“Physics Research in an Era of Global Cyberinfrastructure"

Physics Department Colloquium
UCSD
La Jolla, CA
November 3, 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

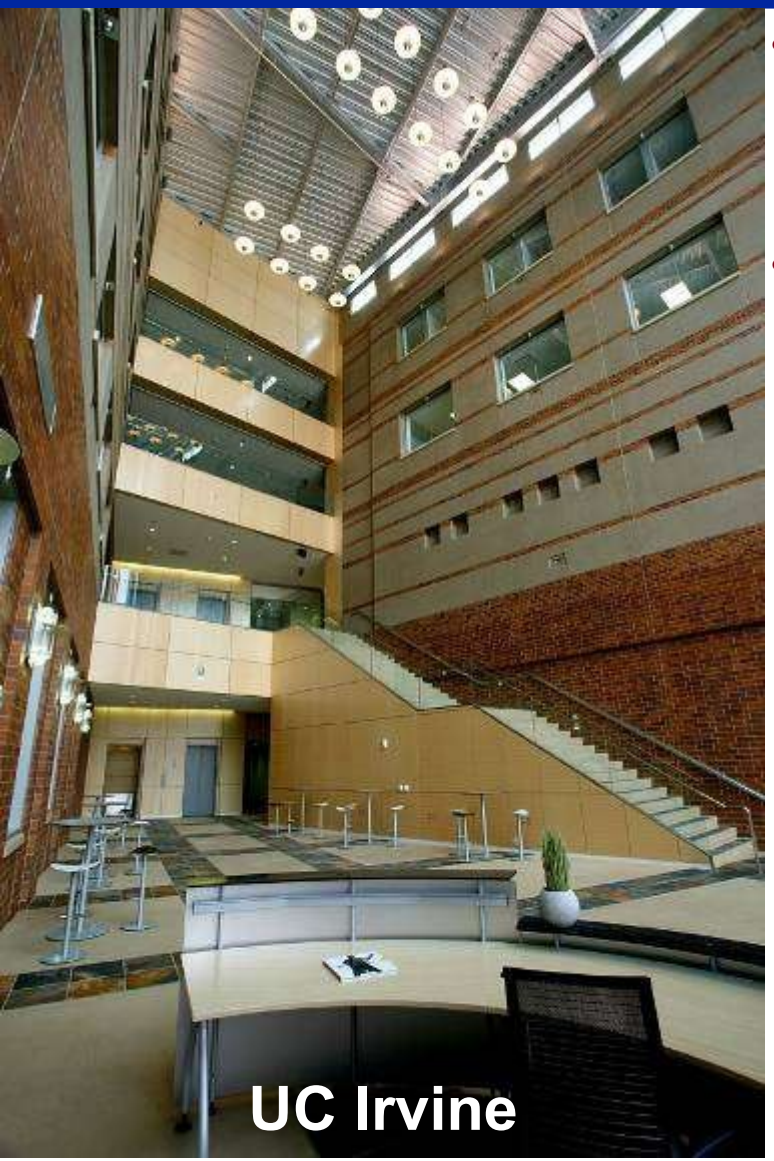
Abstract

Twenty years after the NSFnet launched today's shared Internet, a new generation of optical networks dedicated to single investigators are arising, with the ability to deliver up to 100-fold increase in bandwidth to the end user. The OptIPuter (www.optiputer.net) is one of the largest NSF-funded computer science research projects prototyping this new Cyberinfrastructure. Essentially, the OptIPuter is a “virtual metacomputer" in which the individual “processors” are widely distributed Linux clusters; the “backplane” is provided by Internet Protocol (IP) delivered over multiple dedicated lightpaths or "lambdas" (each 1-10 Gbps); and, the “mass storage systems” are large distributed scientific data repositories, fed by scientific instruments as OptIPuter peripheral devices, operated in near real-time. Furthermore, collaboration will be a defining OptIPuter characteristic; goals include implementing a next-generation Access Grid enabled with multiple HDTV and Super HD streams with photo realism. The OptIPuter extends the Grid program by making the underlying physical network elements discoveable and reservable, as well as the traditional computing and storage assets. Thus, the Grid is transformed into a LambdaGrid. A number of physics and astrophysics data-intensive project are prime candidates to drive this development.

OptIPuter

it2

Two New Calit2 Buildings Will Provide Major New Laboratories to Their Campuses



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

Richard C. Atkinson Hall Dedication Oct. 28, 2005



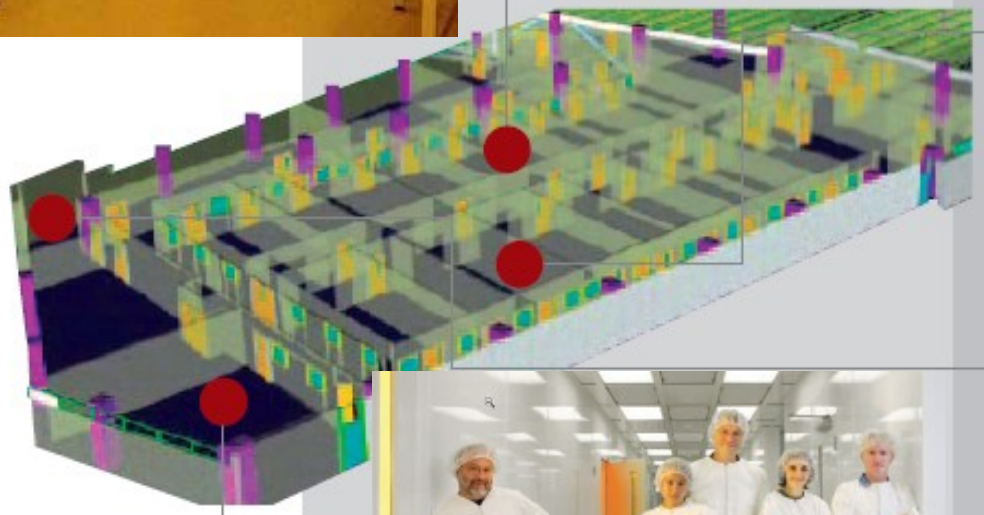
OptIPuter

www.calit2.net

it2

Calit2@UCSD Creates a Dozen Shared Clean Rooms for Nanoscience, Nanoengineering, Nanomedicine







Class 100 Clean Room Modules: Nano/microlithography

Areas with no more than 100 particles larger than 0.5 microns per cubic foot of air. Facilities include e-beam lithography and photolithography modules.

Class 1000 Clean Room Modules: Nano/microfabrication

No more than 1,000 particles larger than 0.5 microns per cubic foot of air. Areas include metallization/thin film deposition, advanced dry etching, metrology, thermal and back-end processes.

Advanced Structural Characterization

Techniques include scanning electron microscopy (using electrons rather than light to form an image) and scanning probe microscopy. Ultrahigh vacuum surface analysis. Optical characterization.

OTHER FACILITIES INCLUDE:

Metal-Organic Chemical Vapor Deposition Facility

MOCVD technique used for thin film growth in advanced, compound semiconductors and nanostructures.

Photo Courtesy of Bernd Fruhberger, Calit2

The Calit2@UCSD Building is Designed for Prototyping Extremely High Bandwidth Applications

1.8 Million Feet of Cat6 Ethernet Cabling

Speed From Here

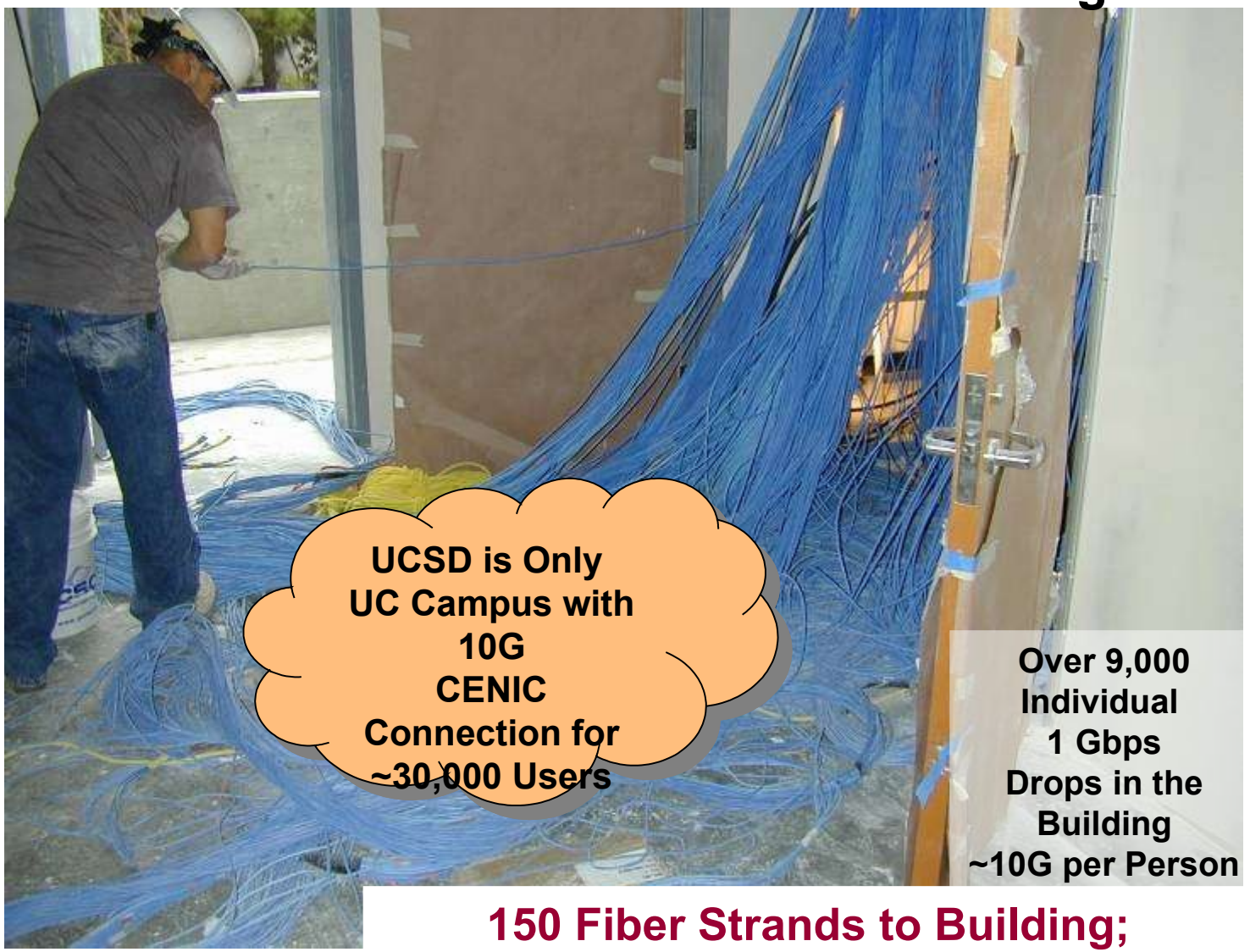


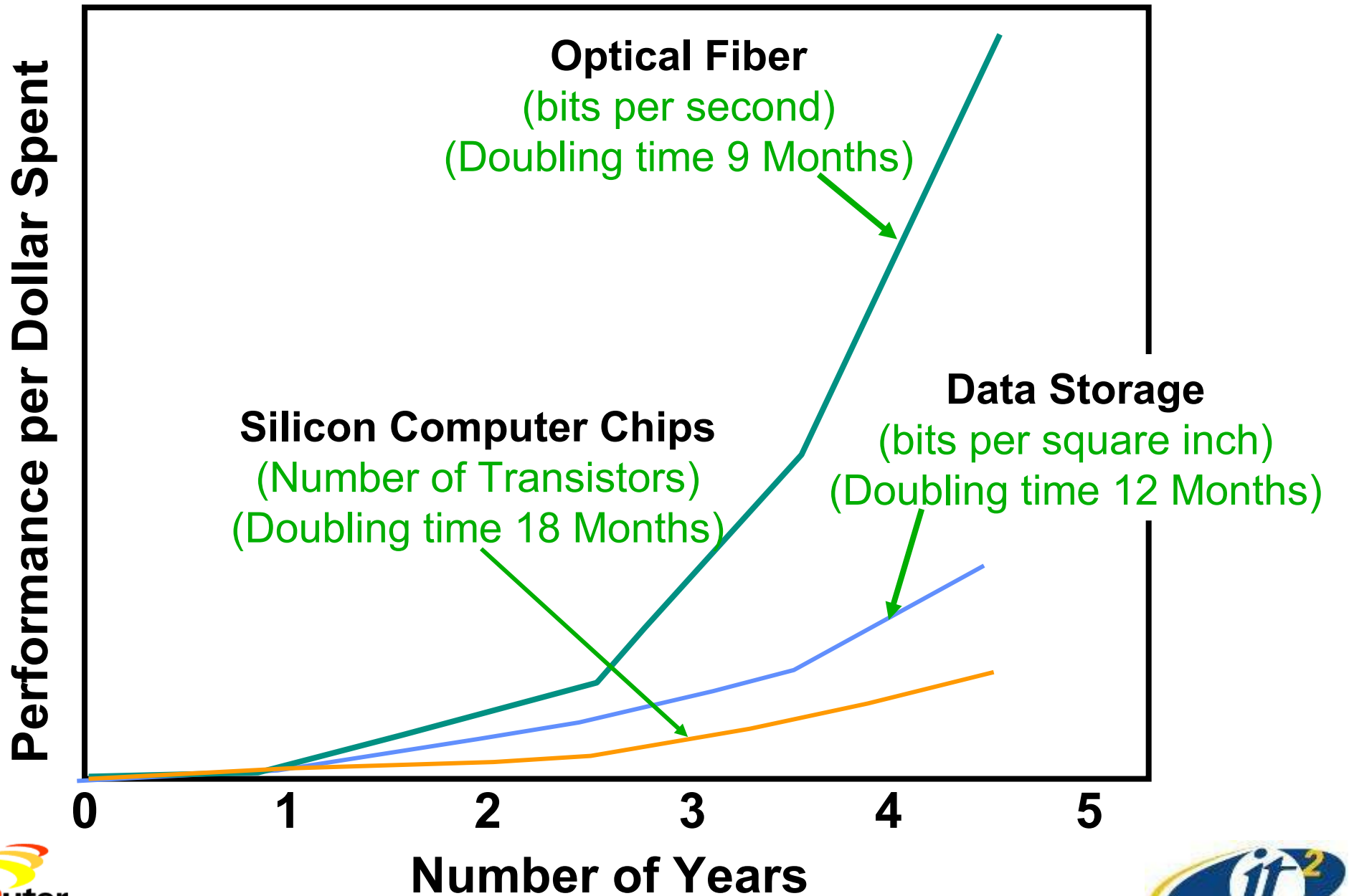
Photo: Tim Beach, Calit2

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

OptIPuter

it2

Why Optical Networks Will Become the 21st Century Driver



Scientific American, January 2001

Calit2@UCSD Is Connected to the World 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

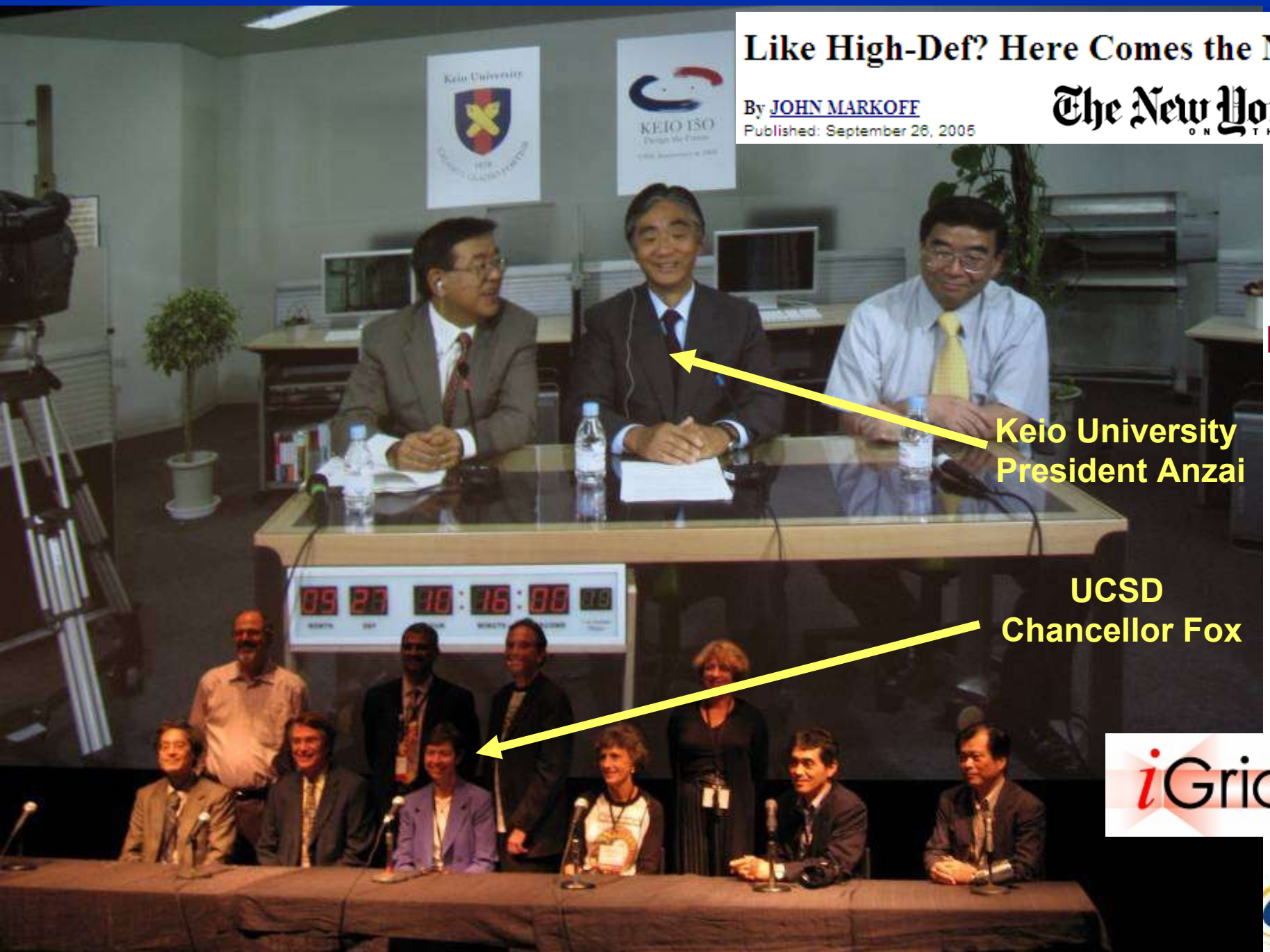
Calit2 @ University of California, San Diego

California Institute for Telecommunications and Information Technology



50 Demonstrations, 20 Counties, 10 Gbps/Demo

First Trans-Pacific Super High Definition Telepresence Meeting in New Calit2 Digital Cinema Auditorium



Used 1Gbps Dedicated

Sony
NTT
SGI

iGrid 2005

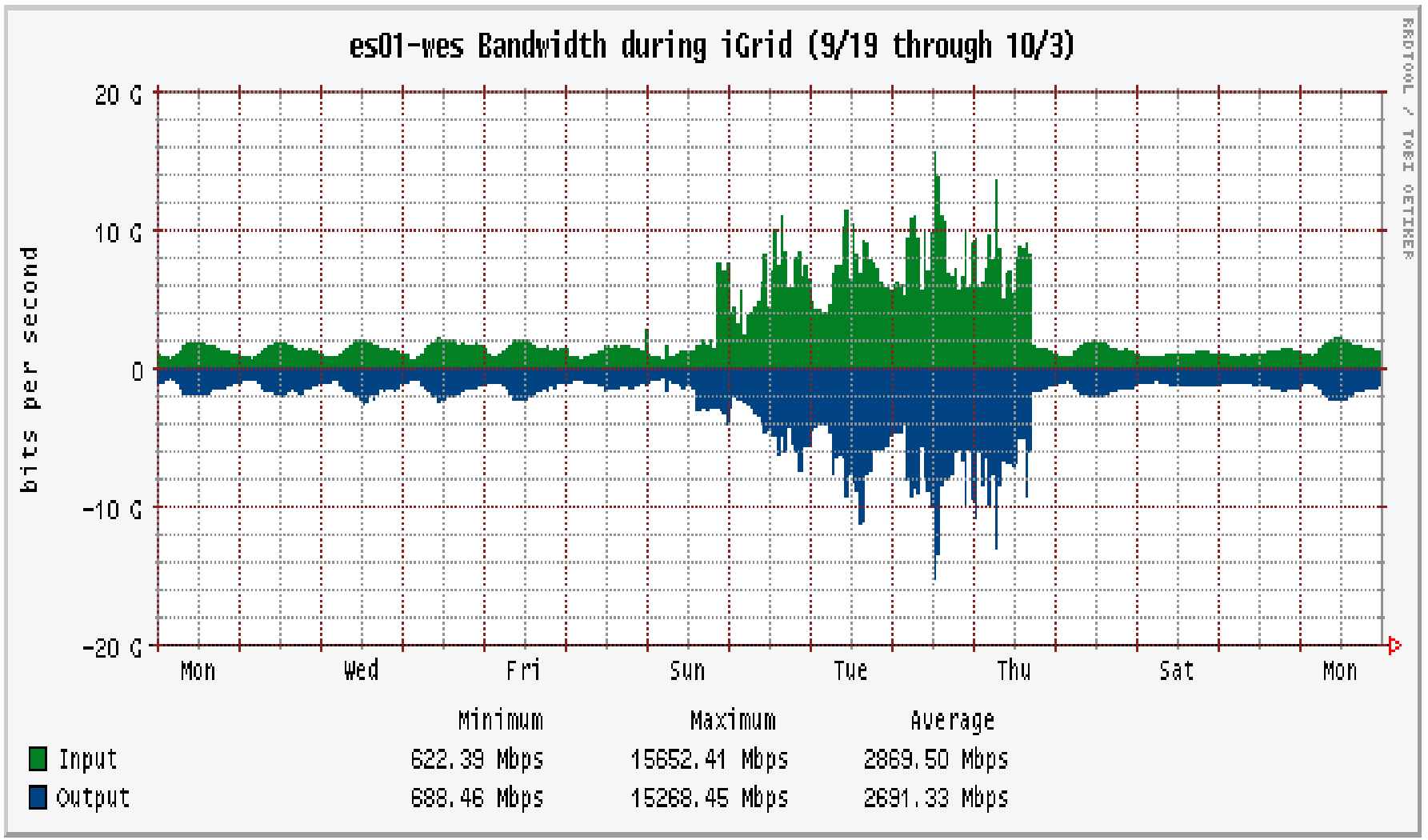
it2

First Remote Interactive High Definition Video Exploration of Deep Sea Vents



Source John Delaney & Deborah Kelley, UWash

iGrid2005 Data Flows Multiplied Normal Flows by Five Fold!

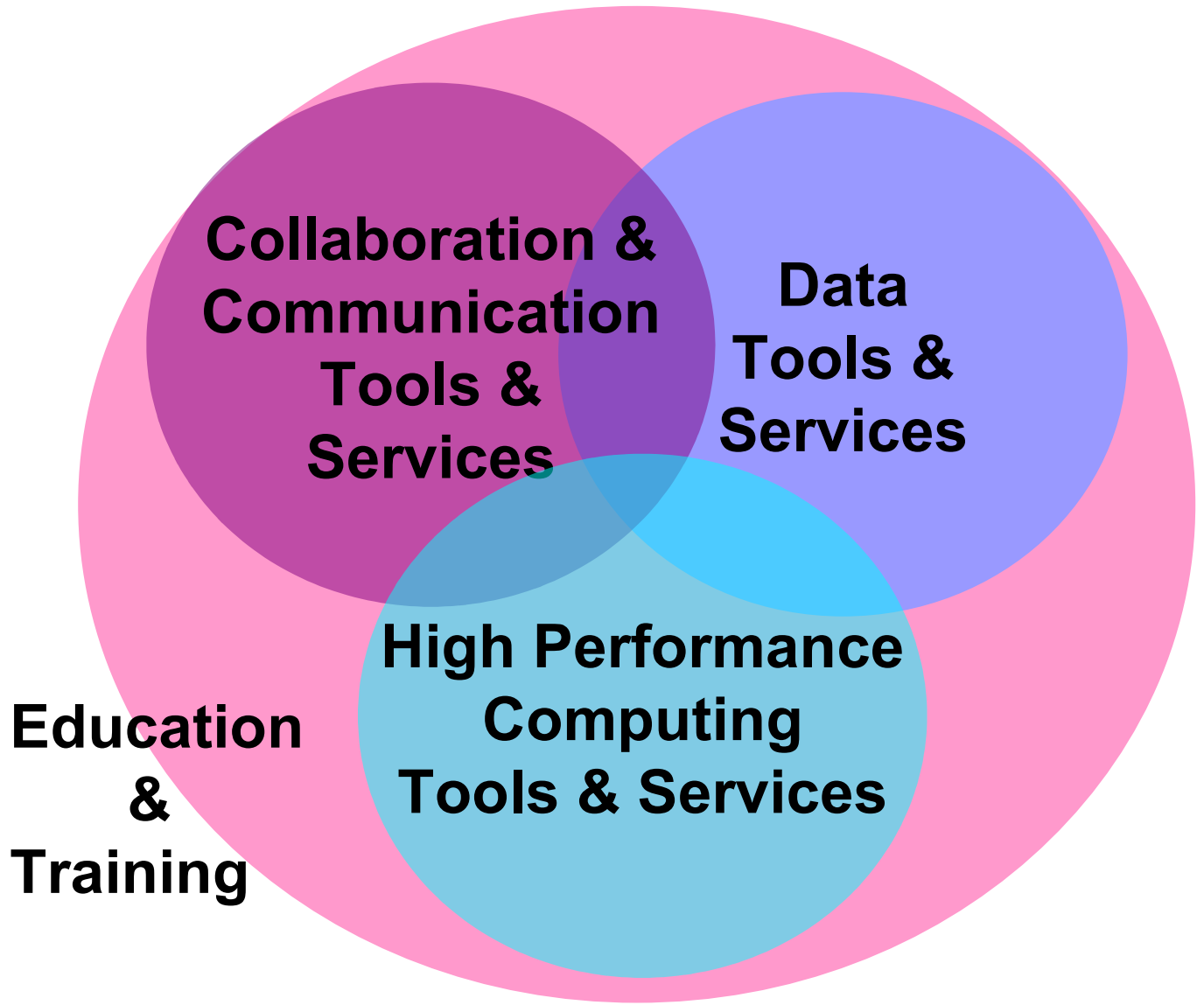


Data Flows Through the Seattle PacificWave International Switch

OptIPuter

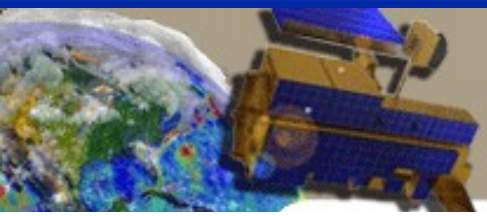
it2

A National Cyberinfrastructure is Emerging for Data Intensive Science



Source: Guy Almes,
Office of Cyberinfrastructure, NSF

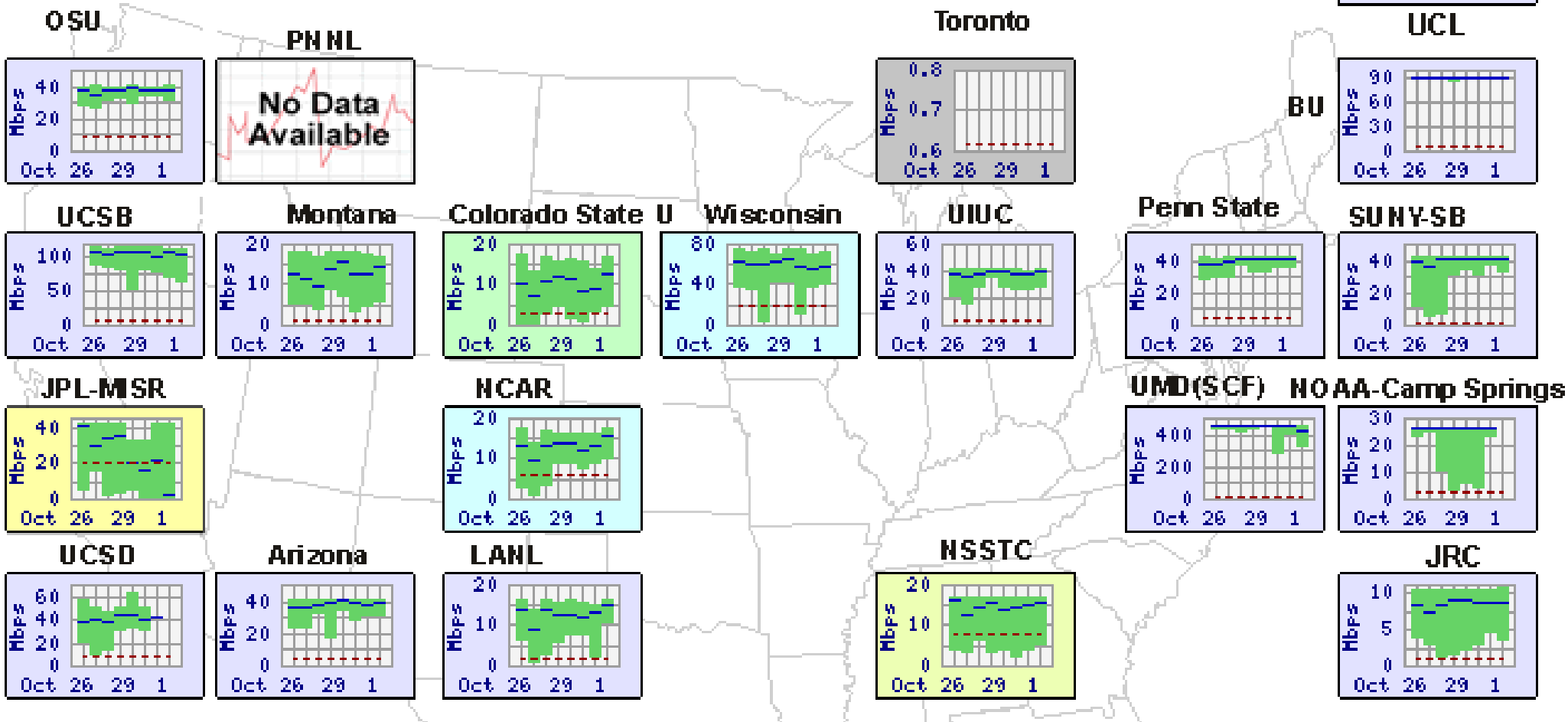
Challenge: Average Throughput of NASA Data Products to End User is < 50 Mbps



ENSIGHT Active Testing TERRA Destinations

Tested October 2005

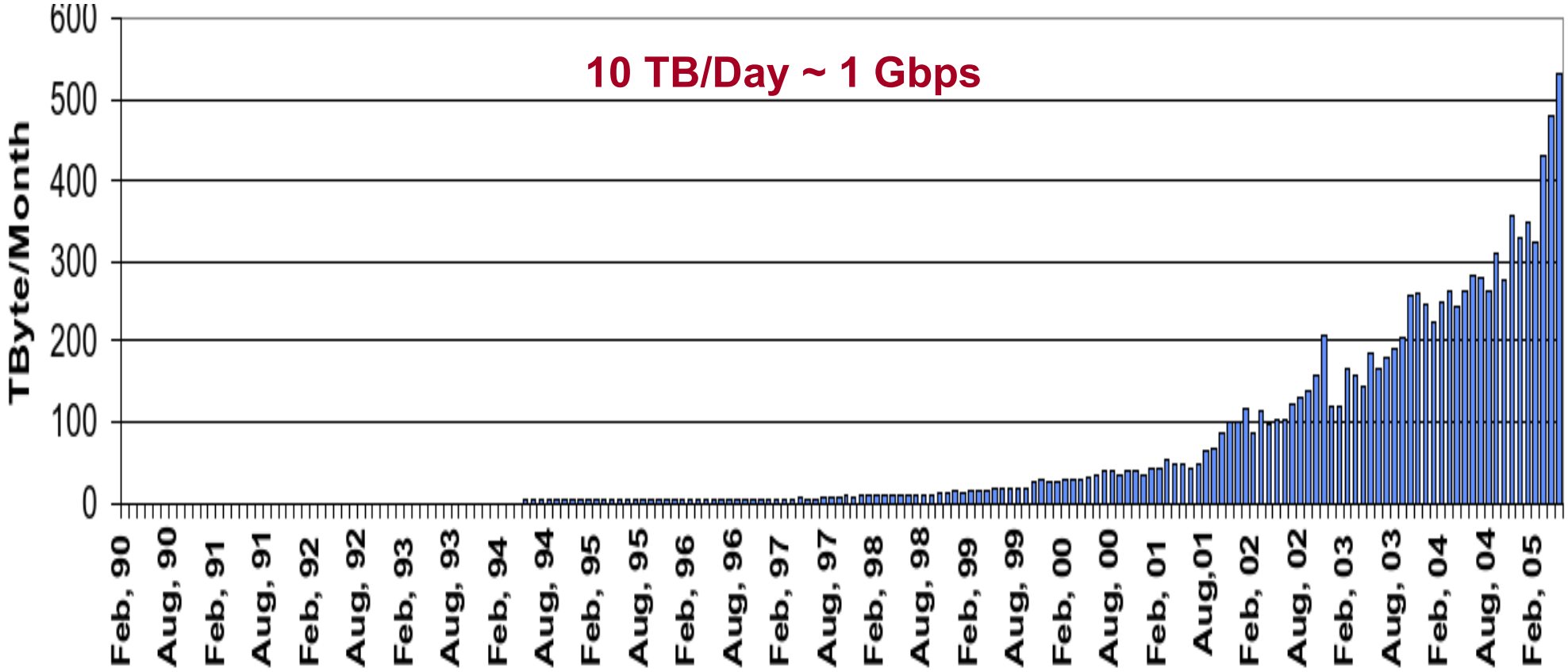
Internet2 Backbone is 10,000 Mbps!
Throughput is < 0.5% to End User



Data Intensive Science is Overwhelming the Conventional Internet

ESnet is Currently Transporting About 20 Terabytes/Day and This Volume is Increasing Exponentially

ESnet Monthly Accepted Traffic
Feb., 1990 – May, 2005



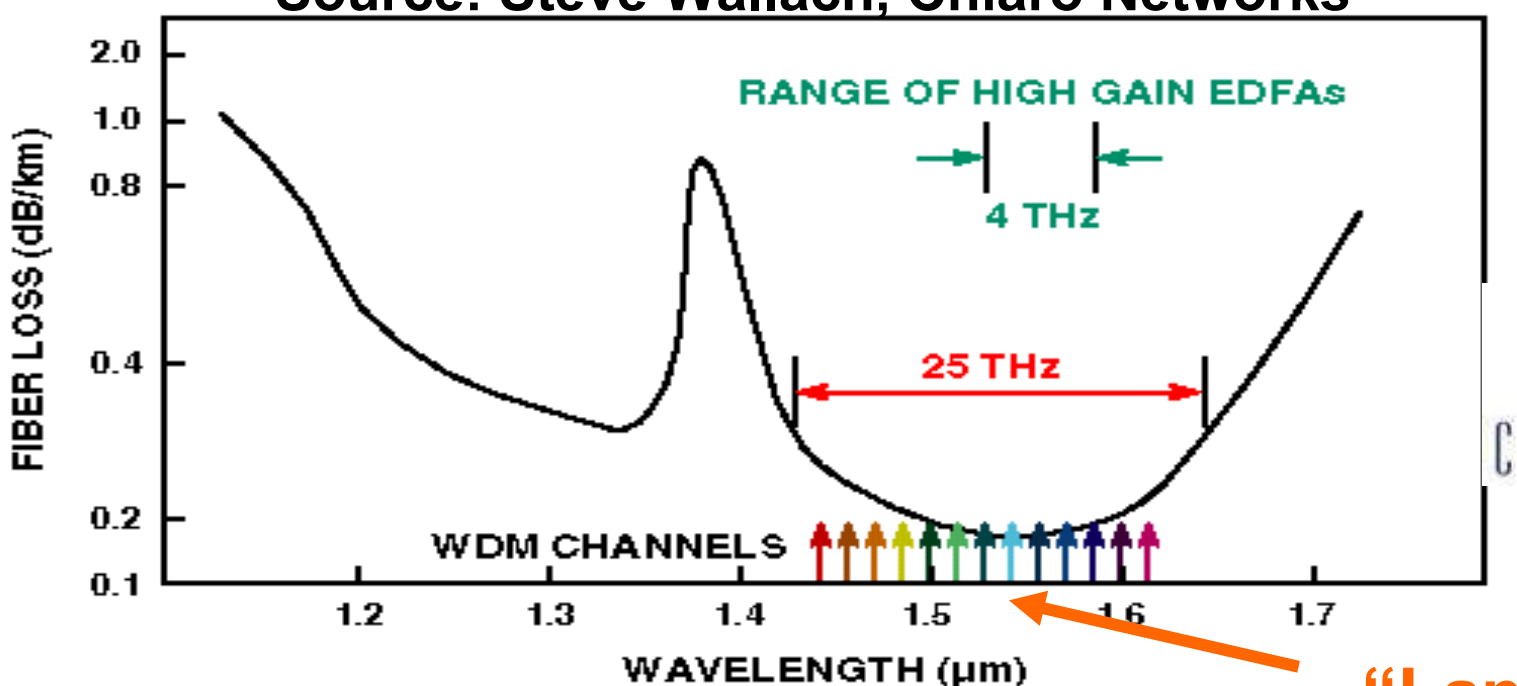
Source: Bill Johnson, DOE

Dedicated Optical Channels Makes High Performance Cyberinfrastructure Possible

WAVELENGTH DIVISION MULTIPLEXING (WDM)

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

Source: Steve Wallach, Chiaro Networks



CHIARO

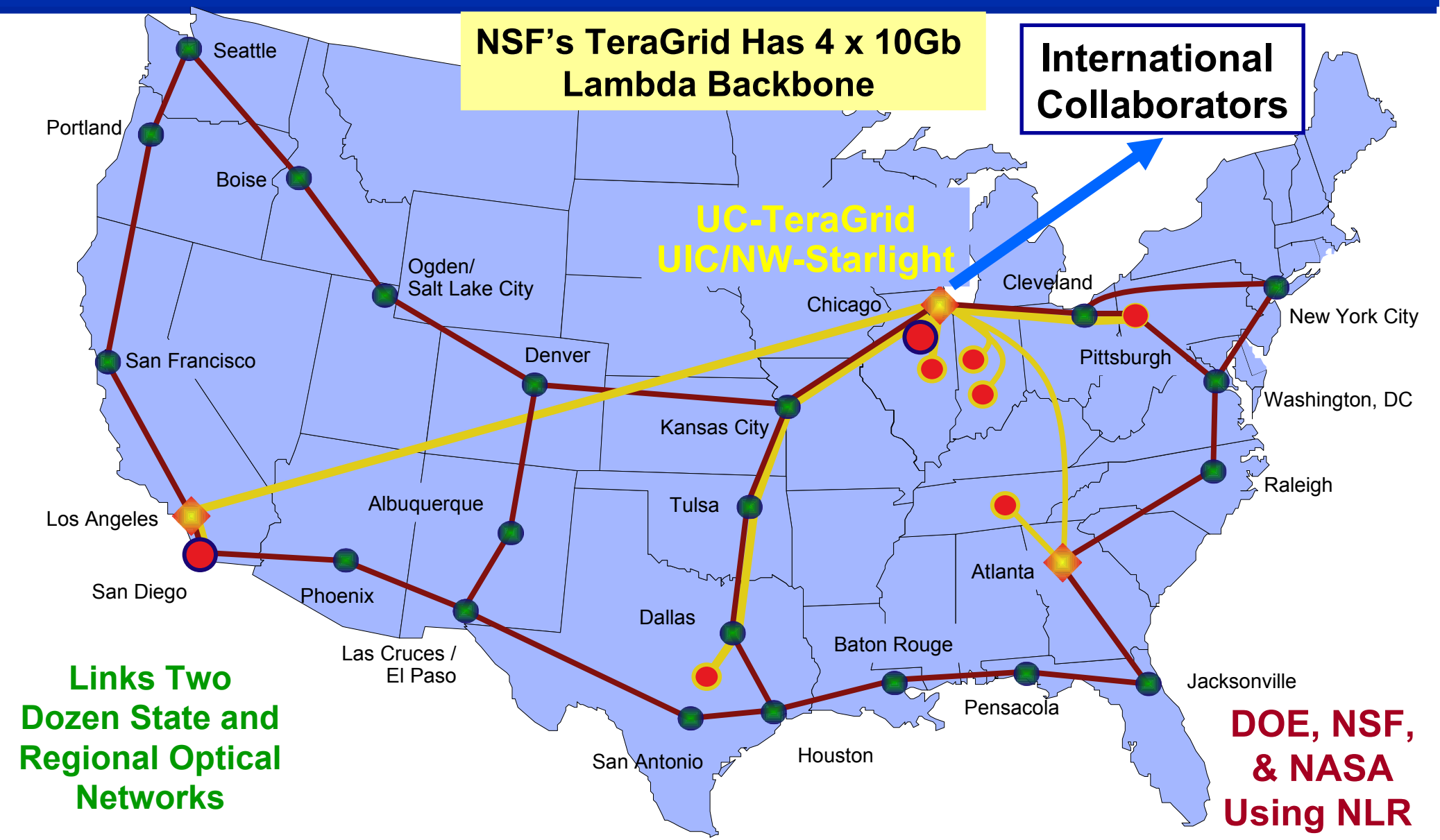
"Lambdas"

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

OptIPuter

it2

National LambdaRail (NLR) and TeraGrid Provides Cyberinfrastructure Backbone for U.S. Researchers



NLR 4 x 10Gb Lambdas Initially Capable of 40 x 10Gb wavelengths at Buildout

OptIPuter

it2

Campus Infrastructure is the Obstacle

"Research is being stalled by 'information overload,'" Mr. Bement said, because data from digital instruments are piling up far faster than researchers can study.

In particular, he said, campus networks need to be improved. High-speed data lines crossing the nation are the equivalent of six-lane superhighways, he said. But networks at colleges and universities are not so capable.

"Those massive conduits are reduced to two-lane roads at most college and university campuses," he said.

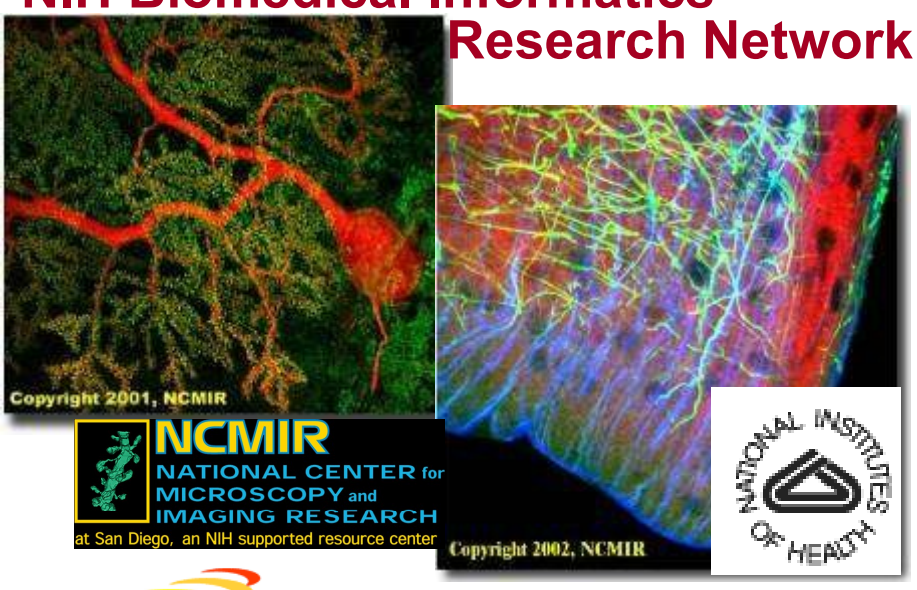
Improving cyberinfrastructure, he said, "will transform the capabilities of campus-based scientists."

--Arden Bement, director National Science Foundation,
Chronicle of Higher Education 51 (36), May 2005.

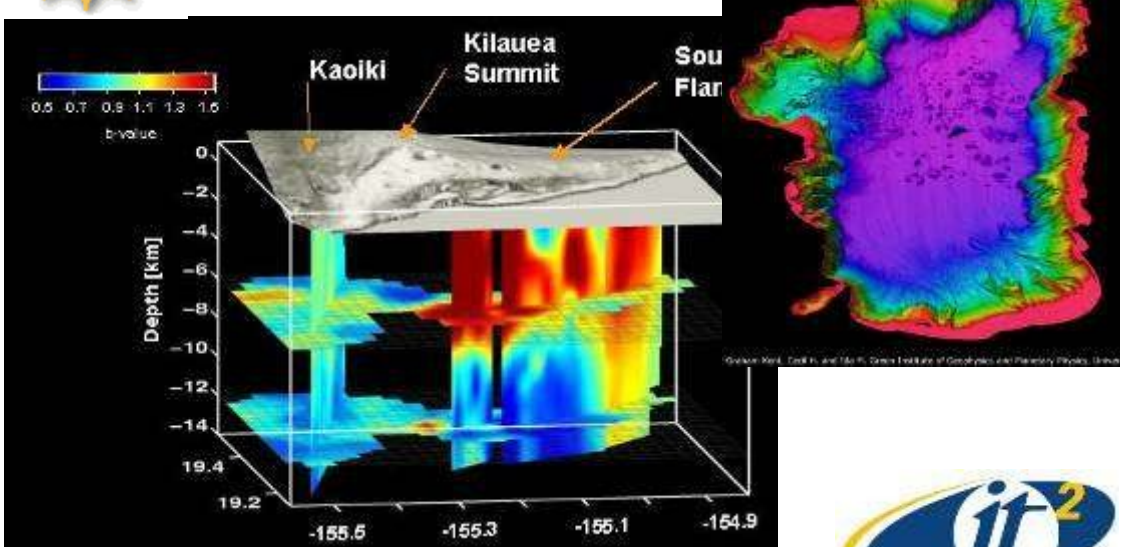
The OptIPuter Project – Linking Global Scale Science Resources to User's Linux Clusters

- **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—Entering 4th Year**
- **Creating a LambdaGrid "Web" for Gigabyte Data Objects**

NIH Biomedical Informatics Research Network



NSF EarthScope and ORION



OptIPuter

SDSC SAN DIEGO SUPERCOMPUTER CENTER

it2

UCSD is Prototyping Campus-Scale National LambdaRail "On-Ramps"



0.320 Tbps Backplane Bandwidth

Juniper T320

20X

6.4 Tbps Backplane Bandwidth



Chiaro Estara

To CENIC

SDSC

SAN DIEGO SUPERCOMPUTER CENTER

SDSC

SDSC Annex

Engineering

High School

Medicine

Collocation

Earth Sciences

½ Mile

Campus Provided Dedicated Fibers Between Sites Linking Linux Clusters

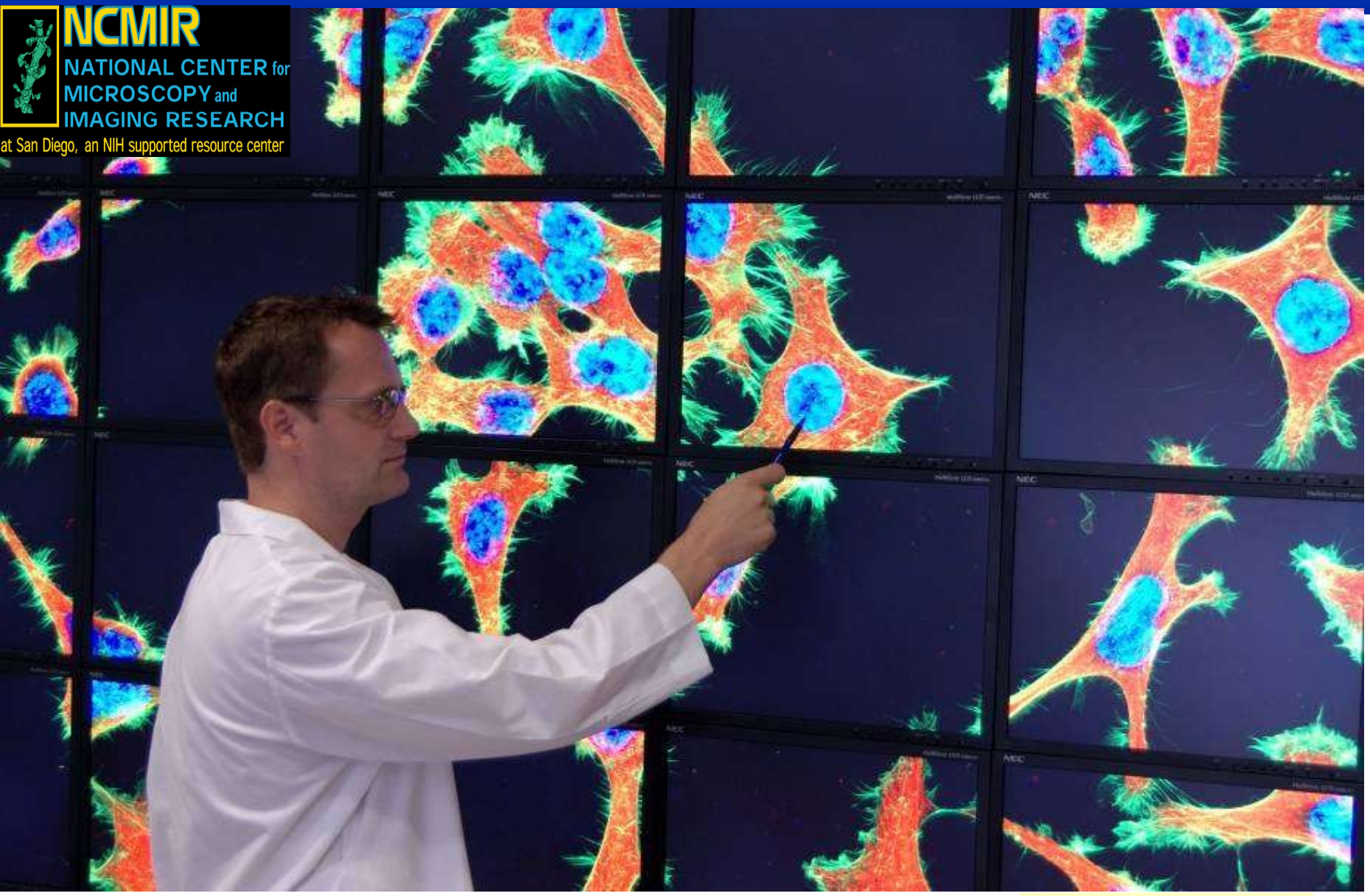
UCSD Has ~ 50 Labs With Clusters

Source: Phil Papadopoulos, SDSC; Greg Hidley, Calit2

OptIPuter

it2

Increasing Data Rate into Lab by 100x, Requires High Resolution Portals to Global Science Data



Source: Mark Ellisman, David Lee, Jason Leigh, Tom Deerinck

OptIPuter

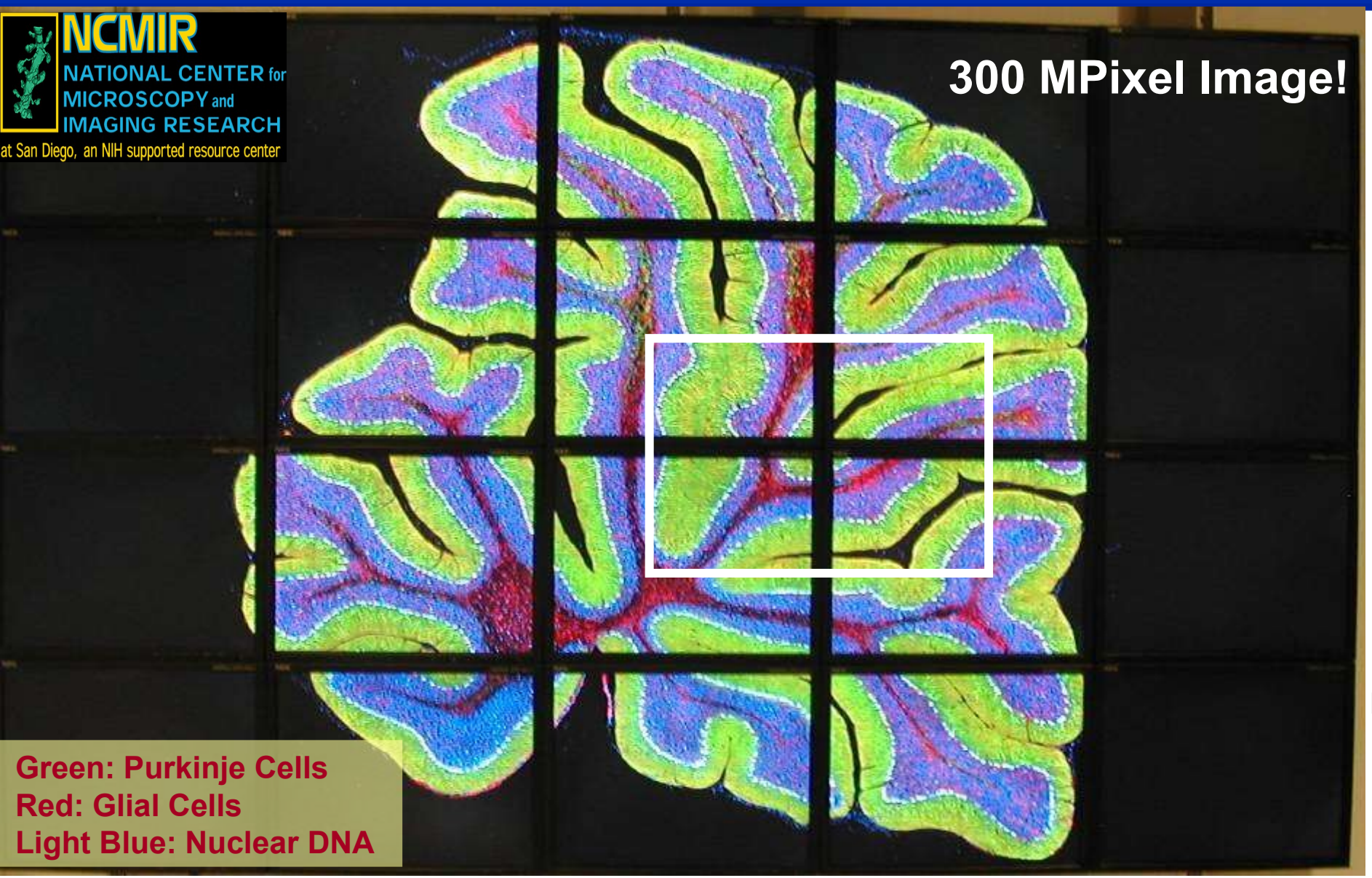
650 Mpixel 2-Photon Microscopy
Montage of HeLa Cultured Cancer Cells

Green: Actin
Red: Microtubles
Light Blue: DNA

it2

OptIPuter Scalable Displays Developed for Multi-Scale Imaging

NCMIR
NATIONAL CENTER for MICROSCOPY and IMAGING RESEARCH
at San Diego, an NIH supported resource center



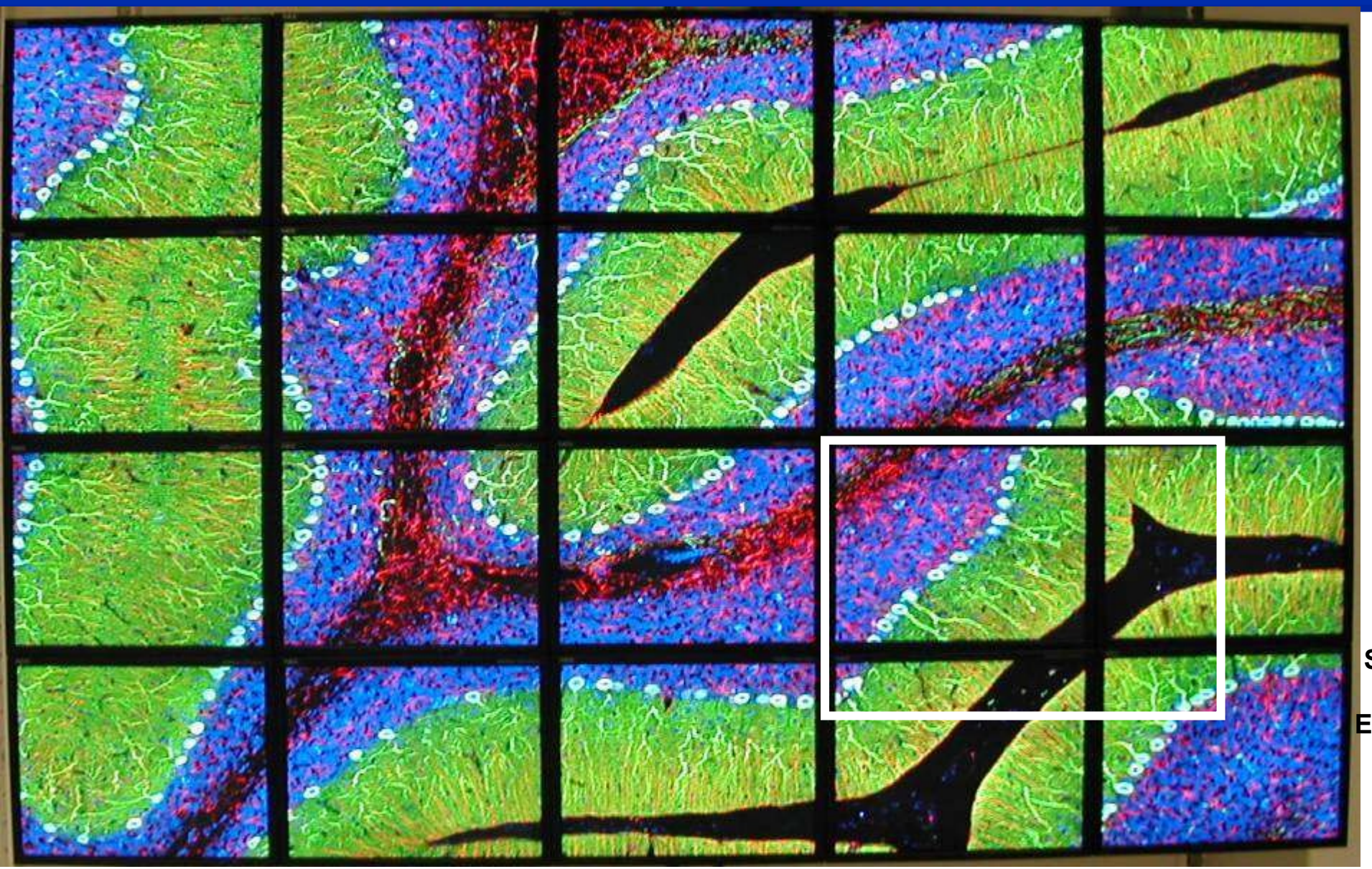
Source: Mark Ellisman, David Lee, Jason Leigh

Two-Photon Laser Confocal Microscope Montage of 40x36=1440 Images in 3 Channels of a Mid-Sagittal Section of Rat Cerebellum Acquired Over an 8-hour Period

OptIPuter

it2

Scalable Displays Allow Both Global Content and Fine Detail



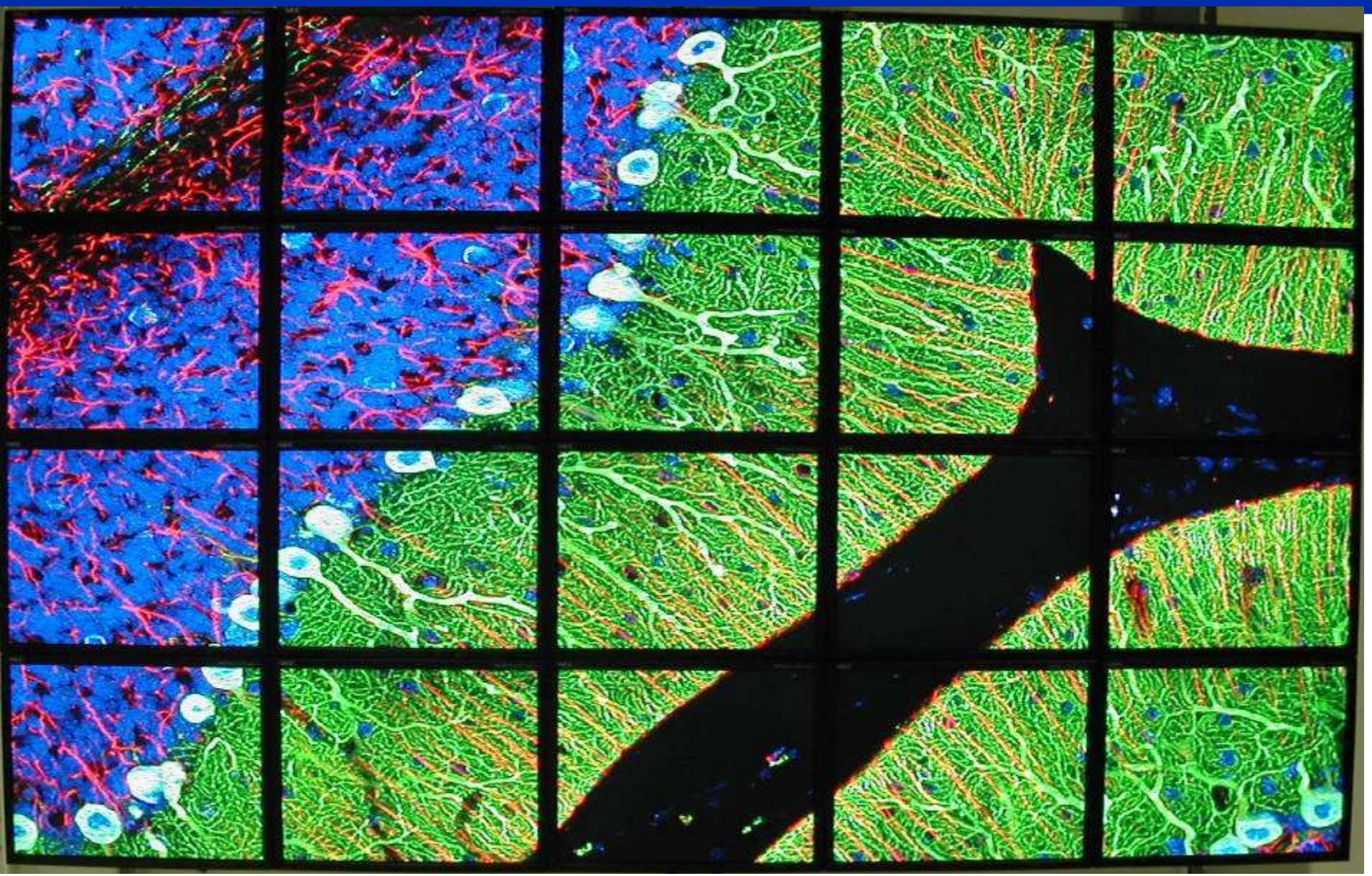
Source: Mark Ellisman, David Lee, Jason Leigh

OptIPuter

30 MPixel SunScreen Display Driven by a 20-node Sun Opteron Visualization Cluster

it2

Allows for Interactive Zooming from Cerebellum to Individual Neurons



OptIPuter

Source: Mark Ellisman, David Lee, Jason Leigh

it²

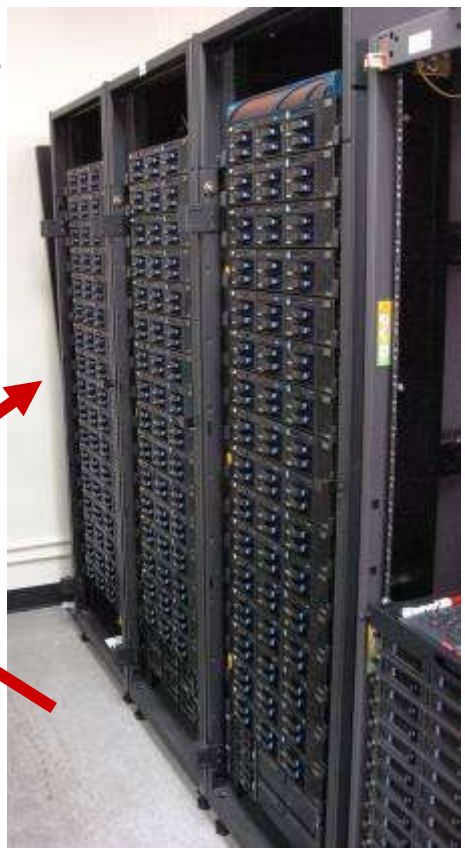
Campuses Must Provide Fiber Infrastructure to End-User Laboratories & Large Rotating Data Stores

SIO Ocean Supercomputer



UCSD Campus LambdaStore Architecture

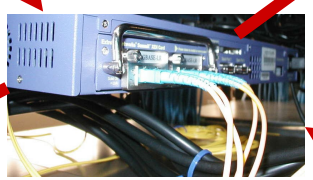
IBM Storage Cluster



Streaming Microscope

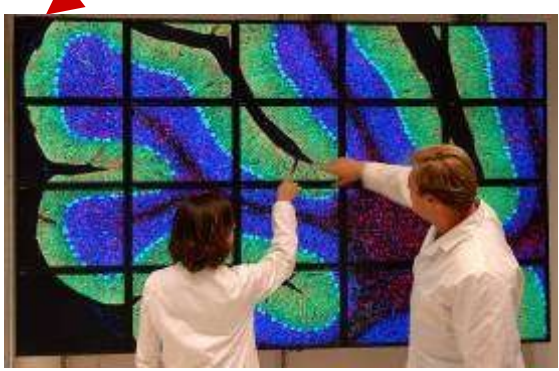






2 Ten Gbps Campus Lambda Raceway







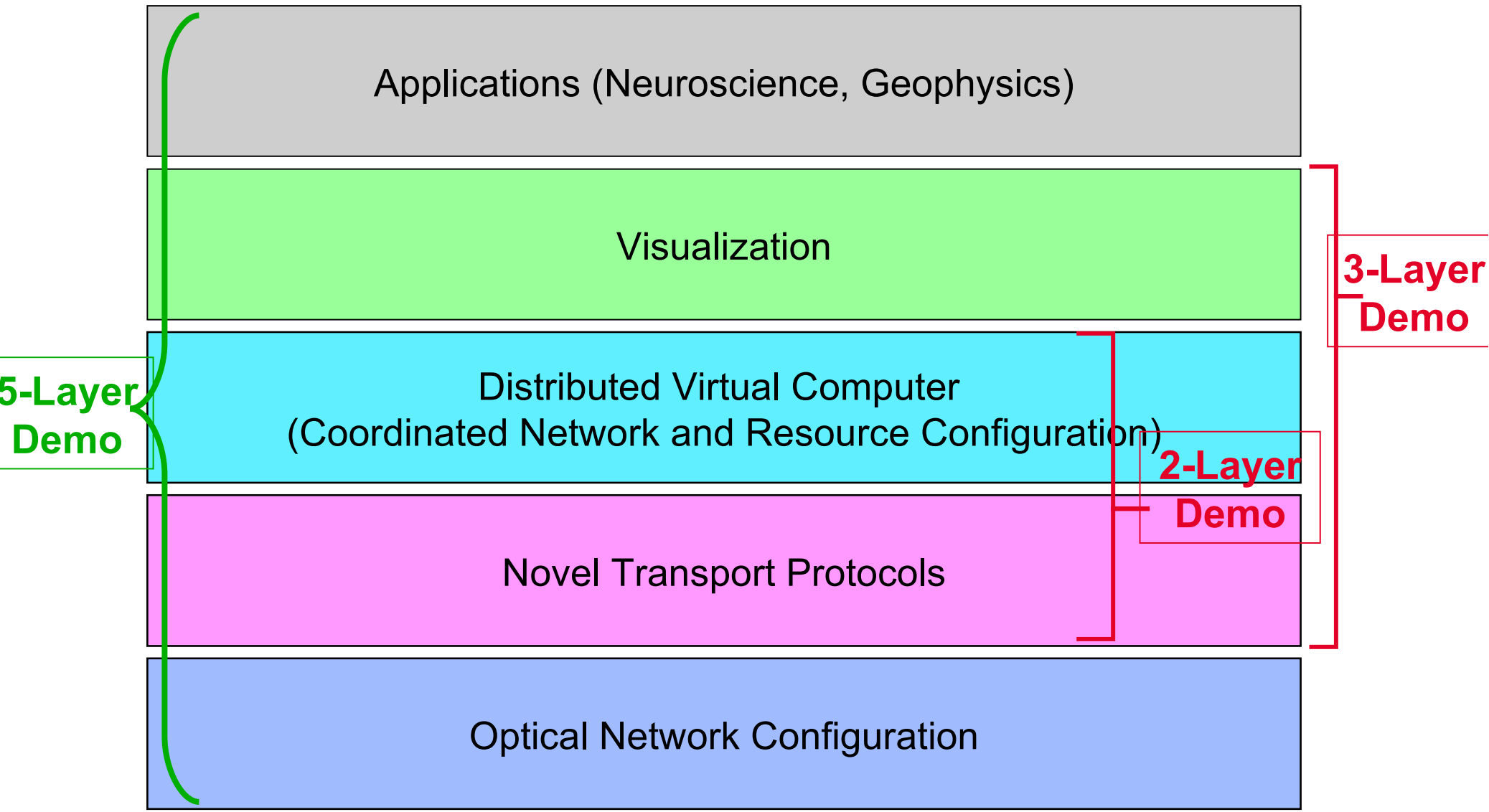
Global LambdaGrid

Source: Phil Papadopoulos, SDSC, Calit2

OptIPuter

it2

Exercising the OptIPuter LambdaGrid Middleware Software "Stack"

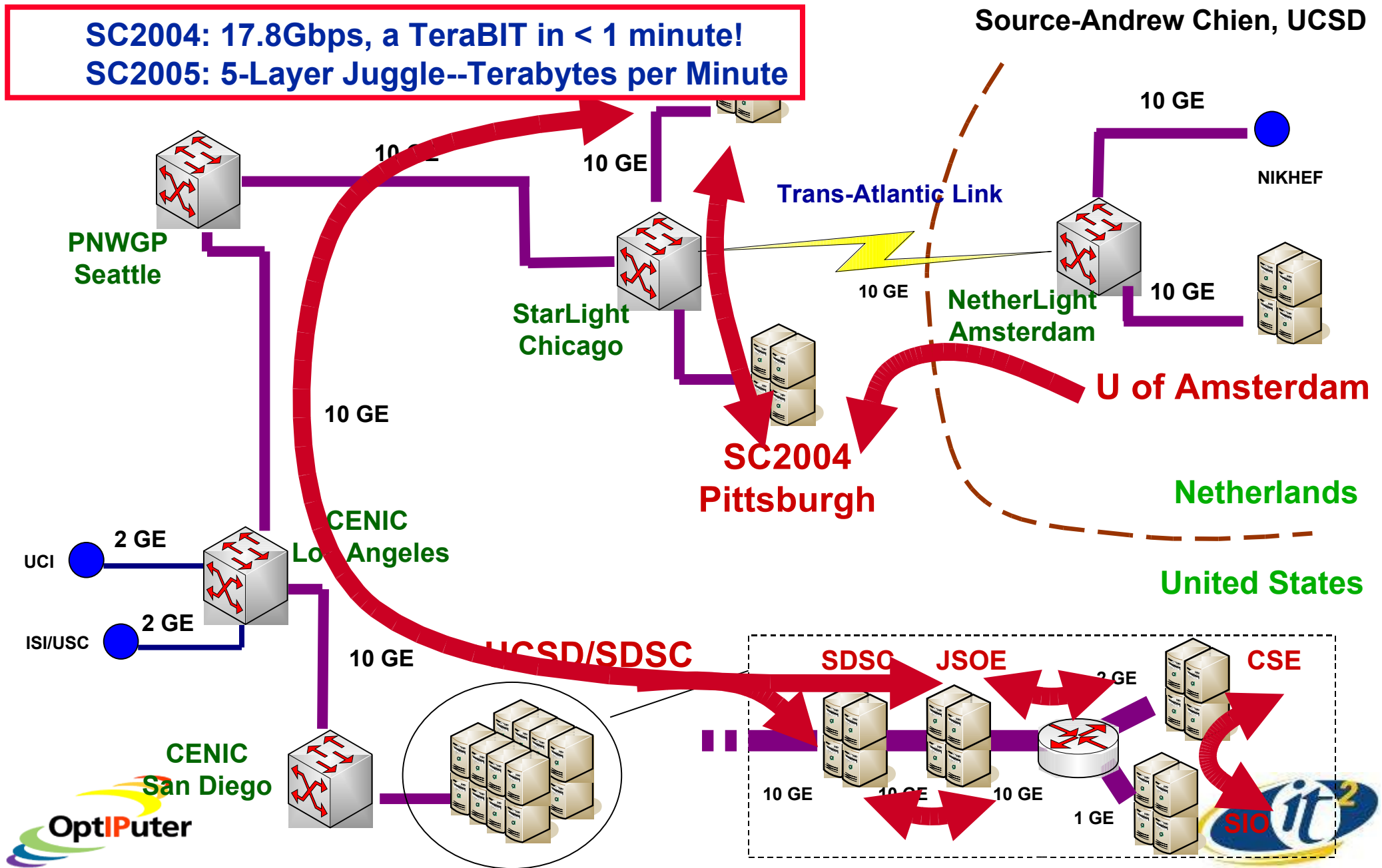


Source-Andrew Chien, UCSD- OptIPuter Software System Architect

First Two-Layer OptIPuter Terabit Juggling on 10G WANs

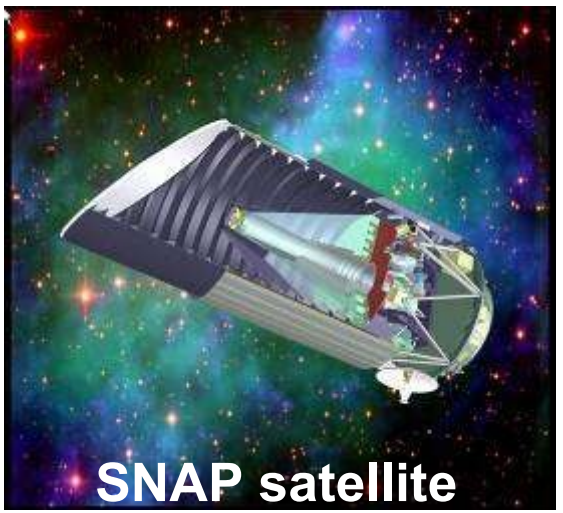
SC2004: 17.8Gbps, a TeraBIT in < 1 minute!
SC2005: 5-Layer Juggle--Terabytes per Minute

Source-Andrew Chien, UCSD

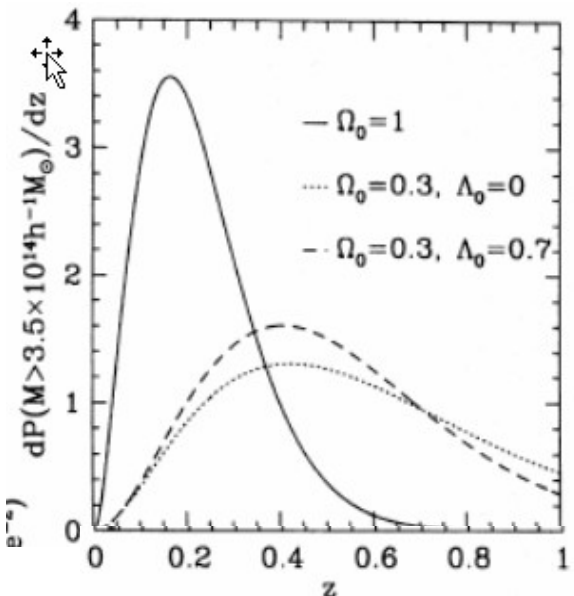


UCSD Physics Department Research That Requires a LambdaGrid — The Universe's Dark Energy Equation of State

- **Principal Goal of NASA-DOE Joint Dark Energy Mission (JDEM)**
- **<u>Approach</u>: Precision Measurements of Expansion History of the Universe Using Type Ia Supernovae *Standardizable Candles***
- **<u>Complimentary Approach</u>: Measure Redshift Distribution of Galaxy Clusters**
  - **Must Have Detailed Simulations of How Cluster Observables Depend on Cluster Mass *On The Lightcone* for Different Cosmological Models**



Cluster abundance vs. z

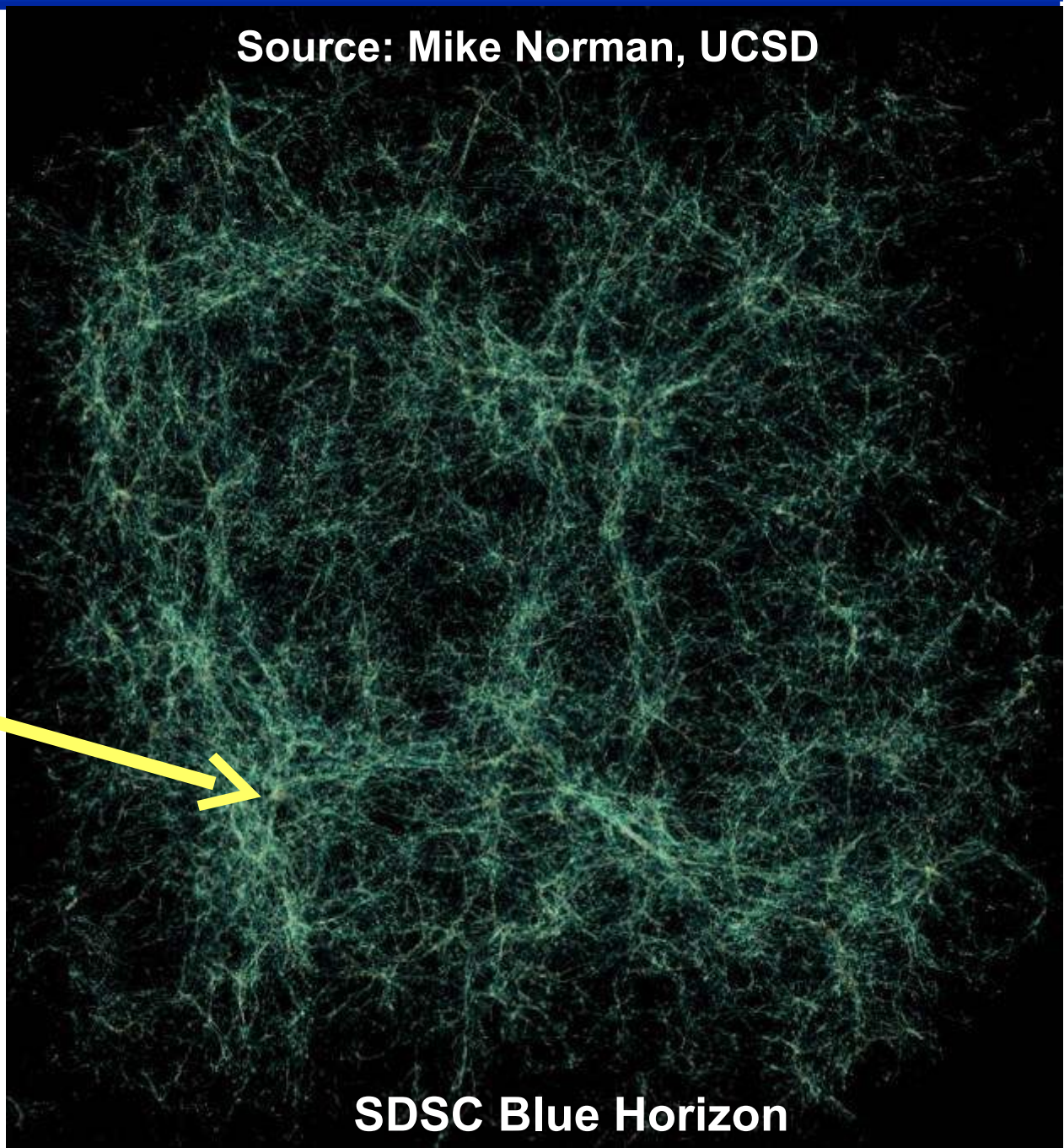


OptIPuter

Source: Mike Norman, UCSD

Cosmic Simulator with a Billion Zone and Gigaparticle Resolution

Problem with Uniform Grid--Gravitation Causes Continuous Increase in Density Until There is a Large Mass in a Single Grid Zone



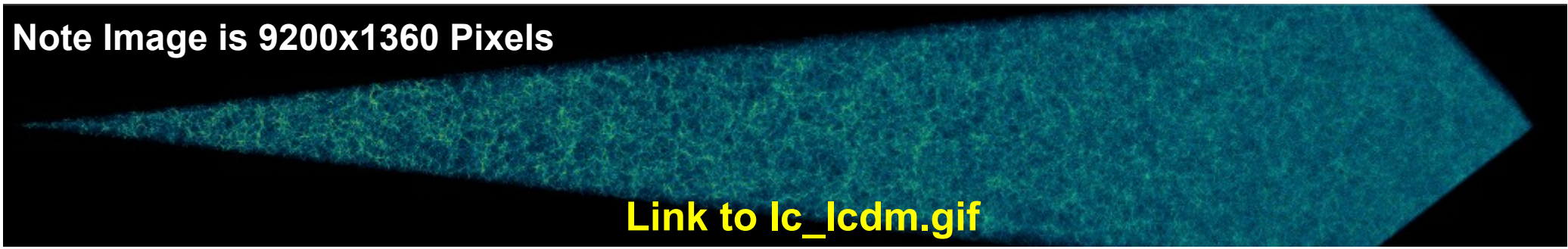
OptIPuter

AMR Allows Digital Exploration of Early Galaxy and Cluster Core Formation

- **Background Image Shows Grid Hierarchy Used**
  - Key to Resolving Physics is More Sophisticated Software
  - Evolution is from 10Myr to Present Epoch
- **Every Galaxy > 10^{11} M_{solar} in 100 Mpc/H Volume Adaptively Refined With AMR**
  - 256^3 Base Grid
    - Over 32,000 Grids At 7 Levels Of Refinement
    - Spatial Resolution of 4 kpc at Finest
    - 150,000 CPU-hr On 128-Node IBM SP
- **512^3 AMR or 1024^3 Unigrid Now Feasible**
  - 8-64 Times The Mass Resolution
  - Can Simulate First Galaxies
  - One Million CPU-Hr Request to LLNL
    - Bottleneck--Network Throughput from LLNL to UCSD

Source: Mike Norman, UCSD

Lightcone Simulation--Computing the Statistics of Galaxy Clustering versus Redshift



redshift

ct (Gyr)

0 -1 -2 -3 -4 -5

- **Evrard et al. (2003)**
  - Single, 1024^3 P^3M
  - $L/\Delta=10^4$
  - Dark matter only
- **Norman/LLNL Project**
  - Multiple, 512^3 AMR
  - Optimal Tiling of Lightcone
  - $L/\Delta=10^5$
  - Dark Matter + Gas

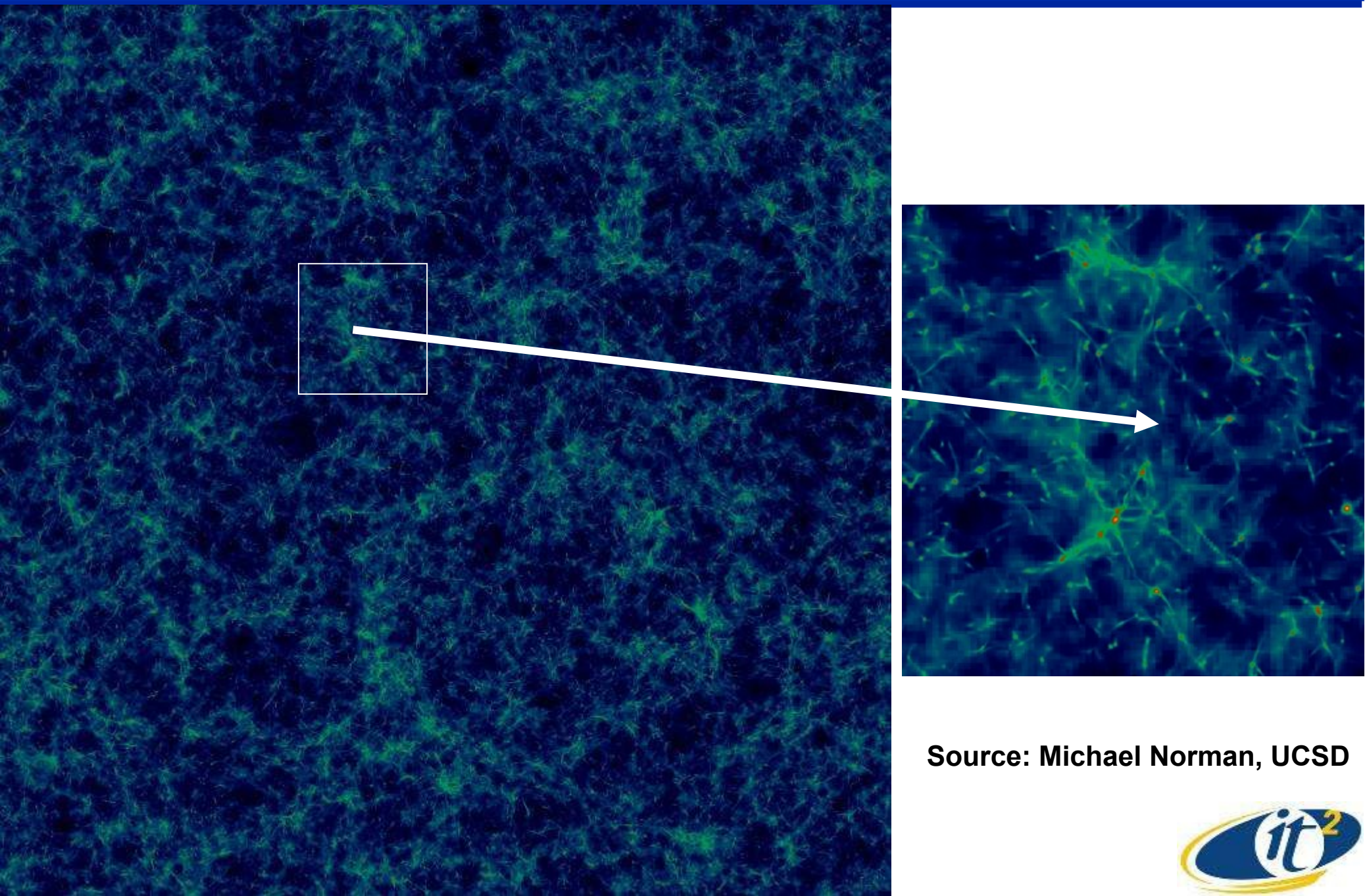
Researchers hope to distinguish between the possibilities by measuring simply how the density of dark energy changed as the universe expanded.

--*Science* Sept. 2, 2005 Vol 309, 1482-1483.

OptIPuter

it2

AMR Cosmological Simulations Generate 4kx4k Images and Needs Interactive Zooming Capability



Source: Michael Norman, UCSD

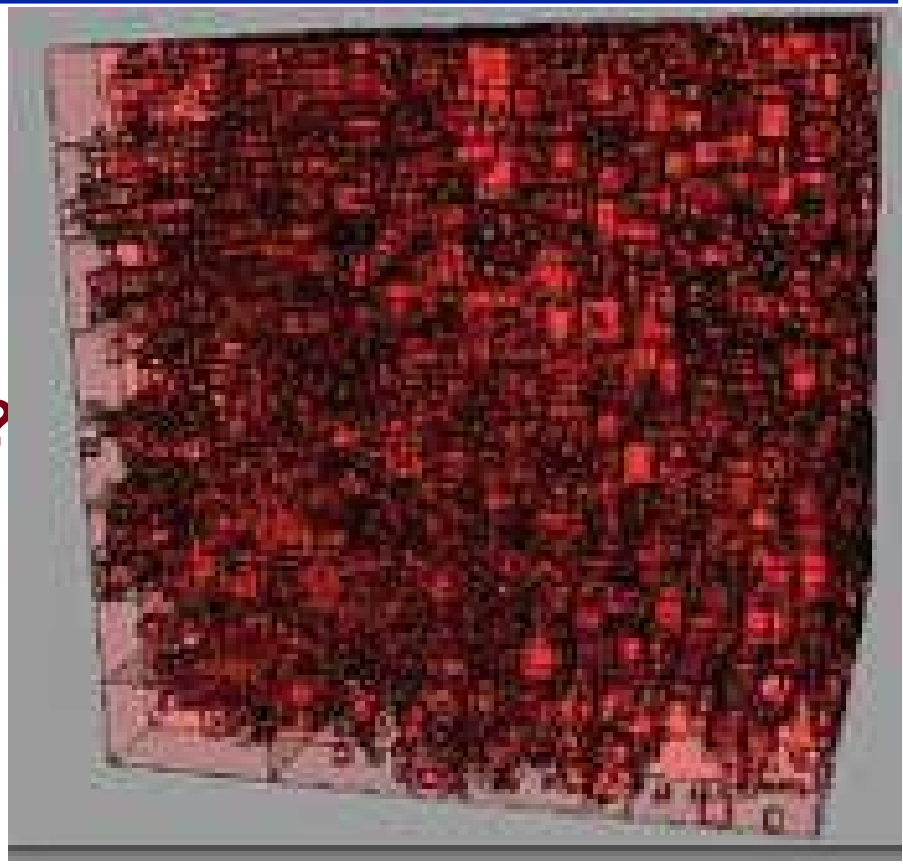
Why Does the Cosmic Simulator Need LambdaGrid Cyberinfrastructure?

- One Gigazone Uniform Grid or 512^3 AMR Run:
  - Generates ~10 TeraByte of Output
  - A "Snapshot" is 100s of GB
  - Need to Visually Analyze as We Create SpaceTimes
- Visual Analysis Daunting
  - Single Frame is About 8GB
  - A Smooth Animation of 1000 Frames is 1000 x 8 GB=8TB
  - Stage on Rotating Storage to High Res Displays
- Can Run Evolutions Faster than We can Archive Them
  - File Transport Over Shared Internet ~50 Mbit/s
    - 4 Hours to Move ONE Snapshot!
- AMR Runs Require Interactive Visualization Zooming Over 16,000x!

Source: Mike Norman, UCSD

Furthermore, Lambdas are Needed to Distribute the AMR Cosmology Simulations

- Uses ENZO Computational Cosmology Code
  - Grid-Based Adaptive Mesh Refinement Simulation Code
  - Developed by Mike Norman, UCSD
- Can One Distribute the Computing?
  - iGrid2005 to Chicago to Amsterdam
- Distributing Code Using Layer 3 Routers Fails
- Instead Using Layer 2, Essentially Same Performance as Running on Single Supercomputer
  - Using Dynamic Lightpath Provisioning



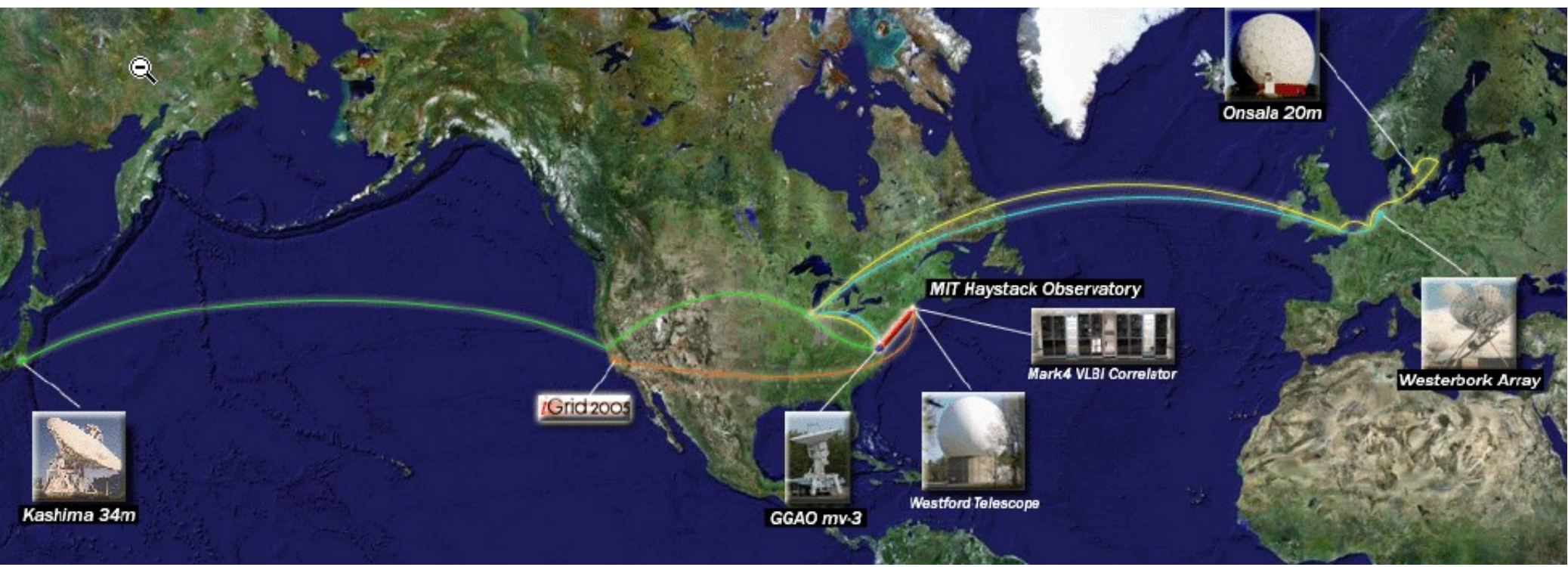
OptIPuter

Source: Joe Mambretti, Northwestern U

it2

Lambdas Enable Real-Time Very Long Baseline Interferometry

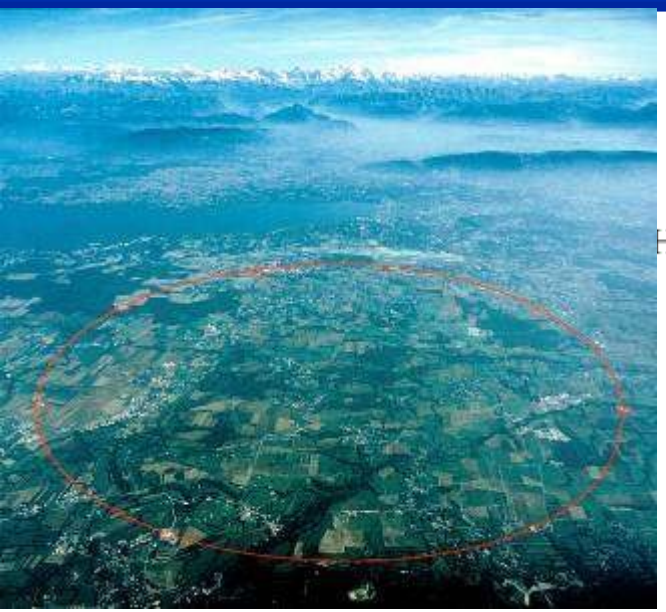
- From Tapes to Real-Time Data Flows
  - Three Telescopes (US, Sweden) Each Generating 0.5 Gbps Data Flow
  - Data Feeds Correlation Computer at MIT Haystack Observatory
  - Transmitted Live to iGrid2005
    - At SC05 will Add in Japan and Netherlands Telescopes
- In Future, e-VLBI Will Allow for Greater Sensitivity by Using 10 Gbps Flows



Global VLBI Network Used for Demonstration

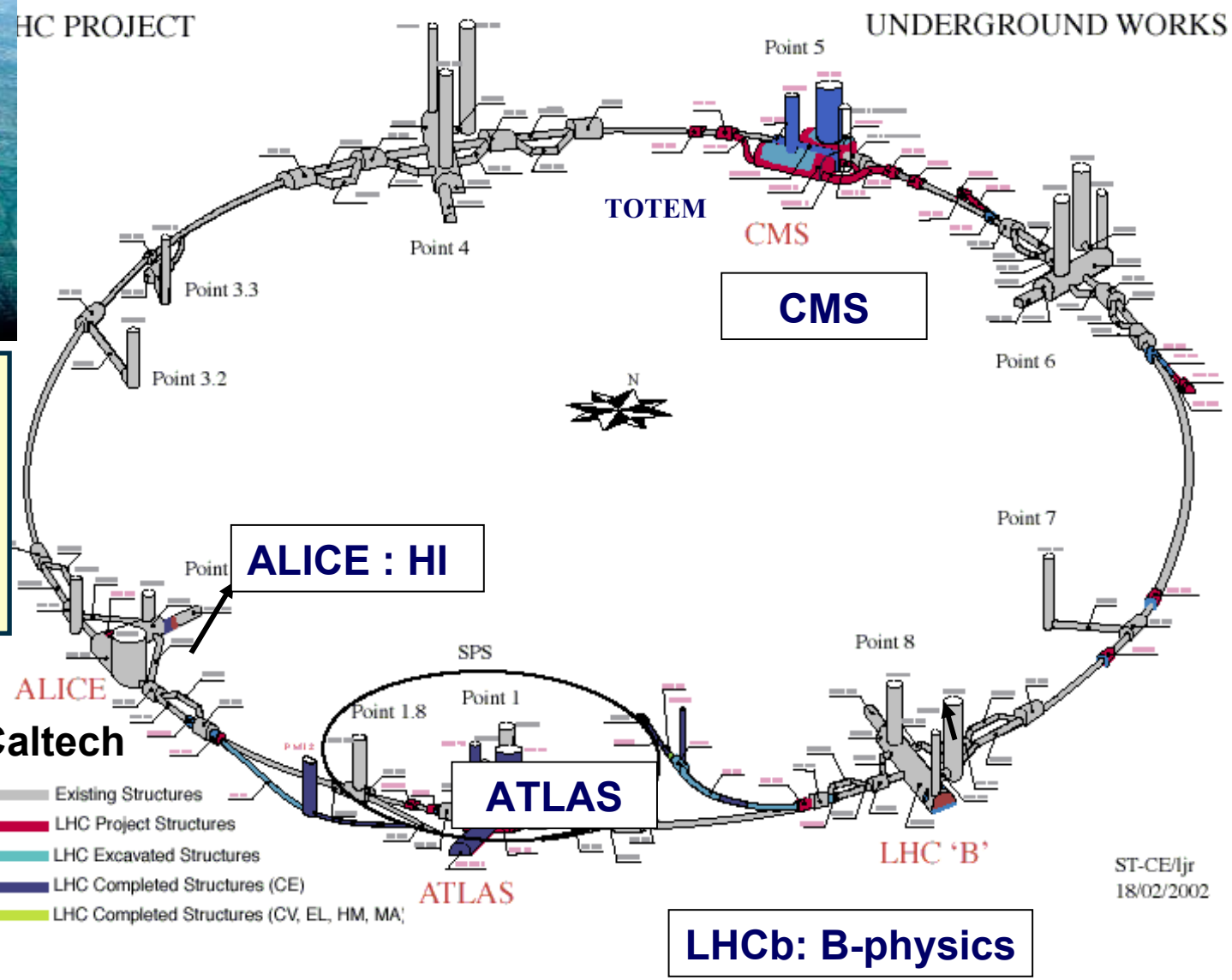
Source: MIT Haystack Observatory

Large Hadron Collider (LHC)
e-Science Driving Global Cyberinfrastructure



- pp $\sqrt{s}$ =14 TeV L=10^{34} cm^{-2} s^{-1}
- 27 km Tunnel in Switzerland & France

First Beams:
April 2007
Physics Runs:
from Summer 2007



Source: Harvey Newman, Caltech

OptIPuter

High Energy and Nuclear Physics
A Terabit/s WAN by 2010!

Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10	DWDM; 1 + 10 GigE Integration
2005	10	2-4 X 10	λ Switch; λ Provisioning
2007	2-4 X 10	~10 X 10; 40 Gbps	1st Gen. λ Grids
2009	~10 X 10 or 1-2 X 40	~5 X 40 or ~20-50 X 10	40 Gbps λ Switching
2011	~5 X 40 or ~20 X 10	~25 X 40 or ~100 X 10	2nd Gen λ Grids Terabit Networks
2013	~Terabit	~MultiTbps	~Fill One Fiber

Source: Harvey Newman, Caltech

Continuing the Trend: ~1000 Times Bandwidth Growth Per Decade; We are Rapidly Learning to Use Multi-Gbps Networks Dynamically

The Optical Core of the UCSD Campus-Scale Testbed -- Evaluating Packet Routing versus Lambda Switching

Goals by 2007:

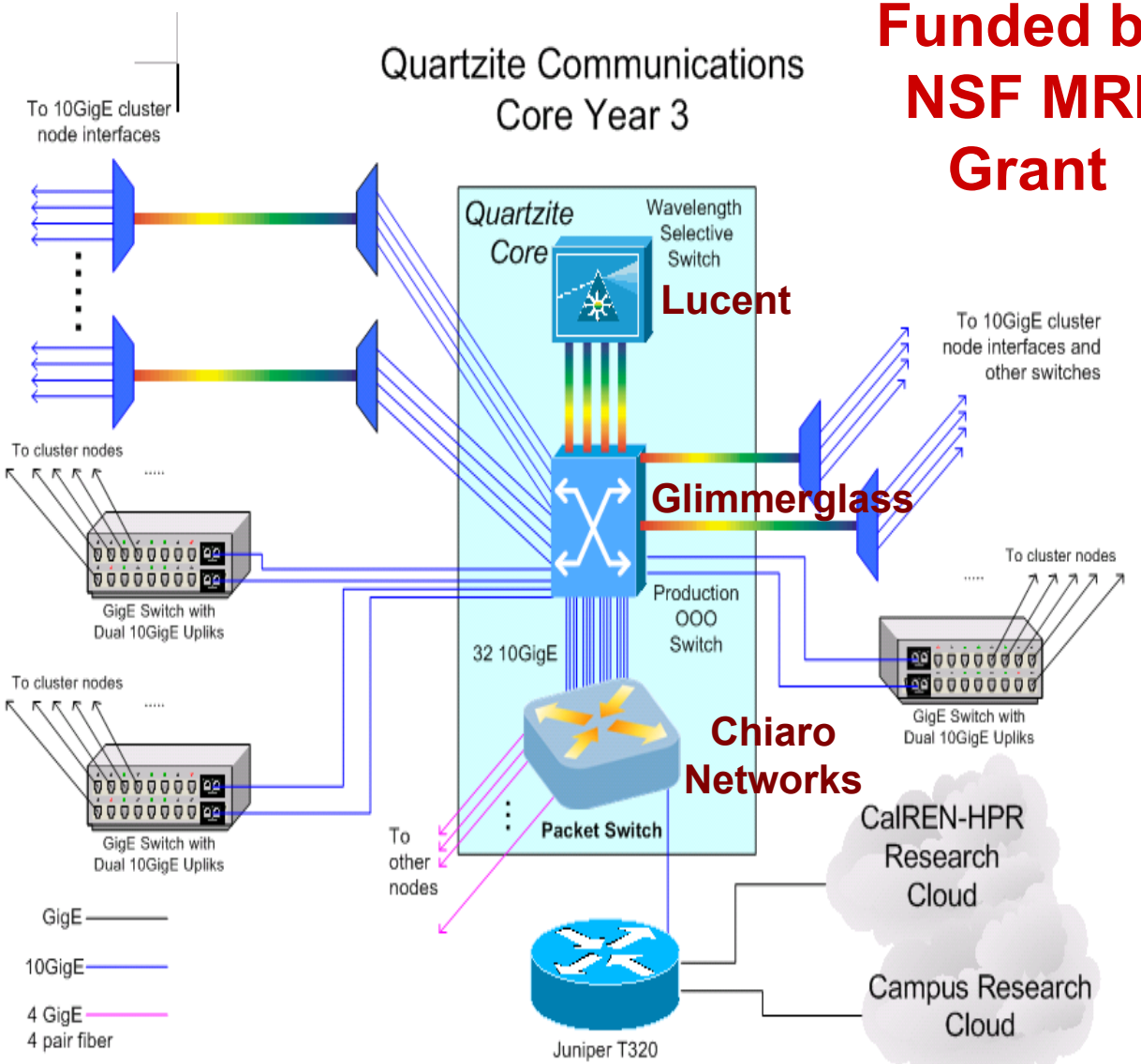
>= 50 endpoints at 10 GigE

>= 32 Packet switched

>= 32 Switched wavelengths

>= 300 Connected endpoints

Approximately 0.5 TBit/s Arrive at the "Optical" Center of Campus
Switching will be a Hybrid Combination of:
Packet, Lambda, Circuit -- OOO and Packet Switches Already in Place



Funded by NSF MRI Grant

Source: Phil Papadopoulos, SDSC, Calit2

Multiple HD Streams Over Lambdas Will Radically Transform Global Collaboration



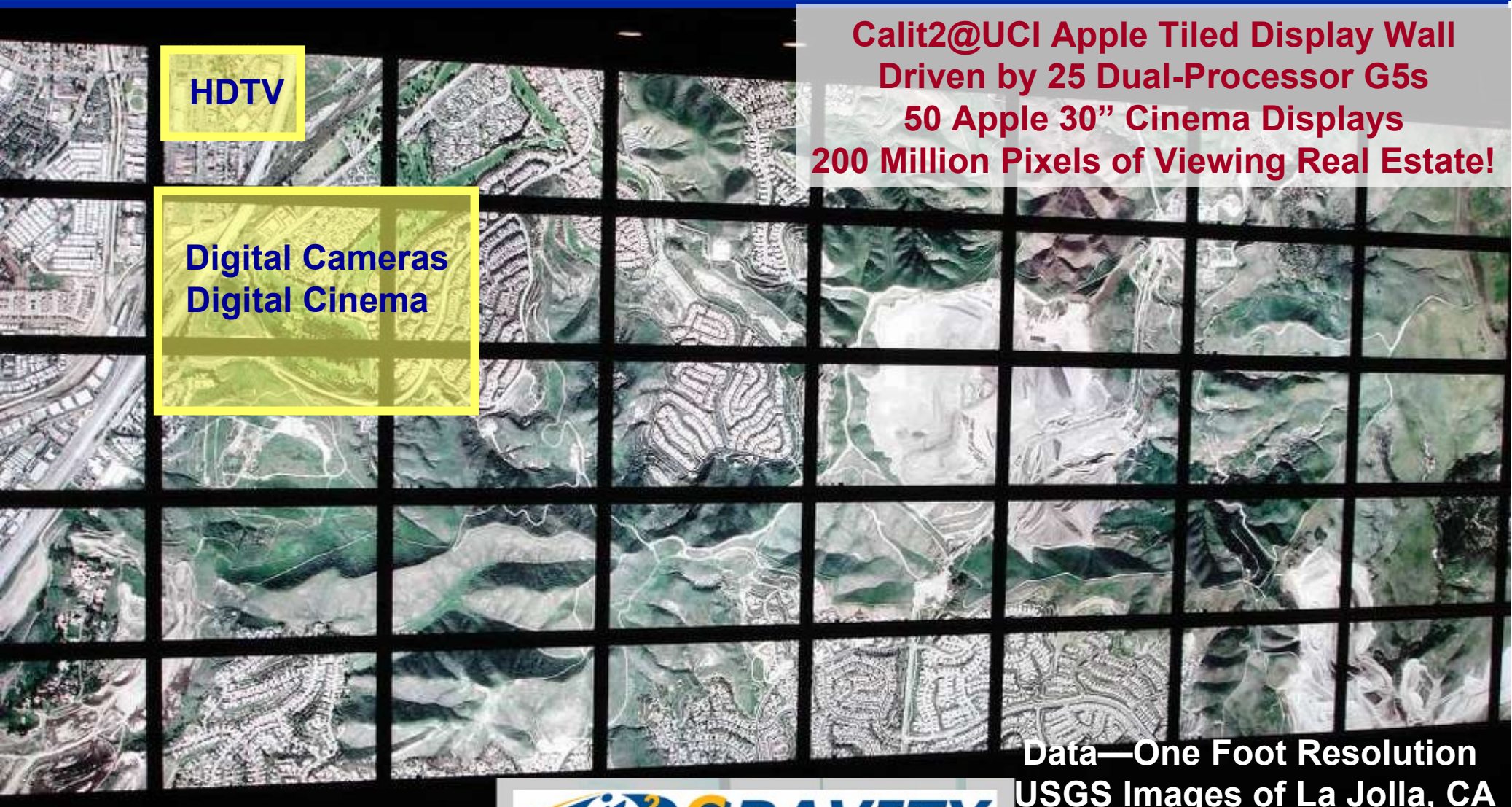


Telepresence Using Uncompressed 1.5 Gbps HDTV Streaming Over IP on Fiber Optics--
75x Home Cable "HDTV" Bandwidth!



Source: U Washington Research Channel

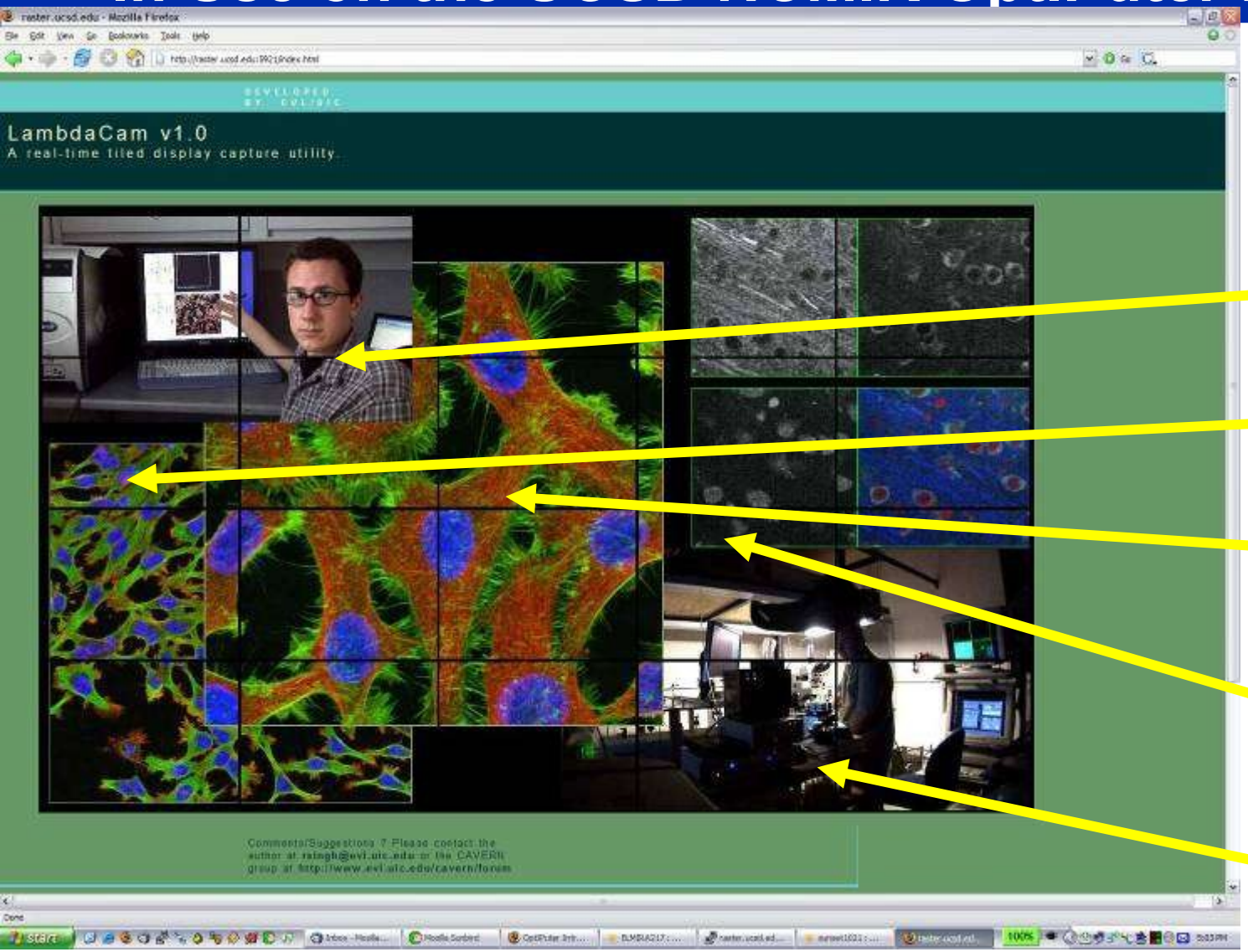
Largest Tiled Wall in the World Enables Integration of Streaming High Resolution Video



Source: Falko Kuester, Calit2@UCI
NSF Infrastructure Grant

OptIPuter Software Enables HD Collaborative Tiled Walls

In Use on the UCSD NCMIR OptIPuter Display Wall



- HD Video from BIRN Trailer
- Macro View of Montage Data
- Micro View of Montage Data
- Live Streaming Video of the RTS-2000 Microscope
- HD Video from the RTS Microscope Room

LambdaCam Used to Capture the Tiled Display on a Web Browser

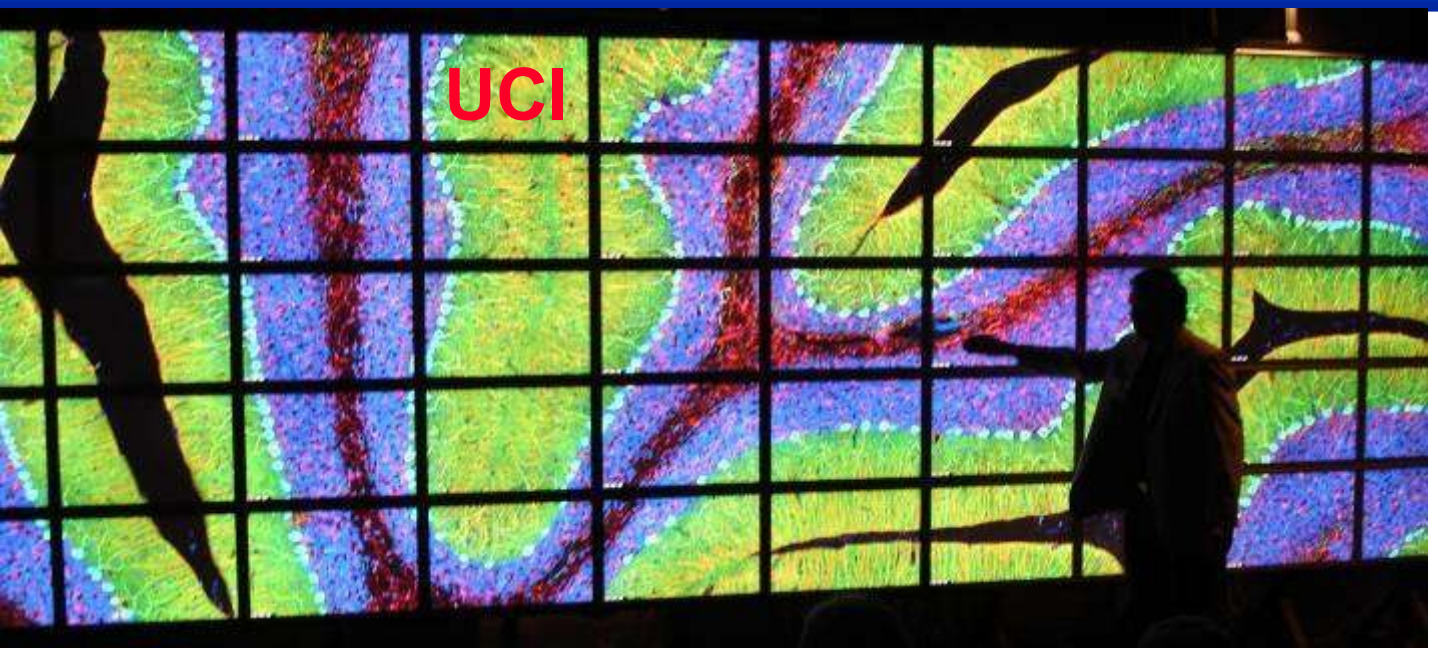
OptIPuter evl electronic visualization laboratory

Source: David Lee, NCMIR, UCSD

NCMIR NATIONAL CENTER for MICROSCOPY and IMAGING RESEARCH at San Diego, an NIH supported resource center

it2

The OptIPuter Enabled Collaboratory: Remote Researchers Jointly Exploring Complex Data



OptIPuter

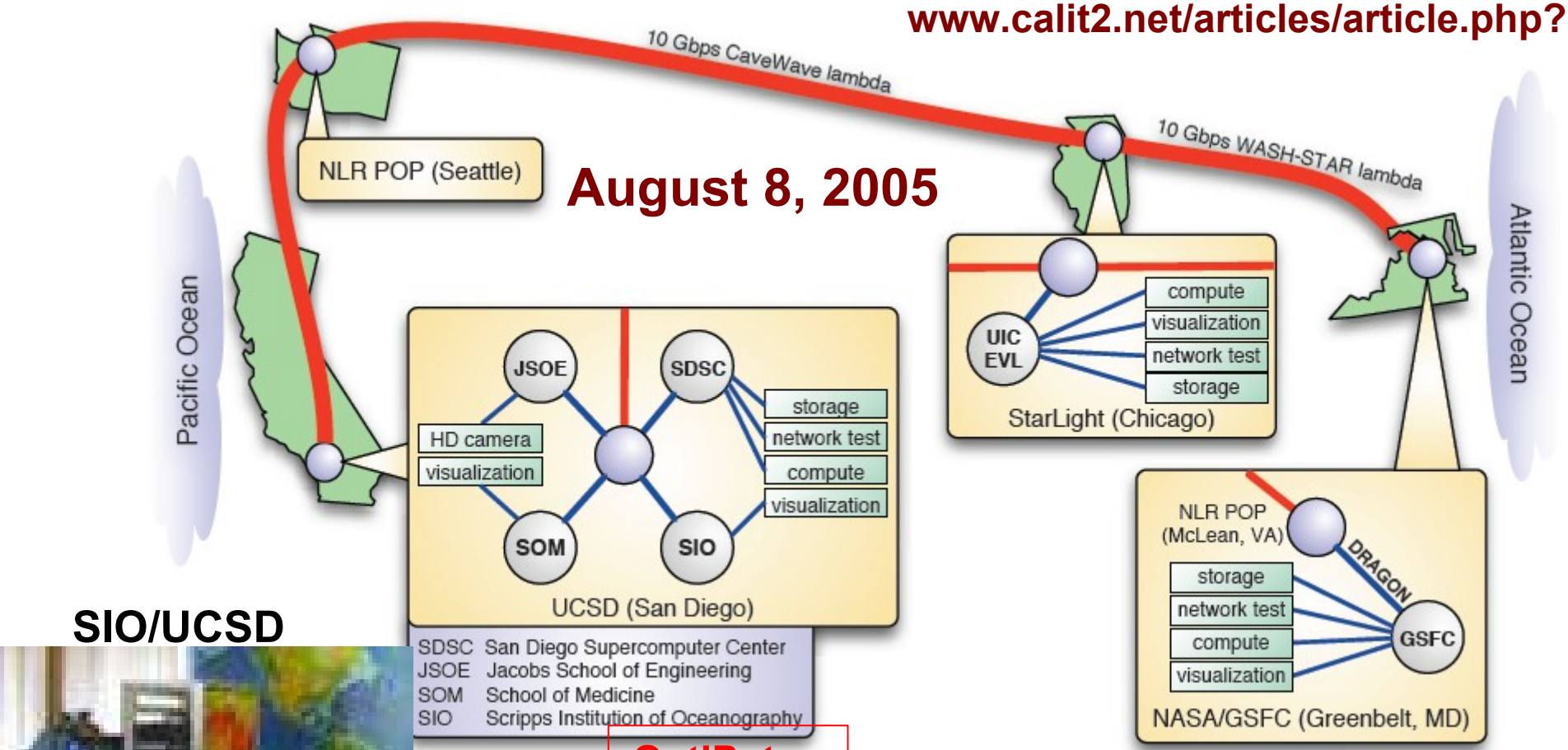
OptIPuter will Connect
The Calit2@UCI
200M-Pixel Wall
to
The Calit2@UCSD
100M-Pixel Display
With Shared Fast Deep Storage



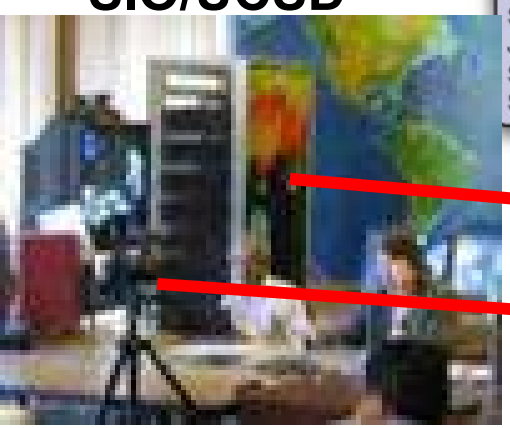
OptIPuter

Combining Telepresence with Remote Interactive Analysis of Data Over NLR

www.calit2.net/articles/article.php?id=660



SIO/UCSD



OptIPuter Visualized Data

HDTV Over Lambda

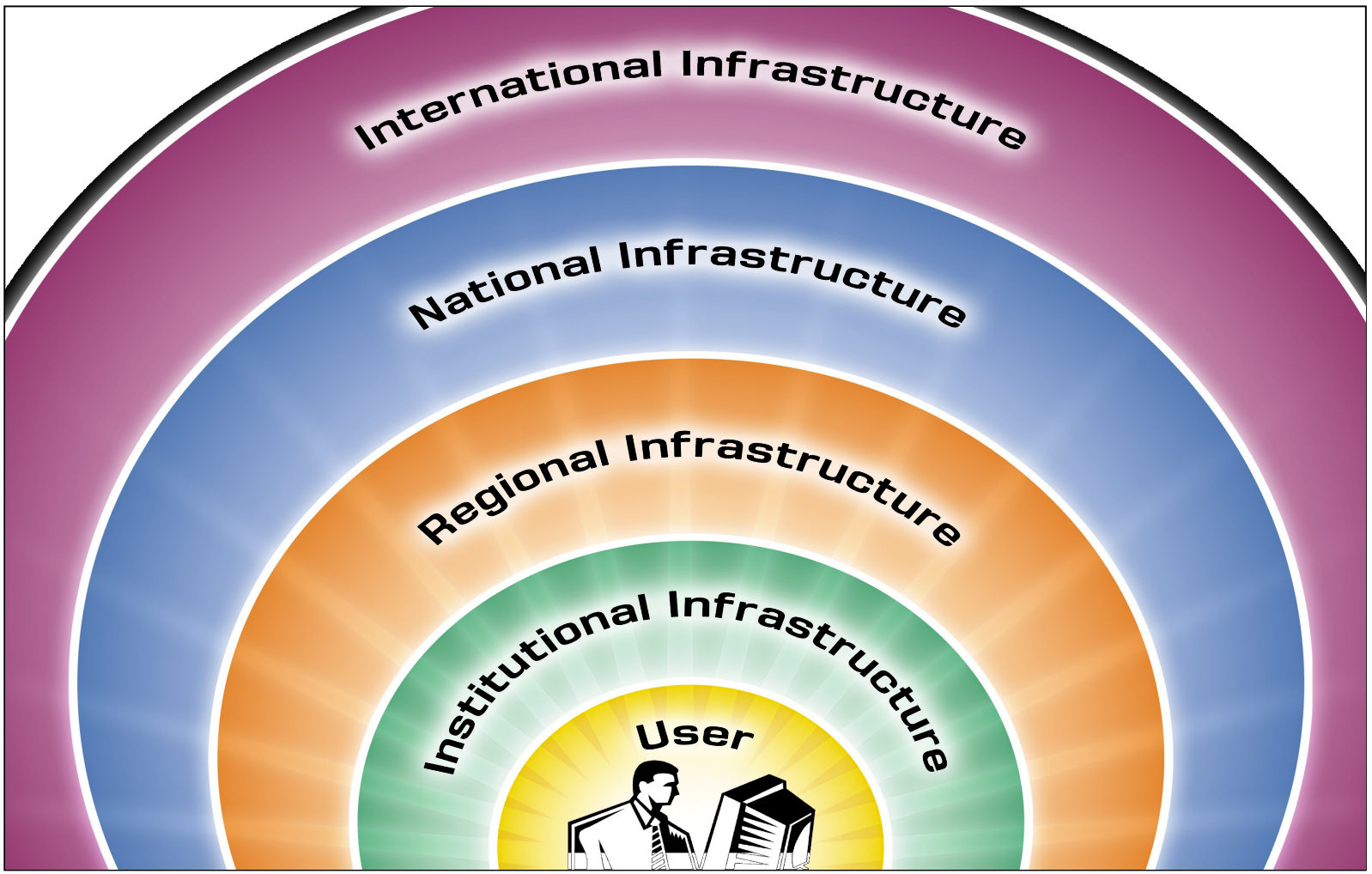


NASA Goddard

OptIPuter

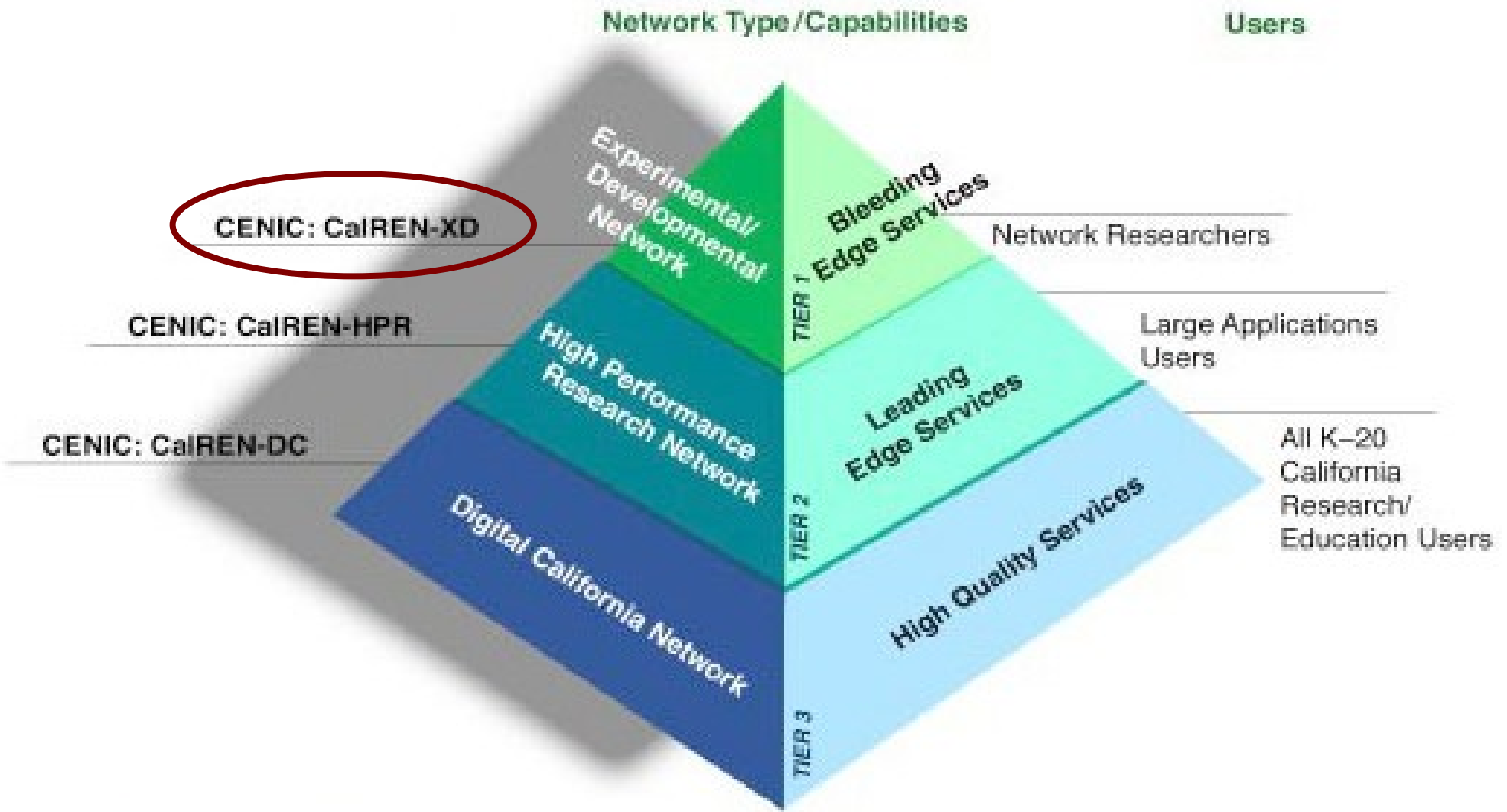
it2

Optical Network Infrastructure Framework Needs to Start with the User and Work Outward



Tom West, NLR

California's CENIC/CalREN Has Three Tiers of Service

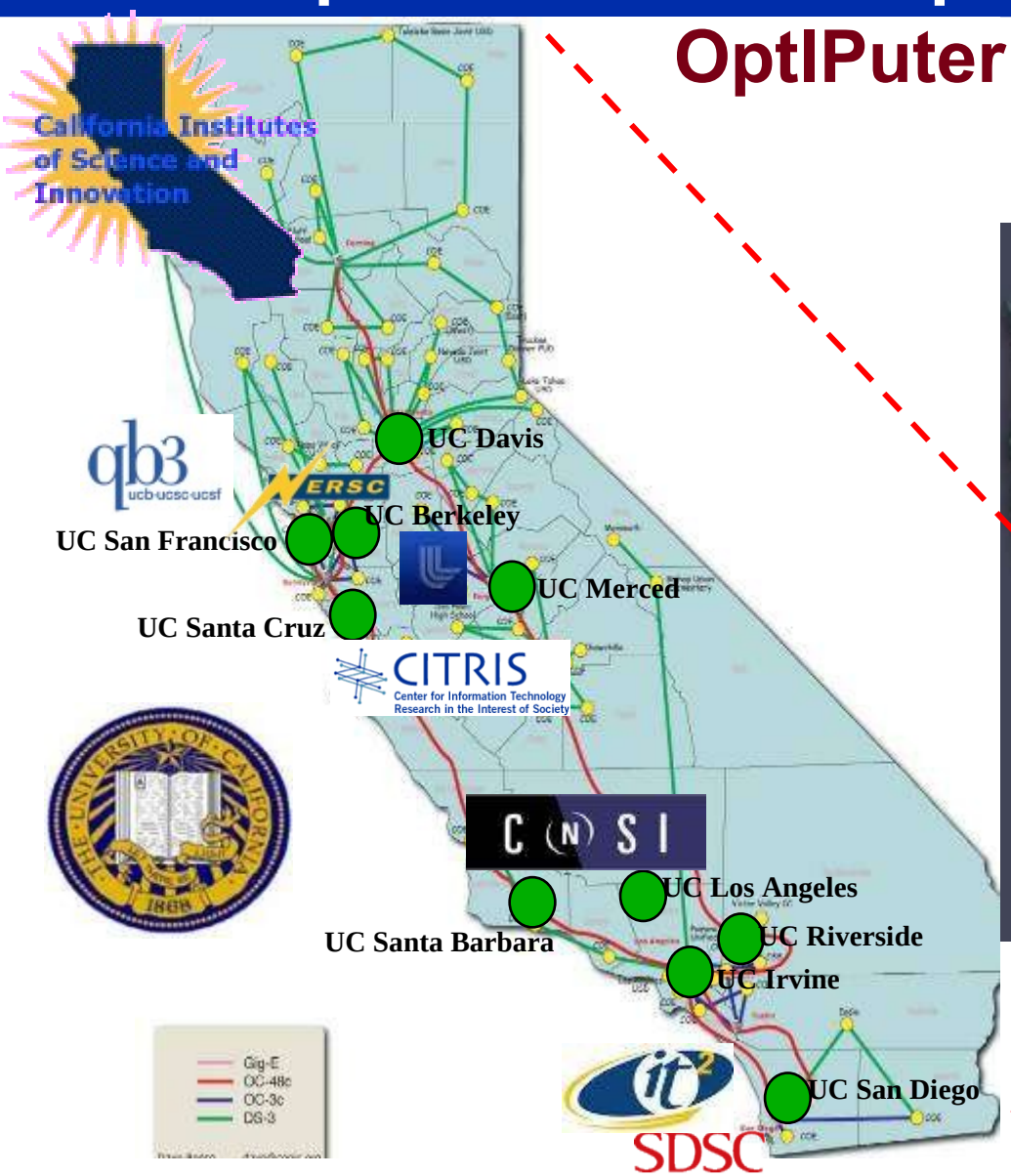


OptIPuter

it2

Calit2/SDSC Proposal to Create a UC Cyberinfrastructure of OptIPuter "On-Ramps" to TeraGrid Resources

OptIPuter + CalREN-XD + TeraGrid = "OptiGrid"





Creating a Critical Mass of End Users on a Secure LambdaGrid

Source: Fran Berman, SDSC