Supercomputer End Users: the OptIPuter Killer Application

Keynote

TeraGrid '08

Las Vegas, NV

June 11, 2008

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

Abstract

During the last few years, a radical restructuring of optical networks supporting e-Science projects has occurred around the world. U.S. universities are beginning to acquire access to high bandwidth lightwaves (termed "lambdas") on fiber optics through the National LambdaRail, Internet2's Circuit Services, and the Global Lambda Integrated Facility. The NSF-funded OptIPuter project explores how user controlled 1- or 10- Gbps lambdas can provide direct access to global data repositories, scientific instruments, and computational resources from the researcher's Linux clusters in their campus laboratories. These end user clusters are reconfigured as "OptIPortals," providing the end user with local scalable visualization, computing, and storage. I will describe how this user configurable OptIPuter global platform opens new frontiers in collaborative work environments, digital cinema, biomedical instruments,and marine microbial metagenomics. However, a major new user community should be end users of TeraGrid, allowing them to directly optically connect to the remote Tera or Peta-scale resource from their local laboratories.

OptIPuter

it²

Interactive Supercomputing Collaboratory Prototype:
Using Analog Communications to Prototype the Fiber Optic Future

"What we really have to do is eliminate distance between individuals who want to interact with other people and with other computers."
— Larry Smarr, Director, NCSA

SIGGRAPH 1989







**"We're using satellite technology…
to demo what It might be like to have
high-speed fiber-optic links between
advanced computers
in two different geographic locations."**

— Al Gore, Senator
Chair, US Senate Subcommittee on Science, Technology and Space

OptIPuter

NCSA™

it²

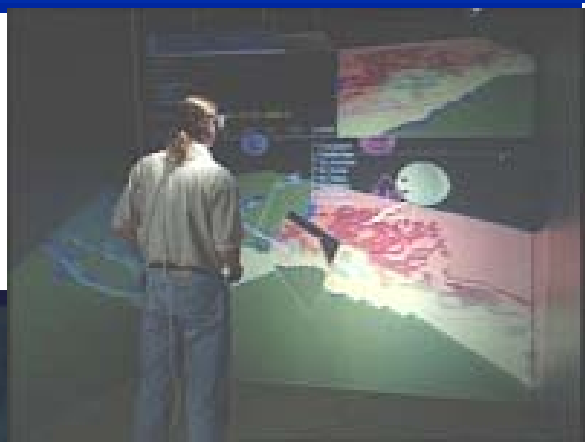
Chesapeake Bay Simulation Collaboratory : vBNS Linked CAVE, ImmersaDesk, Power Wall, and Workstation

Alliance Project: Collaborative Video Production via Tele-Immersion and Virtual Director

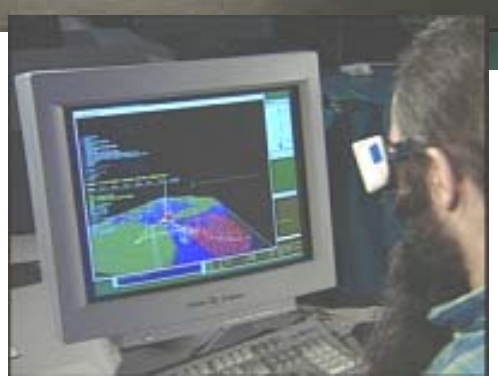
Alliance Application Technologies
Environmental Hydrology Team







Alliance 1997





NCSA™ UIC

sgi

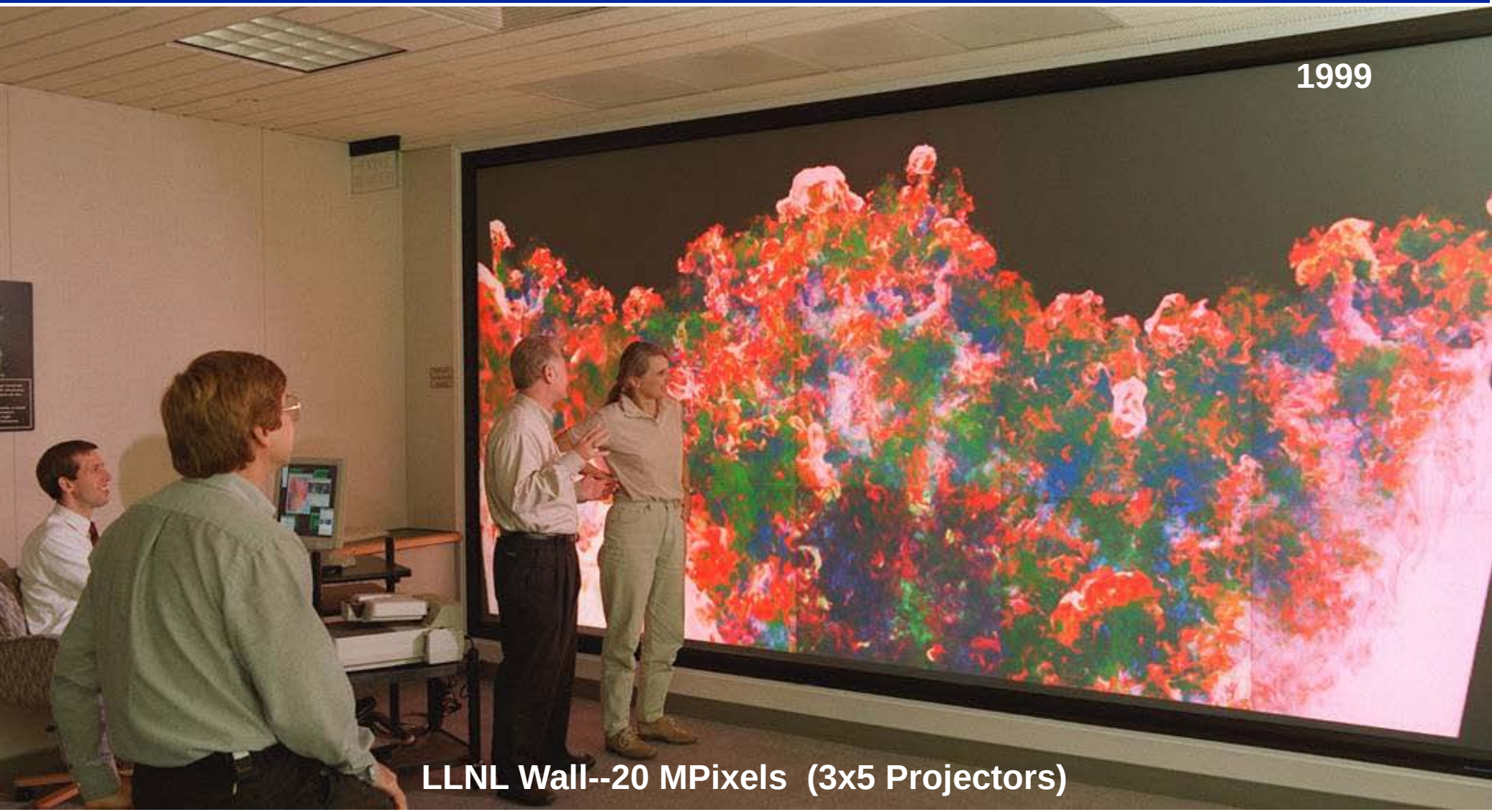
ALLIANCE

OptIPuter

Donna Cox, Robert Patterson, Stuart Levy, NCSA Virtual Director Team
Glenn Wheless, Old Dominion Univ.

it²

ASCI Brought Power Walls to the Frontier of Supercomputing



OptIPuter

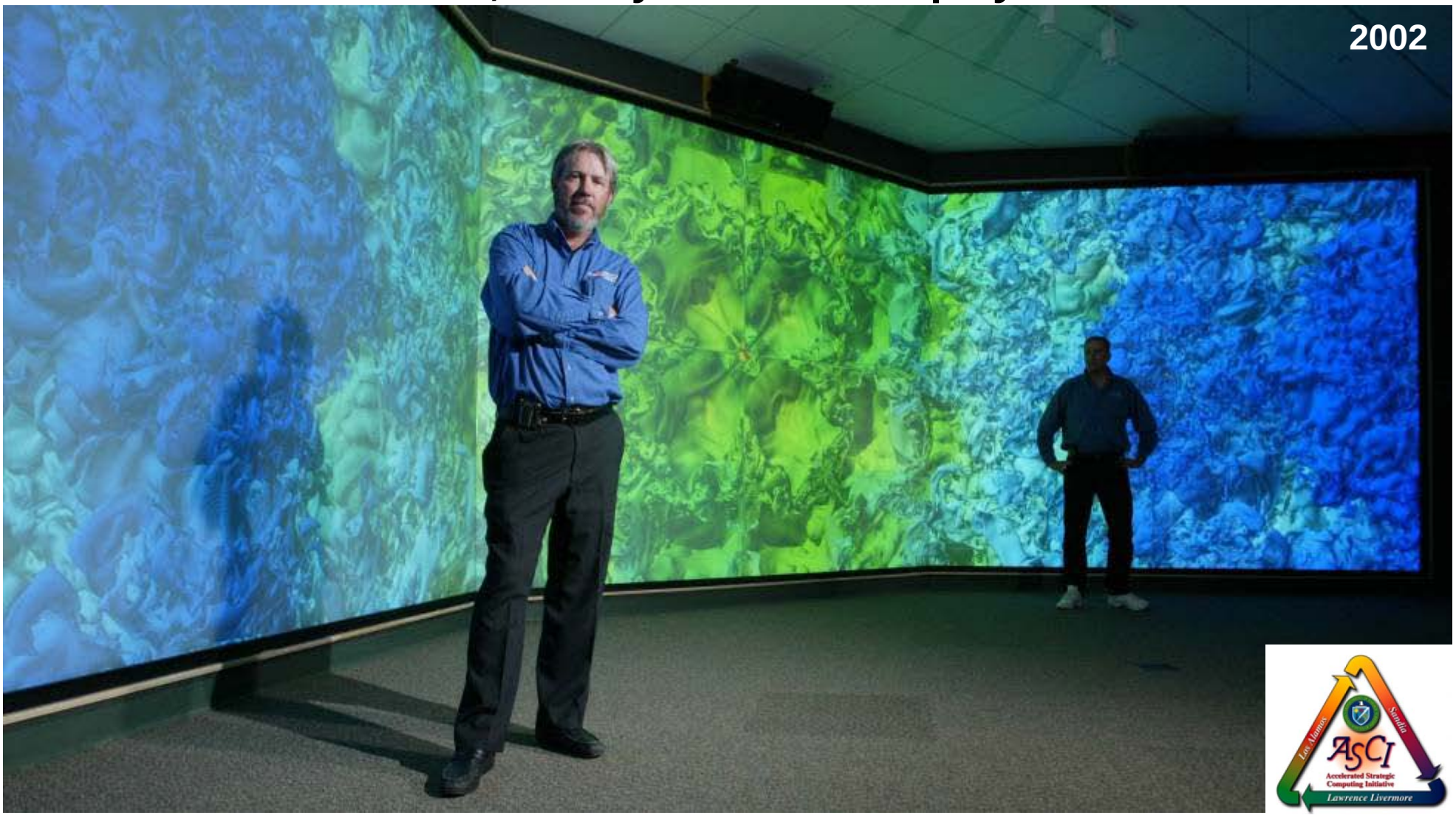
An Early sPPM Simulation Run
Source: LLNL

ASCI Accelerated Strategic Computing Initiative
Los Alamos · Sandia · Lawrence Livermore

it²

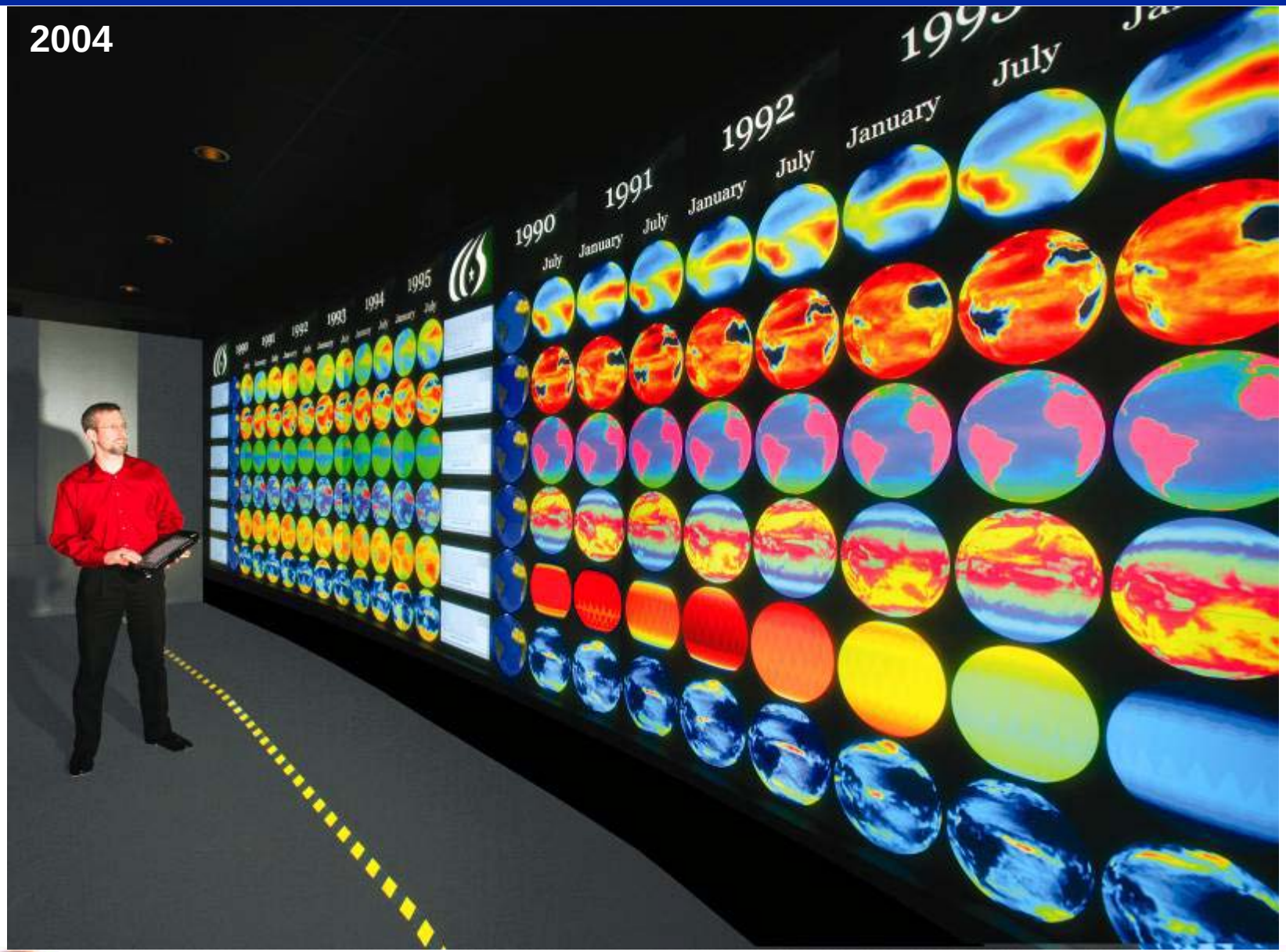
60 Million Pixels Projected Wall Driven By Commodity PC Cluster

At 15 Frames/s, The System Can Display 2.7 GB/Sec



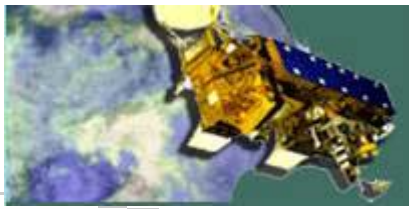
Source: Philip D. Heermann, DOE ASCI Program

Oak Ridge National Laboratory Uses Tiled Projector Walls to Analyze Simulations



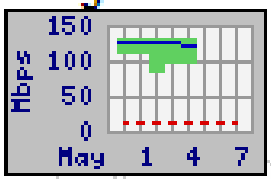
35Mpixel EVEREST Display ORNL

Challenge: Average Throughput of NASA Data Products to End User is ~ 50 Mbps

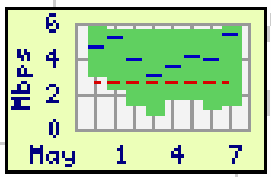


EOS Active Network Testing
AQUA Mission Destinations

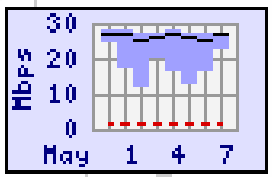
Oregon St.



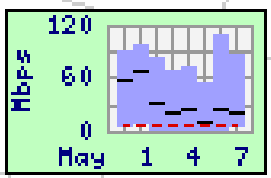
RSS



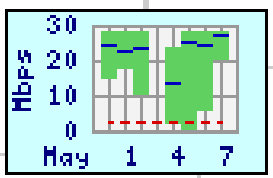
Montana



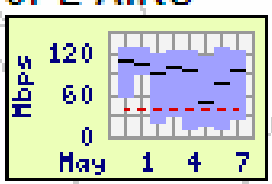
UCSB



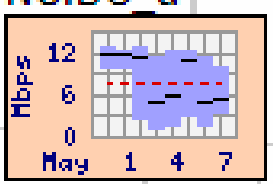
Colo State



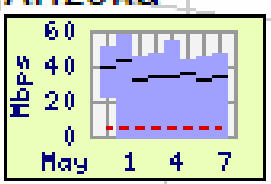
JPL-AIRS



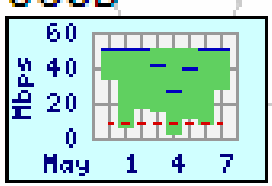
NSIDC_u



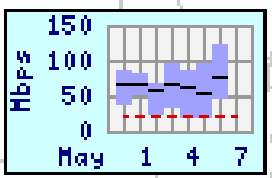
Arizona



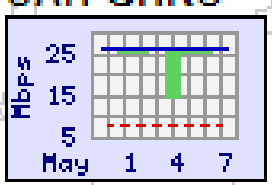
UCSD



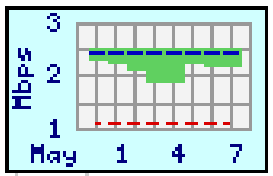
Wisconsin



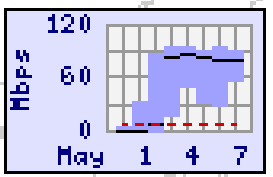
UAH-GHRC



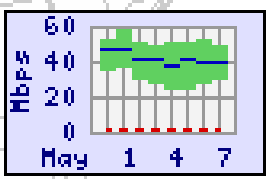
UCL



Boston U



SUNY-SB



Tested
May 2008

Internet2 Backbone is 10,000 Mbps!
Throughput is < 0.5% to End User

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

OptIPuter

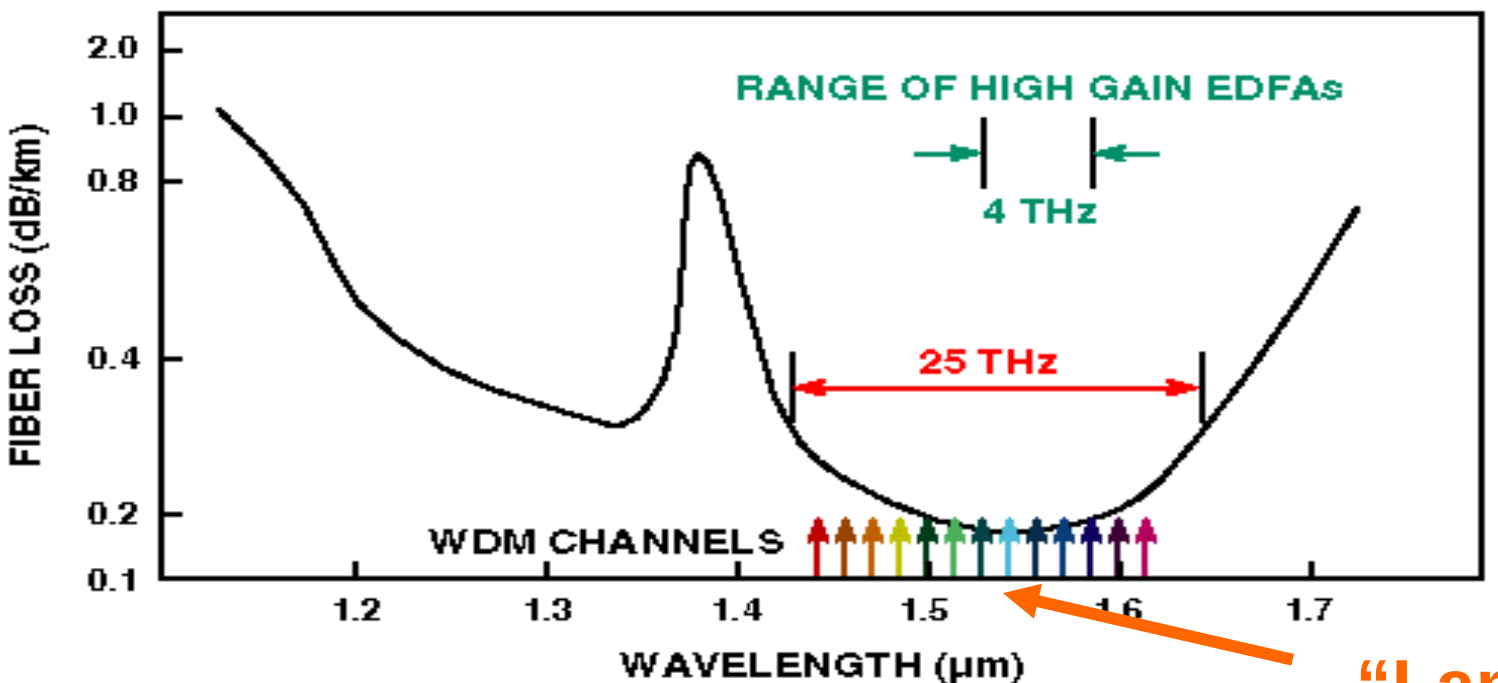
i2

Dedicated Optical Fiber Channels Makes High Performance Cyberinfrastructure Possible

WAVELENGTH DIVISION MULTIPLEXING (WDM)

10 Gbps per User ~ 500x Shared Internet Throughput

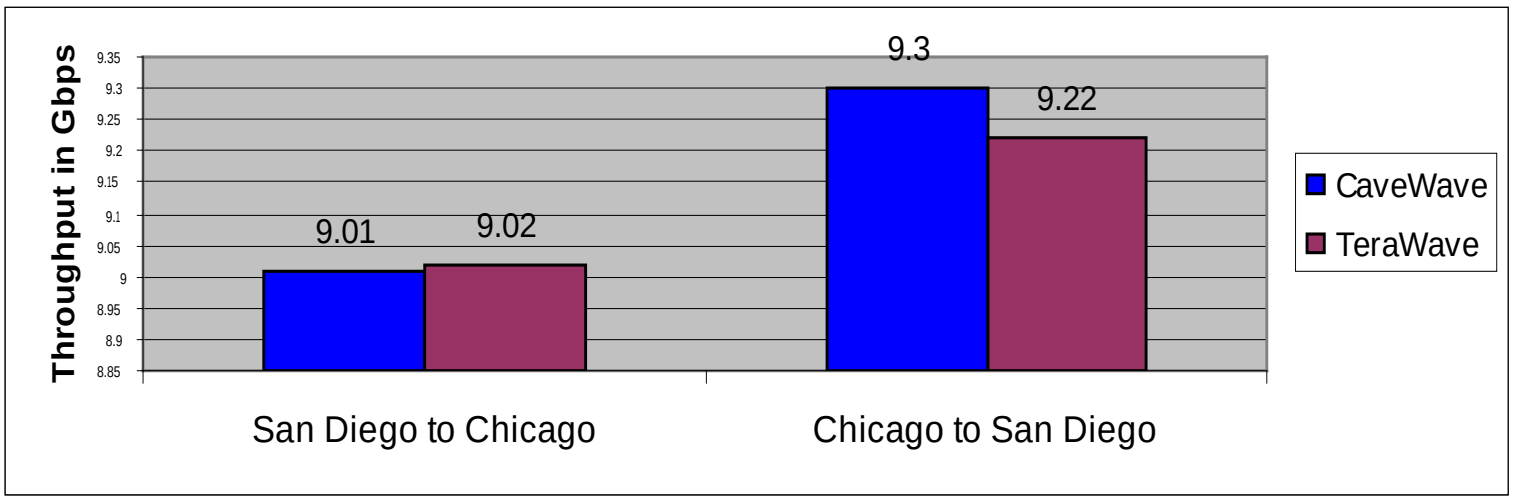
- EXPLOITS
  - ENORMOUS BANDWIDTH OF SILICA FIBER
  - HIGH-GAIN WIDEBAND OPTICAL AMPLIFIERS



"Lambdas"

Parallel Lambdas are Driving Optical Networking The Way Parallel Processors Drove 1990s Computing

9GbpsOut of 10 Gbps Disk-to-Disk Performance Using LambdaStream between EVL and Calit2



CAVEWave:
20 senders to 20 receivers (point to point)

Effective Throughput = **9.01** Gbps
(San Diego to Chicago)
450.5 Mbps disk to disk transfer per stream

Effective Throughput = **9.30** Gbps
(Chicago to San Diego)
465 Mbps disk to disk transfer per stream

TeraGrid:
20 senders to 20 receivers (point to point)

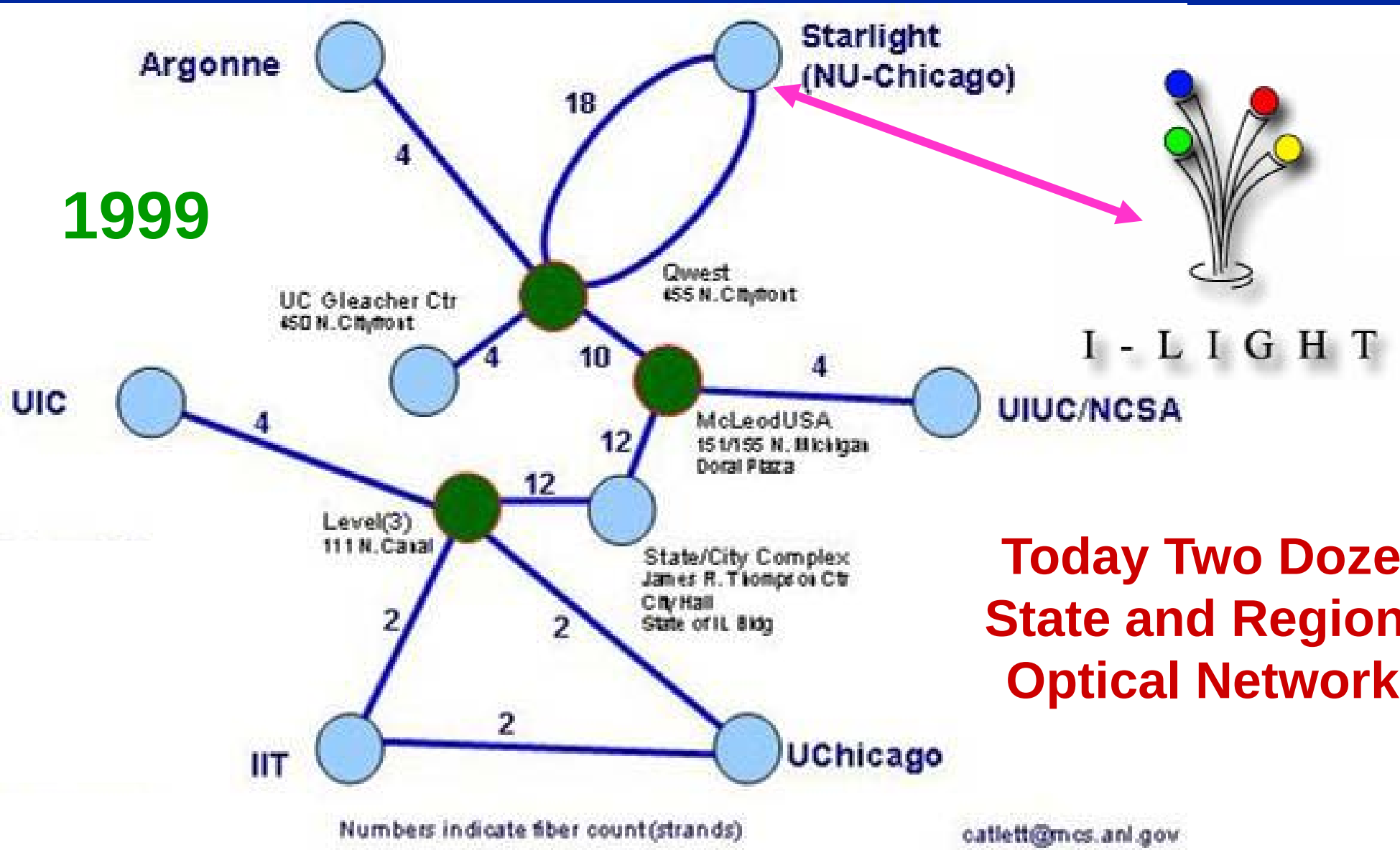
Effective Throughput = **9.02** Gbps
(San Diego to Chicago)
451 Mbps disk to disk transfer per stream

Effective Throughput = **9.22** Gbps
(Chicago to San Diego)
461 Mbps disk to disk transfer per stream

Dataset: 220GB Satellite Imagery of Chicago courtesy USGS.
Each file is 5000 x 5000 RGB image with a size of 75MB i.e ~ 3000 files

Source: Venkatram Vishwanath, UIC EVL

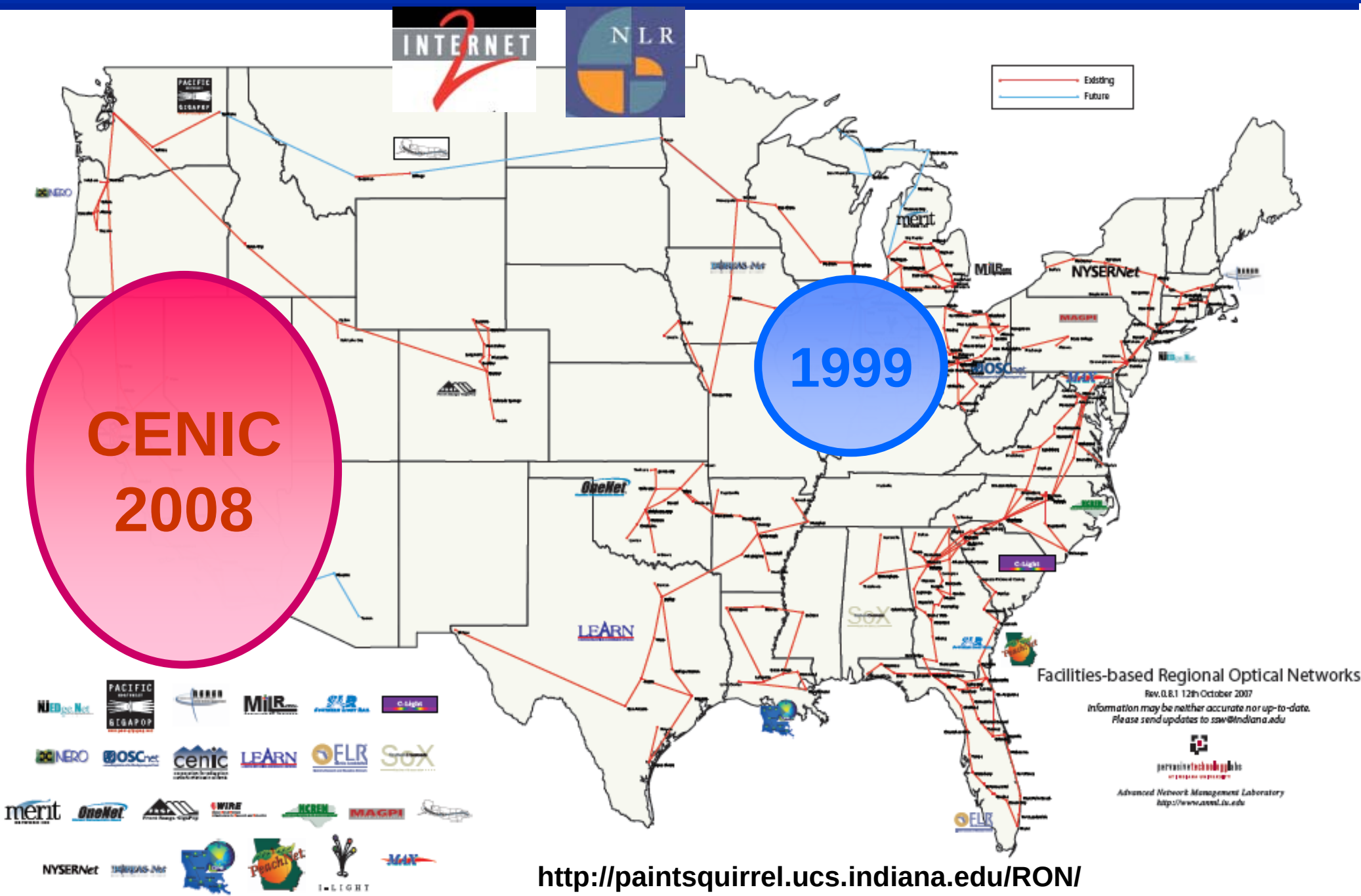
States Have Been Acquiring Their Own Dark Fiber for a Decade -- Illinois's I-WIRE and Indiana's I-LIGHT



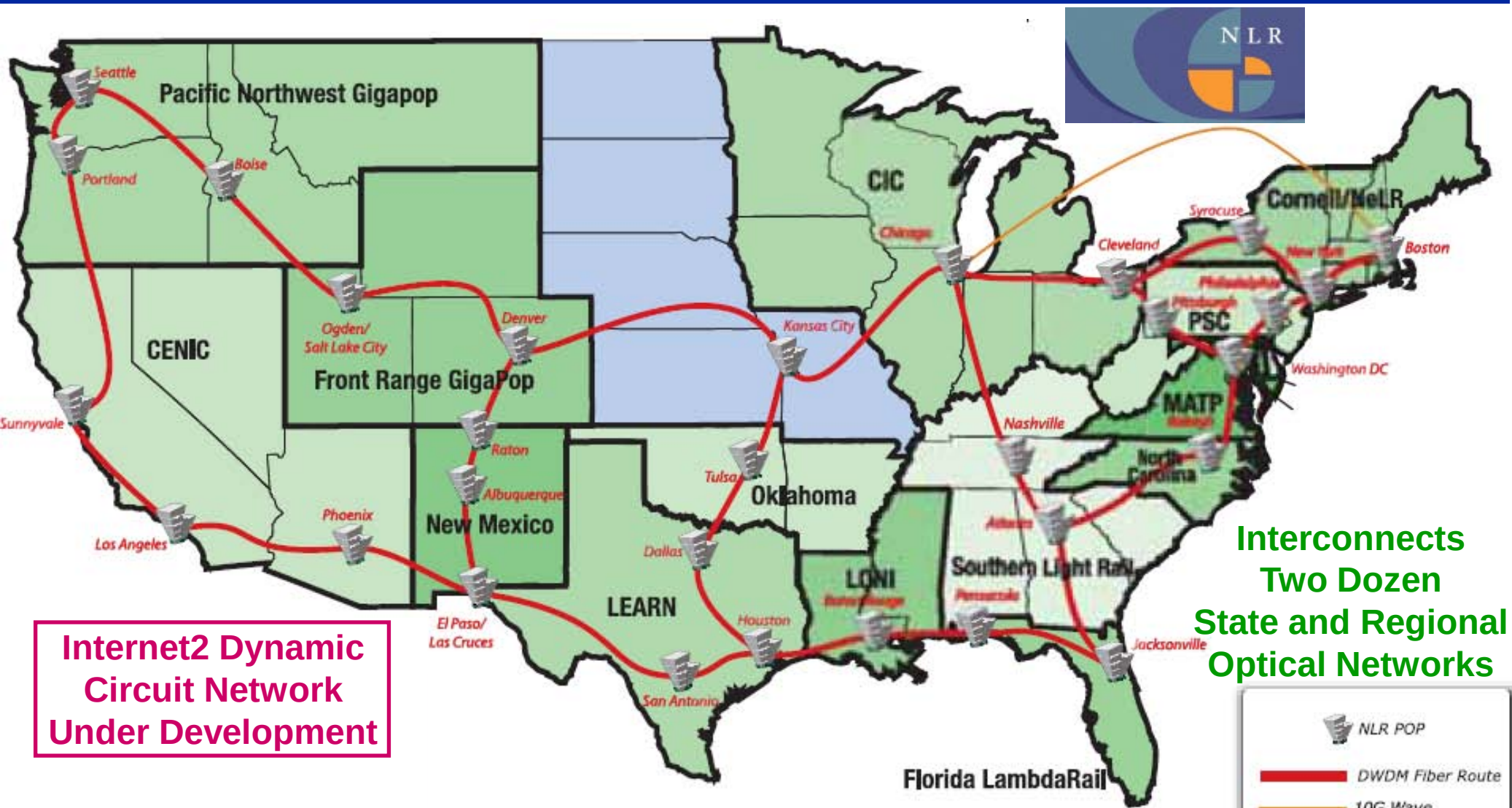
Today Two Dozen State and Regional Optical Networks

Source: Larry Smarr, Rick Stevens, Tom DeFanti, Charlie Catlett

Interconnecting Regional Optical Networks Is Driving Campus Optical Infrastructure Deployment



National Lambda Rail (NLR) Provides Cyberinfrastructure Backbone for U.S. Researchers

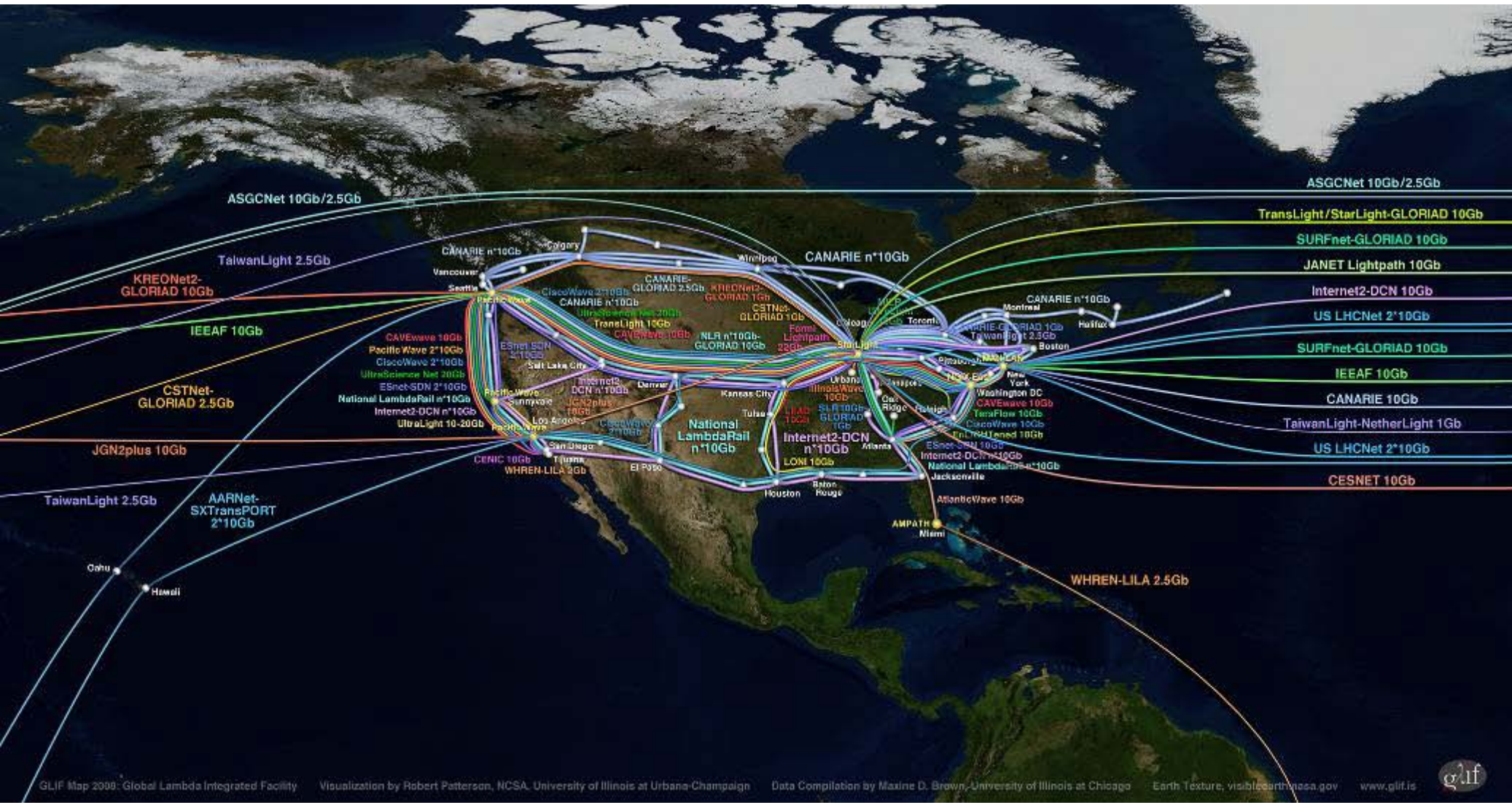


OptIPuter

NLR 40 x 10Gb Wavelengths
Expanding with Darkstrand to 80

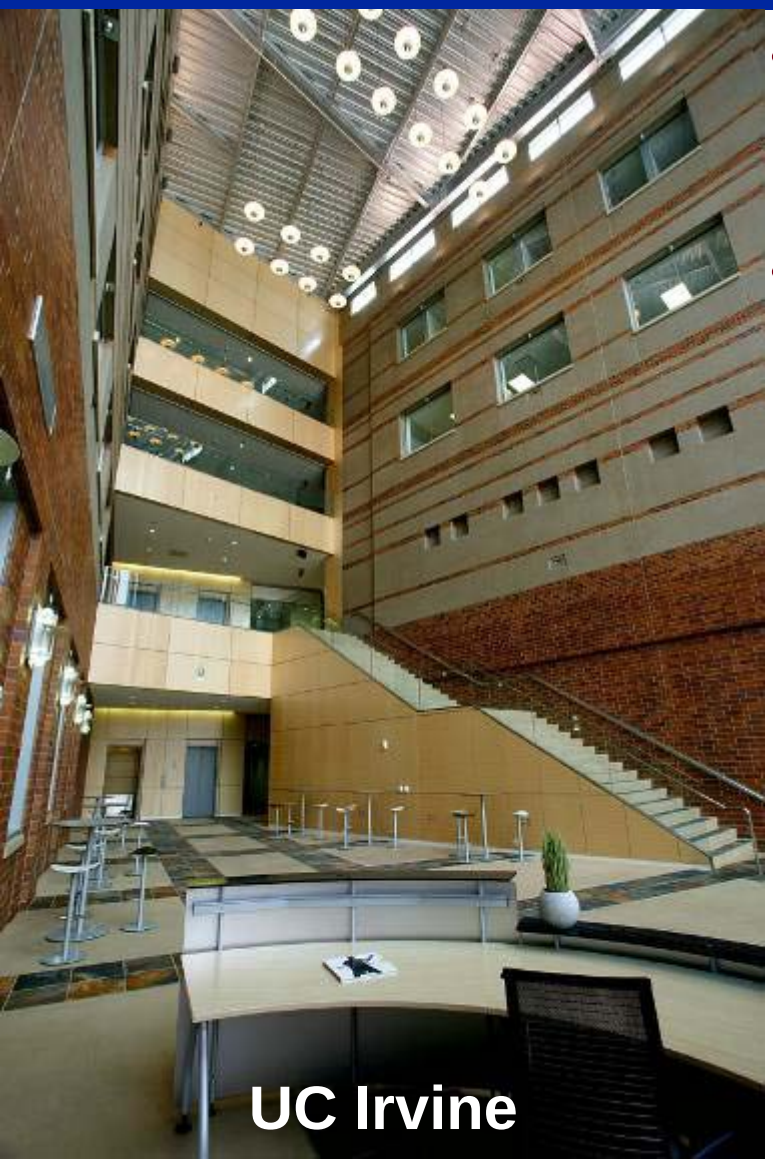
it2

NLR/I2 is Connected Internationally via Global Lambda Integrated Facility



Source: Maxine Brown, UIC and Robert Patterson, NCSA

Two New Calit2 Buildings Provide New Laboratories for "Living in the Future"



- "Convergence" Laboratory Facilities
  - Nanotech, BioMEMS, Chips, Radio, Photonics
  - Virtual Reality, Digital Cinema, HDTV, Gaming
- Over 1000 Researchers in Two Buildings
  - Linked via Dedicated Optical Networks



www.calit2.net

Preparing for a World in Which Distance is Eliminated…

Calit2 Has Become a Global Hub for Optical Connections Between University Research Centers at 10Gbps

Maxine Brown, Tom DeFanti, Co-Chairs

iGrid 2005

THE GLOBAL LAMBDA INTEGRATED FACILITY

www.igrid2005.org

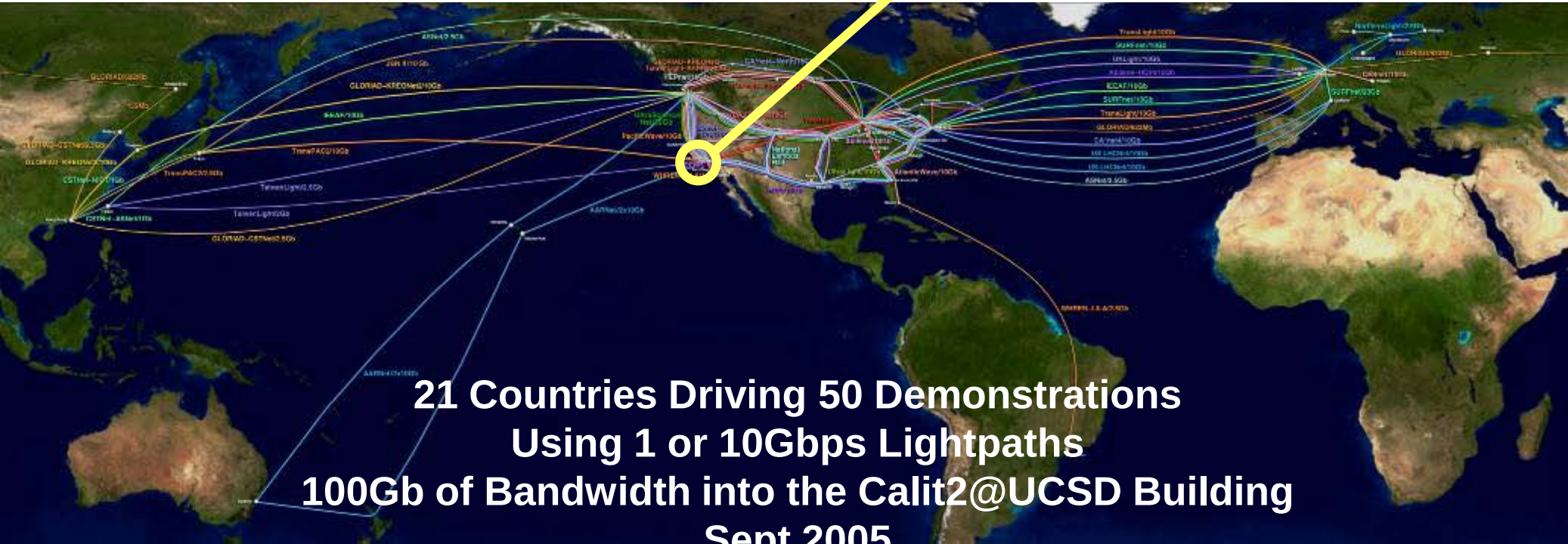


the UC Discovery Grant

September 26-30, 2005

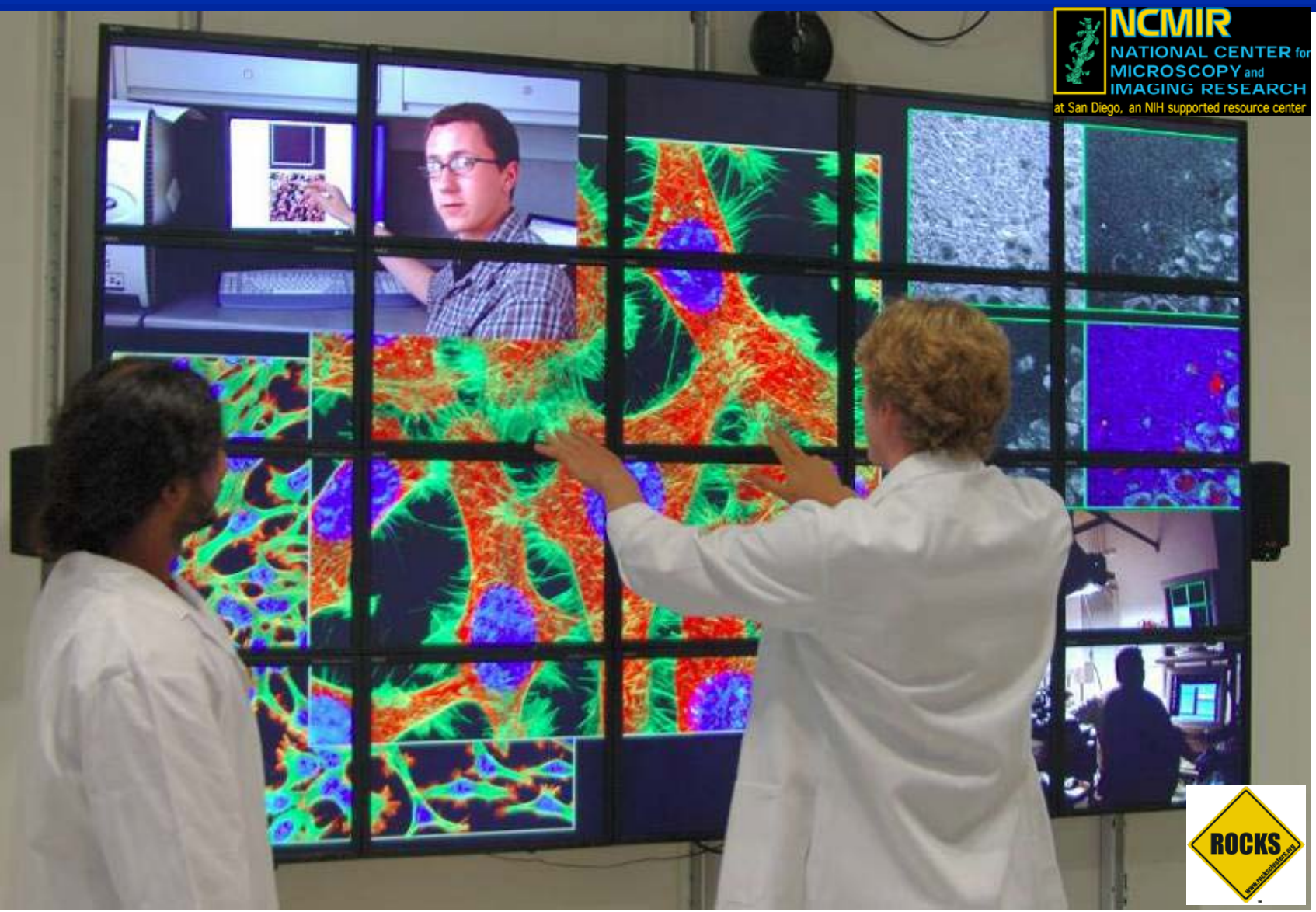
Calit2 @ University of California, San Diego

California Institute for Telecommunications and Information Technology



21 Countries Driving 50 Demonstrations
Using 1 or 10Gbps Lightpaths
100Gb of Bandwidth into the Calit2@UCSD Building
Sept 2005

The OptIPuter Project: Creating High Resolution Portals Over Dedicated Optical Channels to Global Science Data



Scalable Adaptive Graphics Environment (SAGE)

evl electronic visualization laboratory

NSF

Now in Sixth and Final Year

Picture Source: Mark Ellisman, David Lee, Jason Leigh

OptIPuter

Calit2 (UCSD, UCI), SDSC, and UIC Leads—Larry Smarr PI
Univ. Partners: NCSA, USC, SDSU, NW, TA&M, UvA, SARA, KISTI, AIST
Industry: IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent

it2

OptIPuter / OptIPortal

Scalable Adaptive Graphics Environment (SAGE) Applications

evl electronic visualization laboratory

Bitplayer
Streaming animation
of tornado simulation
using UDP.
516 Mbps

MagicCarpet
Streaming Blue Marble
dataset from San Diego
to EVL using UDP.
6.7Gbps

~ 9 Gbps in Total.
SAGE Can Simultaneously Support These
Applications Without Decreasing Their Performance

SVC
Locally streaming
HD camera live
video using UDP.
538Mbps

JuxtaView
Locally streaming the aerial
photography of downtown
Chicago using TCP.
850 Mbps

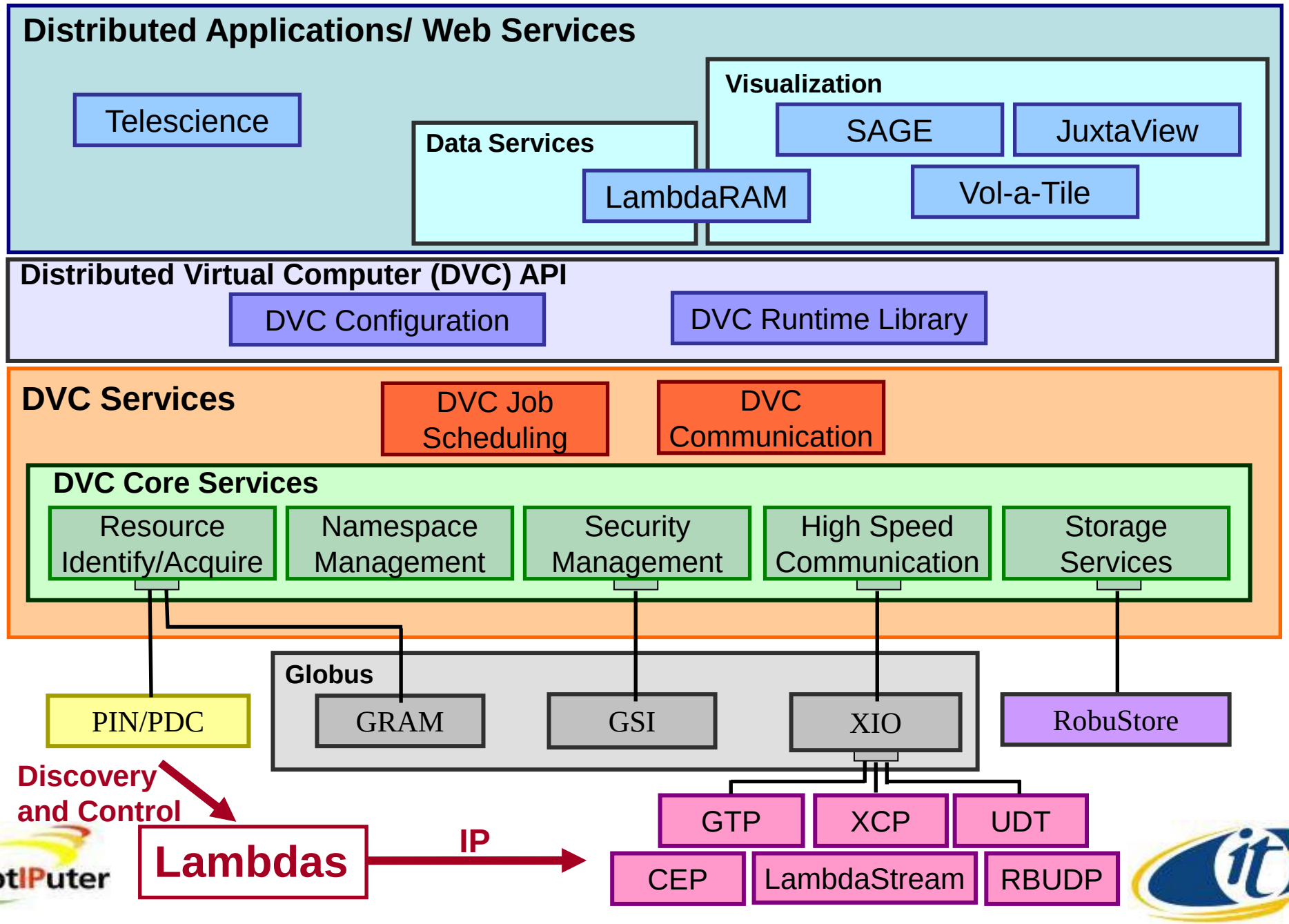
Source: Ji Wang, UIC/EVL

SAGE OptIPortals
Have Been Adopted Worldwide



www.evl.uic.edu/cavern/optiplanet/OptIPortals_Worldwide.html

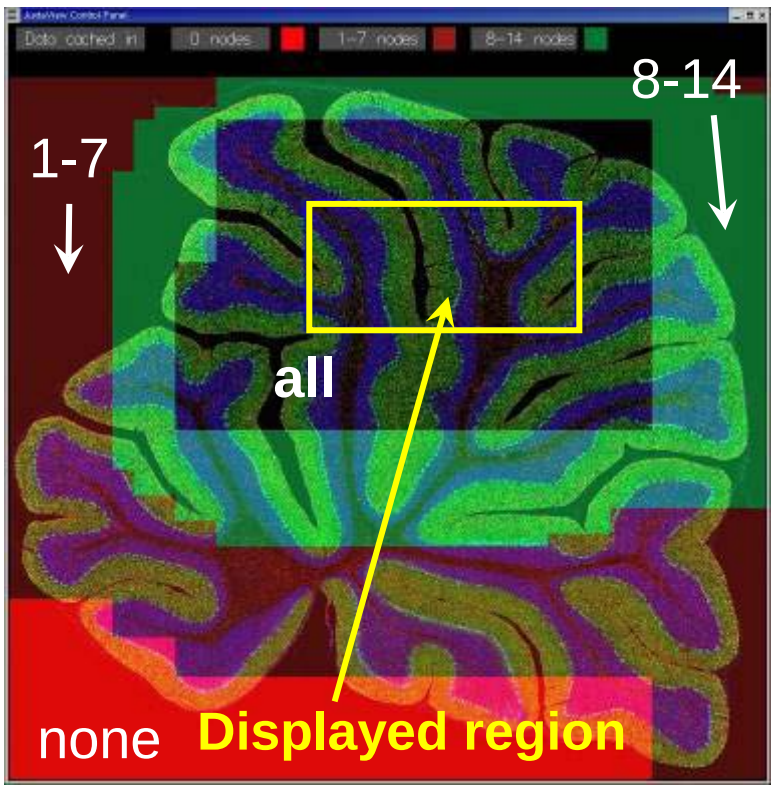
OptIPuter Software Architecture--a Service-Oriented Architecture Integrating Lambdas Into the Grid



LambdaRAM: Clustered Memory To Provide Low Latency Access To Large Remote Data Sets

- **Giant Pool of Cluster Memory Provides Low-Latency Access to Large Remote Data Sets**
  - **Data Is Prefetched Dynamically**
  - **LambdaStream Protocol Integrated into JuxtaView Montage Viewer**
- **3 Gbps Experiments from Chicago to Amsterdam to UIC**
  - **LambdaRAM Accessed Data From Amsterdam Faster Than From Local Disk**

Visualization of the Pre-Fetch Algorithm



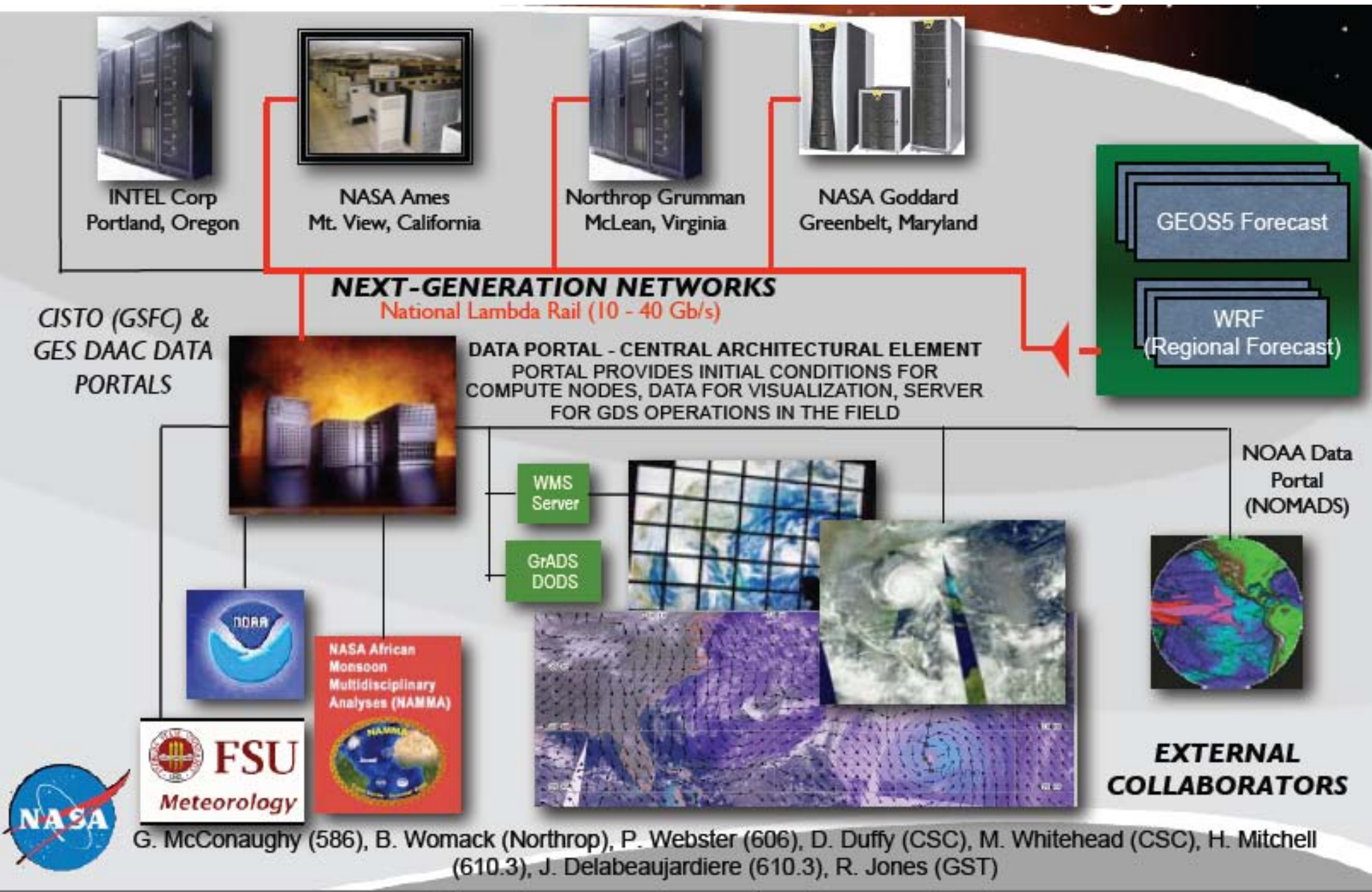
May 2004



Data on Disk in Amsterdam

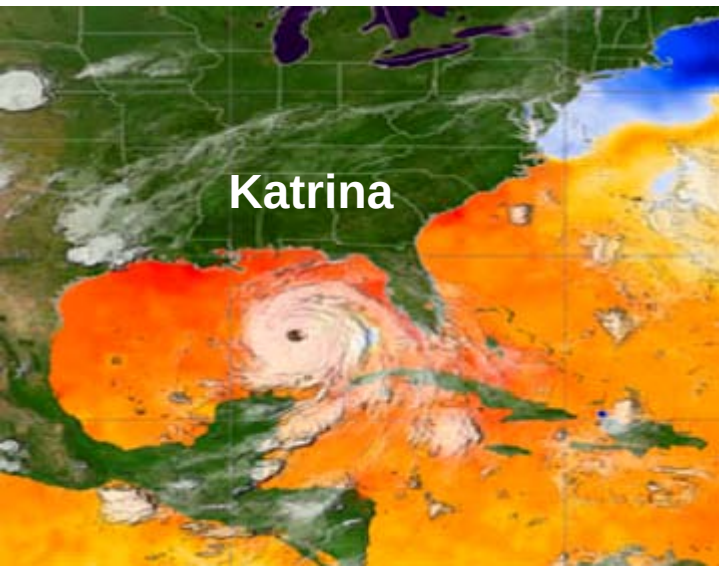
Source: David Lee, Jason Leigh

Distributed Supercomputing: NASA MAP '06 System Configuration Using NLR

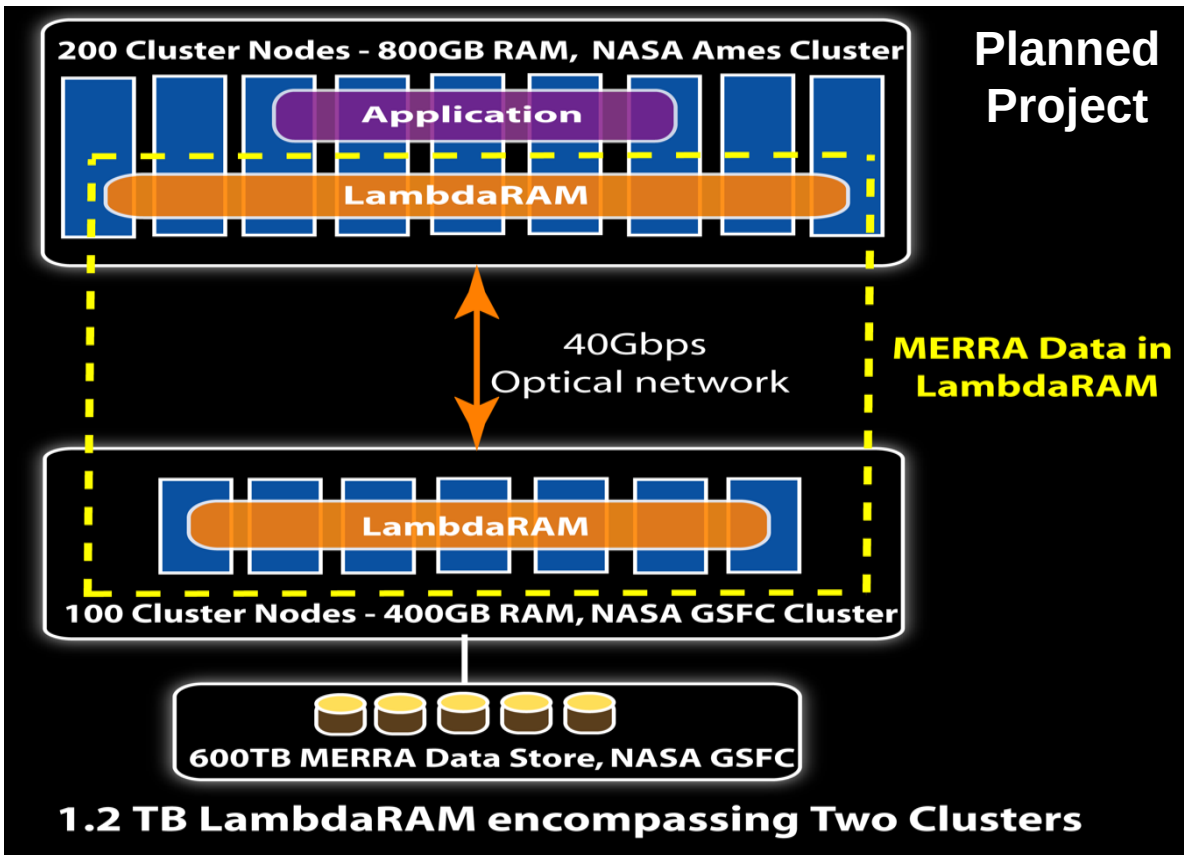


G. McConaughy (586), B. Womack (Northrop), P. Webster (606), D. Duffy (CSC), M. Whitehead (CSC), H. Mitchell (610.3), J. Delabeaujardiere (610.3), R. Jones (GST)

LambdaRAM for Data Pre-Fetching: LambdaGrid Enables Low-Latency Remote Data Access



EVL is working with NASA Goddard and its Modeling, Analysis and Prediction (MAP) Program on Tropical Hurricane Analysis



LambdaRAM uses the entire memory of one or more clusters to mitigate latency. In current trials, LambdaRAM has achieved a 5-fold improvement in accessing remote data. Also, LambdaRAM provides transparent access; i.e., application codes do not need to be modified.

Source:Venkatram Vishwanath, EVL, UIC

CAMERA: Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis



Paul Gilna Ex. Dir.

National LambdaRail Direct Connect Computation and Storage Complex

PI Larry Smarr

Funded by: Gordon and Betty Moore Foundation

Joint Partnership of:

Announced January 17, 2006—
$24.5M Over Seven Years

Building On Centers of Excellence and Achievement









Powered by Rocks









The Scripps Genome Center

Supported by:















CAMERA's Direct Access Core Architecture: An OptIPuter Metagenomics Metacomputer

Gordon and Betty MOORE FOUNDATION

Sargasso Sea Data

Sorcerer II Expedition (GOS)

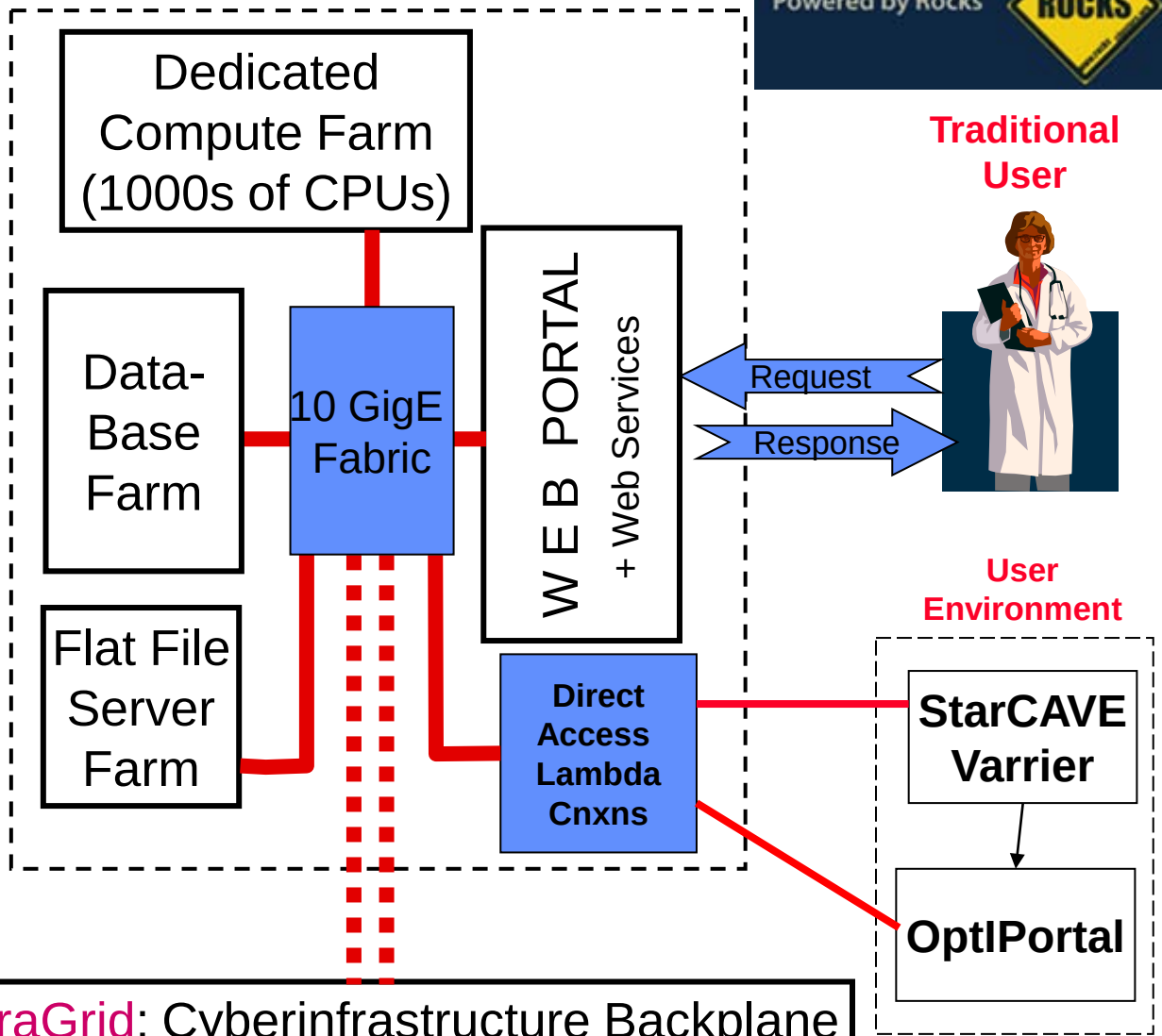
JGI Community Sequencing Project

Moore Marine Microbial Project

NASA and NOAA Satellite Data

Community Microbial Metagenomics Data

Powered by Rocks ROCKS



CENTER FOR EARTH OBSERVATIONS & APPLICATIONS

SCRIPPS INSTITUTION OF OCEANOGRAPHY

SDSC SAN DIEGO SUPERCOMPUTER CENTER

OptIPuter

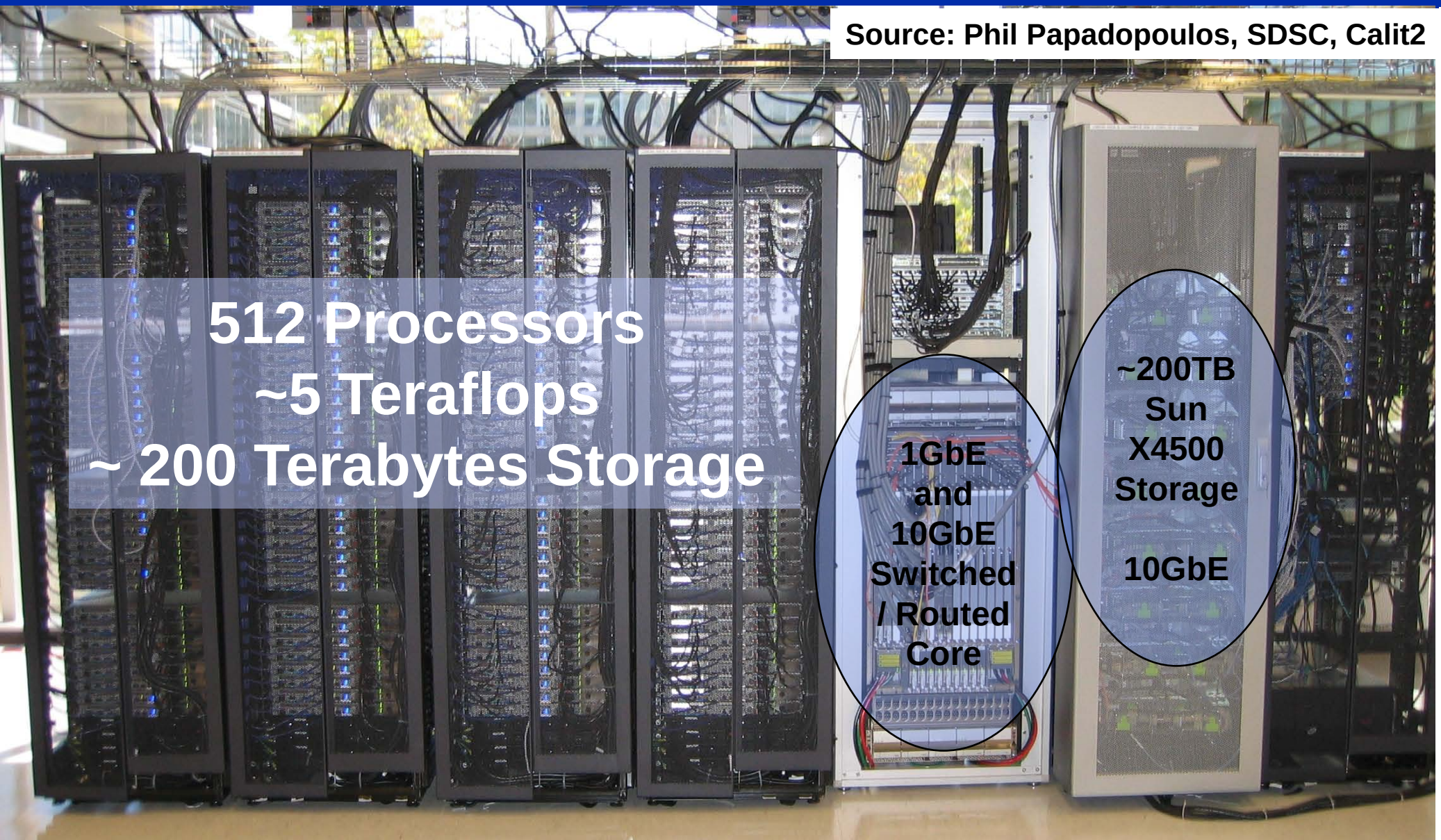
J. Craig Venter INSTITUTE

it2

Source: Phil Papadopoulos, SDSC, Calit2

Calit2 Microbial Metagenomics Cluster- Next Generation Optically Linked Science Data Server

Source: Phil Papadopoulos, SDSC, Calit2



OptIPuter SDSC ROCKS camera Marine Microbial Ecology J. Craig Venter INSTITUTE it²

CAMERA's
Global Microbial Metagenomics CyberCommunity

Over 2010 Registered Users From Over 50 Countries



camera
Marine Microbial Ecology

J. Craig Venter
INSTITUTE

it²

e-Science Collaboratory Without Walls Enabled by Uncompressed HD Telepresence Over 10Gbps

iHDTV: 1500 Mbits/sec Calit2 to UW Research Channel Over NLR



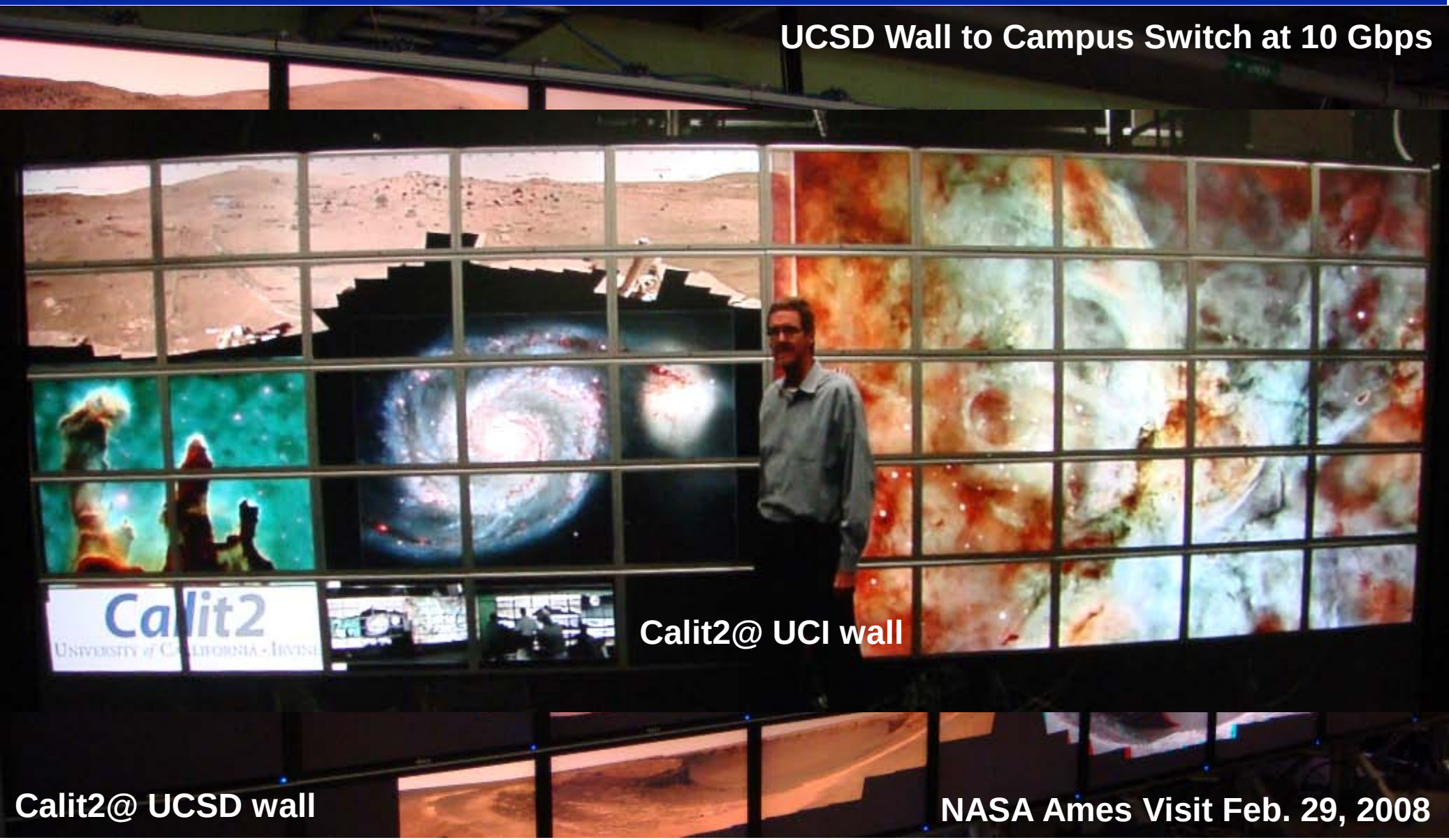
John Delaney, PI LOOKING, Neptune

Photo: Harry Ammons, SDSC

OptIPuter

it2

The Calit2 1/4 Gigapixel OptIPortals at UCSD and UCI Are Joined to Form a Gbit/s HD Collaboratory



OptIPuter

UCSD cluster: 15 x Quad core Dell XPS with Dual nVIDIA 5600s
UCI cluster: 25 x Dual Core Apple G5

it²

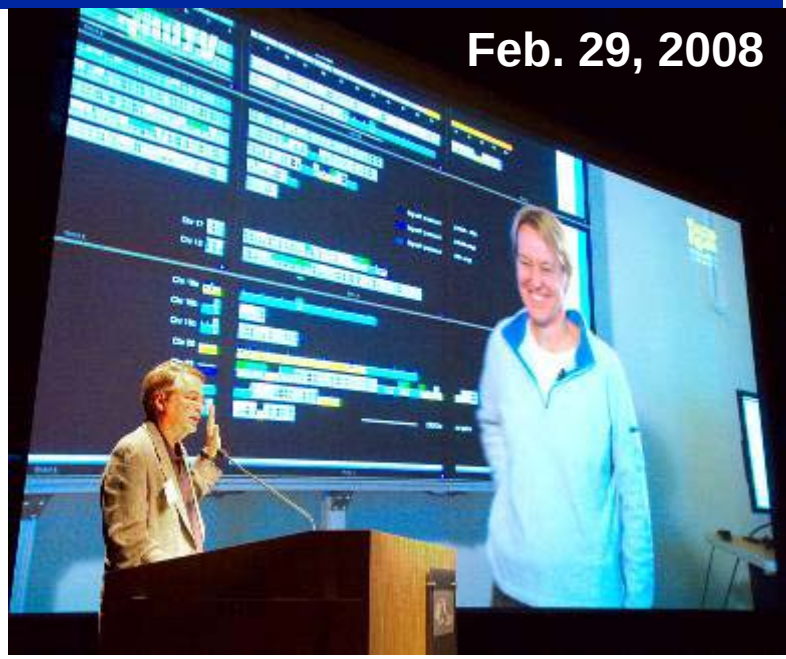
OptIPlanet Collaboratory Persistent Infrastructure Supporting Microbial Research



OptIPuter

Ginger Armbrust's Diatoms: Micrographs, Chromosomes, Genetic Assembly

ROCKS







UW's Research Channel
Michael Wellings

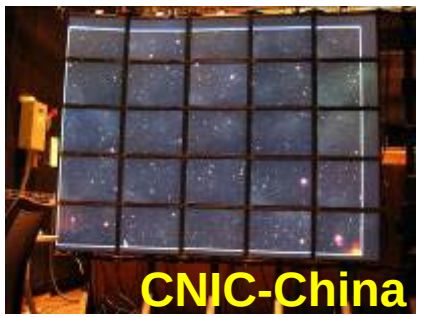
it2

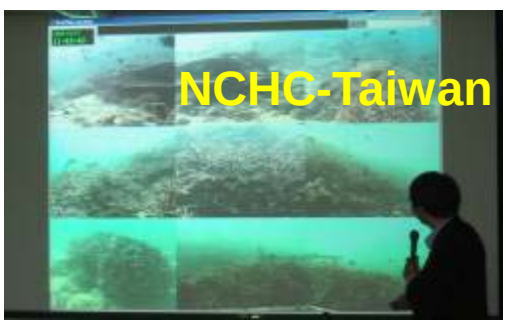
OptIPortals
Are Being Adopted Globally























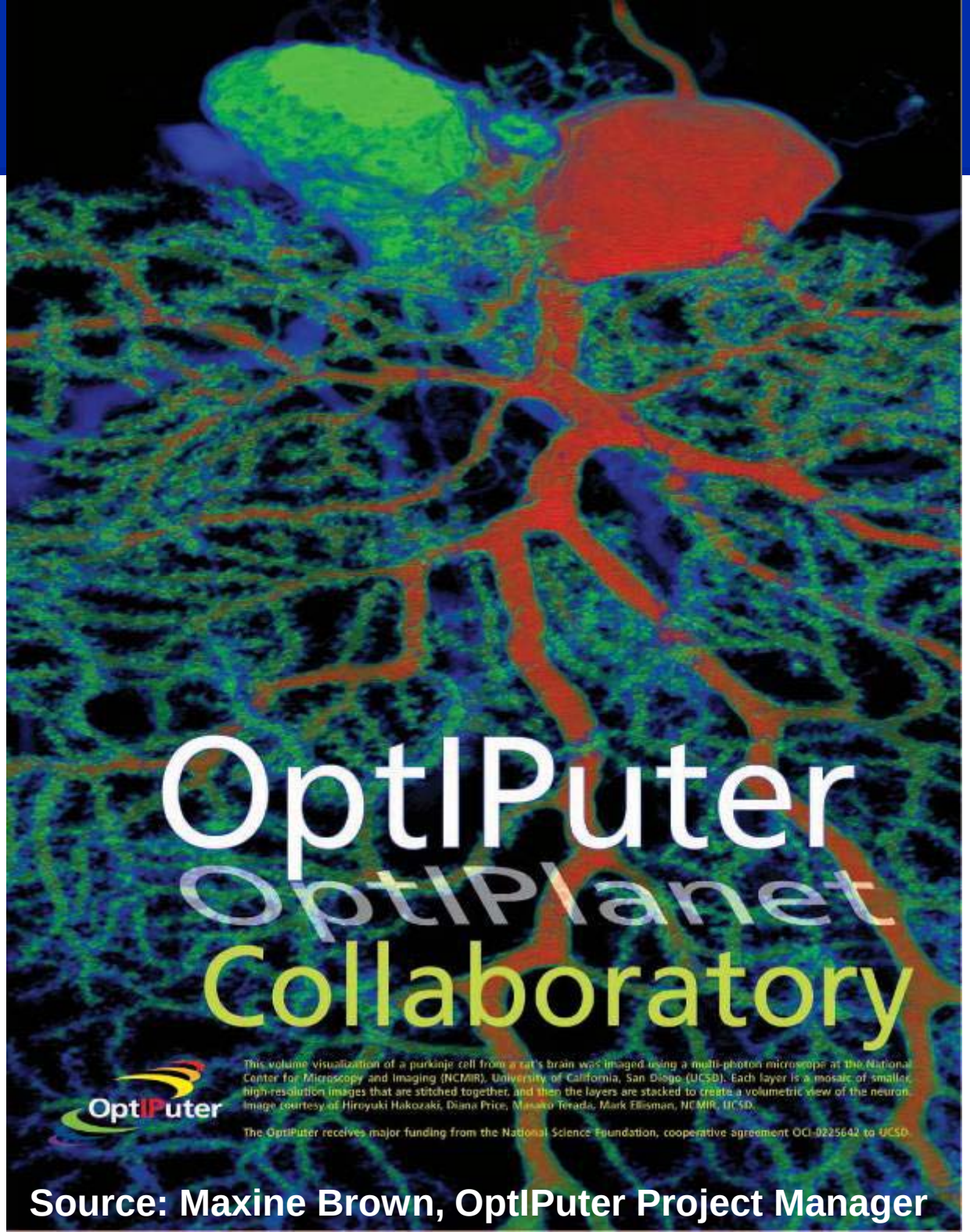


OptIPuter

Powered by Rocks ROCKS

evl electronic visualization laboratory

it2



Source: Maxine Brown, OptIPuter Project Manager

Green Initiative:

Can Optical Fiber Replace Airline Travel for Continuing Collaborations ?

Launch of the 100 Megapixel OzIPortal Over Qvidium Compressed HD on 1 Gbps CENIC/PW/AARNet Fiber



January 15, 2008

OptIPuter

No Calit2 Person Physically Flew to Australia to Bring This Up!

www.calit2.net/newsroom/release.php?id=1219

it2

OptIPuterizing Australian Universities in 2008: CENIC Coupling to AARNet



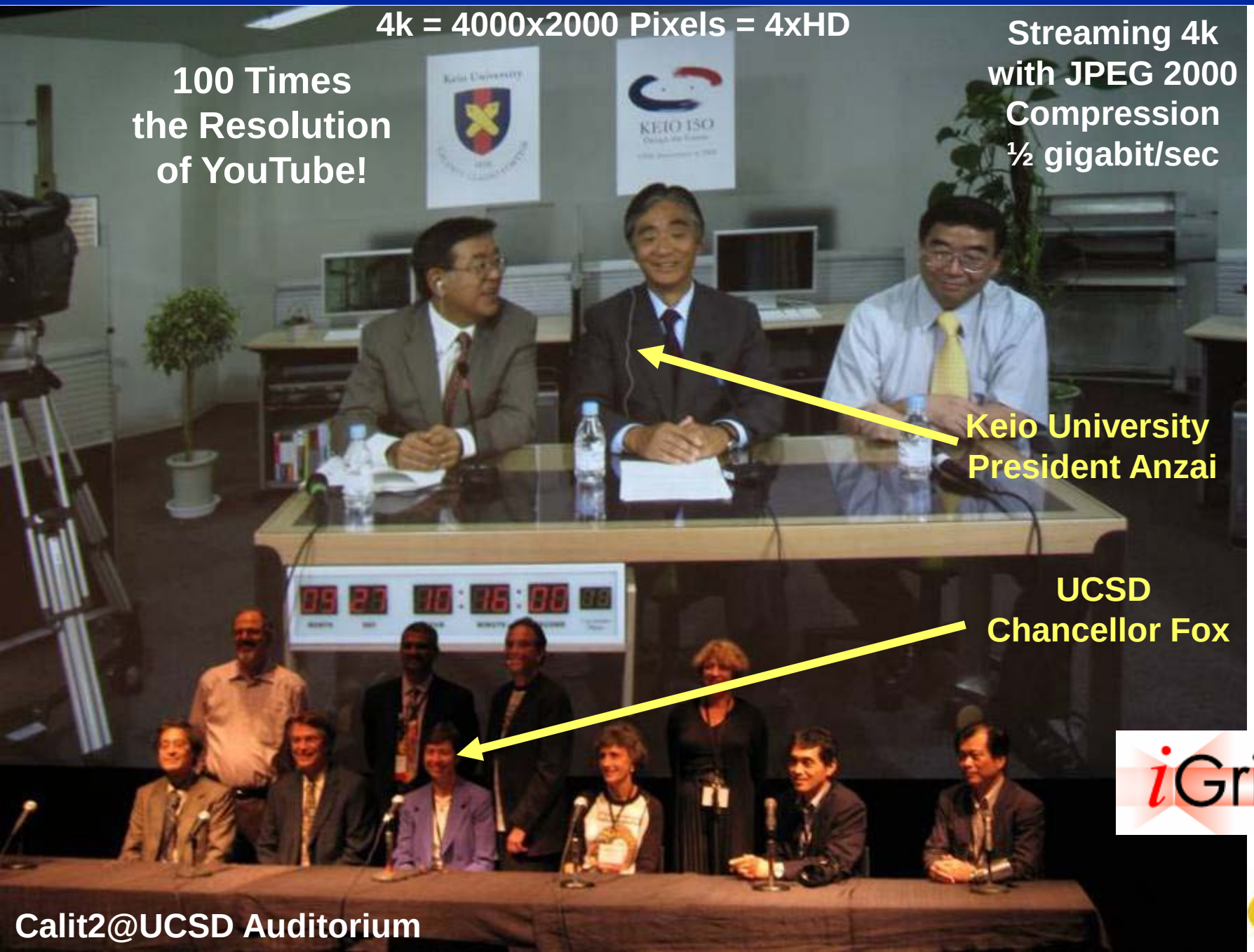
UMelbourne/Calit2 Telepresence Session May 21, 2008

Augmented by Many Physical Visits This Year
Culminating in Two Week Lecture Tour
of Australian Research Universities
by Larry Smarr October 2008

Phil Scanlan
Founder-
Australian American
Leadership Dialogue
www.aald.org



First Trans-Pacific Super High Definition Telepresence Meeting Using Digital Cinema 4k Streams



CineGrid

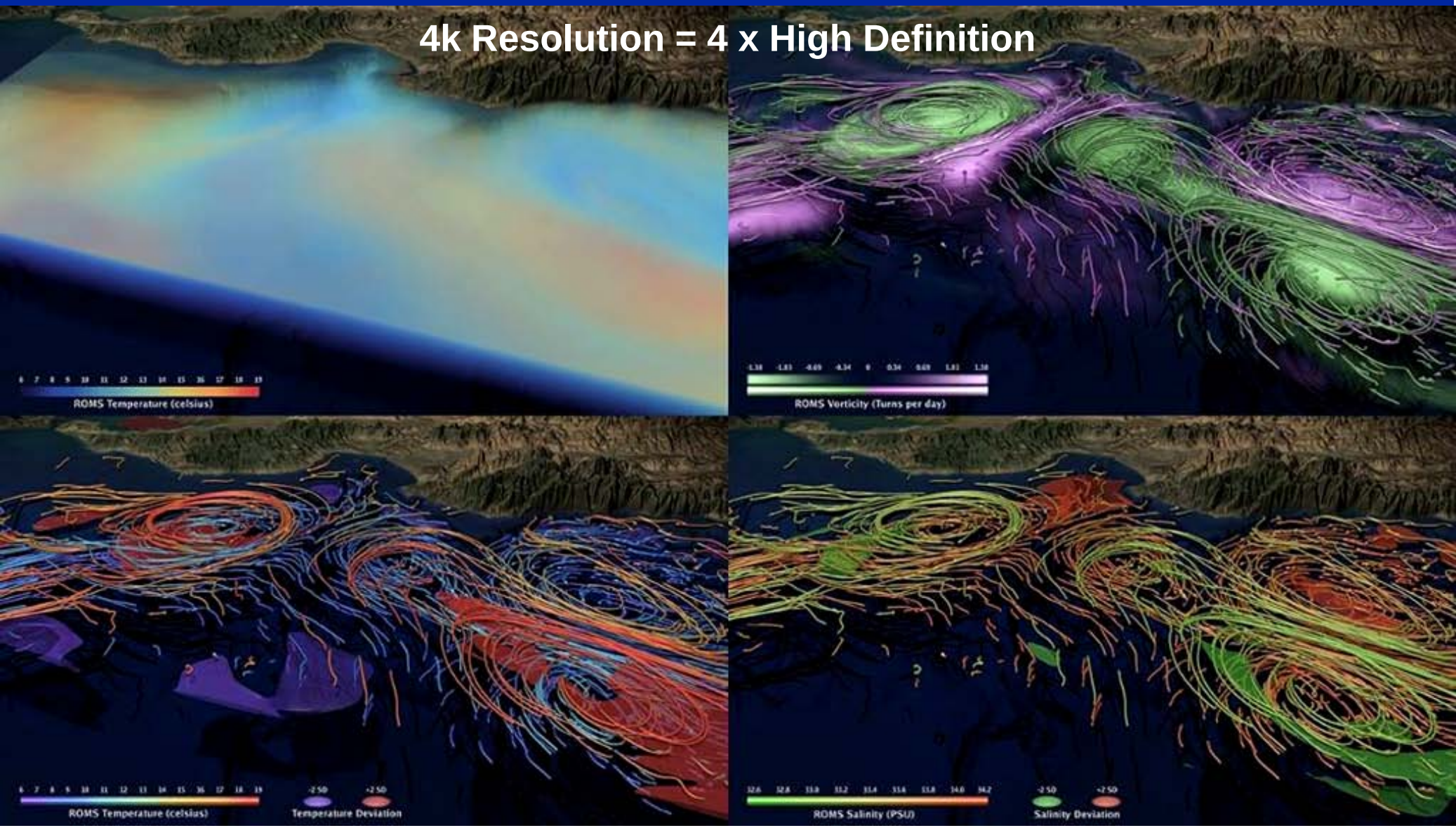
Lays Technical Basis for Global Digital Cinema

Sony
NTT
SGI

iGrid 2005

it²

From Digital Cinema to Scientific Visualization: JPL Simulation of Monterey Bay

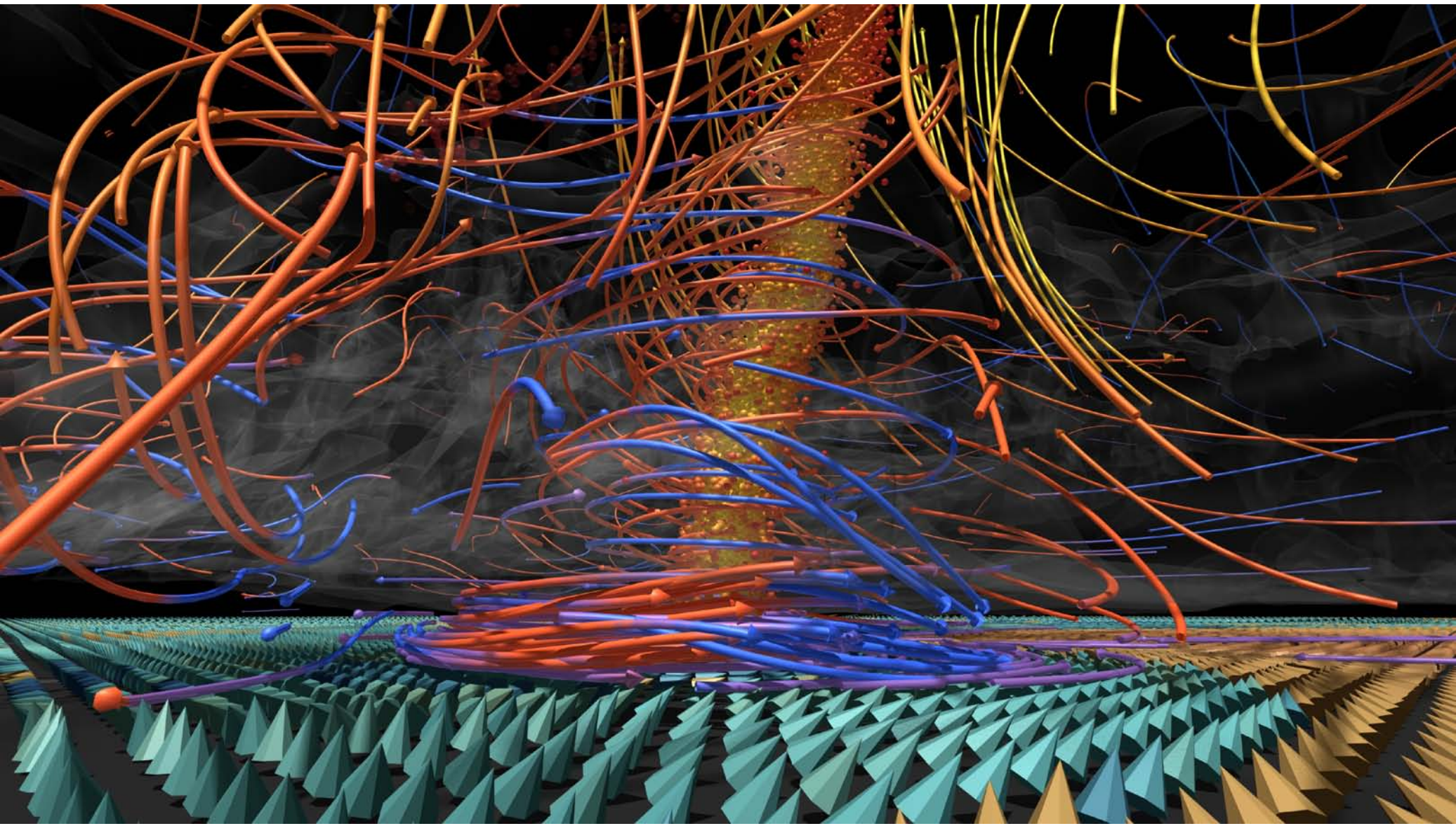


OptIPuter

Source: Donna Cox, Robert Patterson, NCSA
Funded by NSF LOOKING Grant

it2

Rendering Supercomputer Data at Digital Cinema Resolution



OptIPuter

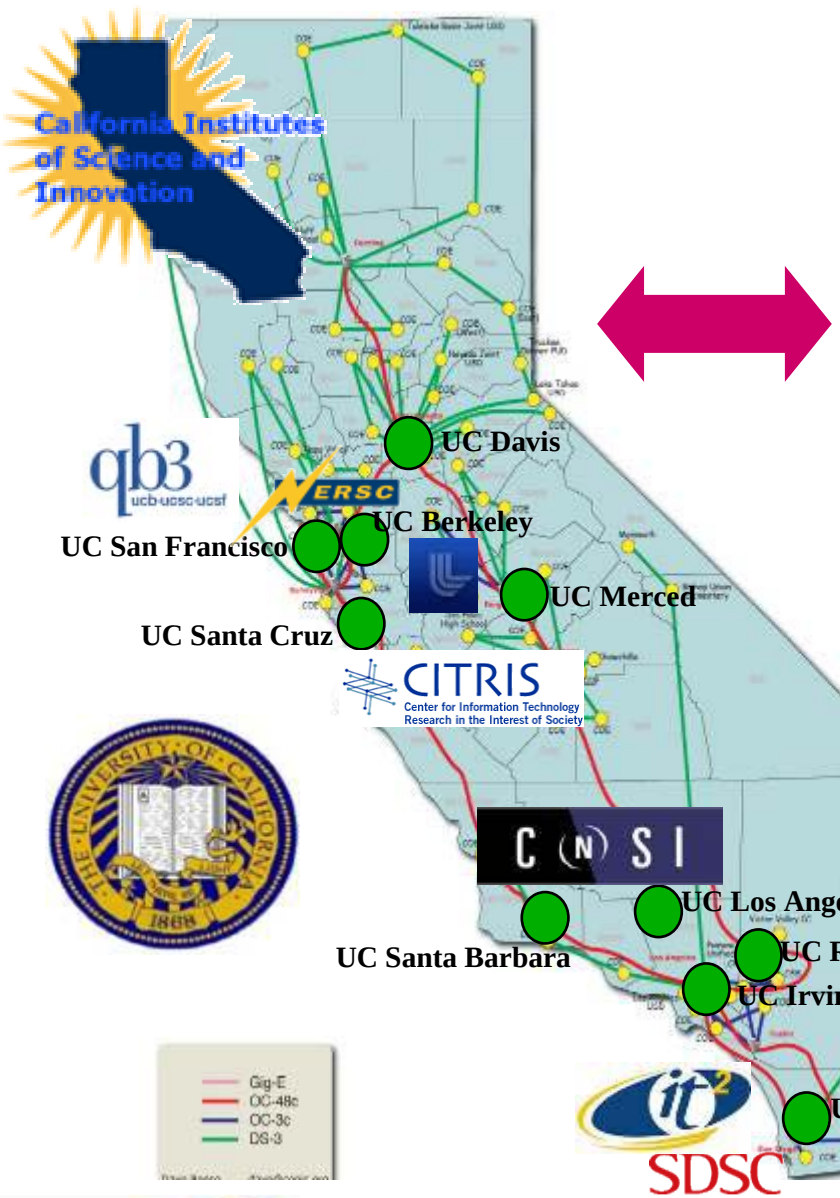
Source: Donna Cox, Robert Patterson, Bob Wilhelmson, NCSA

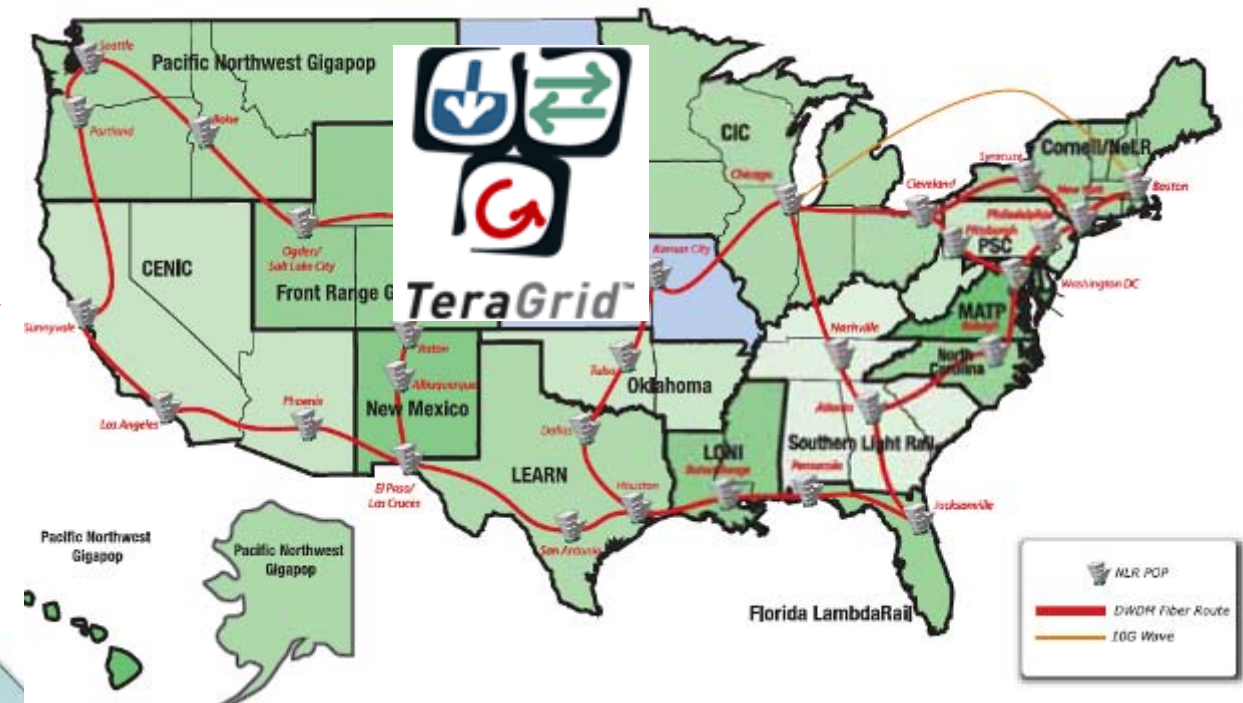
it2

SAGE Global Visualcasting to Europe September 2007



Creating a California Cyberinfrastructure of OptIPuter "On-Ramps" to NLR & TeraGrid Resources



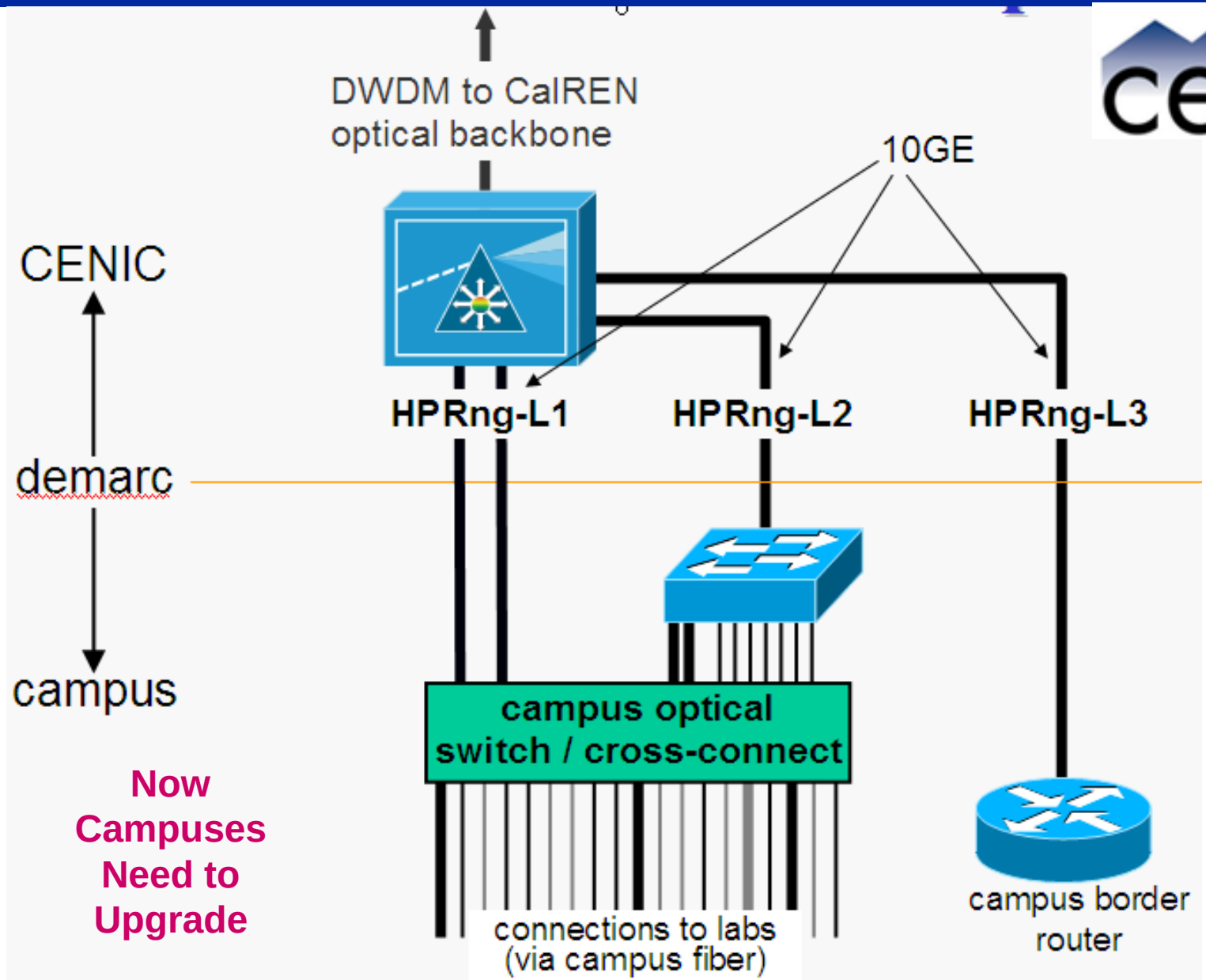


Creating a Critical Mass of OptIPuter End Users on a Secure LambdaGrid

Source: Fran Berman, SDSC , Larry Smarr, Calit2

CENIC's New "Hybrid Network" - Traditional Routed IP and the New Switched Ethernet and Optical Services

~ $14M Invested in Upgrade



Now Campuses Need to Upgrade

Source: Jim Dolgonas, CENIC

The "Golden Spike" UCSD Experimental Optical Core: Ready to Couple Users to CENIC L1, L2, L3 Services



Goals by 2008:

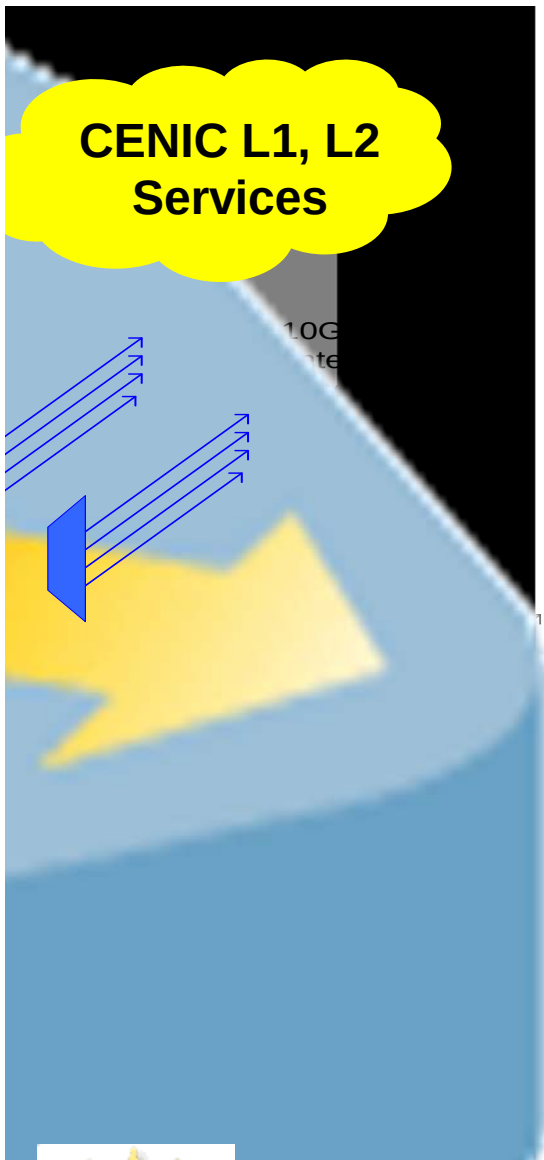
>= 60 endpoints at 10 GigE

>= 30 Packet switched

>= 30 Switched wavelengths

>= 400 Connected endpoints

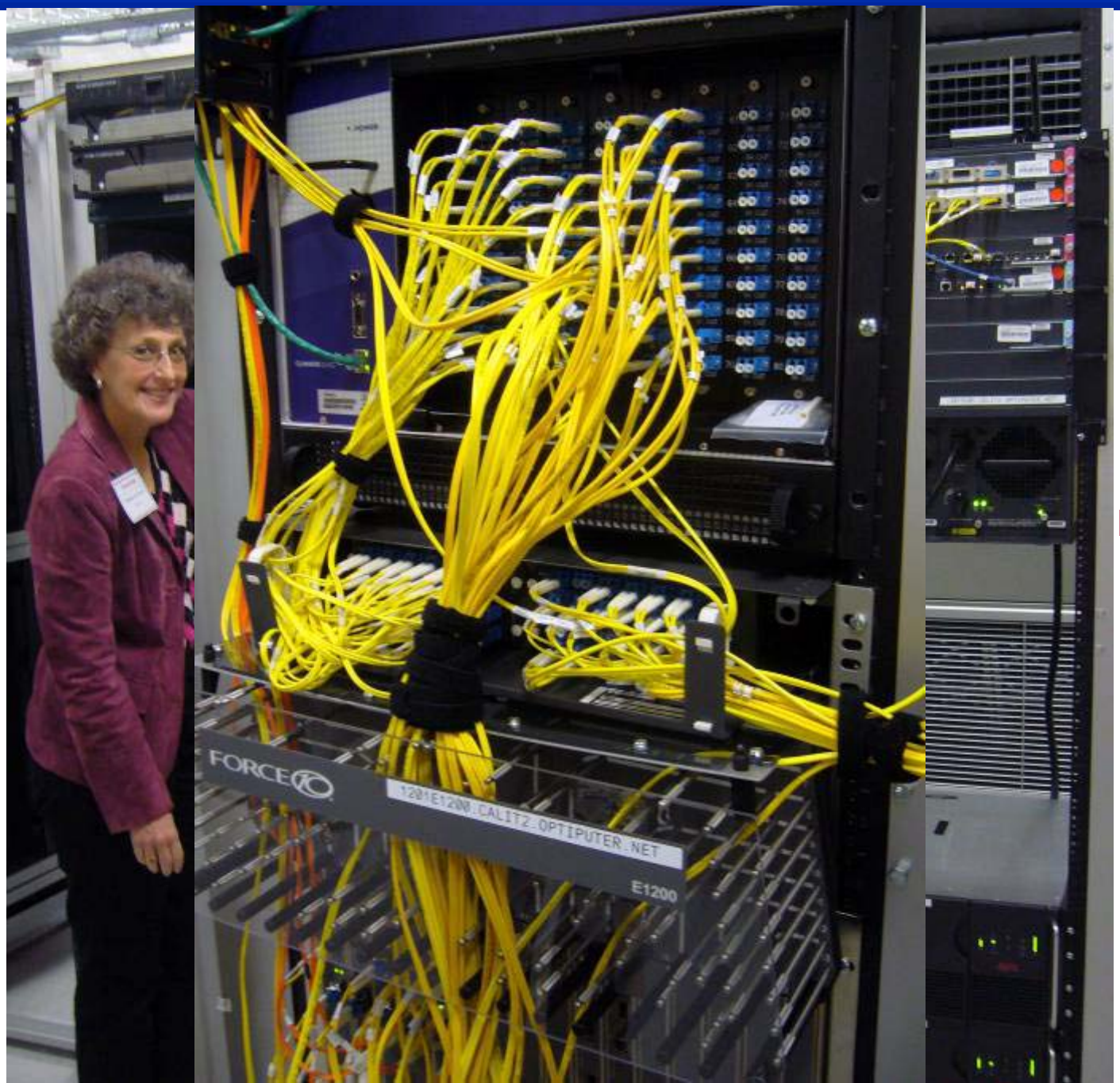
Approximately 0.5 Tbps Arrive at the "Optical" Center of Hybrid Campus Switch



Funded by NSF MRI Grant

Source: Phil Papadopoulos, SDSC/Calit2 (Quartzite PI, OptIPuter co-PI)

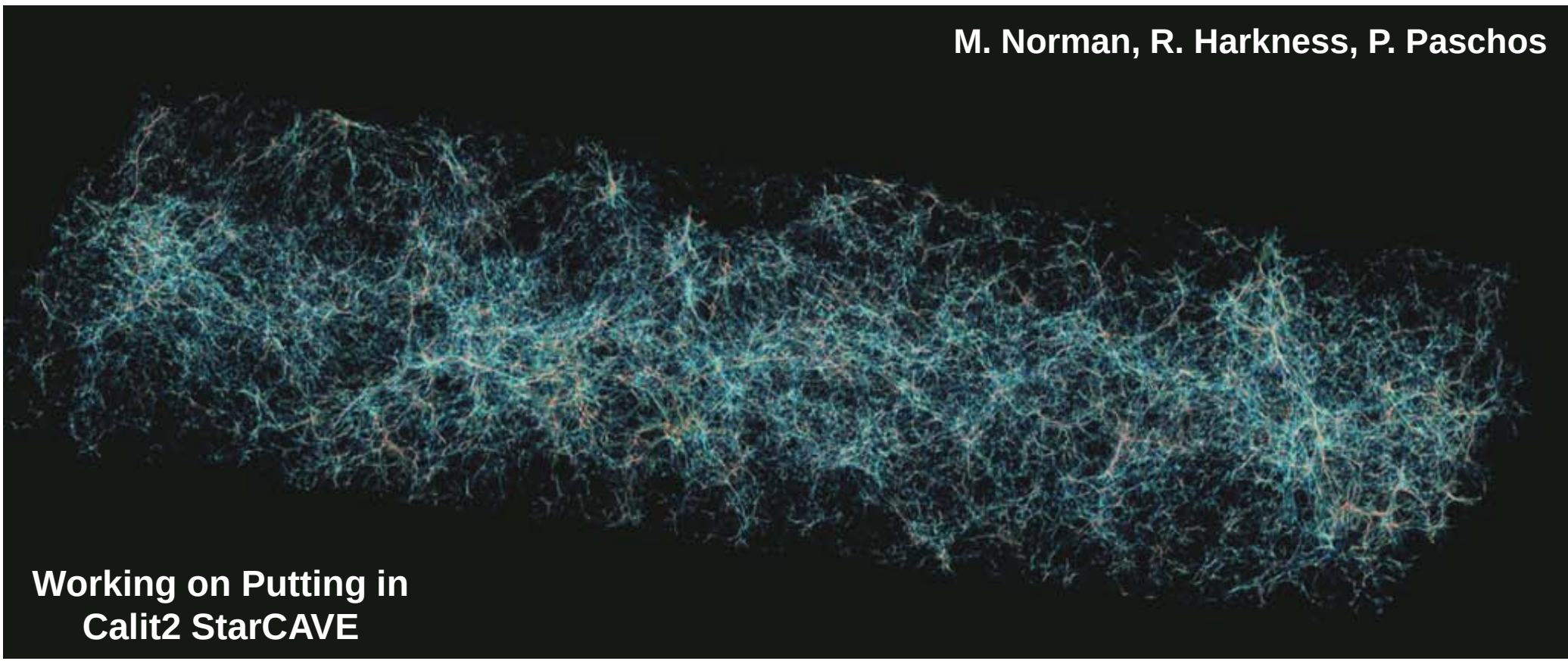
Calit2 Sunlight
Optical Exchange Contains Quartzite



10:45 am
Feb. 21, 2008

Next Step: Experiment on OptIPuter/OptIPortal with Remote Supercomputer Power User

1 Billion Light-year Pencil
From a 20483 Hydro/N-Body Simulation



Structure of the Intergalactic Medium

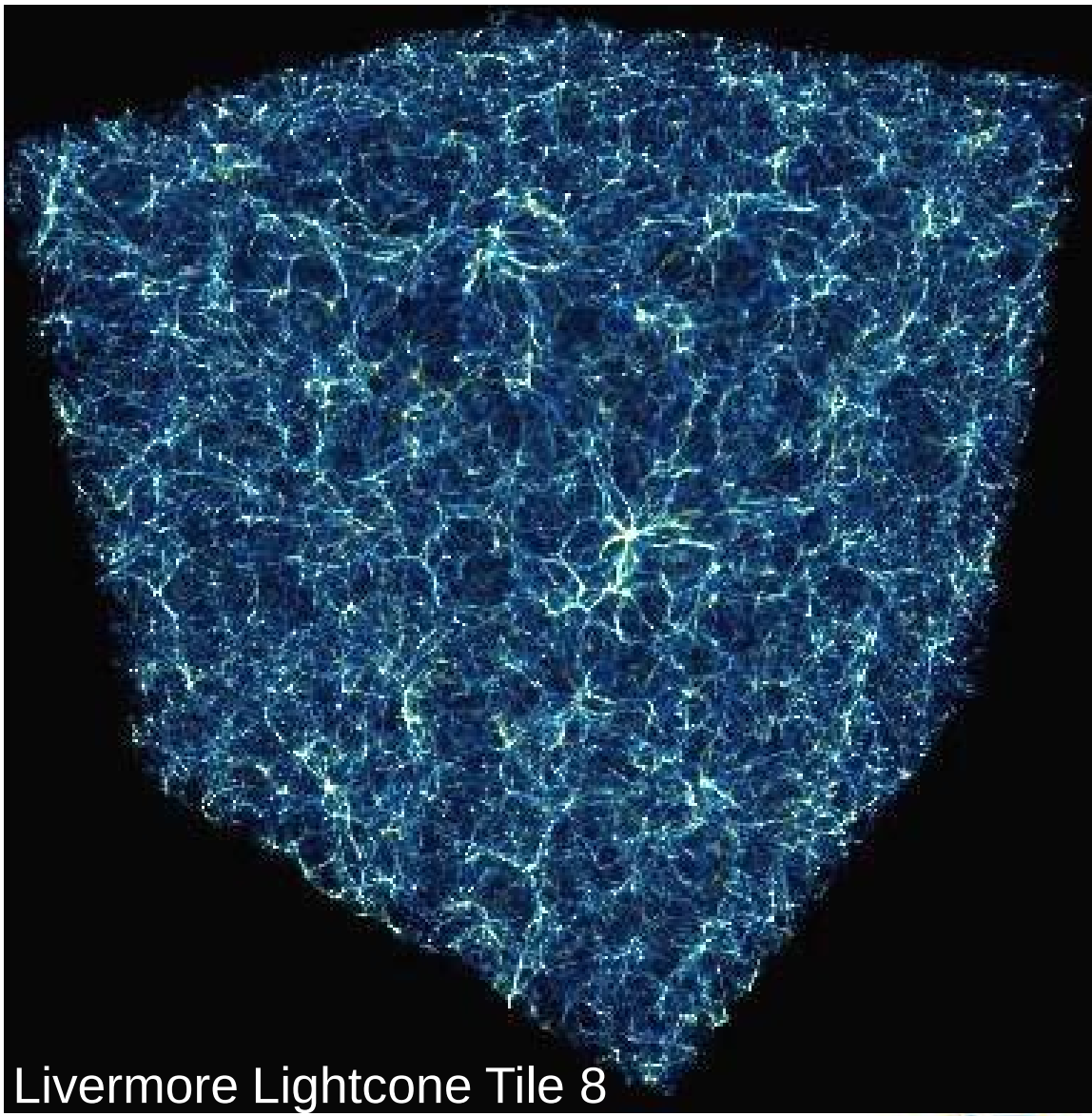
1.3 M SUs, NERSC Seaborg
170 TB output

Source: Michael Norman, SDSC, UCSD

The Livermore Lightcone: 8 Large AMR Simulations Covering 10 Billion Years "Look Back Time"

- **1.5 M SU on LLNL Thunder**
- **Generated 200 TB Data**
- **0.4 M SU Allocated on SDSC DataStar for Data Analysis Alone**

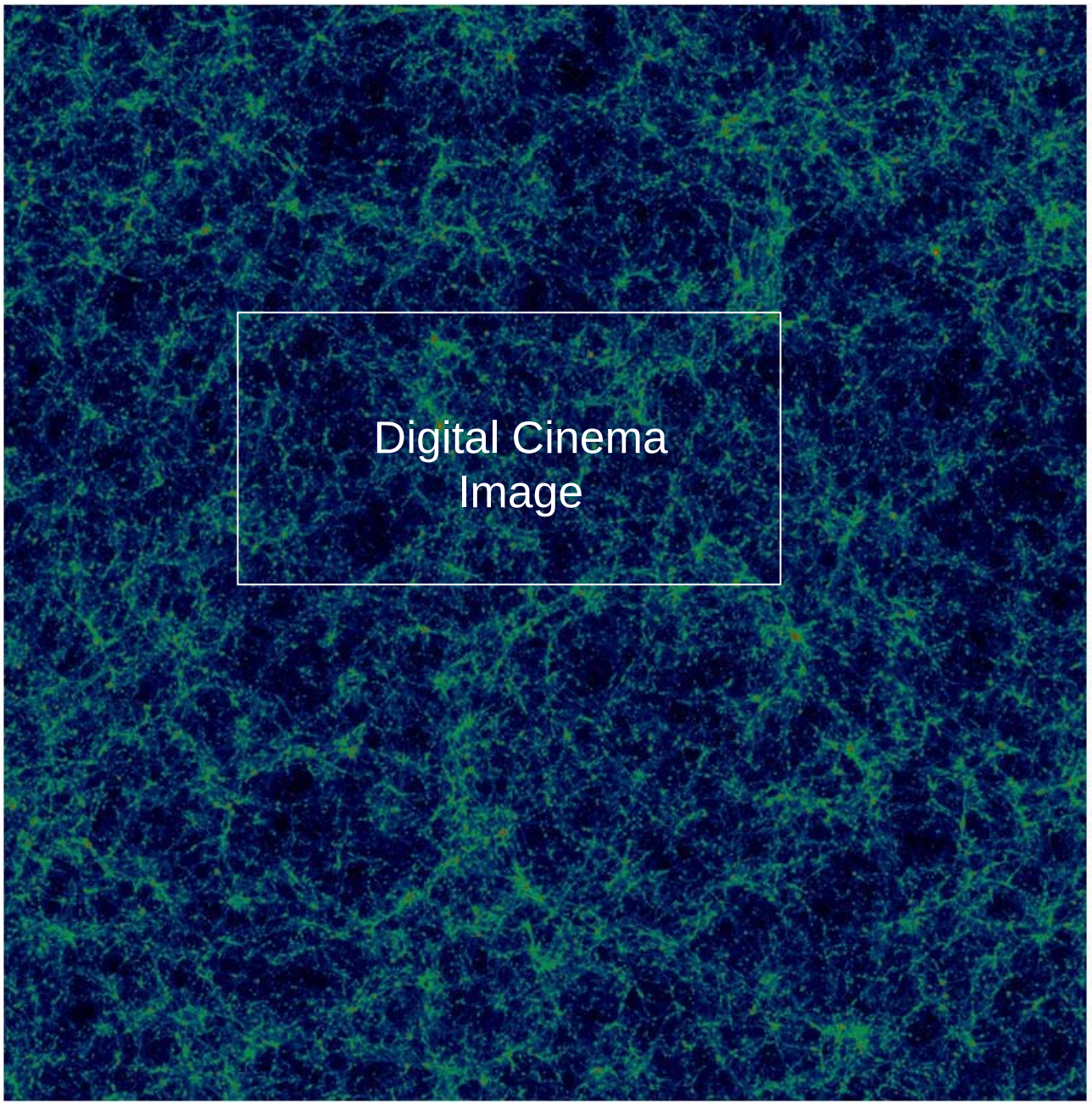
512^3 Base Grid, 7 Levels of Adaptive Refinement➔65,000 Spatial Dynamic Range



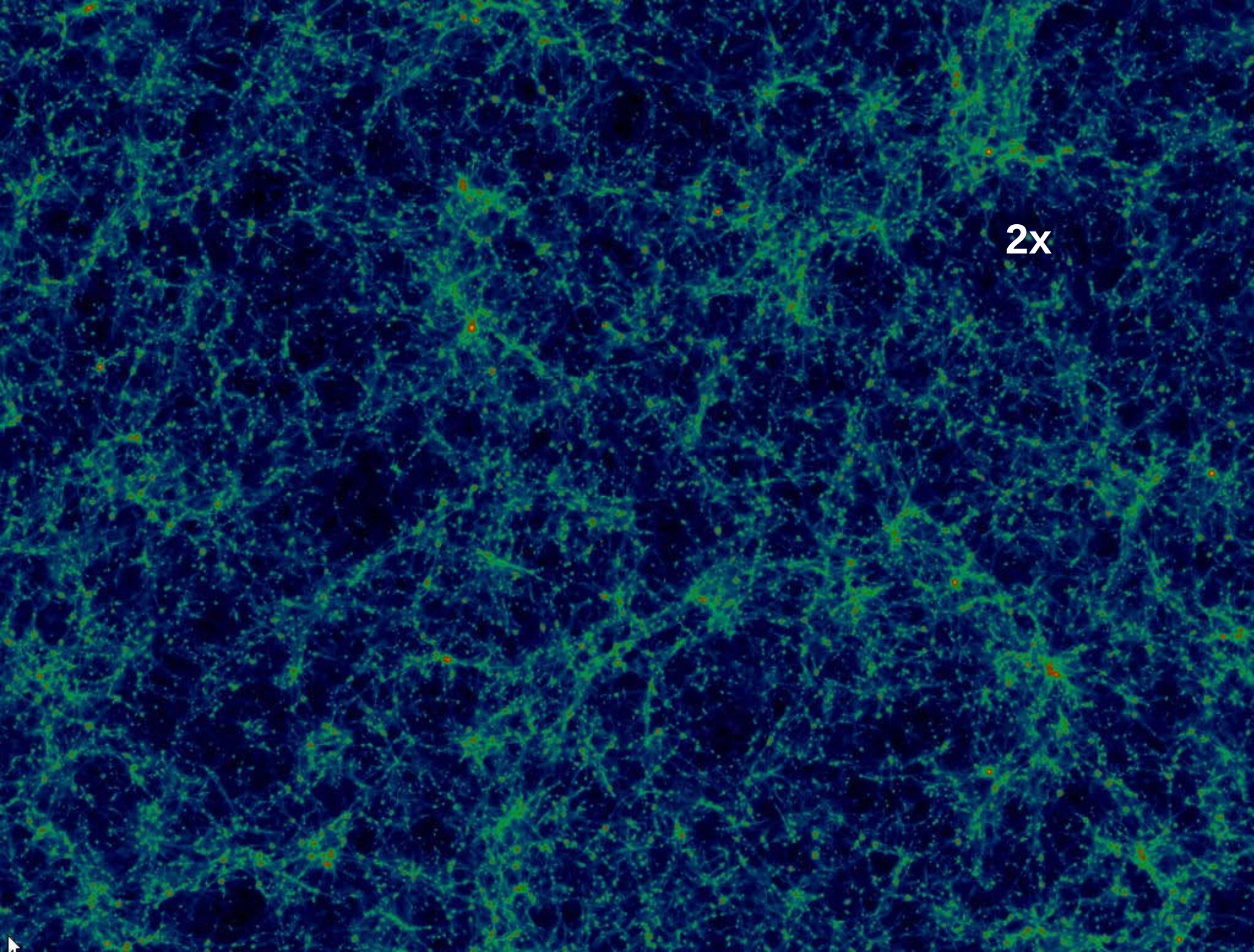
Livermore Lightcone Tile 8

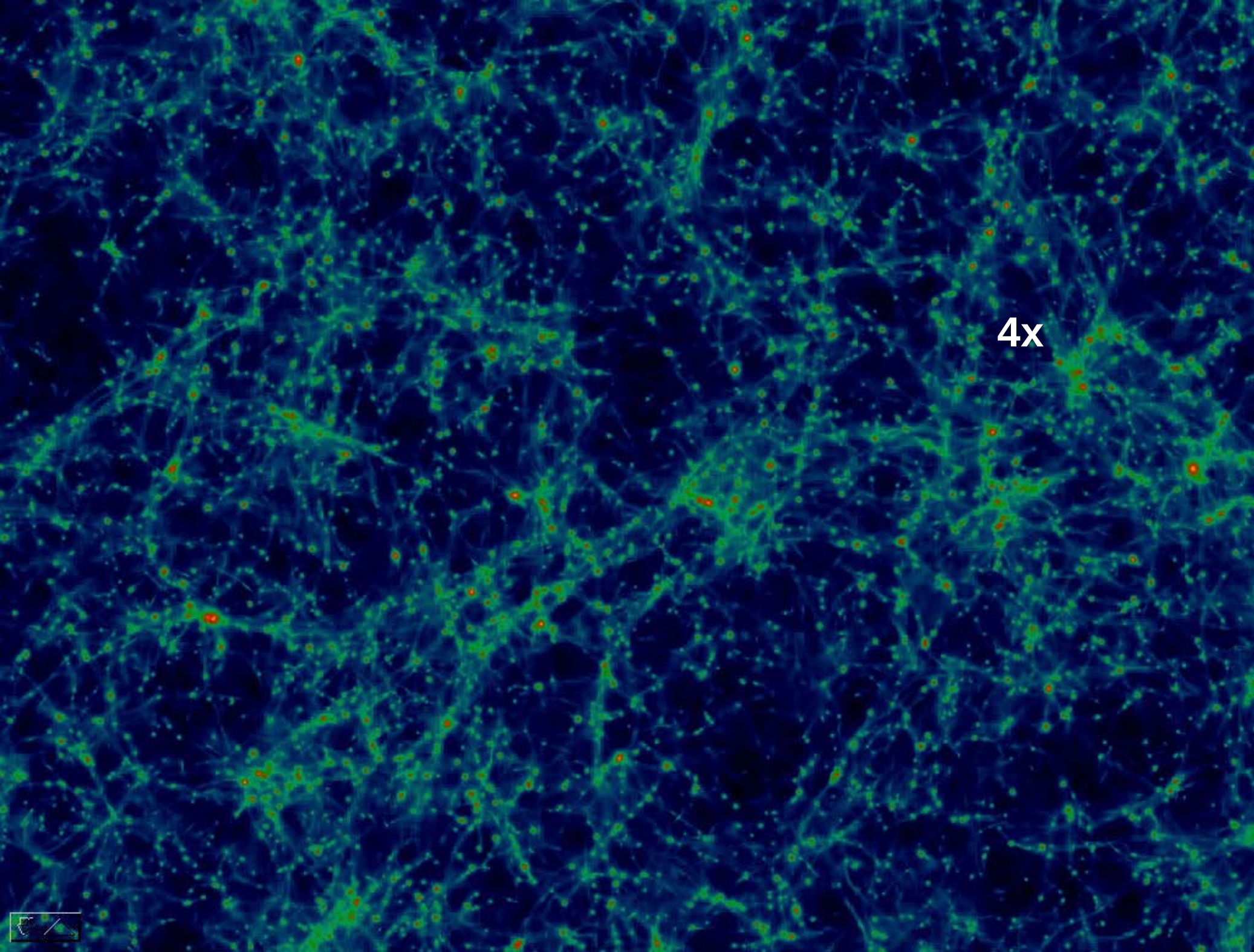
Source: Michael Norman, SDSC, UCSD

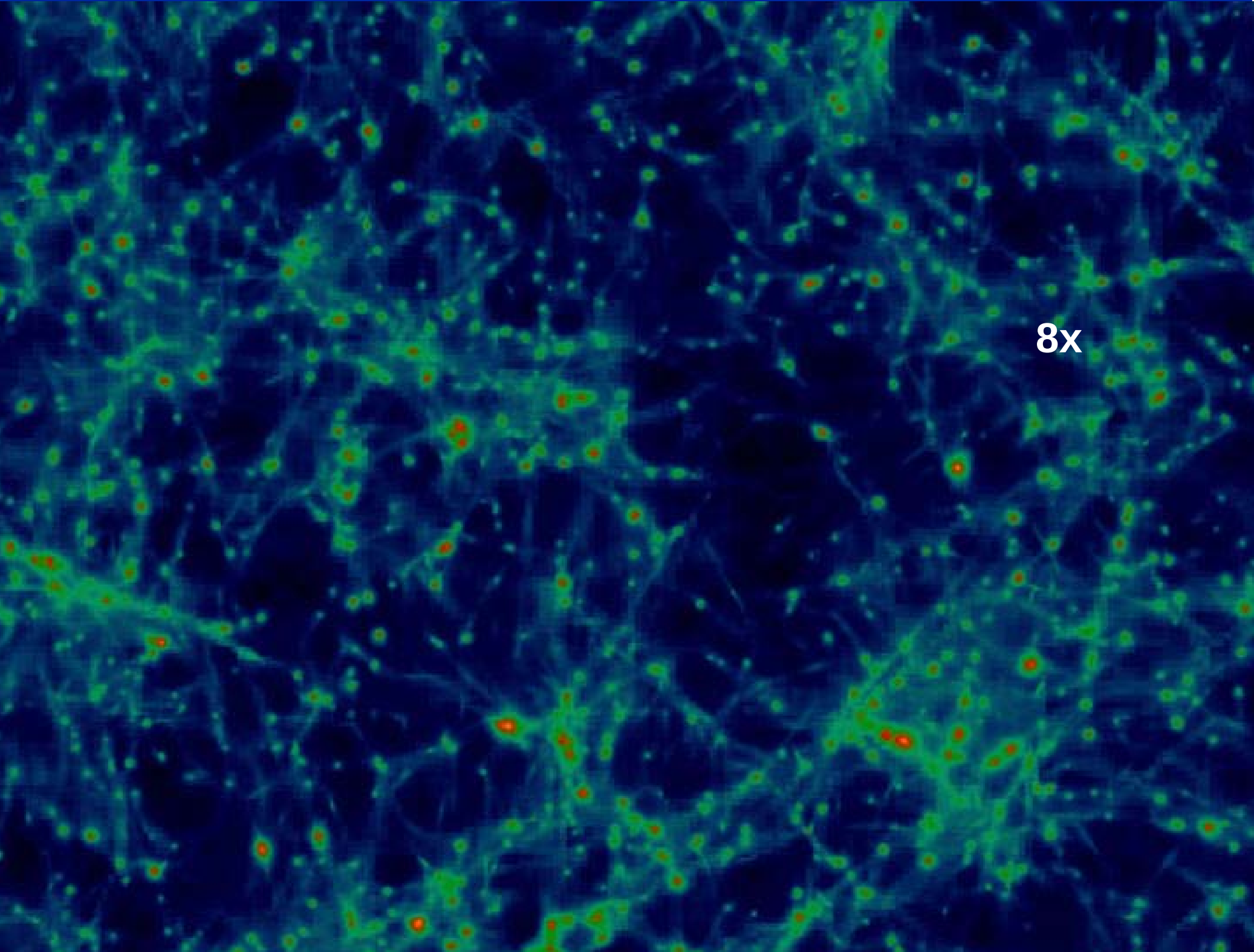
An 8192 x 8192 Image Extracted from Tile 8: How to Display/Explore?

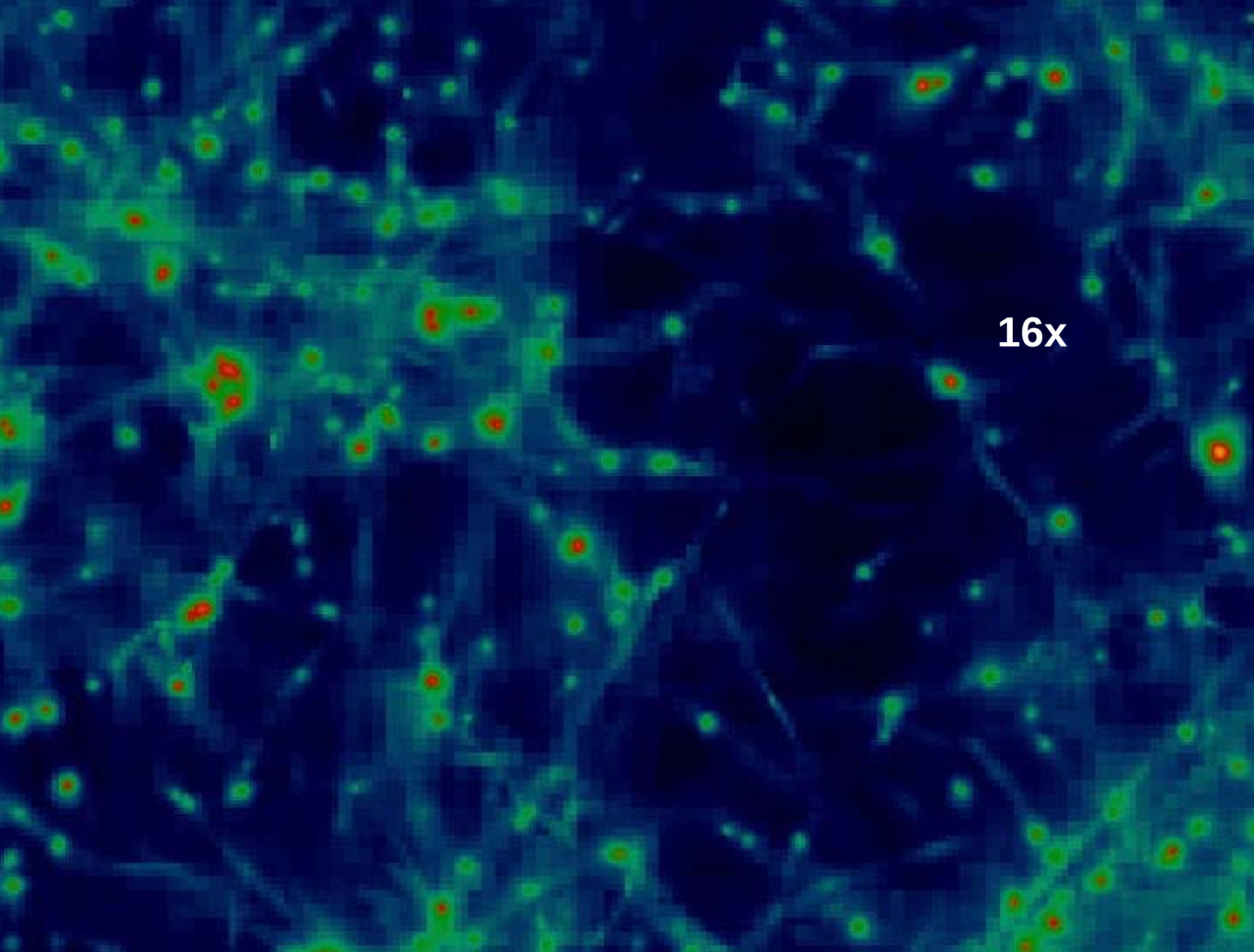


Working on Putting it on Calit2 HIPerWall OptIPortal

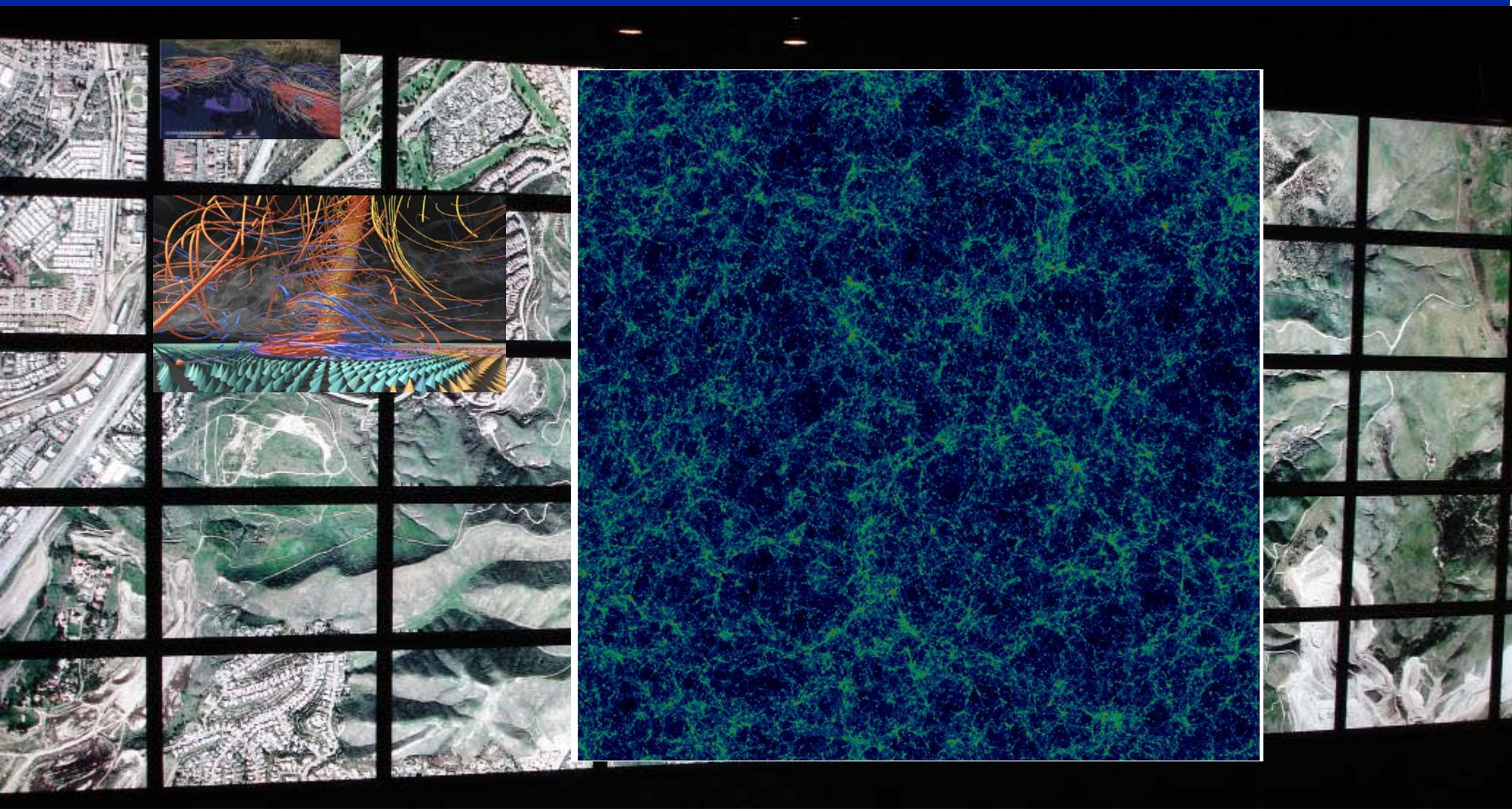

2x


4x


8x


16x

200 Million Pixels of Viewing Real Estate For Visually Analyzing Supercomputer Datasets



Goal: Link Norman's Lab OptIPortal Over Quartzite, CENIC, NLR/TeraGrid to Petascale Track 2 at Ranger@TACC and Kraken@NICS by October 2008

OptIPuter

it²