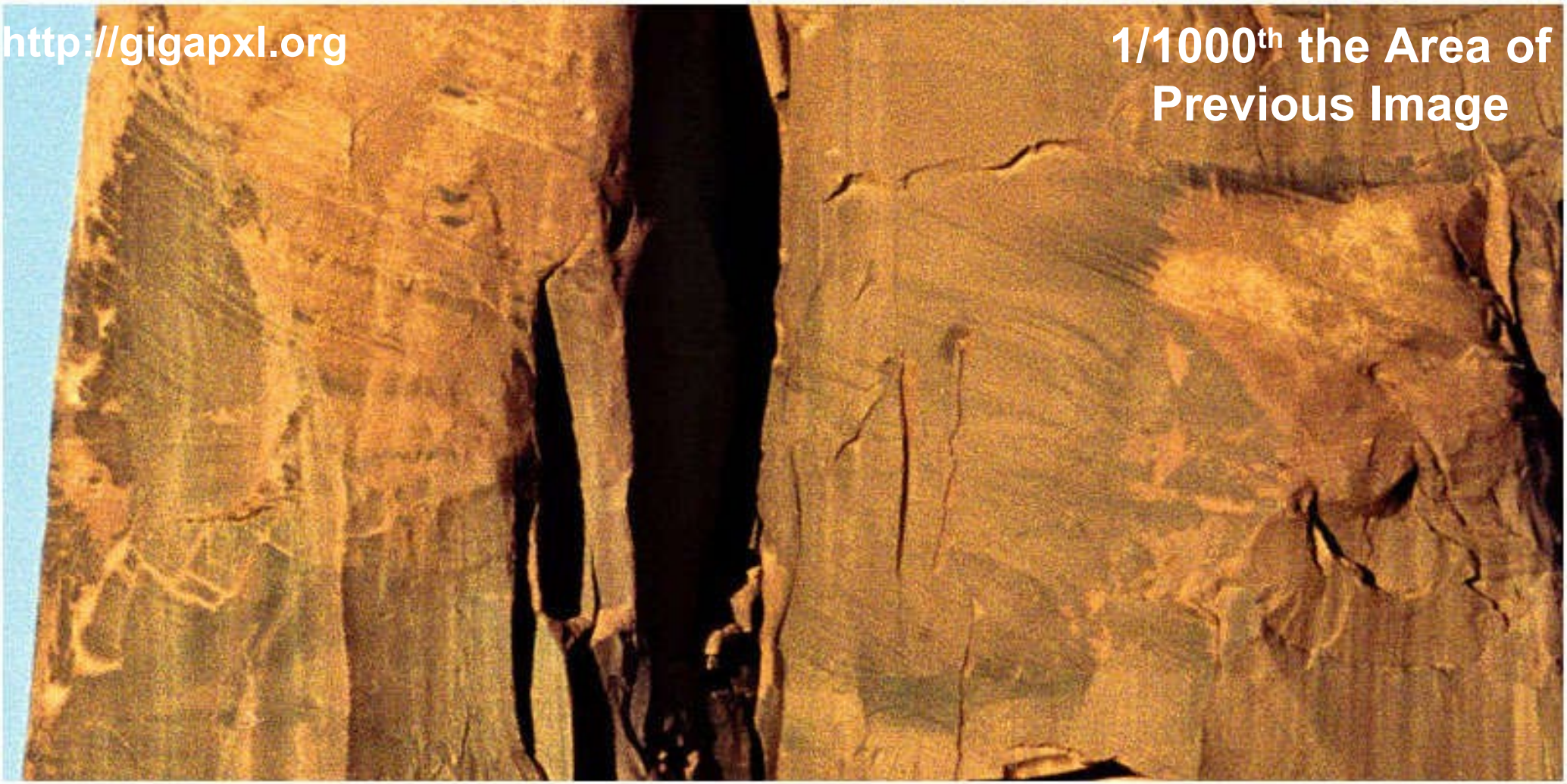


OptIPuter

The Gigapxl Project
http://gigapxl.org

it²

Large Image with Enormous Detail Requires Interactive Hundred Million Pixel Systems



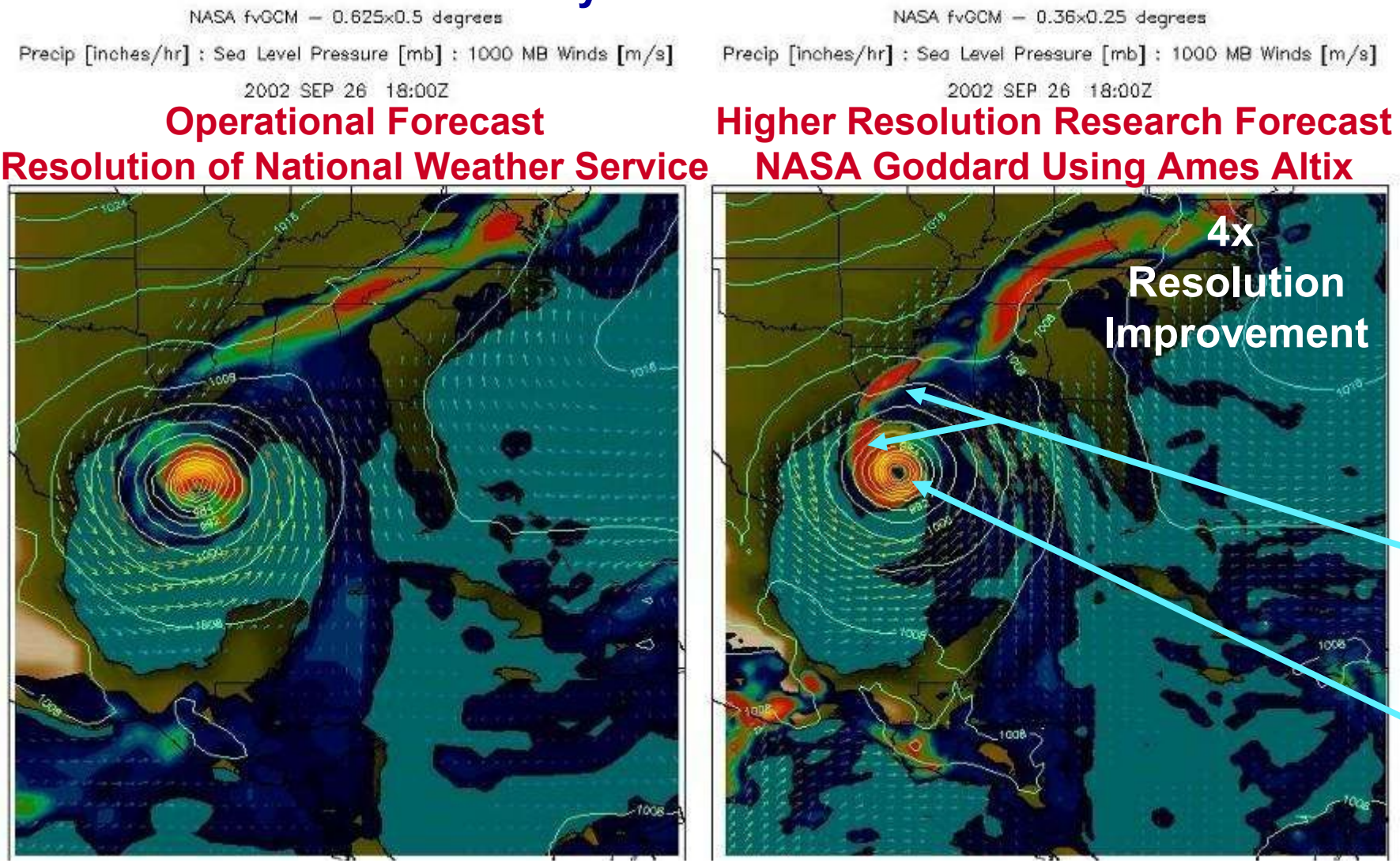
OptIPuter

it2

Increasing Accuracy in Hurricane Forecasts

Real Time Diagnostics in GSFC of Ensemble Runs on ARC Project Columbia

5.75 Day Forecast of Hurricane Isidore



Operational Forecast Resolution of National Weather Service

Higher Resolution Research Forecast NASA Goddard Using Ames Altix

How to Remove the InterCenter Networking Bottleneck?

Intense Rain-Bands

Resolved Eye Wall

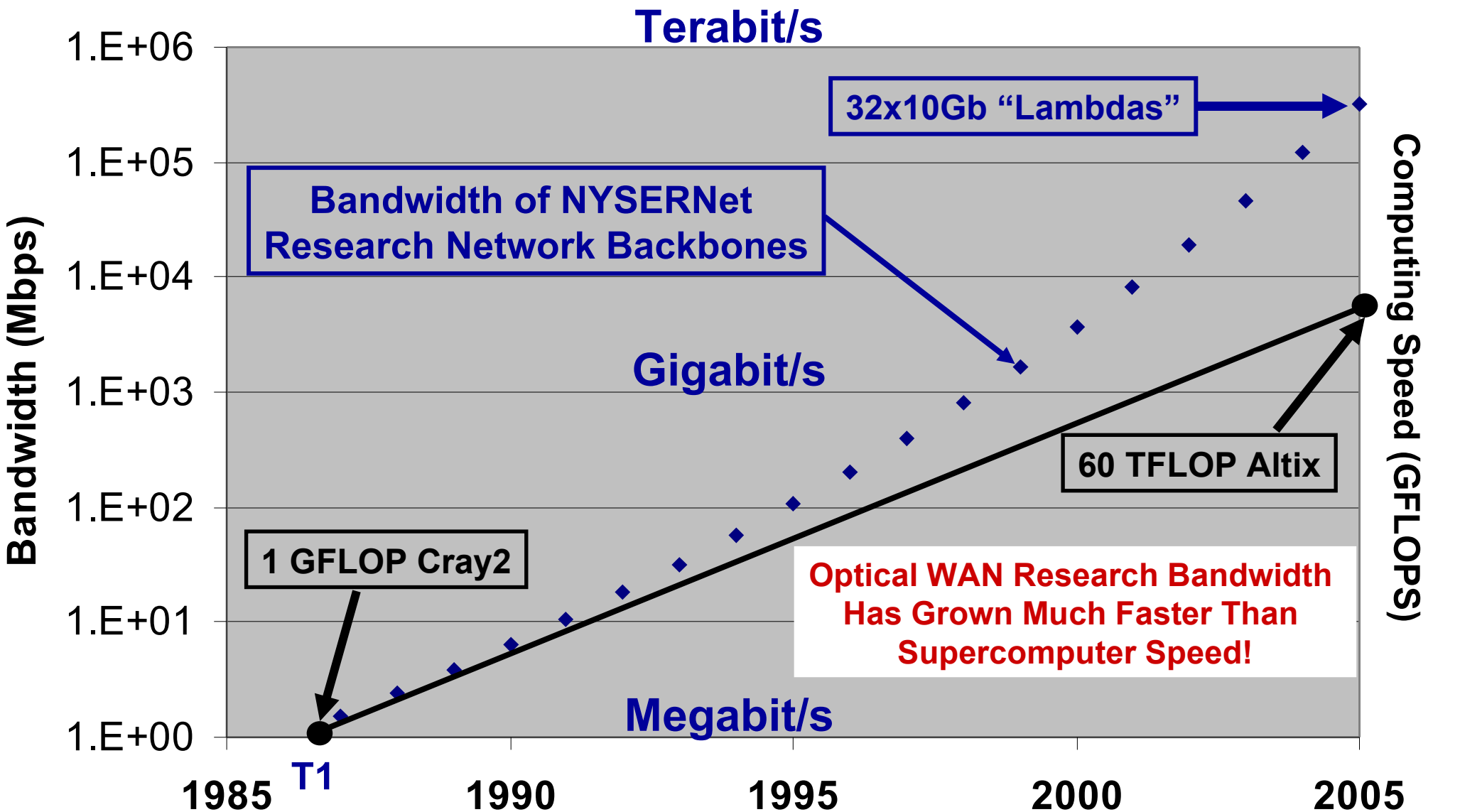
Source: Bill Putman, Bob Atlas, GFSC

OptIPuter

Project Contacts: Ricky Rood, Bob Atlas, Horace Mitchell, GSFC; Chris Henze, ARC

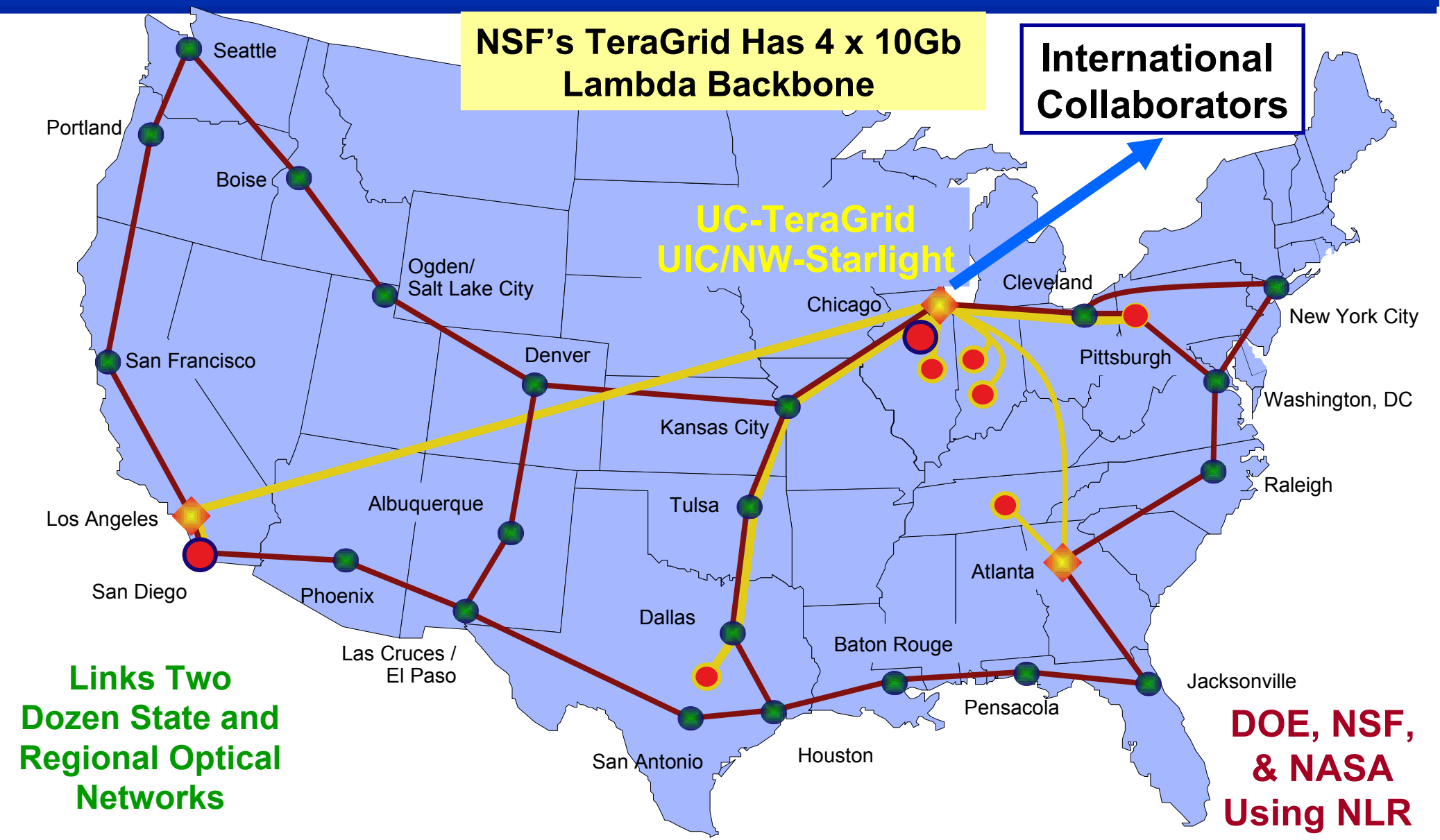
it2

From "Supercomputer–Centric" to "Supernetwork-Centric" Cyberinfrastructure



Network Data Source: Timothy Lance, President, NYSERNet

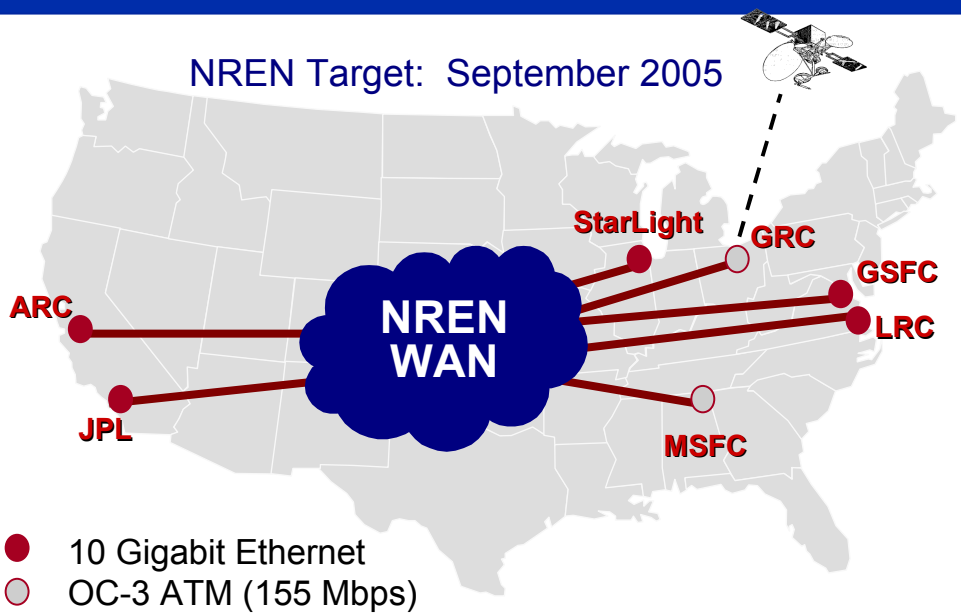
National Lambda Rail (NLR) and TeraGrid Provides Researchers a Cyberinfrastructure Backbone



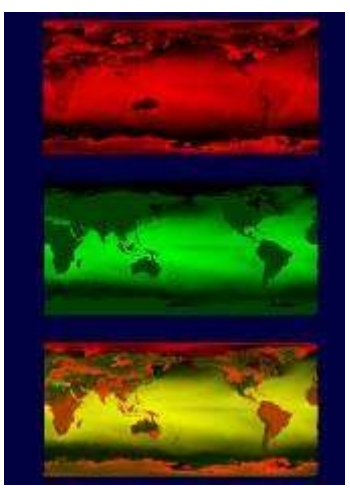
NASA Research and Engineering Network (NREN) Overview

- NREN Goal
  - Provide a Wide Area, High-speed Network for Large Data Distribution and Real-time Interactive Applications
  - Provide Access to NASA Research & Engineering Communities - Primary Focus: Supporting Distributed Data Access to/ from Project Columbia





- Sample Application: Estimating the Circulation and Climate of the Ocean (ECCO)
  - ~78 Million Data Points
  - 1/6 Degree Latitude-Longitude Grid
  - Decadal Grids ~ 0.5 Terabytes / Day
  - Sites: NASA JPL, MIT, NASA Ames



- **Next Steps**
  - **1 Gbps (JPL to ARC) Across CENIC (February 2005)**
  - **10 Gbps ARC, JPL & GSFC Across NLR (May 2005)**
  - **StarLight Peering (May 2005)**
  - **10 Gbps LRC (Sep 2005)**

OptIPuter

Source: Kevin Jones, Walter Brooks, ARC



The Networking Double Header of the Century Will Be Driven by LambdaGrid Applications

Maxine Brown, Tom DeFanti, Co-Organizers



iGrid 2005

THE GLOBAL LAMBDA INTEGRATED FACILITY



www.startap.net/igrid2005/

September 26-30, 2005

Calit2 @ University of California, San Diego

California Institute for Telecommunications and Information Technology



SC|05 Gateway to Discovery

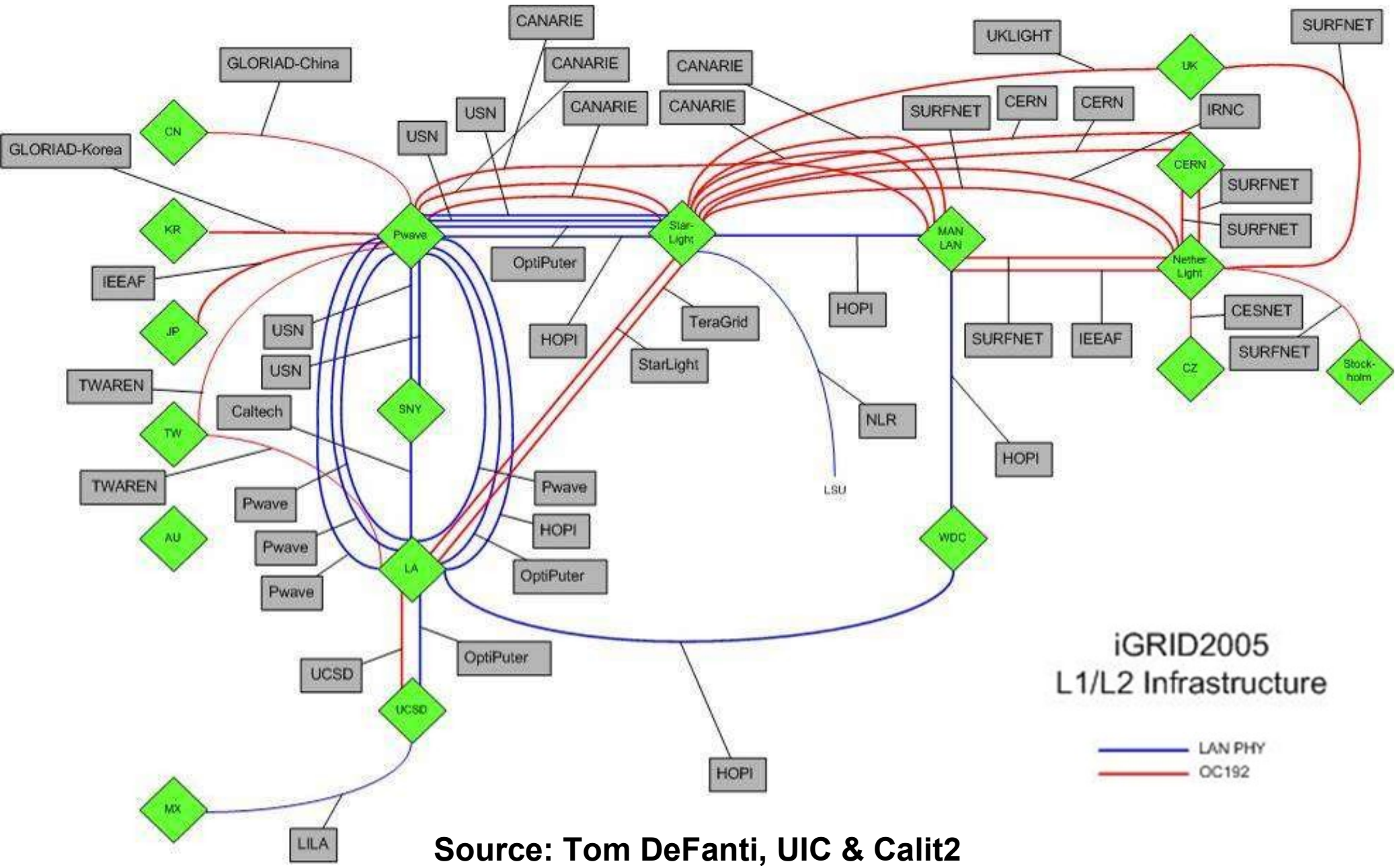
Washington State Convention and Trade Center Seattle, WA
November 12-18 2005

http://sc05.supercomp.org

OptIPuter

it²

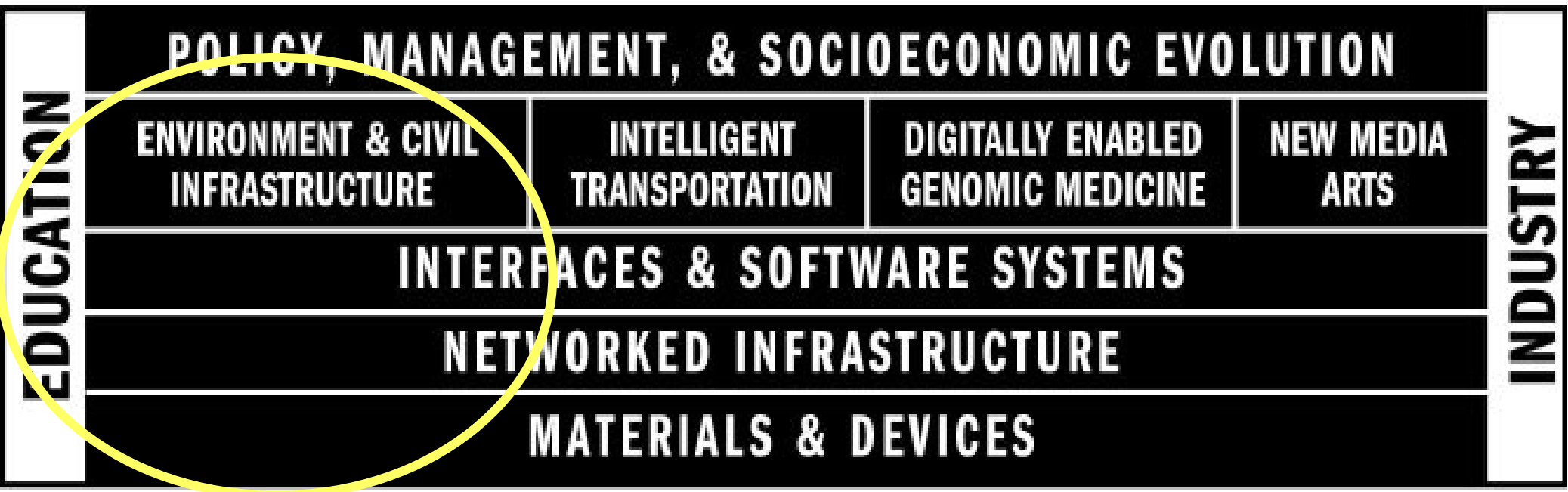
The International Lambda Fabric
Being Assembled to Support iGrid Experiments



Source: Tom DeFanti, UIC & Calit2

Calit2 -- Research and Living Laboratories on the Future of the Internet

UC San Diego & UC Irvine Faculty
Working in Multidisciplinary Teams
With Students, Industry, and the Community



www.calit2.net

Two New Calit2 Buildings Will Provide a Persistent Collaboration "Living Laboratory"



- Over 1000 Researchers in Two Buildings
  - Linked via Dedicated Optical Networks
  - International Conferences and Testbeds
- New Laboratory Facilities
  - Virtual Reality, Digital Cinema, HDTV
  - Nanotech, BioMEMS, Chips, Radio, Photonics



OptIPuter

California Provided $100M for Buildings
Industry Partners $85M, Federal Grants $250M

it²

The Calit2@UCSD Building is Designed for Extremely High Bandwidth

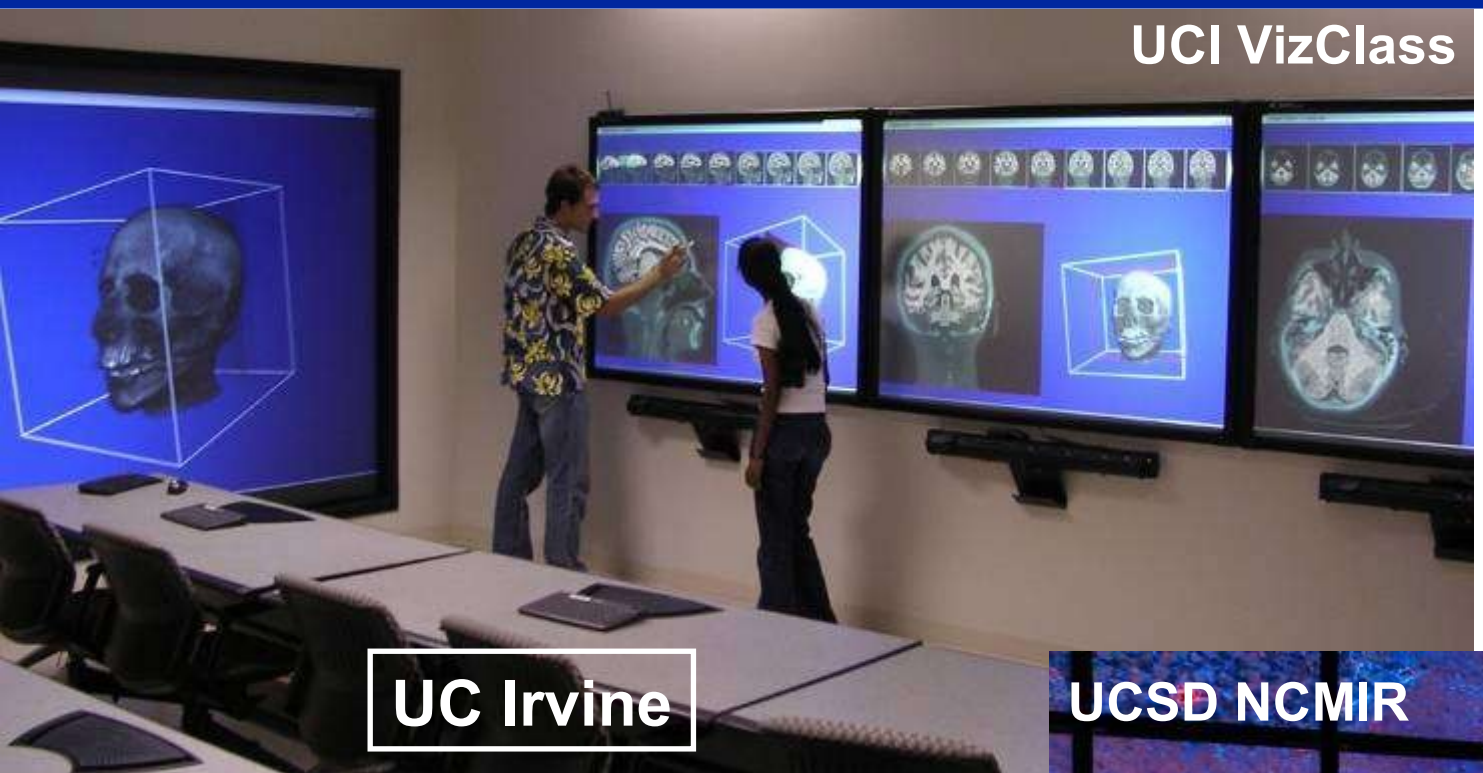
1.8 Million Feet of Cat6 Ethernet Cabling



150 Fiber Strands to Building
Experimental Roof Radio Antenna Farm
Building Radio Transparent
Ubiquitous WiFi

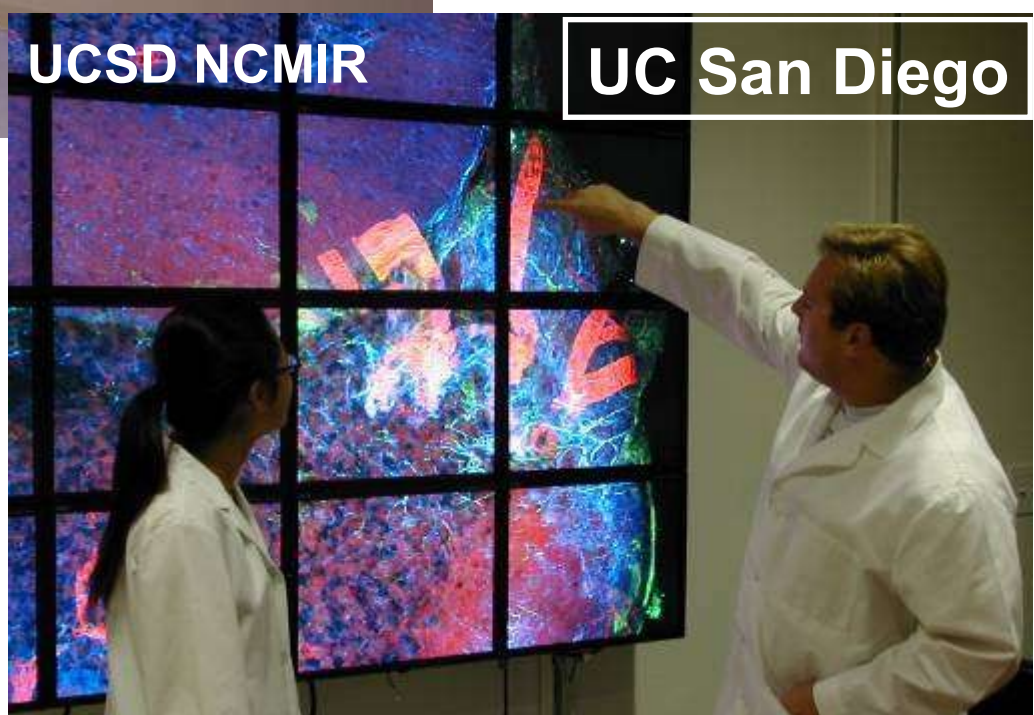
Photo: Tim Beach, Calit2

Calit2 Collaboration Rooms Testbed
UCI to UCSD



Source: Falko Kuester, UCI & Mark Ellisman, UCSD

In 2005 Calit2 will Link Its Two Buildings via CENIC-XD Dedicated Fiber over 75 Miles to Create a Distributed Collaboration Laboratory

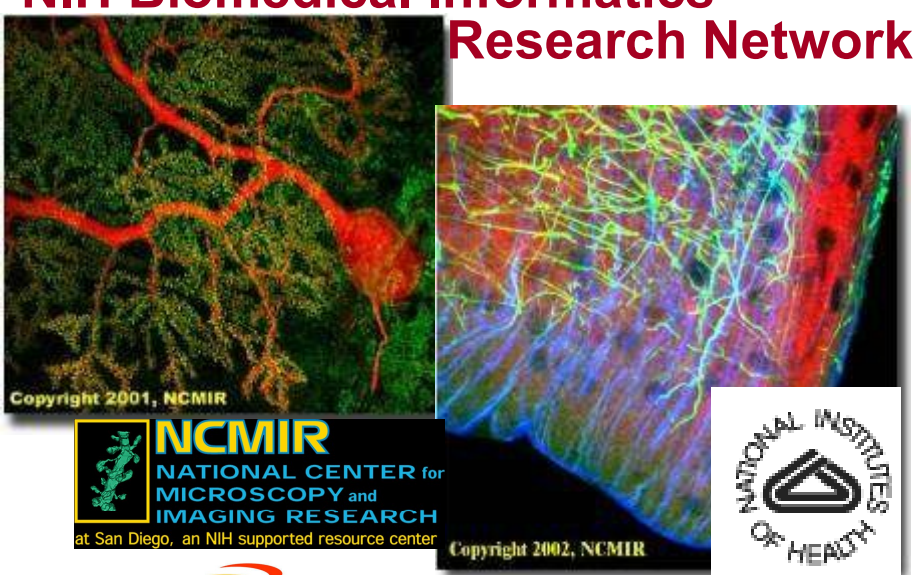


OptIPuter

The OptIPuter Project –
Creating a LambdaGrid "Web" for Gigabyte Data Objects

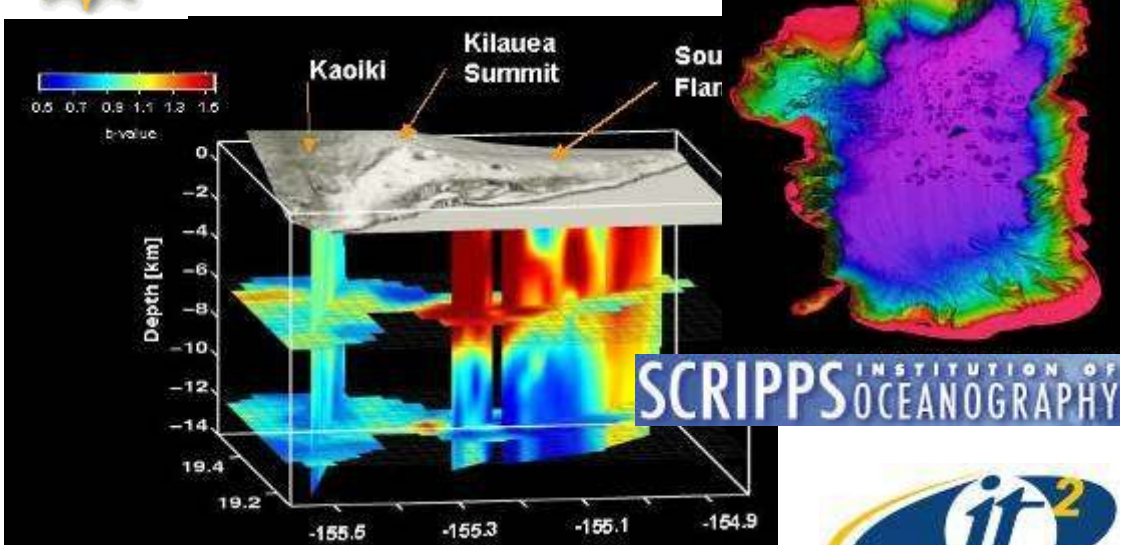
- **NSF Large Information Technology Research Proposal**
  - Calit2 (UCSD, UCI) and UIC Lead Campuses—Larry Smarr PI
  - Partnering Campuses: USC, SDSU, NW, TA&M, UvA, SARA, NASA
- **Industrial Partners**
  - IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent
- **$13.5 Million Over Five Years**
- **Linking User's Linux Clusters to Remote Science Resources**

NIH Biomedical Informatics Research Network



http://ncmir.ucsd.edu/gallery.html

NSF EarthScope and ORION



siovizcenter.ucsd.edu/library/gallery/shoot1/index.shtml

OptIPuter SDSC it²

Optical Networking, Internet Protocol, Computer Bringing the Power of Lambdas to Users

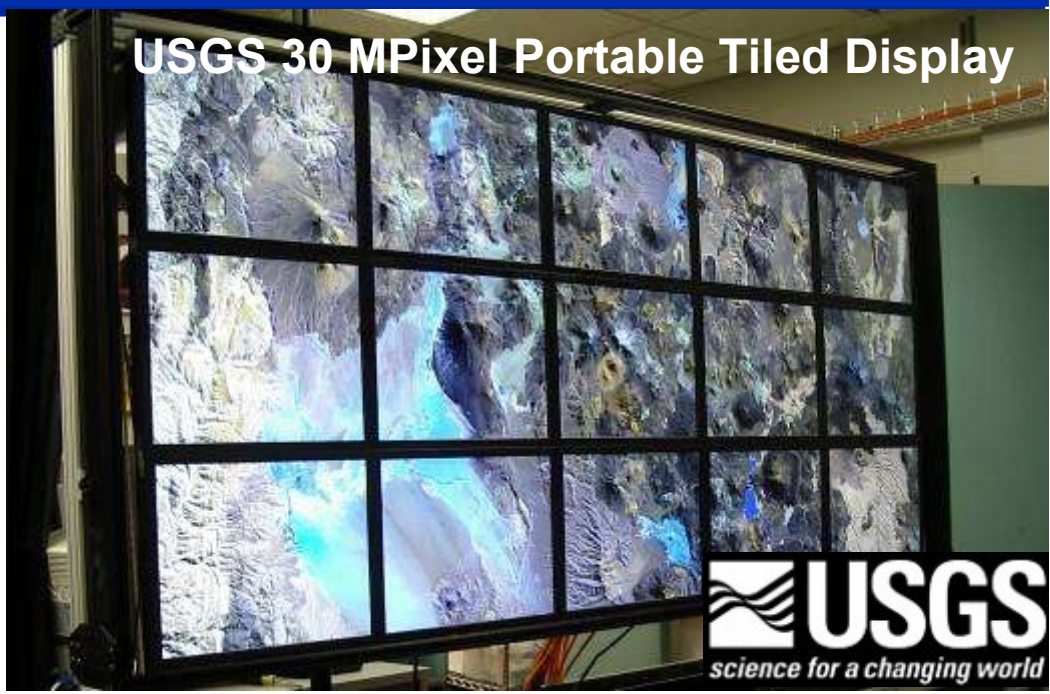
- **Complete the Grid Paradigm by Extending Grid Middleware to Control Jitter-Free, Fixed Latency, Predictable Optical Circuits**
  - **One or Parallel Dedicated Light-Pipes**
    - **1 or 10 Gbps WAN Lambdas**
  - **Uses Internet Protocol, But Does NOT Require TCP**
  - **Exploring Both Intelligent Routers and Passive Switches**
- **Tightly Couple to End User Clusters Optimized for Storage, Visualization, or Computing**
  - **Linux Clusters With 1 or 10 Gbps I/O per Node**
  - **Scalable Visualization Displays with OptIPuter Clusters**
- **Applications Drivers:**
  - **Earth and Ocean Sciences**
  - **Biomedical Imaging**
  - **Designed to Work with any Discipline Driver**

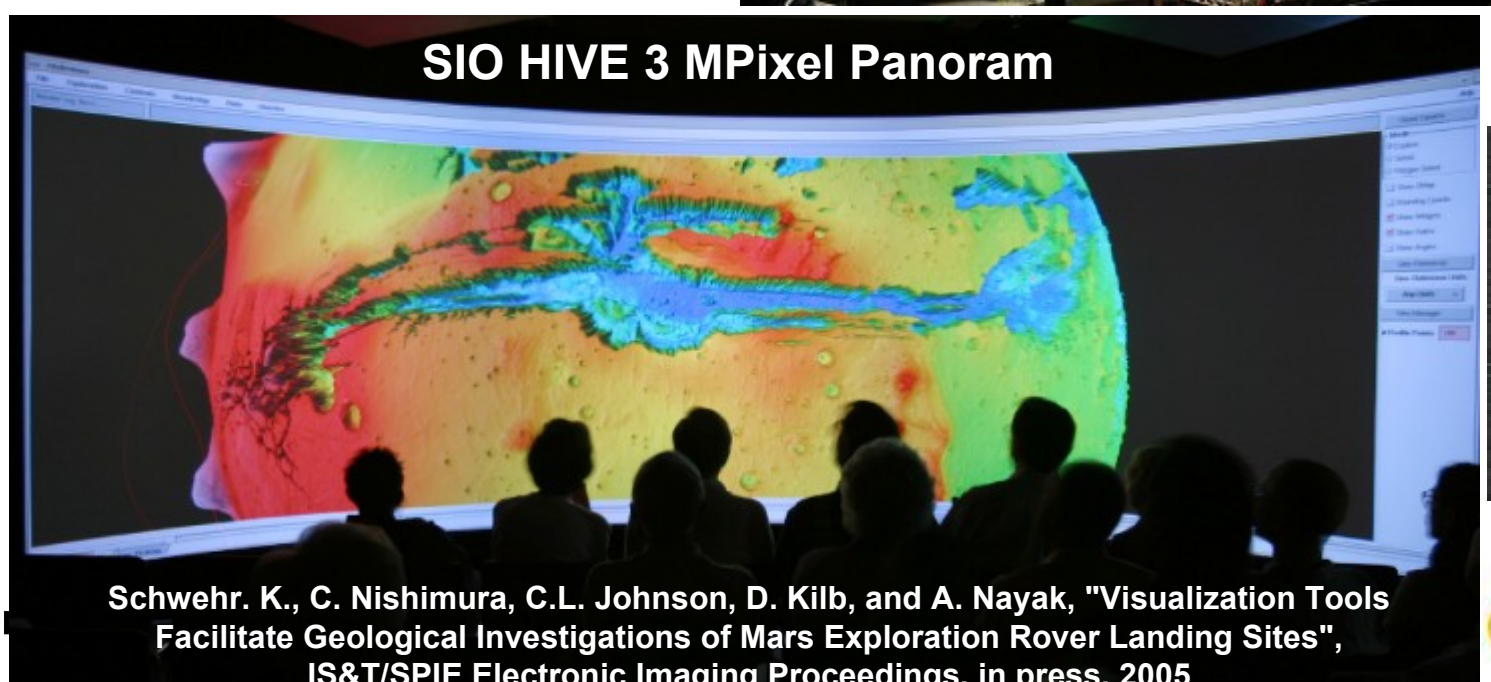
OptIPuter

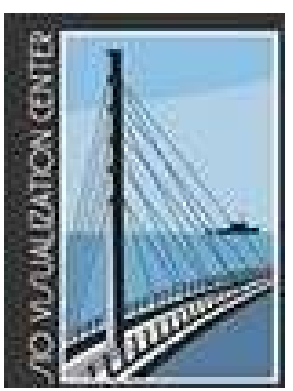
it²

Earth and Planetary Sciences: High Resolution Portals to Global Earth Sciences Data







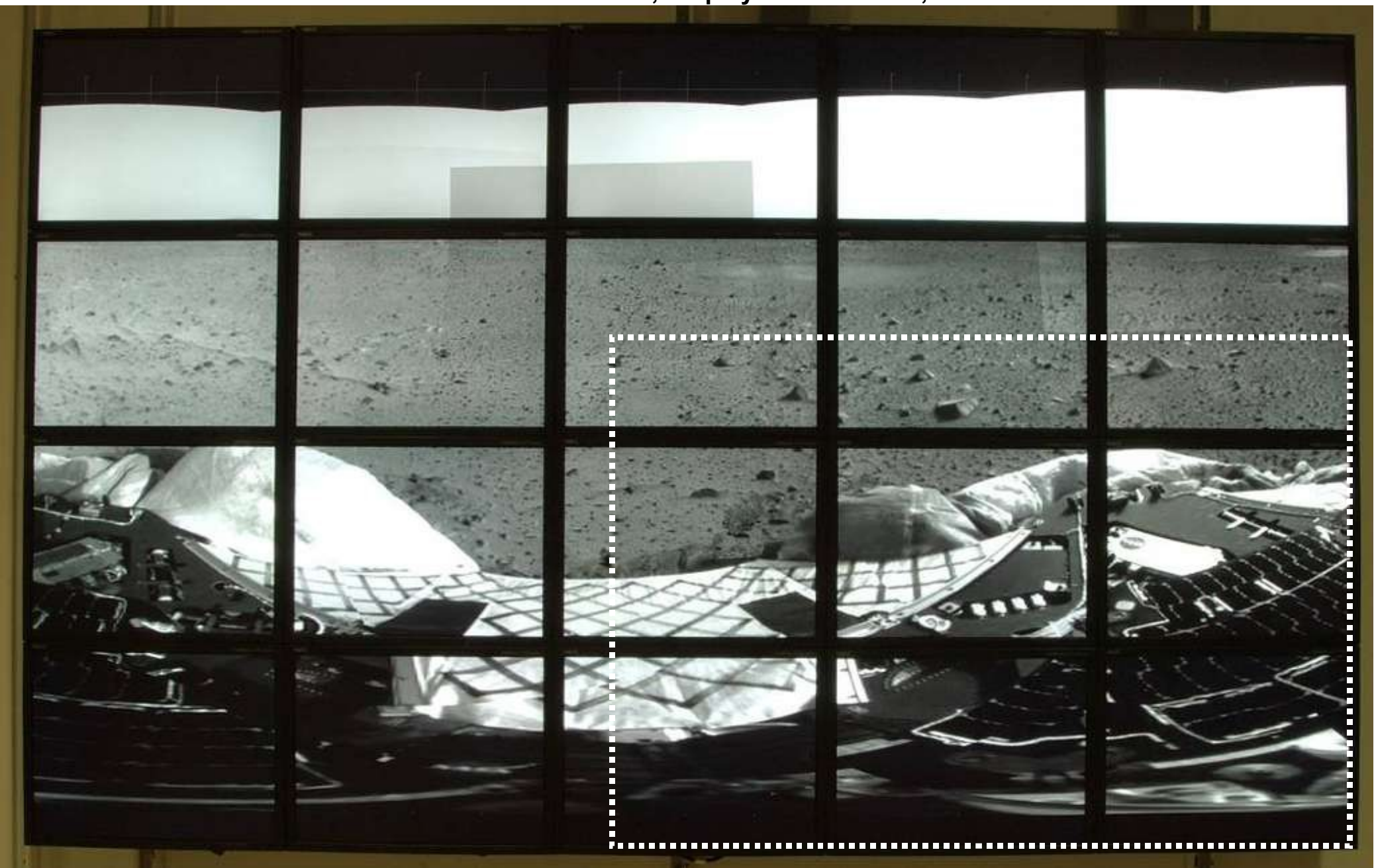


OptIPuter

it²

Tiled Displays Allow for Both Global Context and High Levels of Detail—150 MPixel Rover Image on 40 MPixel OptIPuter Visualization Node Display

"Source: Data from JPL/Mica; Display UCSD NCMIR, David Lee"



Interactively Zooming In Using UIC's Electronic Visualization Lab's JuxtaView Software

"Source: Data from JPL/Mica; Display UCSD NCMIR, David Lee"



Highest Resolution Zoom

"Source: Data from JPL/Mica; Display UCSD NCMIR, David Lee"



Toward an Interactive Gigapixel Display

- Scalable Adaptive Graphics Environment (SAGE) Controls:
- 100 Megapixels Display
  - 55-Panel
- 1/4 TeraFLOP
  - Driven by 30-Node Cluster of 64-bit Dual Opterons
- 1/3 Terabit/sec I/O
  - 30 x 10GE interfaces
  - Linked to OptIPuter
- 1/8 TB RAM
- 60 TB Disk

NSF LambdaVision MRI@UIC

Calit2 is Building a LambdaVision Wall in Each of the UCI & UCSD Buildings

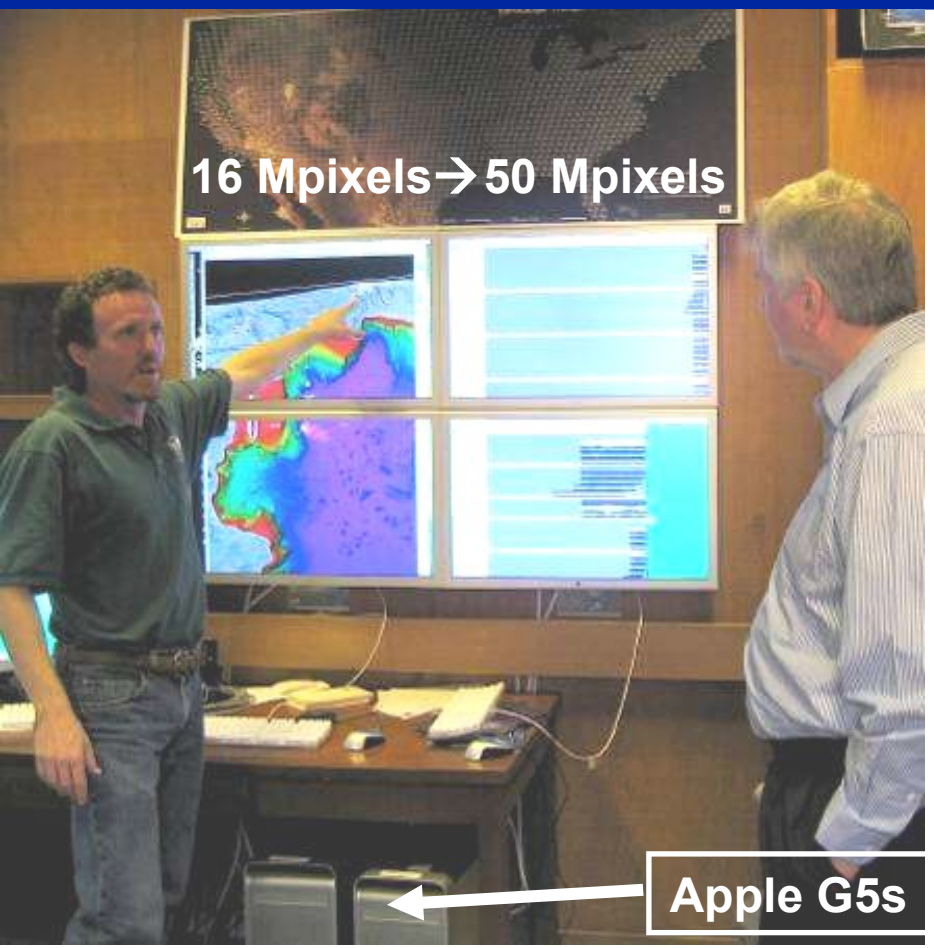


OptIPuter

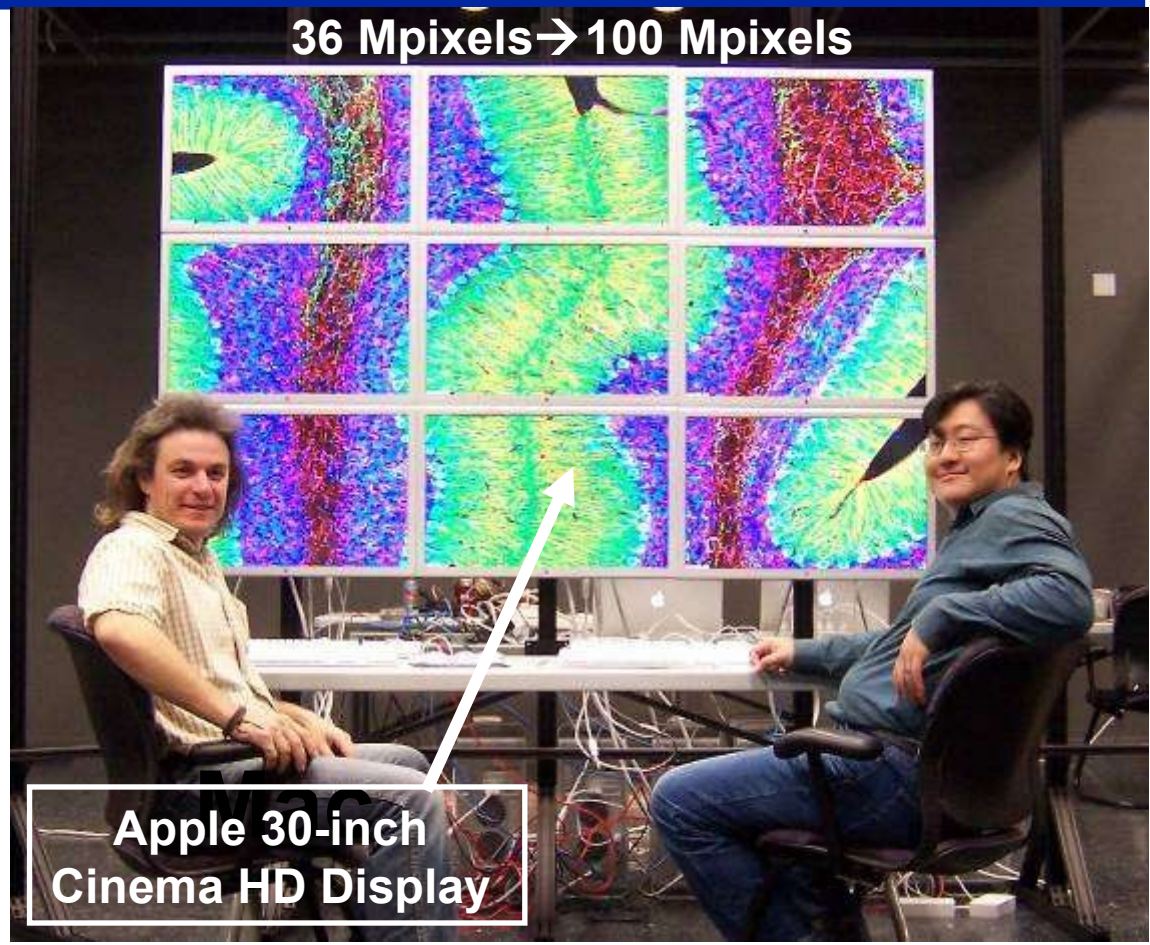
Source: Jason Leigh, Tom DeFanti, EVL@UIC
OptIPuter Co-PIs

it2

OptIPuter Scalable Displays Have Been Extended to Apple-Based Systems "iWall Driven by iCluster"



Source: Atul Nayak, SIO
Collaboration of
Calit2/SIO/OptIPuter/USArray



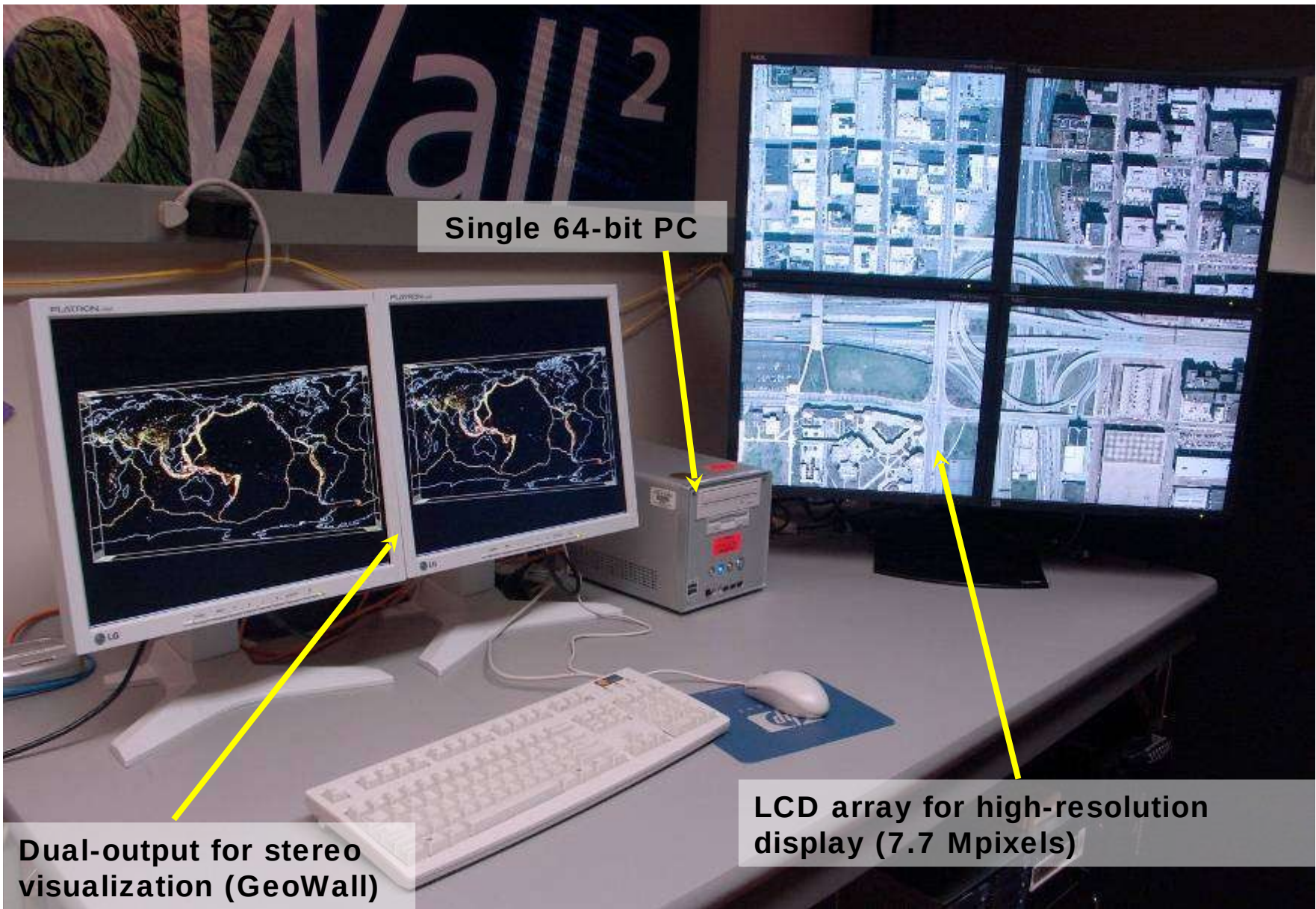
Source: Falko Kuester, Calit2@UCI
NSF Infrastructure Grant

OptIPuter

See GEON Poster: iCluster : Visualizing USArray Data on a Scalable High Resolution Tiled Display Using the OptIPuter

Personal GeoWall 2 (PG2): Individual OptIPuter User Node

Demonstrated by EVL (UIC) at 4th GeoWall Consortium Meeting



OptIPuter

SDSC/Calit2 Synthesis Center
You Will Be Visiting This Week

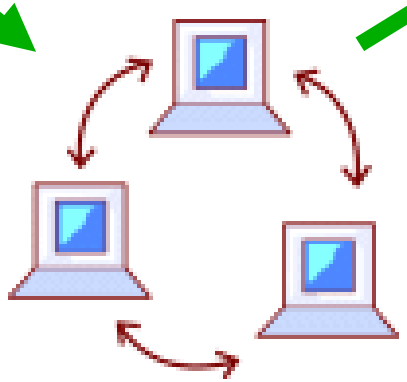


Collaboration to Set Up Experiments





Collaboration to Study Experimental Results



Collaboration to Run Experiments

NSF

GEON

Cyberinfrastructure for the Geosciences

www.geongrid.org

OptIPuter

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

UC San Diego

it²

The Synthesis Center is an Environment Designed for Collaboration with Remote Data Sets

- **Environment With …**
  - **Large-scale, Wall-sized Displays**
  - **Links to On-Demand Cluster Computer Systems**
  - **Access to Networks of Databases and Digital Libraries**
  - **State-of-the-Art Data Analysis and Mining Tools**
- **Linked, "Smart" Conference Rooms Between SDSC and Calit2 Buildings on UCSD and UCI Campuses**
- **Coupled to OptIPuter Planetary Infrastructure**

Currently in SDSC Building
Future Expansion into Calit2@UCSD Building

OptIPuter

it²

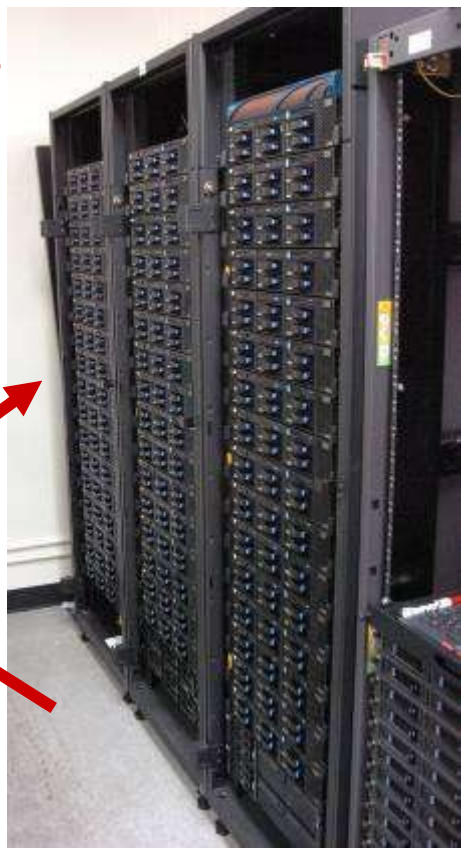
Campuses Must Provide Fiber Infrastructure to End-User Laboratories & Large Rotating Data Stores

SIO Ocean Supercomputer



UCSD Campus LambdaStore Architecture

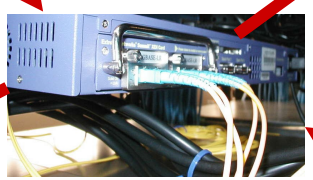
IBM Storage Cluster



Streaming Microscope







2 Ten Gbps Campus Lambda Raceway





Global LambdaGrid

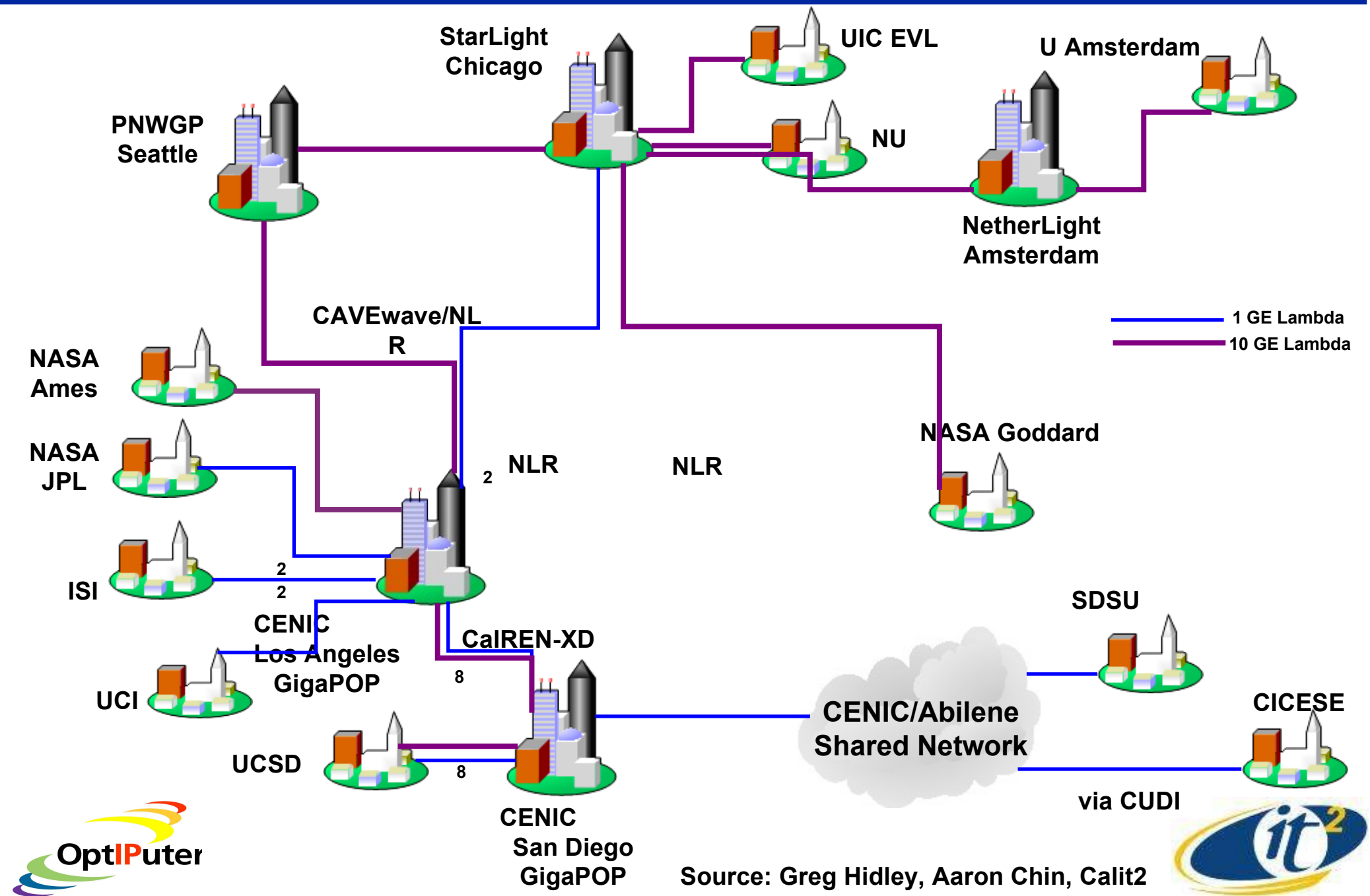


OptIPuter

Source: Phil Papadopoulos, SDSC, Calit2

it2

The OptIPuter LambdaGrid is Rapidly Expanding



Interactive Retrieval and Hyperwall Display of Earth Sciences Images Using NLR

Enables Scientists To Perform Coordinated Studies Of Multiple Remote-Sensing 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

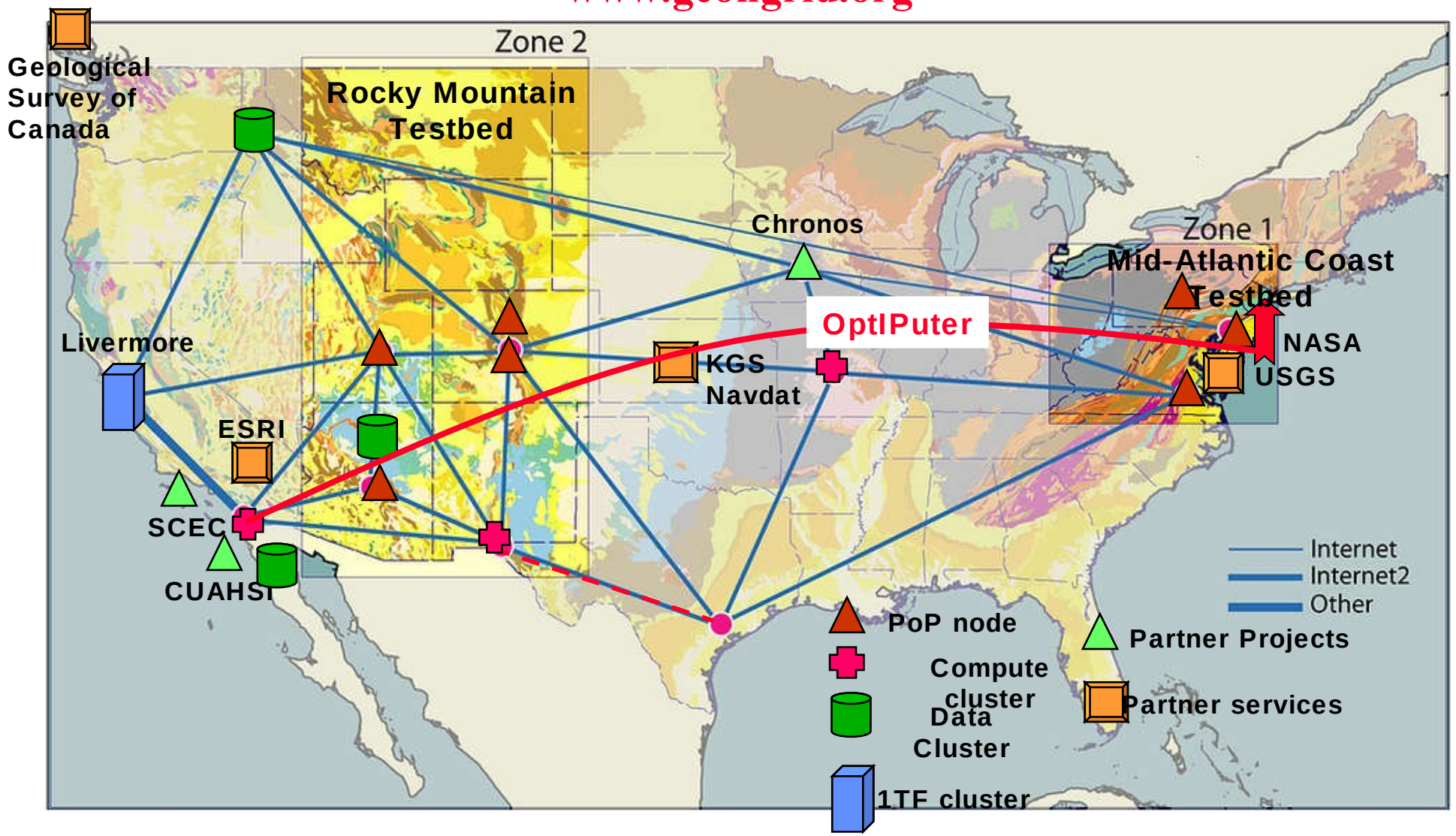
OptIPuter

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

it2

The GEONgrid: Building on the OptIPuter with NASA Goddard

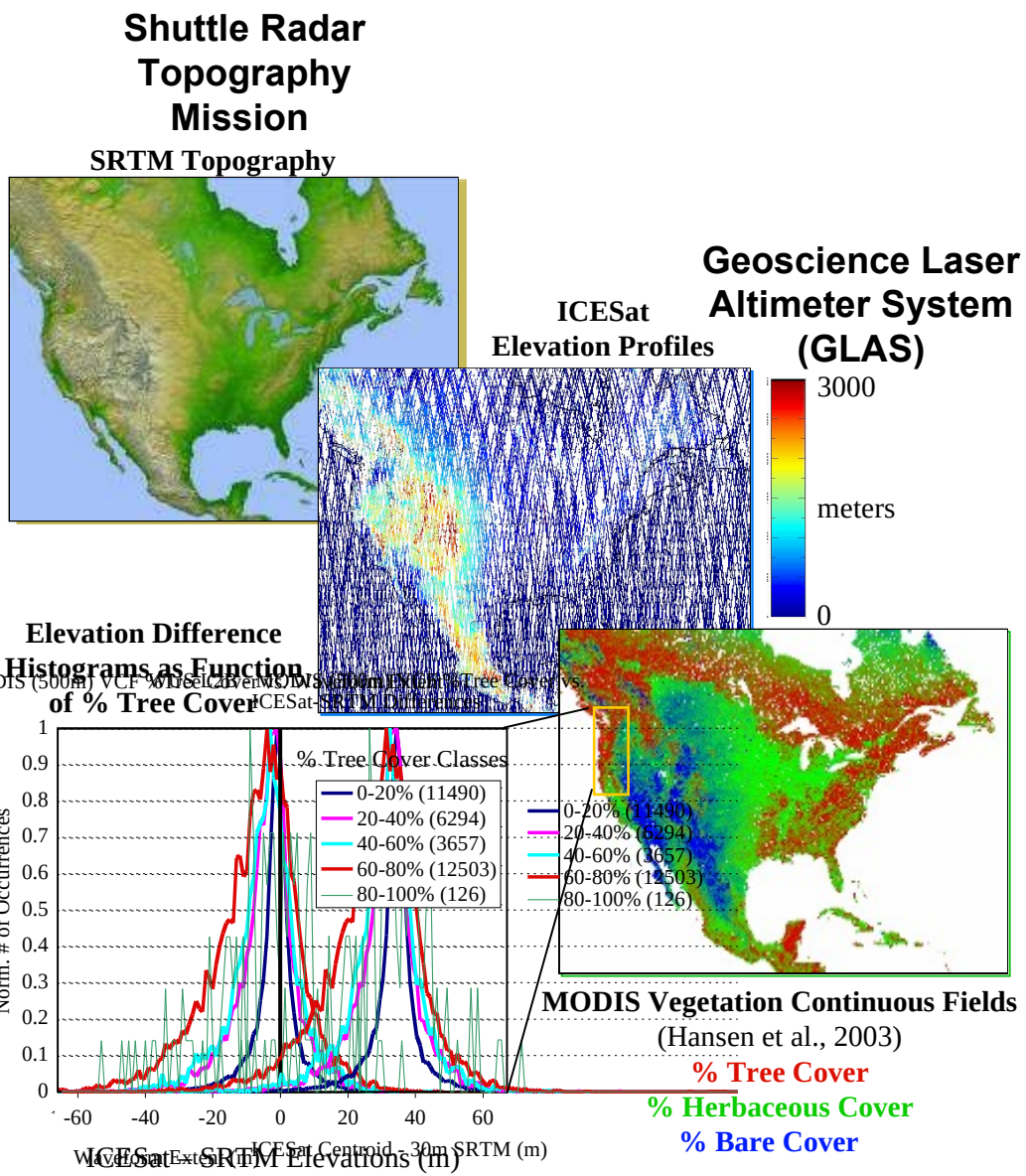
www.geongrid.org



Source: Chaitan Baru, SDSC

NLR GSFC/JPL/SIO Application: Integration of Laser and Radar Topographic Data with Land Cover Data

- **Merge the 2 Data Sets, Using SRTM to Achieve Good Coverage & GLAS to Generate Calibrated Profiles**
- **Interpretation Requires Extracting Land Cover Information from Landsat, MODIS, ASTER, and Other Data Archived in Multiple DAACs**
- **Use of the OptIPuter over NLR and Local Data Mining and Sub-Setting Tools on NASA ECHO Data Pools will Permit Systematic Fusion Of Global Data Sets, Which are Not Possible with Current Bandwidth**



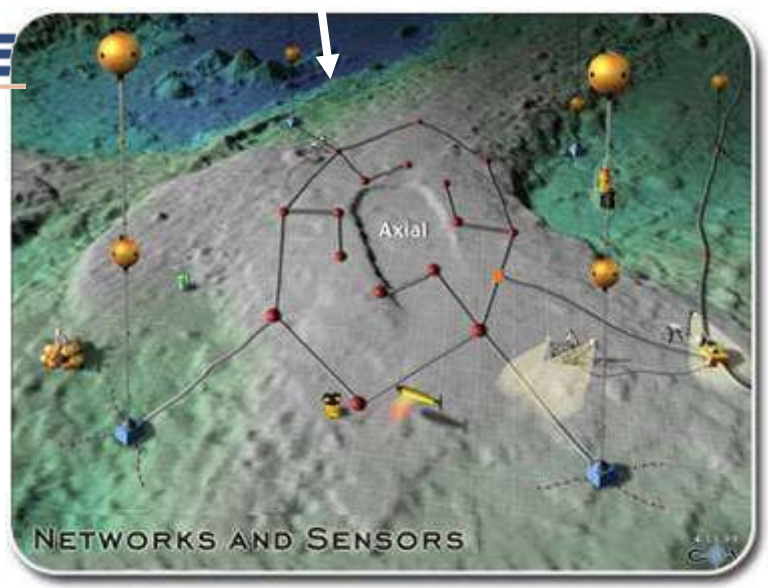
Key Contacts: H.K. Ramapriyan, R. Pfister, C. Carabajal, C. Lynn, D. Harding, M. Seablom, P. Gary GSFC; T. Yunck, JPL; B. Minster, SIO; L. Smarr, UCSD, S. Graves, UTA

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

NSF's Ocean Observatories Initiative (OOI) Envisions Global, Regional, and Coastal Scales



NEPTUNE



LEO15 Inset Courtesy of Rutgers University, Institute of Marine and Coastal Sciences

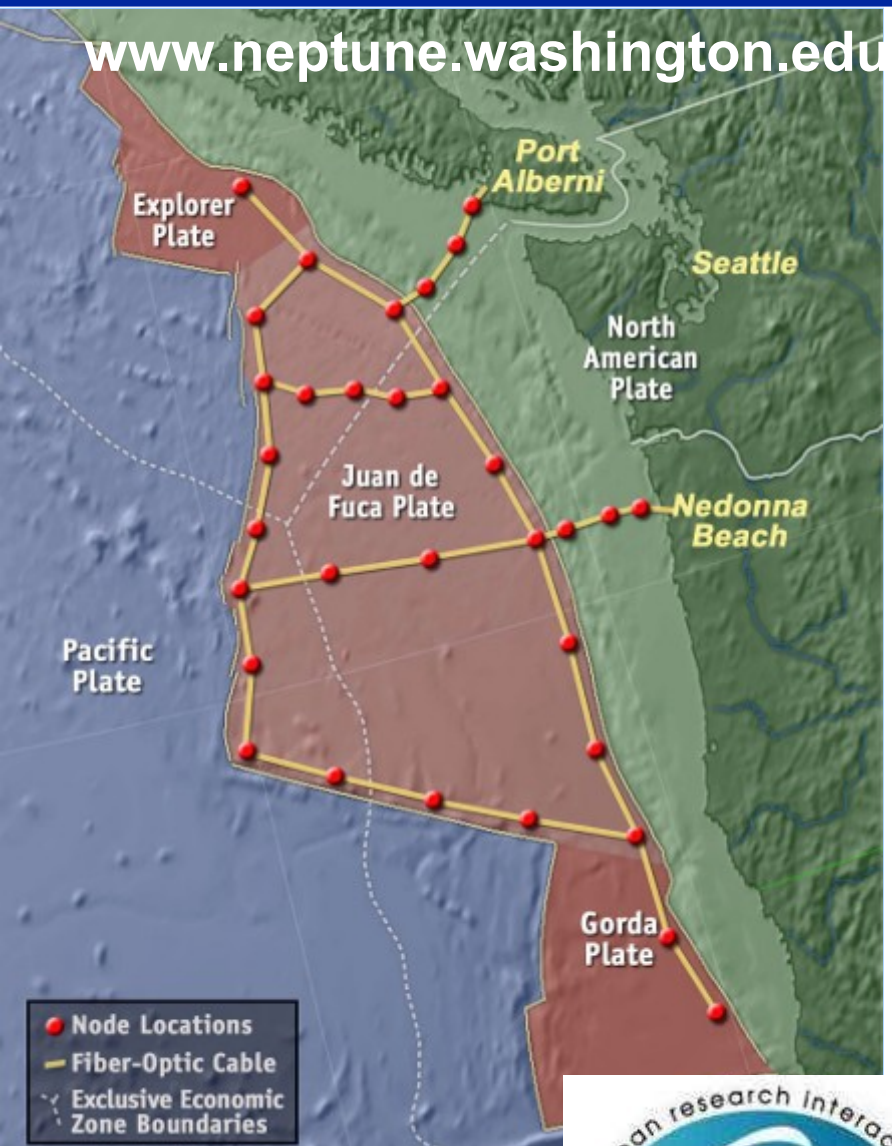
NSF

ocean research interactive ORION observatory networks

OptIPuter

it2

Adding Web and Grid Services to Lambdas to Provide Real Time Control of Ocean Observatories



LOOKING:
(Laboratory for the Ocean Observatory Knowledge Integration Grid)

http://lookingtosea.ucsd.edu/

- **Goal:**
  - Prototype Cyberinfrastructure for NSF's Ocean Research Interactive Observatory Networks (ORION)
- **LOOKING NSF ITR with PIs:**
  - John Orcutt & Larry Smarr - UCSD
  - John Delaney & Ed Lazowska –UW
  - Mark Abbott – OSU
- **Collaborators at:**
  - MBARI, WHOI, NCSA, UIC, CalPoly, UVic, CANARIE, Microsoft, NEPTUNE-Canarie

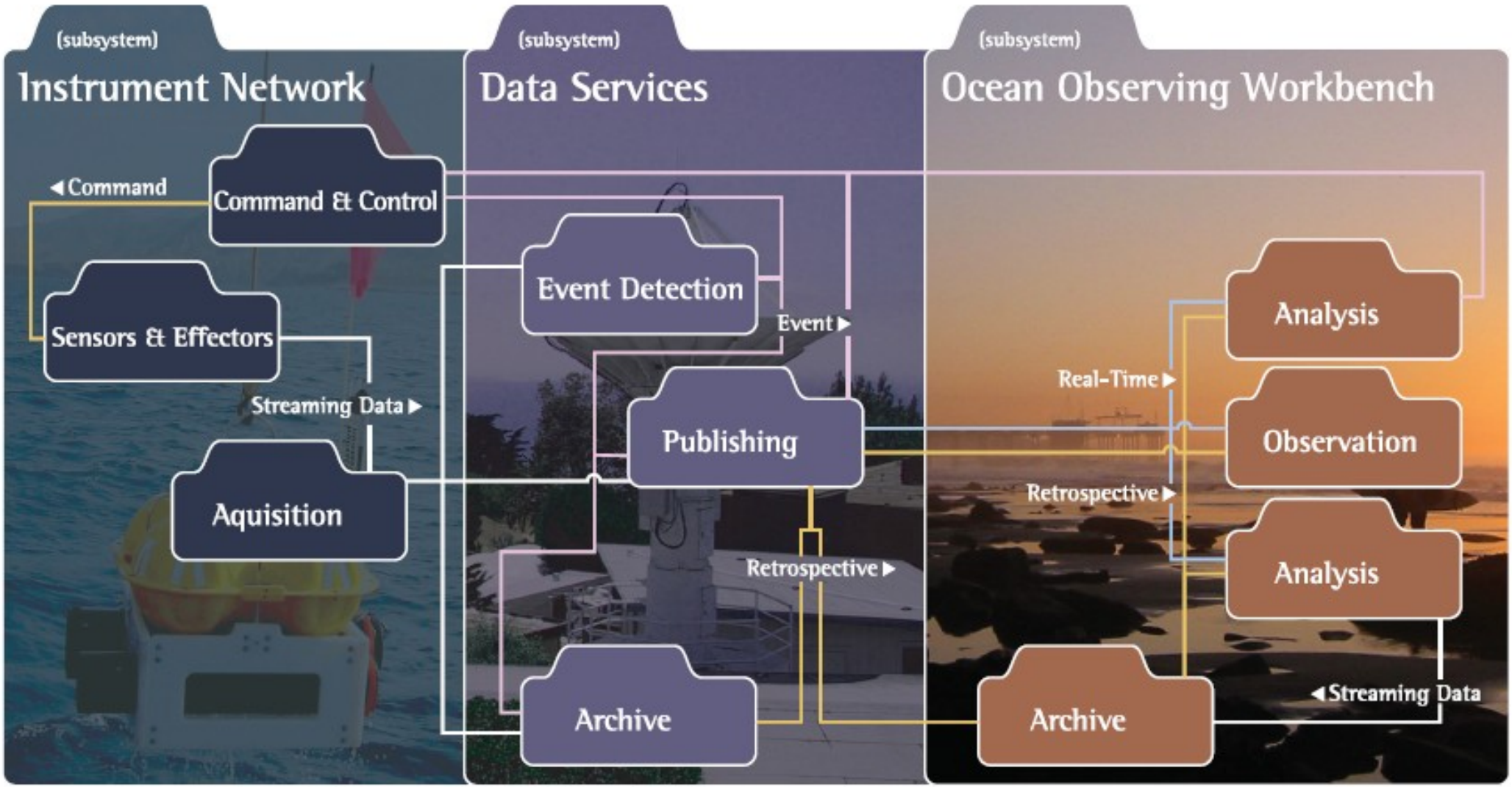
OptIPuter

NSF

ORION ocean research interactive observatory networks

it²

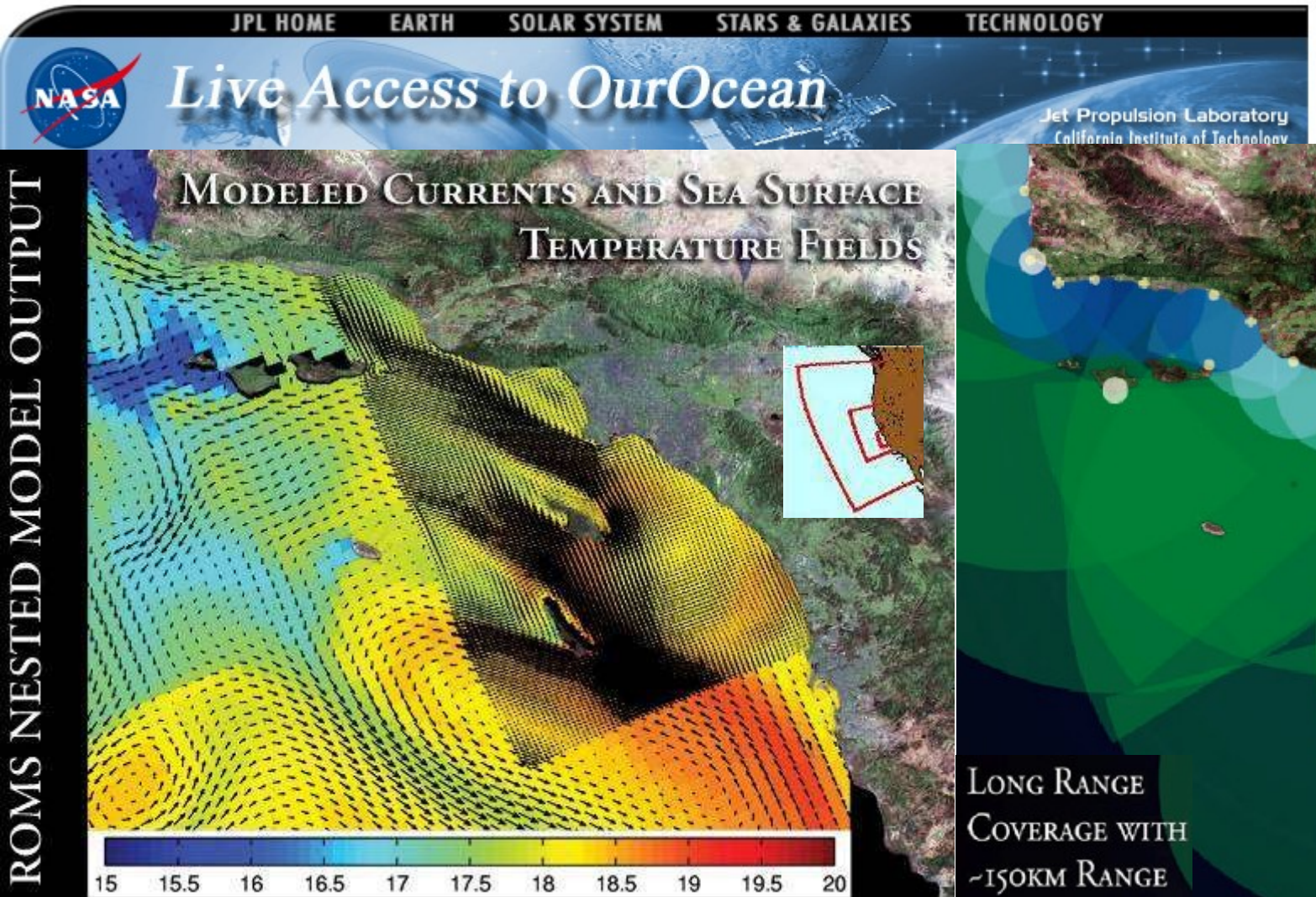
Looking High Level
LOOKING Service System Architecture

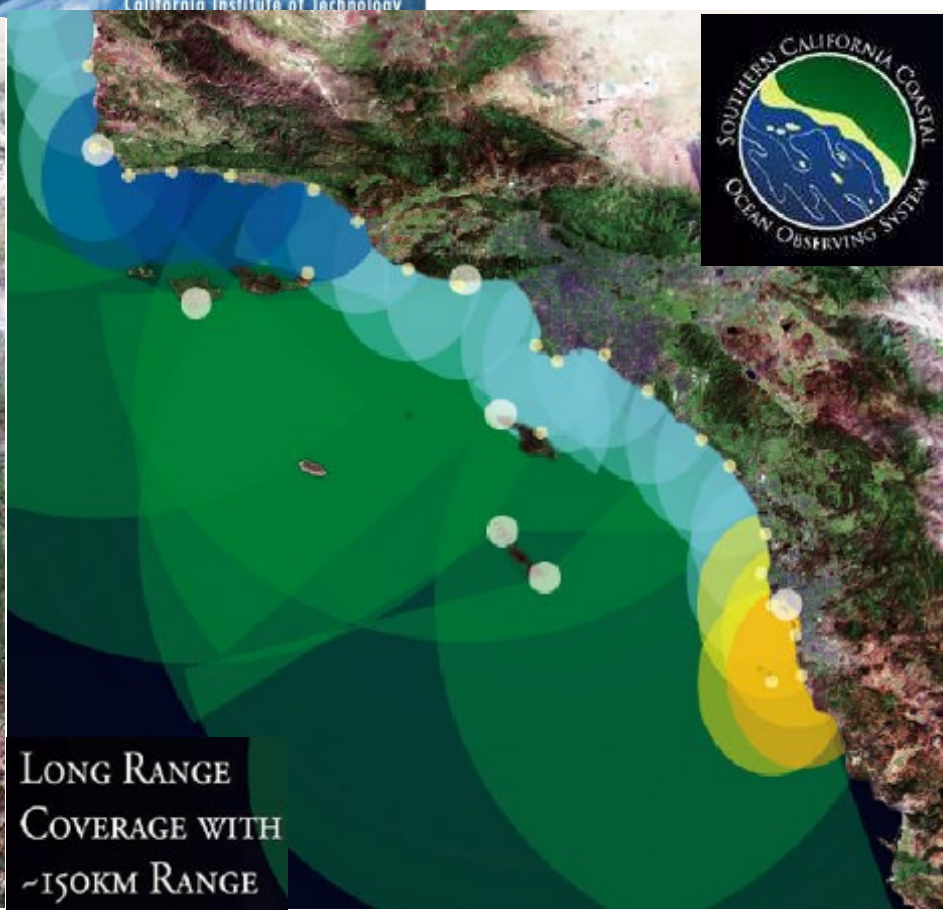


OptIPuter

it2

Use OptIPuter to Couple Data Assimilation Models to Remote Data Sources and Analysis





JPL ONR MBARI California Institute of Technology

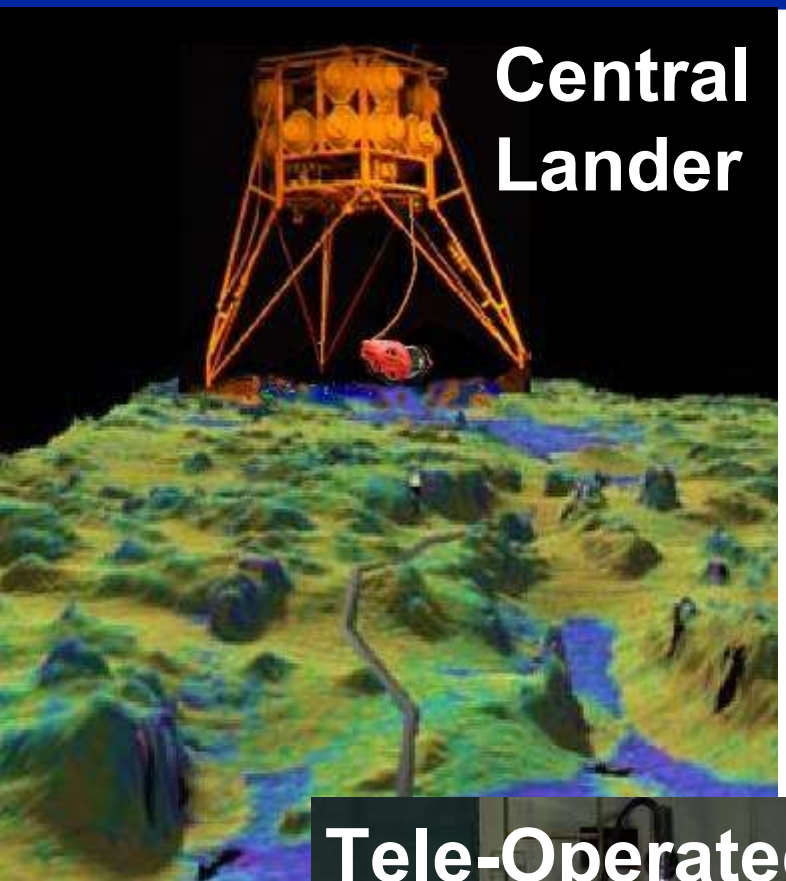
OptIPuter

Regional Ocean Modeling System (ROMS)

http://ourocean.jpl.nasa.gov/

it2

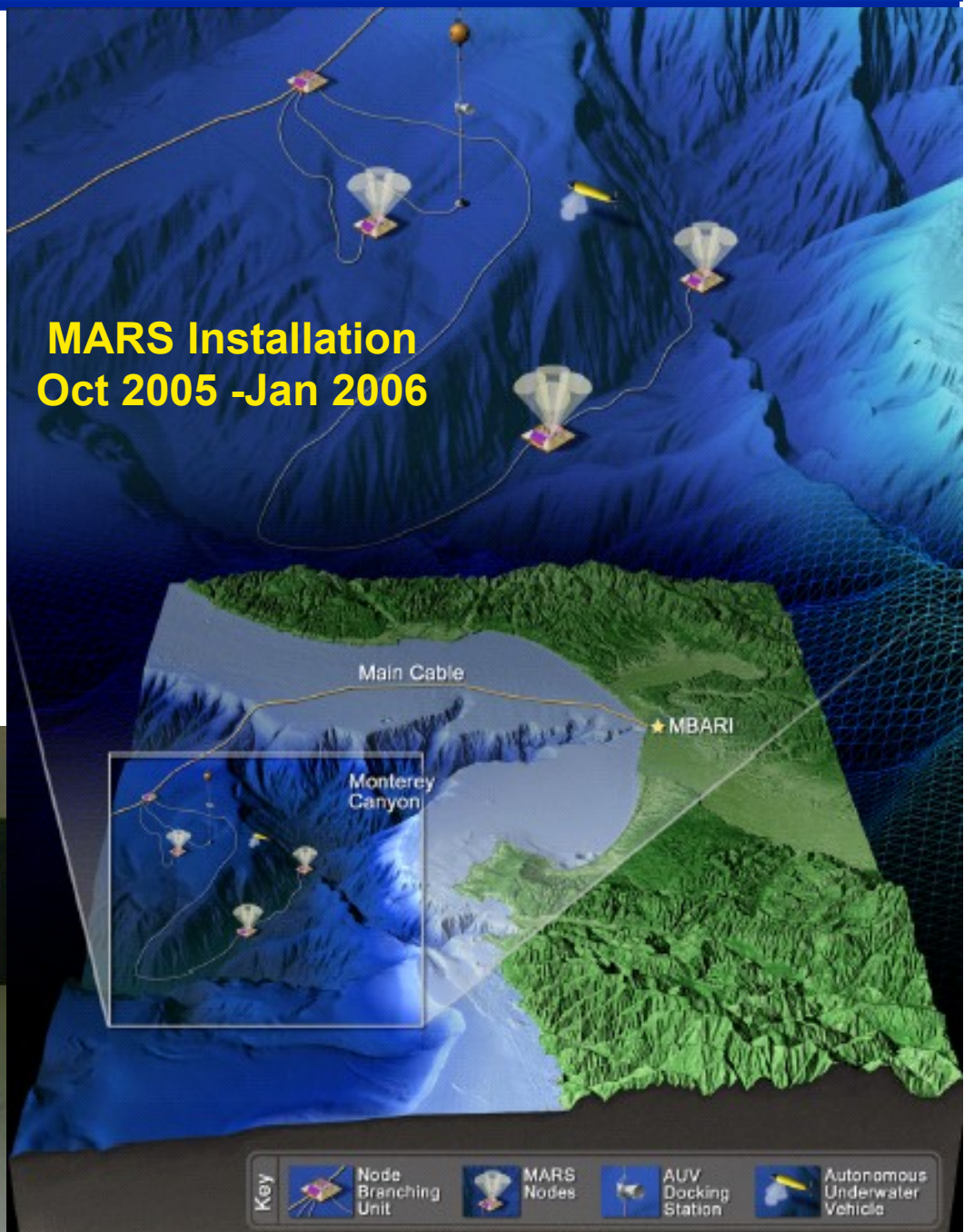
MARS Cable Observatory Testbed – LOOKING Living Laboratory



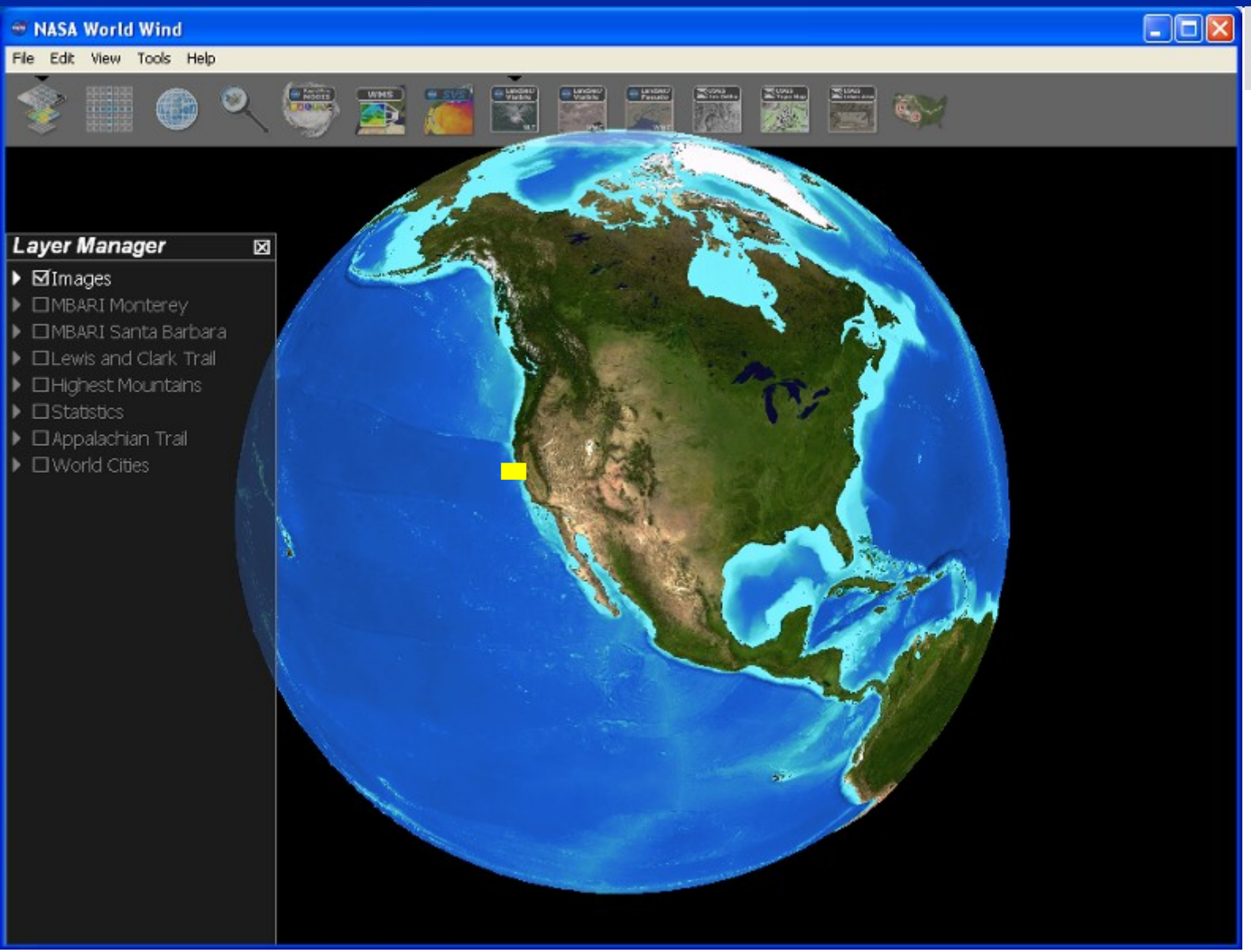


Source:
Jim
Bellingham,
MBARI

OptIPuter



Using NASA's World Wind to Integrate Ocean Observing Data Sets



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

SDSU and SDSC are Increasing the WW Data Access Bandwidth

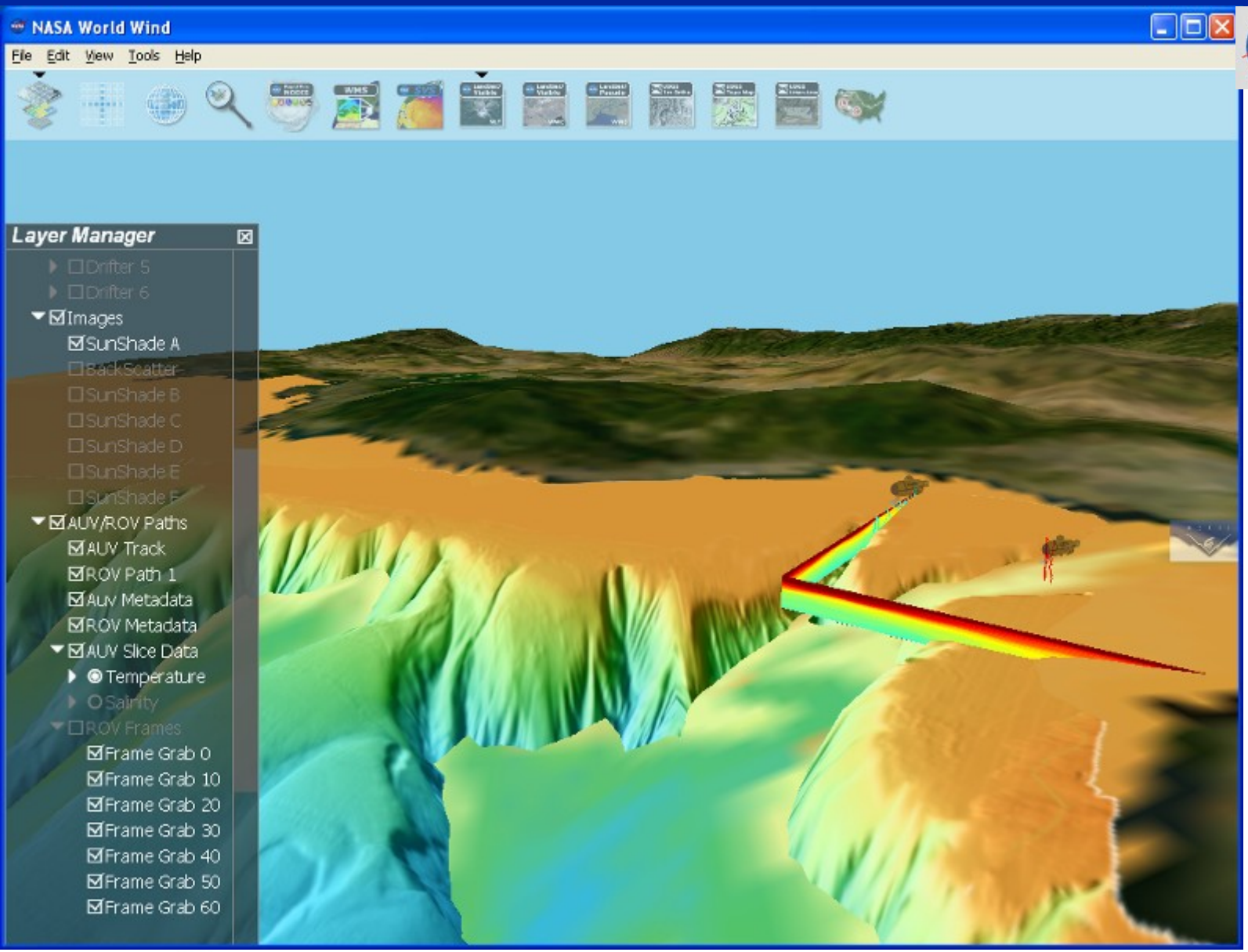
SDSC will be Serving as a National Data Repository for WW Datasets

OptIPuter

Source: Ed Lazowska, Keith Grochow, UWash

it2

Zooming Into Monterey Bay Showing Temperature Profile of an MBARI Remotely Operated Vehicle



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

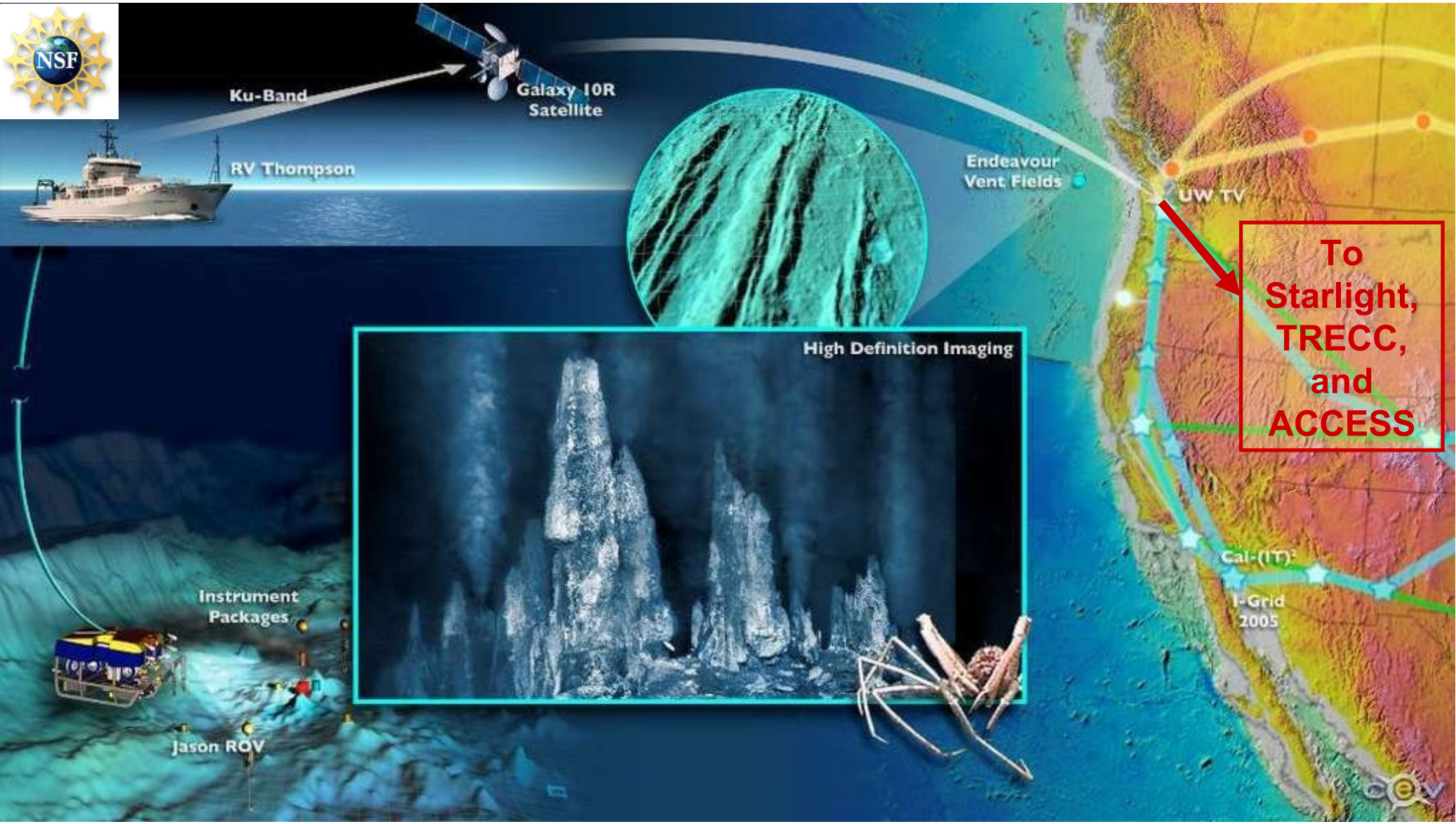
UW, as part of LOOKING, is Enhancing the WW Client to Allow Oceanographic Data to be Visualized

OptIPuter

Source: Ed Lazowska, Keith Grochow, UWash

it2

Proposed Experiment for iGrid 2005 – Remote Interactive HD Imaging of Deep Sea Vent



OptIPuter NEPTUNE

Source John Delaney & Deborah Kelley, UWash

it²