"An End-to-End Campus-Scale High Performance Cyberinfrastructure for Data-Intensive Research"

The Annual Robert Stewart Distinguished Lecture

Iowa State University

Ames, Iowa

April 19, 2012

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
http://lsmarr.calit2.net

Abstract

Campuses are experiencing an enormous increase in the quantity of data generated by scientific instruments and computational clusters. The shared Internet, engineered to enable interaction with megabyte-sized data objects is not capable of dealing with the typical gigabytes to terabytes of modern scientific data. Instead, a high performance end-to-end cyberinfrastructure built on 10,000 Mbps optical fibers is emerging to support data-intensive research. I will give examples of early prototypes which integrate scalable data generation, transmission, storage, analysis, visualization, and sharing, driven by applications as diverse as genomics, medical imaging, cultural analytics, earth sciences, and cosmology.

The Data-Intensive Discovery Era Requires High Performance Cyberinfrastructure

- Growth of Digital Data is Exponential
  - "Data Tsunami"
- Driven by Advances in Digital Detectors, Computing, Networking, & Storage Technologies
- Shared Internet Optimized for Megabyte-Size Objects
- Need Dedicated Photonic Cyberinfrastructure for Gigabyte/Terabyte Data Objects
- Finding Patterns in the Data is the New Imperative
  - Data-Driven Applications
  - Data Mining
  - Visual Analytics
  - Data Analysis Workflows



SDSC

Source: SDSC

it²

News

White House launches big data R&D effort

NSF and DARPA are among the agencies looking to harness the power of large volumes of data

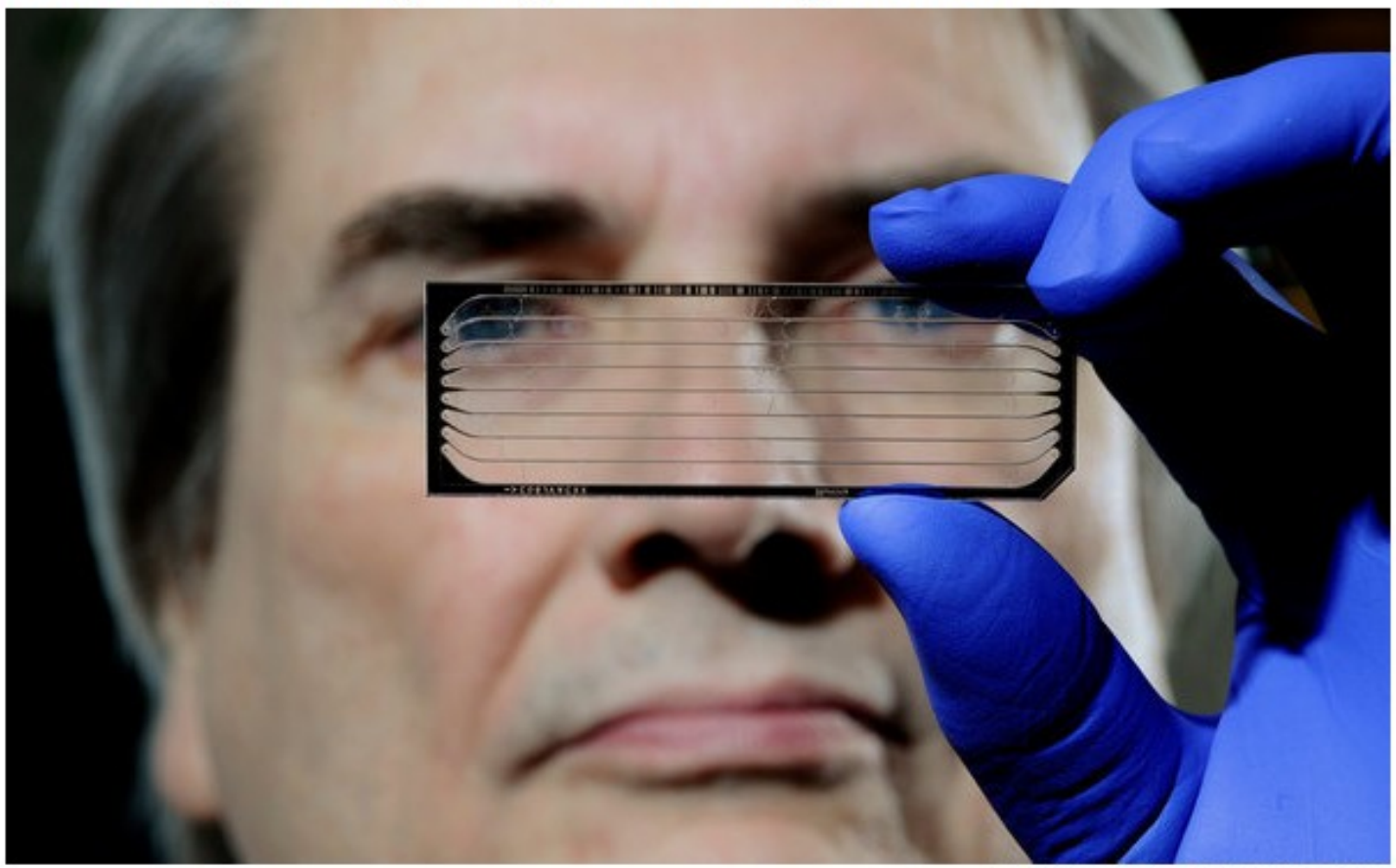
By Grant Gross

March 29, 2012 12:34 PM ET

5 Comments + Briefcase What's this?

Genomic Sequencing is Driving Big Data

DNA Sequencing Caught in Deluge of Data



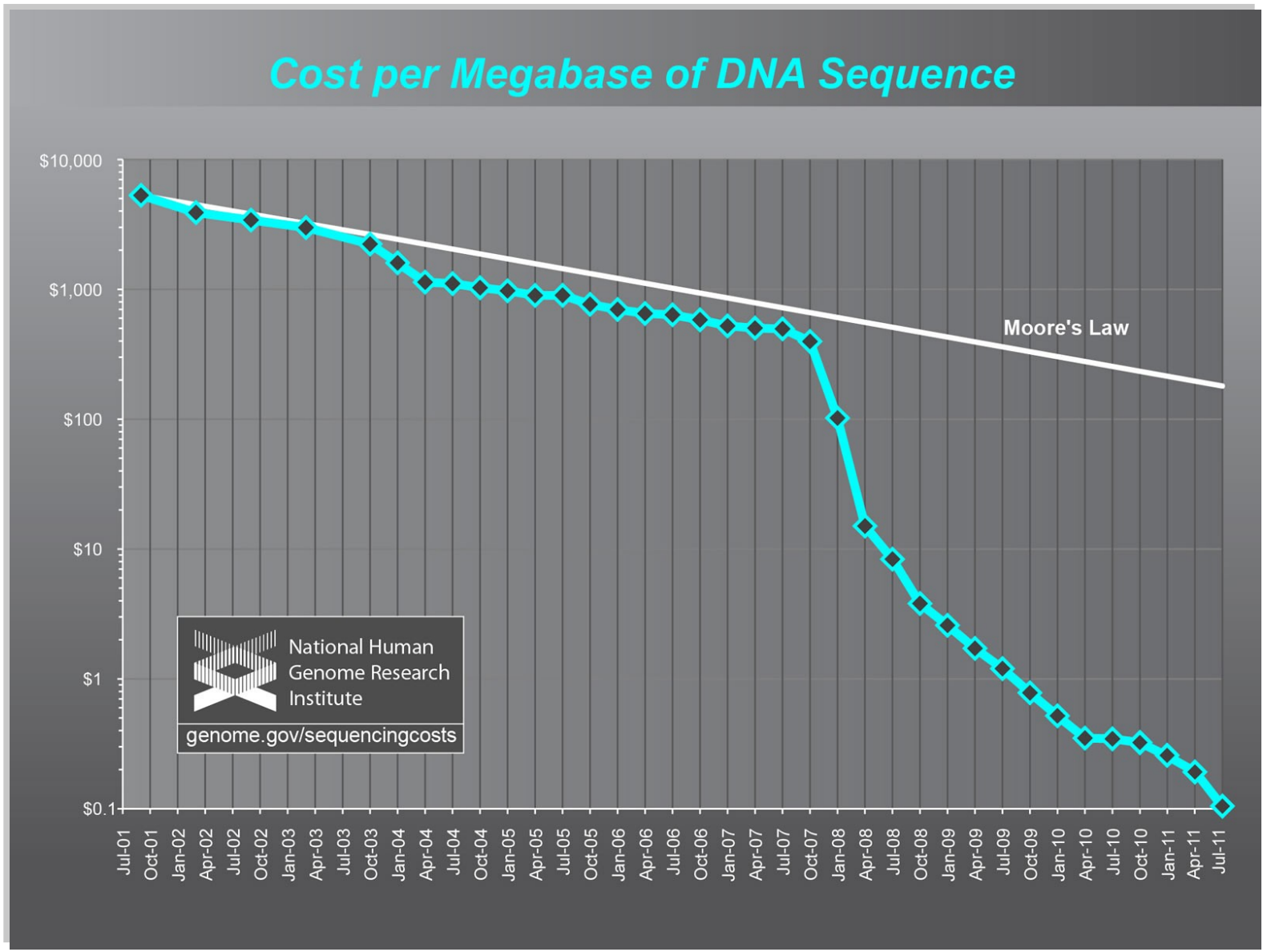
Kathy Kmonicek for The New York Times

The New York Times

Business Day

November 30, 2011

Cost Per Megabase in Sequencing DNA is Falling Much Faster Than Moore's Law



www.genome.gov/sequencingcosts/

BGI—The Beijing Genome Institute is the World's Largest Genomic Institute





- **Main Facilities in Shenzhen and Hong Kong, China**
  - **Branch Facilities in Copenhagen, Boston, UC Davis**
- **137 Illumina HiSeq 2000 Next Generation Sequencing Systems**
  - **Each Illumina Next Gen Sequencer Generates 25 Gigabases/Day**
- **Supported by High Performance Computing and Storage**
  - **~160TF, 33TB Memory**
  - **Large-Scale (12PB) Storage**

From 10,000 Human Genomes Sequenced in 2011 to 1 Million by 2015 in Less Than 5,000 sq. ft.!

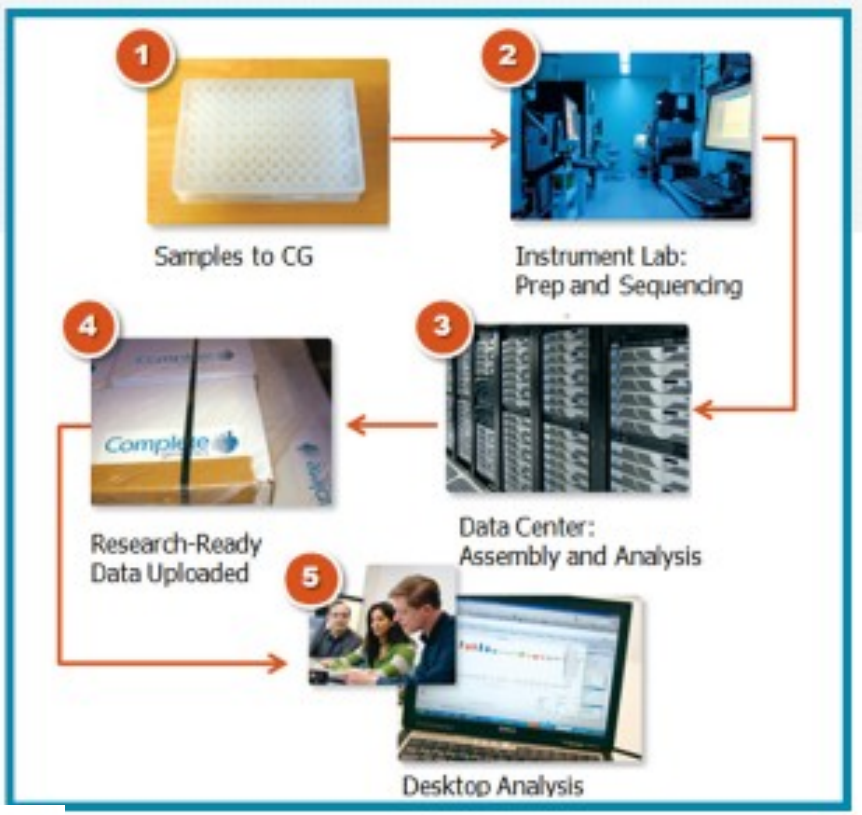
Proprietary Sequencing Technology

Designed and Optimized for Whole Human Genome Sequencing **Quality**, **Cost** and **Scale**



Outsourced Business Model

A Turnkey Service Enabling Customers to **Outsource** WGS – Samples In, **Research Ready Data** Out



Complete genomics

4 Million Newborns / Year in U.S.

Needed: Interdisciplinary Teams Made From Computer Science, Data Analytics, and Genomics

The $1,000 Genome, The $1M Interpretation

Dr. Kevin Davies



Dr. Kevin Davies is the Editor in Chief at Bio-IT World. He will be presenting *The $1,000 Genome, The $1,000,000 Interpretation.*

Copenhagenomics

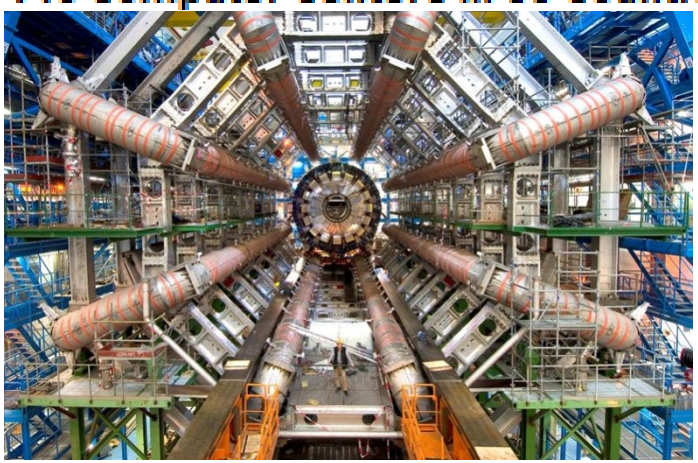
CPH$_x$ Genomics Conference on Improving Life • June 14-15, 2012

The Large Hadron Collider Uses a Global Fiber Infrastructure To Connect Its Users

CERN Unveils Global Grid For Particle Physics Research

The network can pull in the IT power of more than 140 computer centers in 33 countries to crunch an expected 15 million GB of data every year.

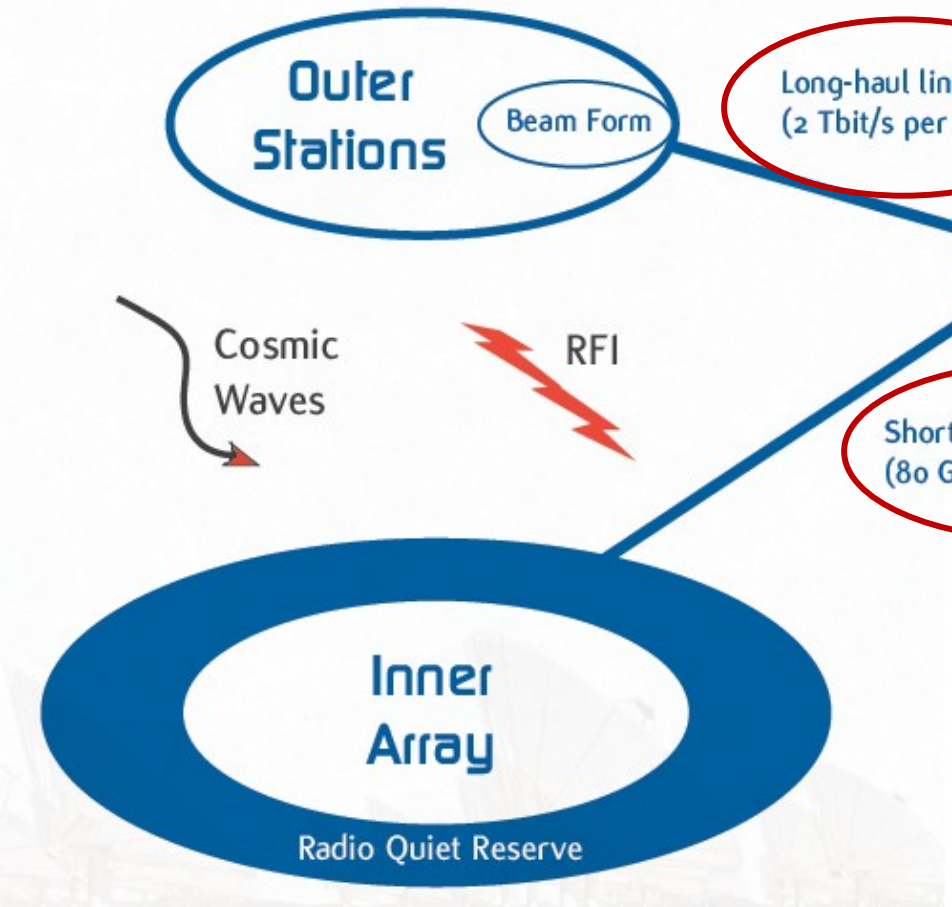
By Antone Gonsalves
InformationWeek
October 3, 2008 04:57 PM



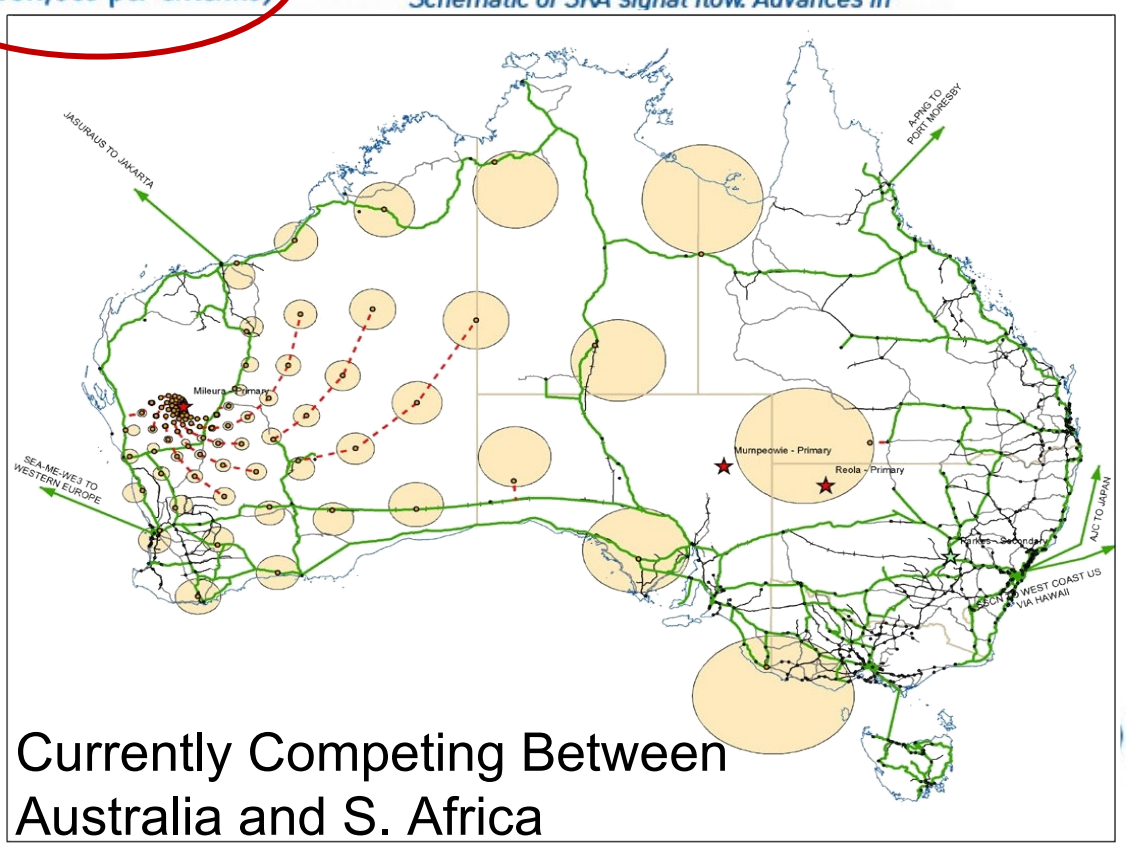
- The grid relies on optical fiber networks to distribute data from CERN to 11 major computer centers in Europe, North America, and Asia
- The grid is capable of routinely processing 250,000 jobs a day
- The data flow will be ~6 Gigabits/sec or 15 million gigabytes a year for 10 to 15 years

Next Great Planetary Instrument: The Square Kilometer Array Requires Dedicated Fiber

www.skatelescope.org



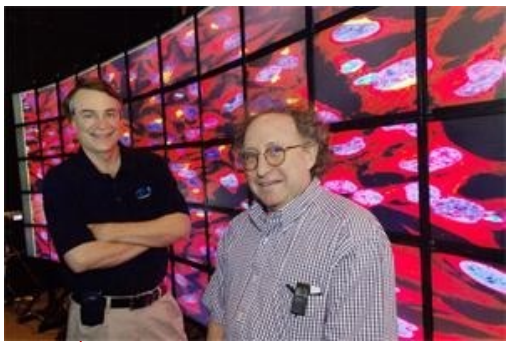
Schematic of SKA signal flow. Advances in



Transfers Of
1 TByte Images
World-wide
Will Be Needed
Every Minute!

Currently Competing Between
Australia and S. Africa

A Big Data Global Collaboratory Built on a 10Gbps "End-to-End" Lightpath Cloud



End User OptIPortal





Local or Remote Instruments



HD/4k Live Video

HPC

TeraGrid





10G Lightpaths

Campus Optical Switch

amazon web services



Data Repositories & Clusters

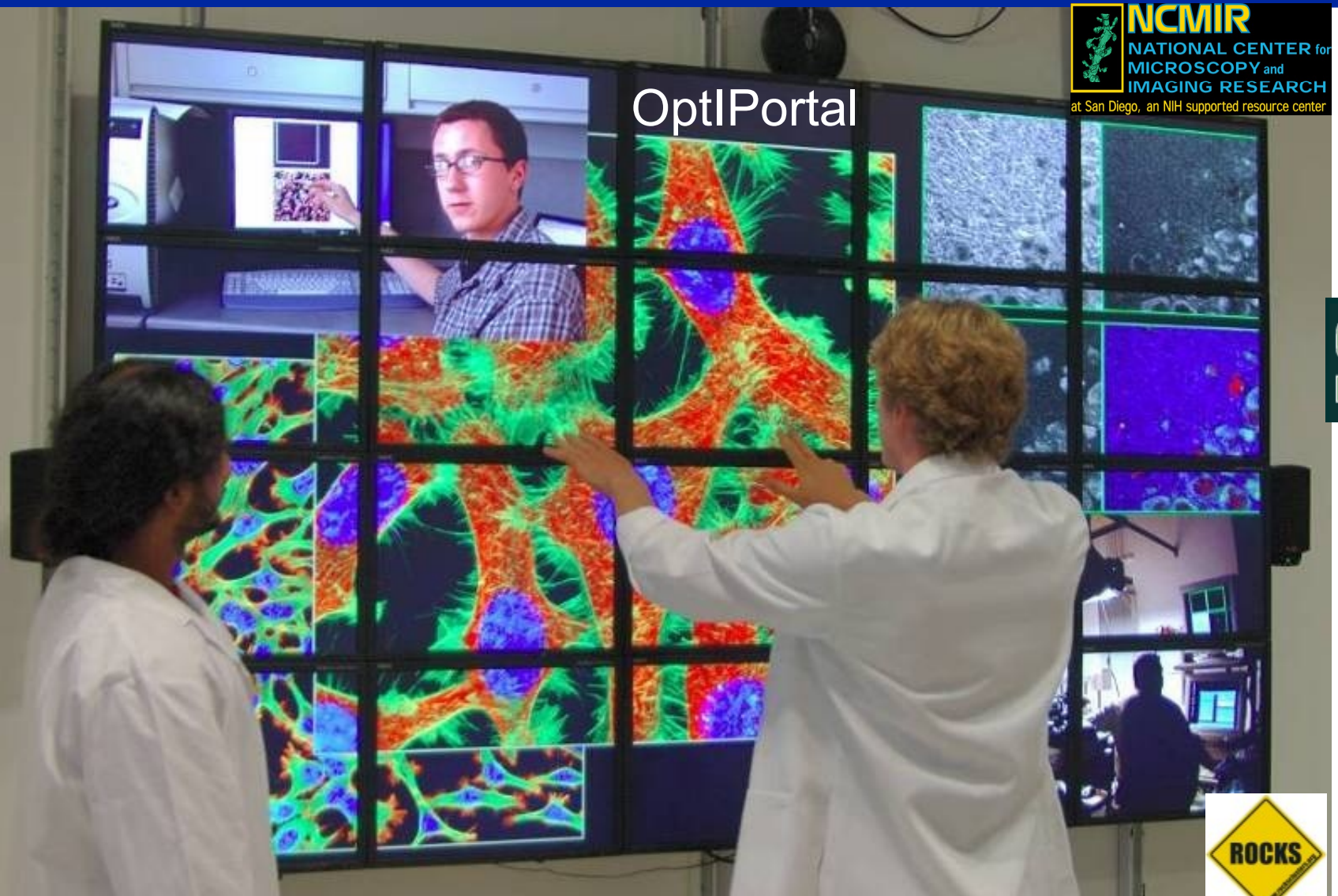


HD/4k Video Repositories

SDSC

it²

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



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

Scalable Adaptive Graphics Environment (SAGE)

SAGE

evl electronic visualization laboratory

NSF

ROCKS

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

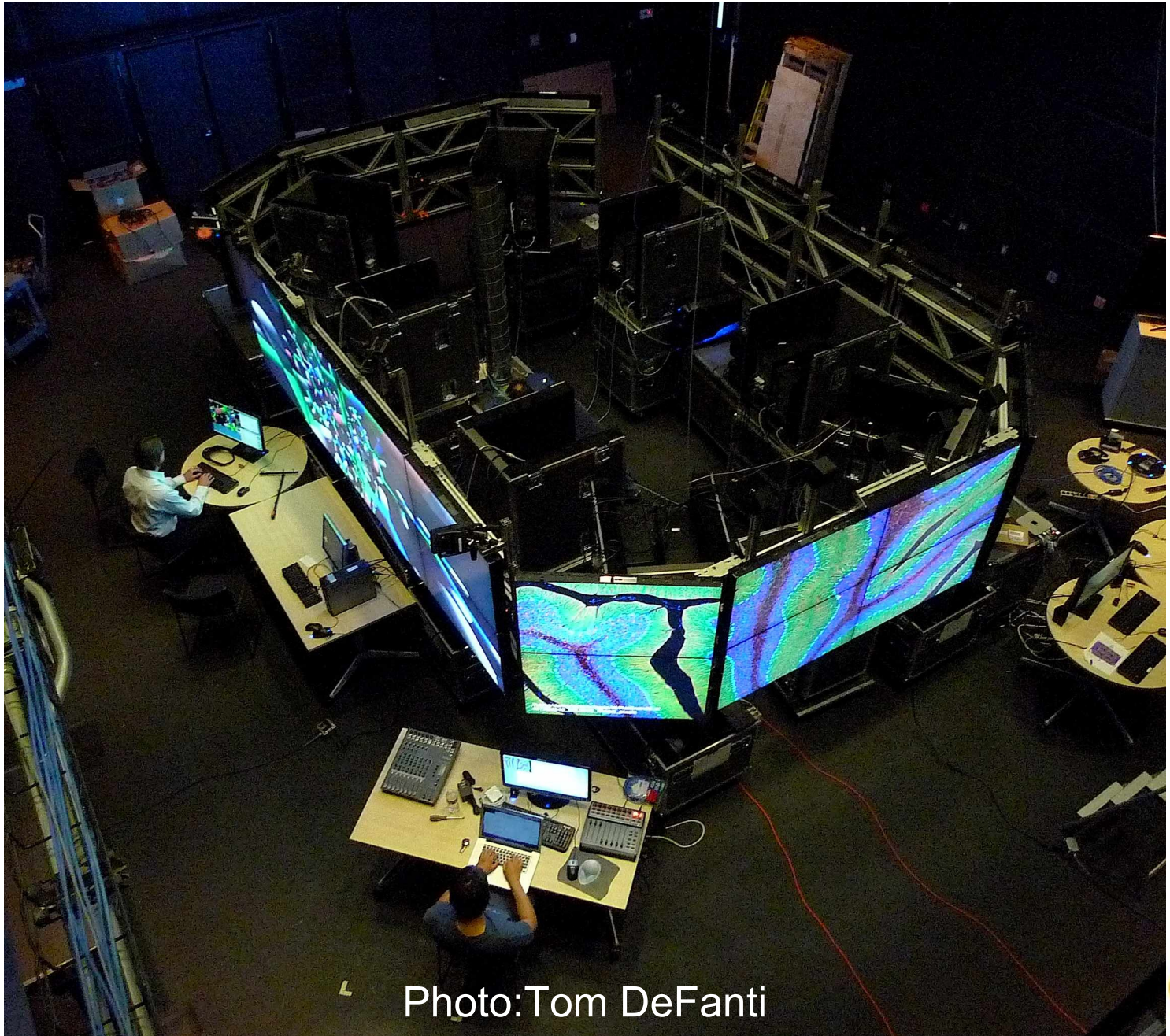
The Latest OptIPuter Innovation: Quickly Deployable Nearly Seamless OptIPortables

45 minute setup, 15 minute tear-down with two people (possible with one)



Image From the Calit2 KAUST Lab

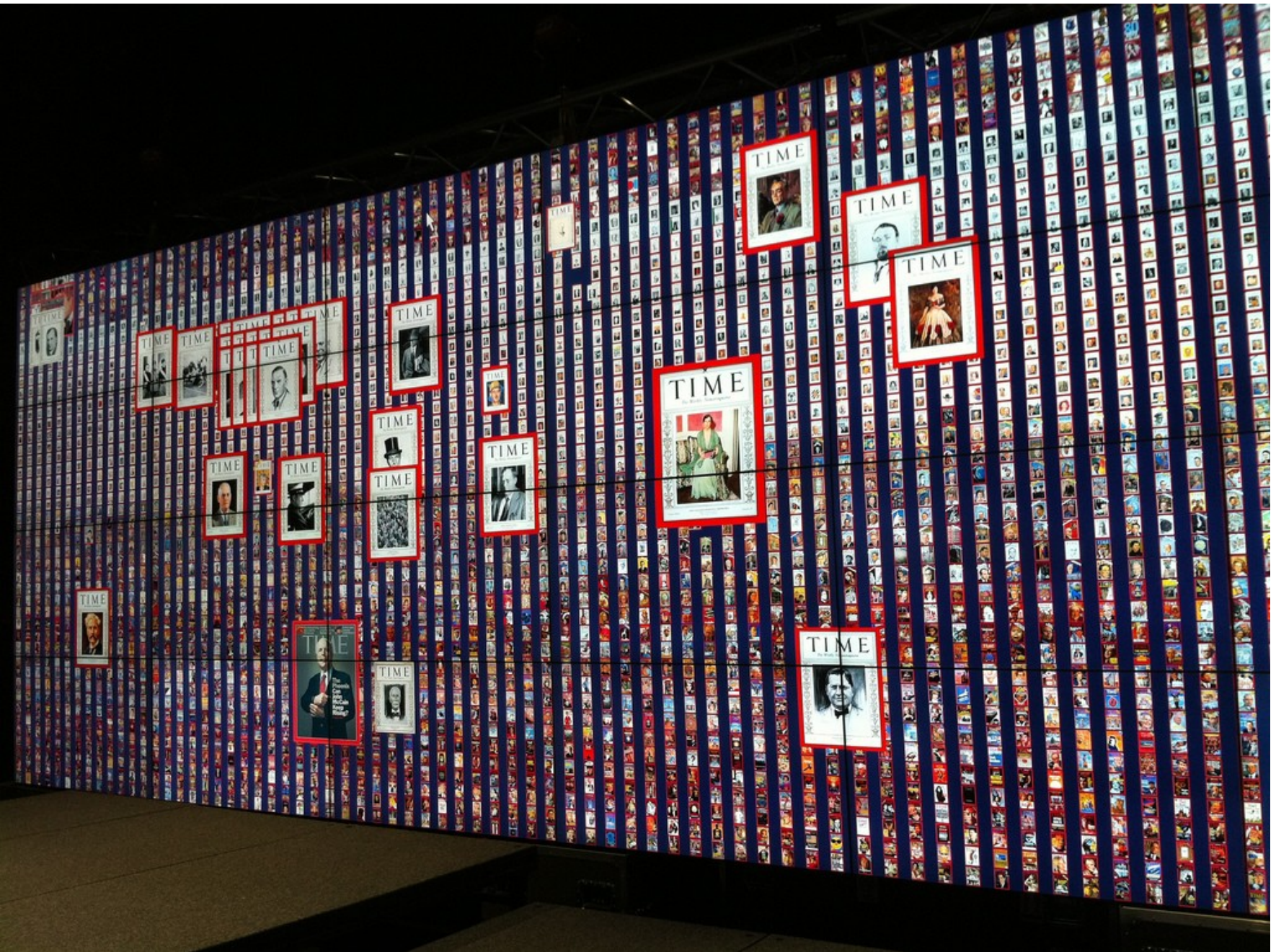
The OctIPortable Being Checked Out Prior to Shipping to the Calit2/KAUST Booth at SIGGRAPH 2011



Hubble Space Telescope Collage of 48 Frames (30,000x 14,000 pixels) on Calit2's Vroom

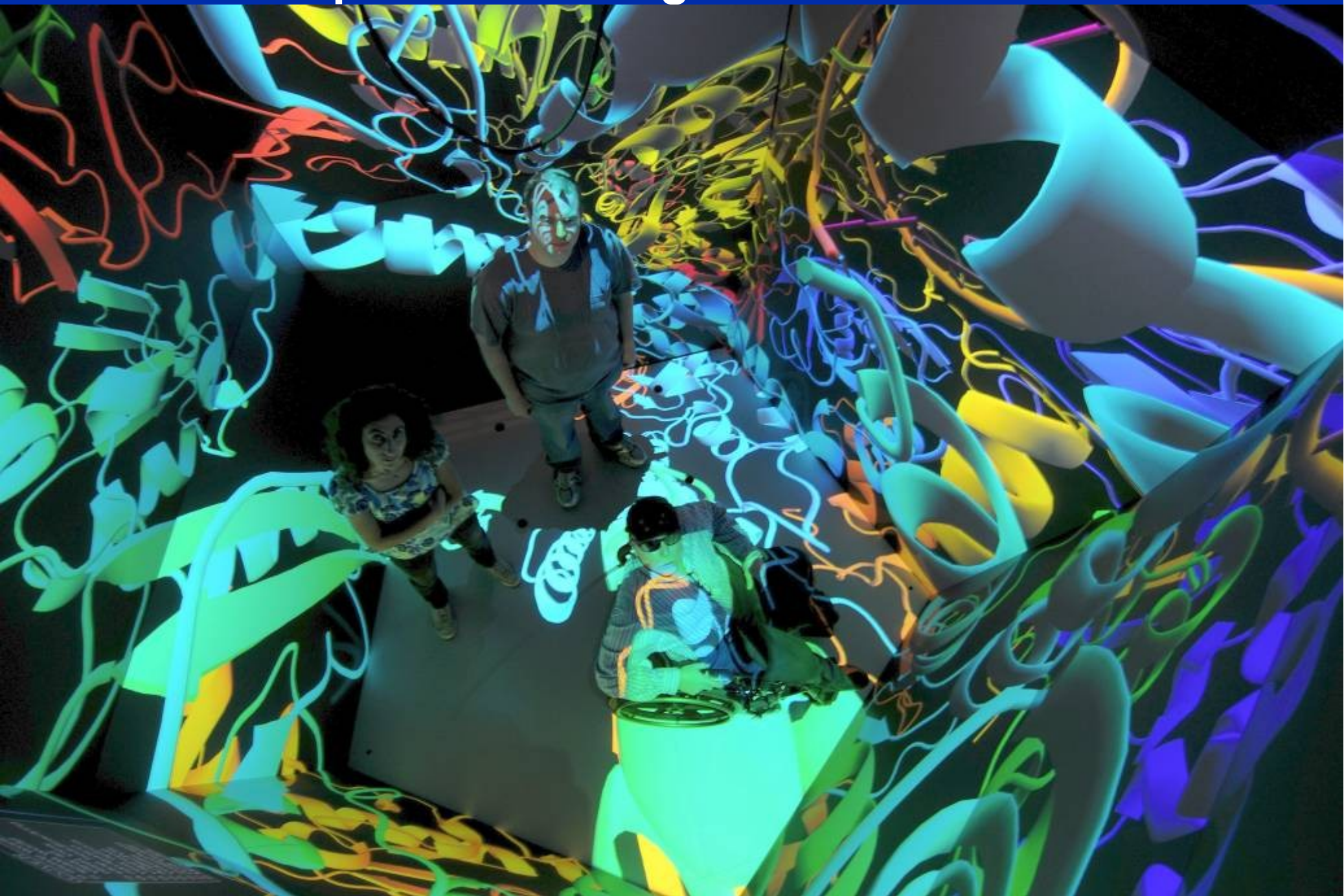


Scalable Cultural Analytics: 4535 Time magazine covers (1923-2009)



Source: Software Studies Initiative, Prof. Lev Manovich, UCSD

Calit2 3D Immersive StarCAVE OptIPortal: Enables Exploration of High Resolution Simulations



3D Stereo Head Tracked OptIPortal: NexCAVE



Array of JVC HDTV 3D LCD Screens
KAUST NexCAVE = 22.5MPixels

www.calit2.net/newsroom/article.php?id=1584

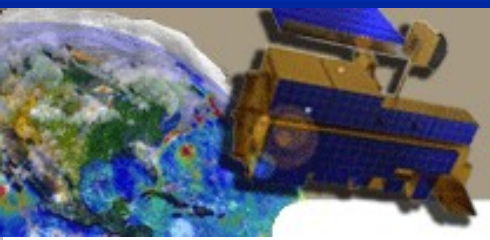
Source: Tom DeFanti, Calit2@UCSD

TourCAVE



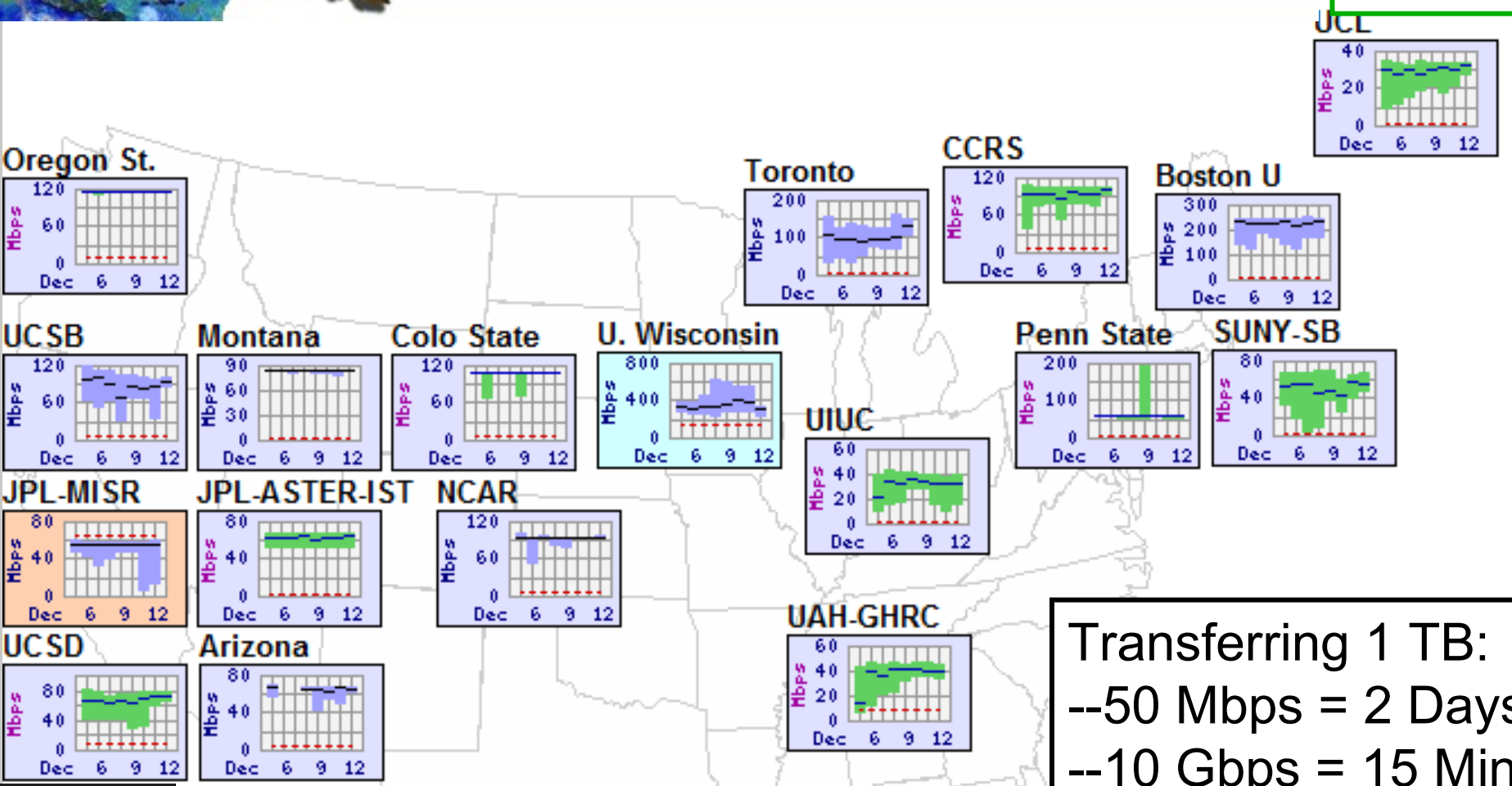
Five 65” LG 3D HDTVs, PC, Tracker--~$33,000

Large Data Challenge: Average Throughput to End User on Shared Internet is 10-100 Mbps



ENSIGHT Active Testing TERRA Destinations

Tested December 2011



Transferring 1 TB:
--50 Mbps = 2 Days
--10 Gbps = 15 Minutes

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

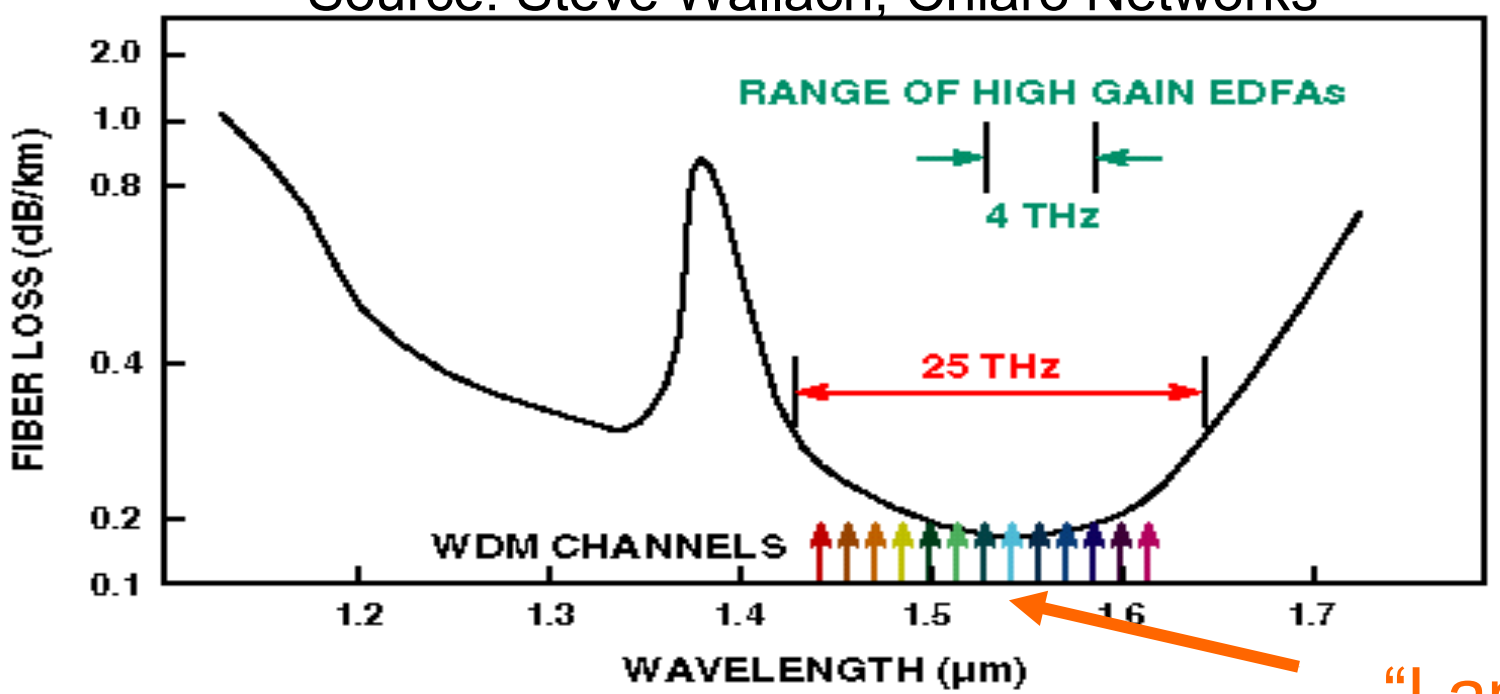
OptIPuter Solution:
Give Dedicated Optical Channels to Data-Intensive Users

WAVELENGTH DIVISION MULTIPLEXING (WDM)

10 Gbps per User ~ 100x
Shared Internet Throughput

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

Source: Steve Wallach, Chiaro Networks

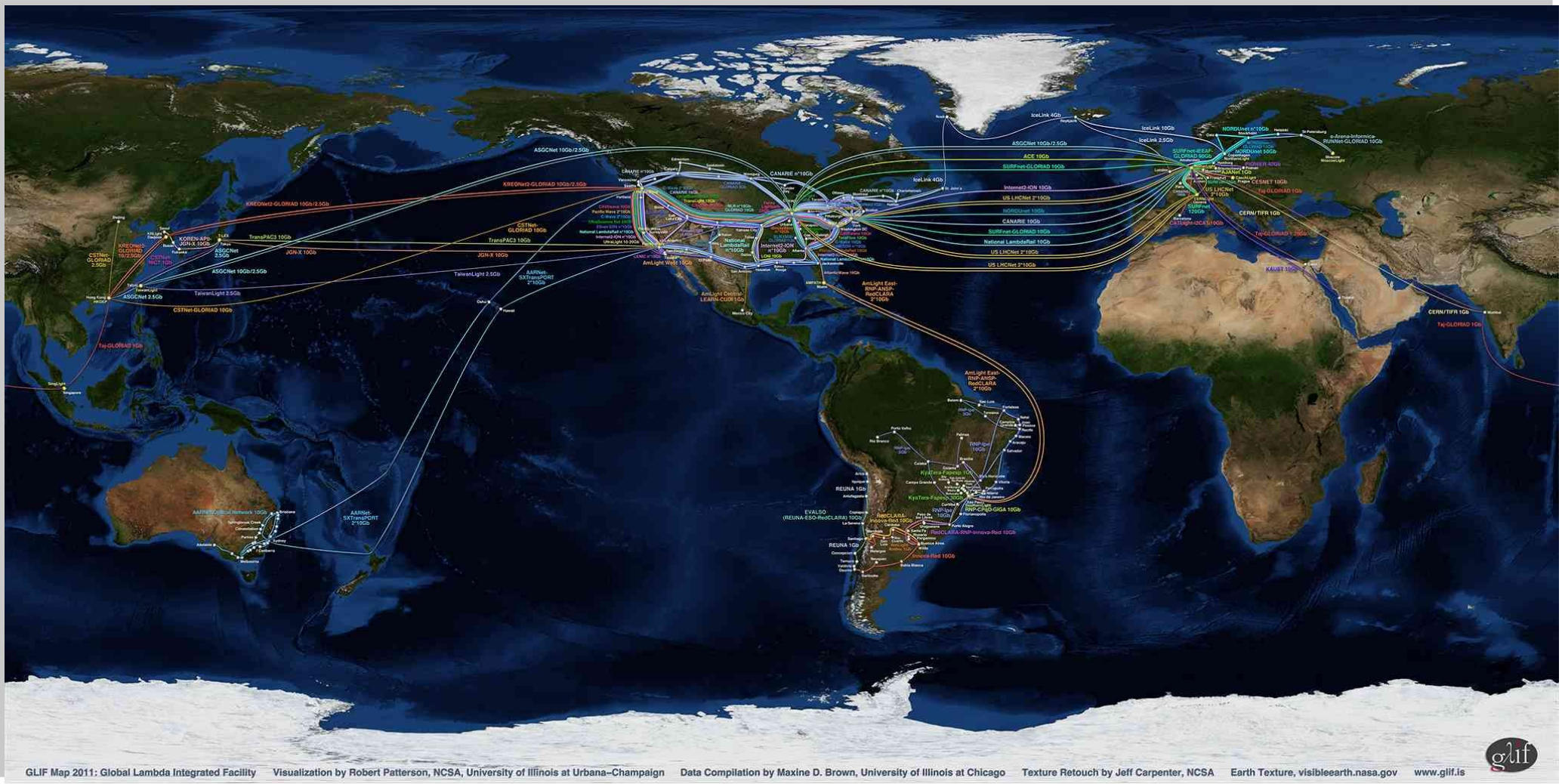


"Lambdas"

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

The Global Lambda Integrated Facility--
Creating a Planetary-Scale High Bandwidth Collaboratory

Research Innovation Labs Linked by 10G Dedicated Lambdas



www.glif.is/publications/maps/GLIF_5-11_World_2k.jpg

High Definition Video Connected OptIPortals: Virtual Working Spaces for Data Intensive Research



NASA Supports Two Virtual Institutes

CGLX

LifeSize HD

Calit2@UCSD 10Gbps Link to
NASA Ames Lunar Science Institute, Mountain View, CA

OptIPuter

Source: Falko Kuester, Kai Doerr Calit2;
Michael Sims, Larry Edwards, Estelle Dodson NASA

it2

Launch of the 100 Megapixel OzIPortal Kicked Off a Rapid Build Out of Australian OptIPortals



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

Covise, Phil Weber, Jurgen Schulze, Calit2
CGLX, Kai-Uwe Doerr , Calit2
http://www.calit2.net/newsroom/release.php?id=1421

Prototyping Next Generation User Access and Large Data Analysis-Between Calit2 and U Washington



OptIPuter

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

ROCKS



NSF

OptIPuter



GORDON AND BETTY MOORE FOUNDATION

it²

Dedicated Optical Fiber Collaboratory: Remote Researchers Jointly Exploring Complex Data



Deploy Throughout Mexico
After CICESE Test

OptIPuter



Proposal:

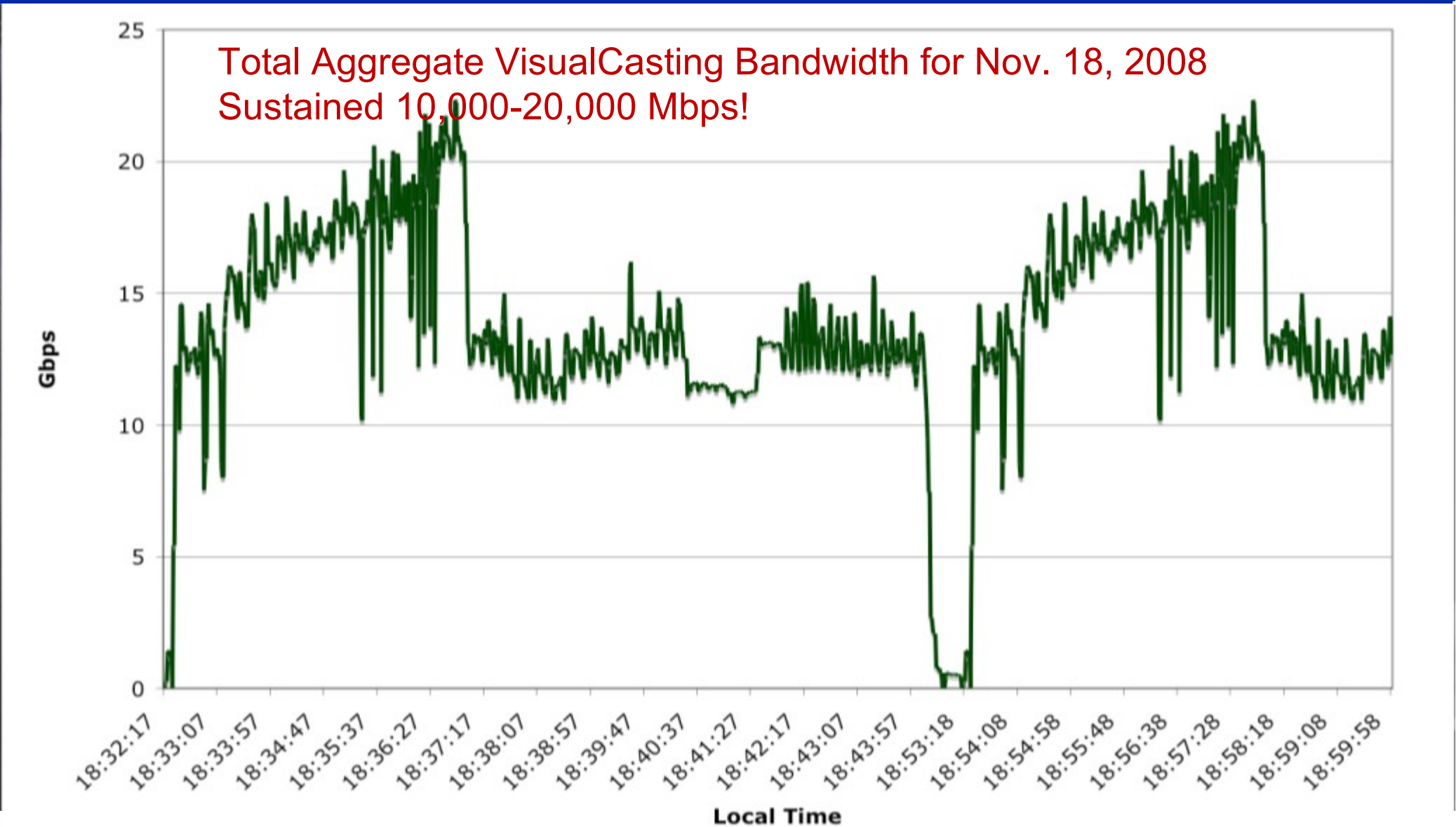
Connect OptIPortals
Between CICESE
and Calit2@UCSD
with 10 Gbps Lambda

CENIC 2012 Award:
End-to-End 10Gbps Calit2 to CICESE



LS is holding the glass award (very cool looking!), flanked by CUDI (Mexico's R&E network) director Carlos Casasus on my right and CICESE (largest Mexican science institute funded by CONACYT) director-general Federico Graef on my left. The CENIC award was presented by Louis Fox, President of CENIC (right of Carlos) and Doug Hartline, UC Santa Cruz, CENIC Conference Committee Chair (left of Federico). The Calit2/CUDI/CICESE technical team is on the right.

EVL's SAGE OptIPortal VisualCasting Multi-Site OptIPuter Collaboratory



Source: Jason Leigh, Luc Renambot, EVL, UI Chicago

Globally 10Gbp Optically Connected Digital Cinema Collaboratory



CineGrid 4K Digital Video Projects: Global Streaming of 4 x HD Over Fiber Optics

CineGrid @ iGrid 2005



CineGrid @ AES 2006





CineGrid @ Holland Festival 2007



CineGrid @ GLIF 2007

First Tri-Continental Premier of a Streamed 4K Feature Film With Global HD Discussion



4K Transmission Over 10Gbps--
4 HD Projections from One 4K Projector

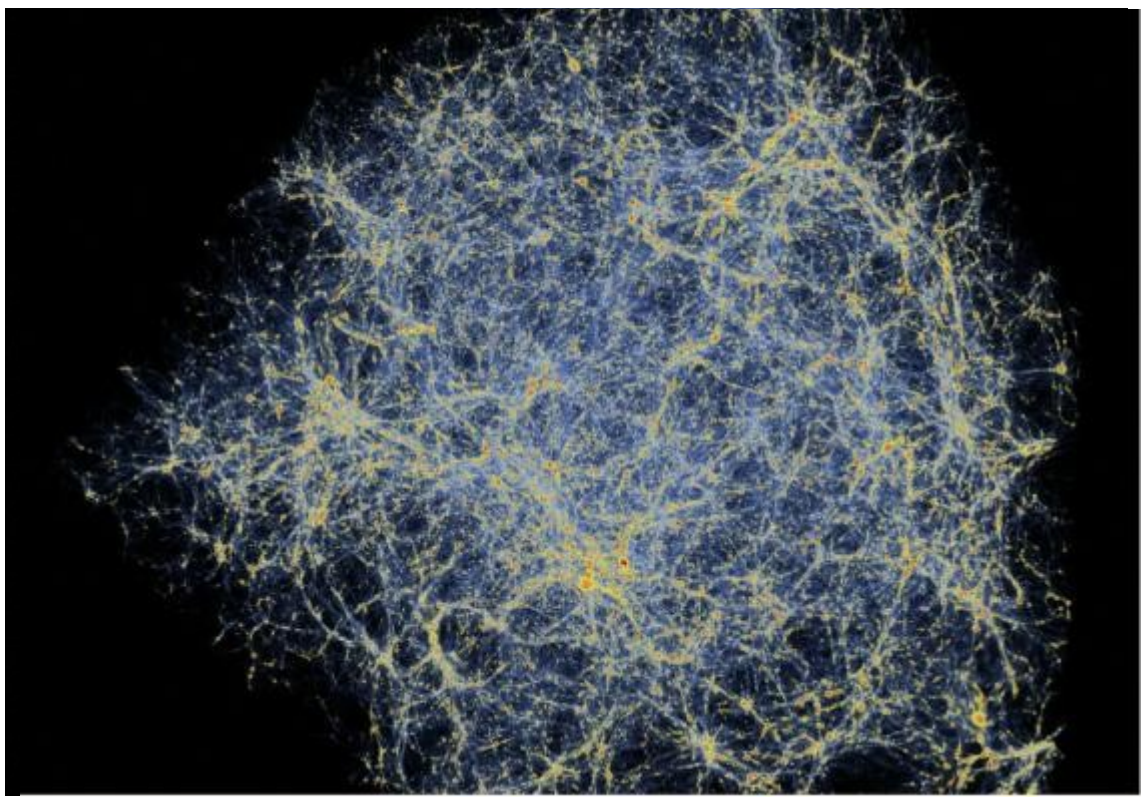
4K Digital Cinema From Keio University to Calit2's VROOM



Exploring Cosmology With Supercomputers, Supernetworks, and Supervisualization

Source: Mike Norman, SDSC

Intergalactic Medium on 2 GLyr Scale

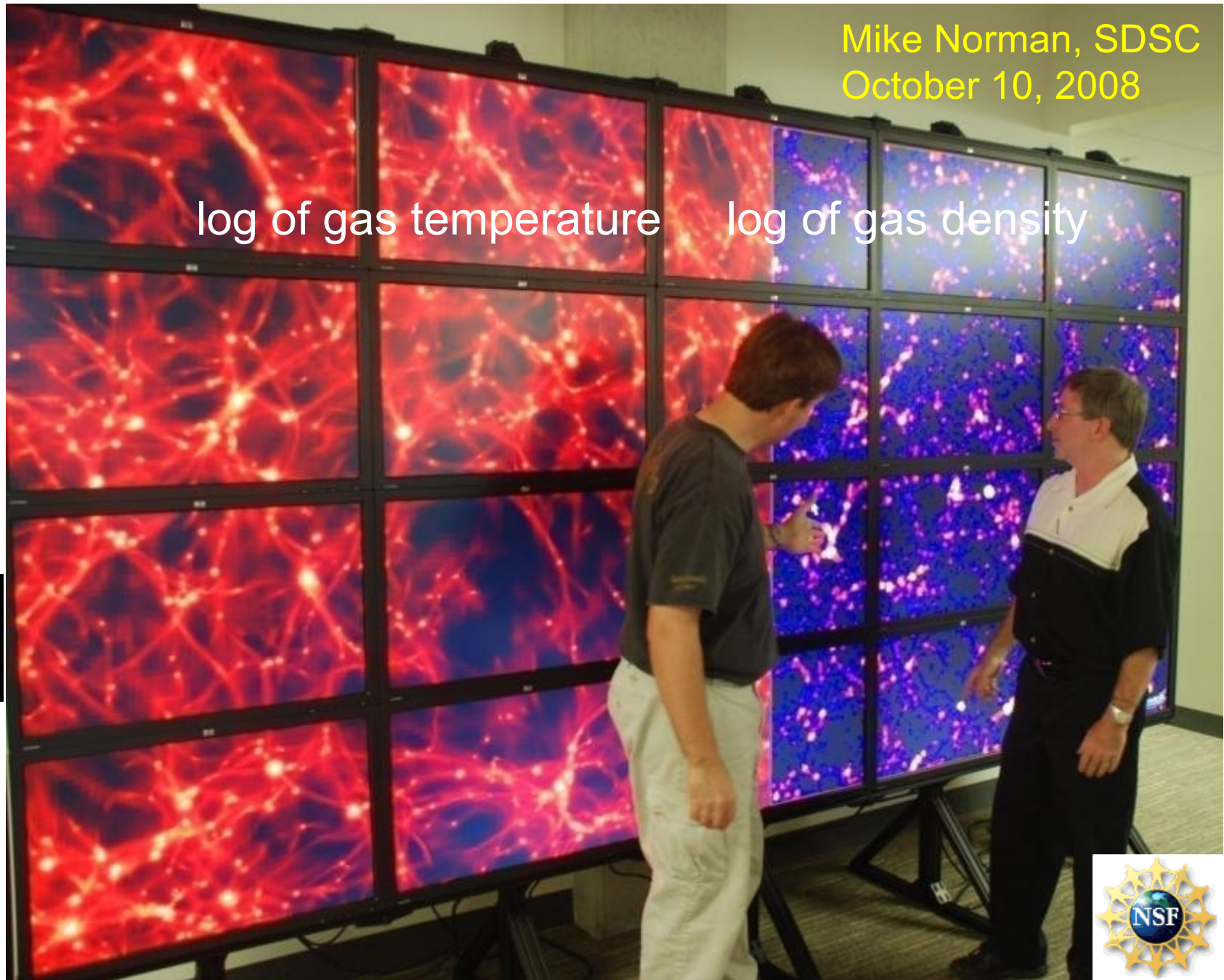


Science: Norman, Harkness,Paschos SDSC
Visualization: Insley, ANL; Wagner SDSC

- **4096^3 Particle/Cell Hydrodynamic Cosmology Simulation**
- **NICS Kraken (XT5)**
  - **16,384 cores**
- **Output**
  - **148 TB Movie Output (0.25 TB/file)**
  - **80 TB Diagnostic Dumps (8 TB/file)**

Providing End-to-End CI for Petascale End Users

Two 64K Images From a Cosmological Simulation of Galaxy Cluster Formation



Using Supernetworks to Couple End User's OptIPortal to Remote Supercomputers and Visualization Servers

Source: Mike Norman, Rick Wagner, SDSC

Argonne NL

DOE Eureka
100 Dual Quad Core Xeon Servers
200 NVIDIA Quadro FX GPUs in 50 Quadro Plex S4 1U enclosures
3.2 TB RAM



Real-Time Interactive Volume Rendering Streamed from ANL to SDSC

ESnet
10 Gb/s fiber optic network

SDSC



Calit2/SDSC OptIPortal1
20 30" (2560 x 1600 pixel) LCD panels
10 NVIDIA Quadro FX 4600 graphics cards > 80 megapixels
10 Gb/s network throughout

NICS ORNL

NSF TeraGrid Kraken
Cray XT5
8,256 Compute Nodes
99,072 Compute Cores
129 TB RAM

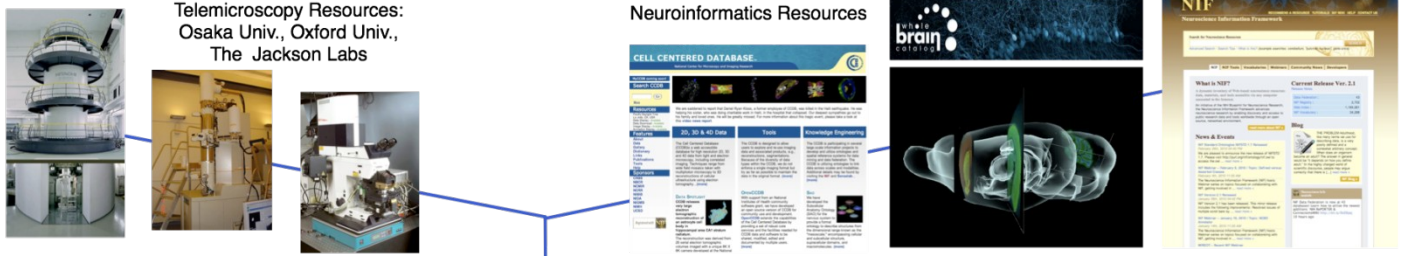


SDSC

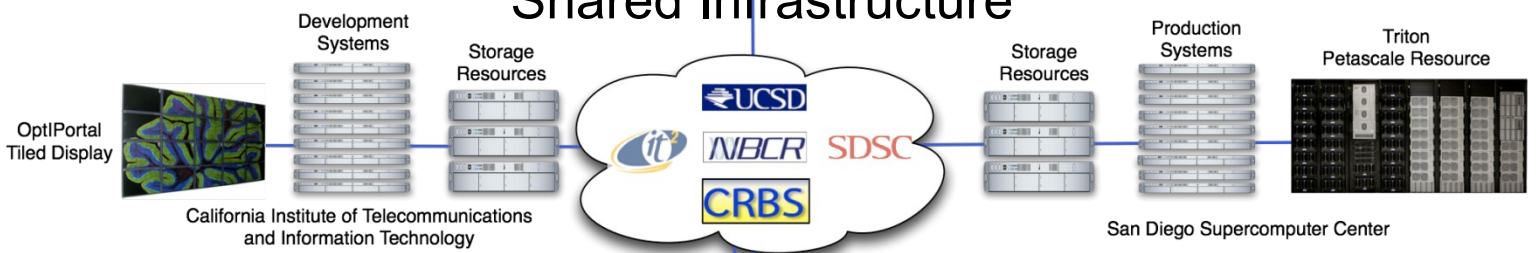
*ANL * Calit2 * LBNL * NICS * ORNL * SDSC

NIH National Center for Microscopy & Imaging Research
Integrated Infrastructure of Shared Resources

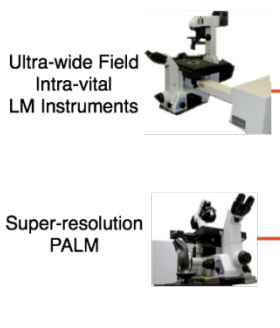
NCMIR

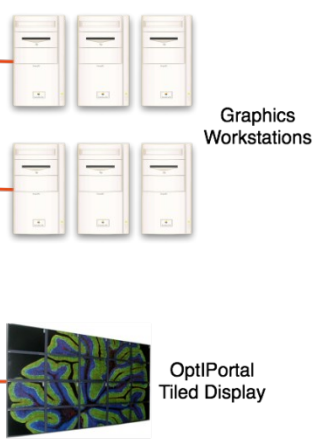


Shared Infrastructure

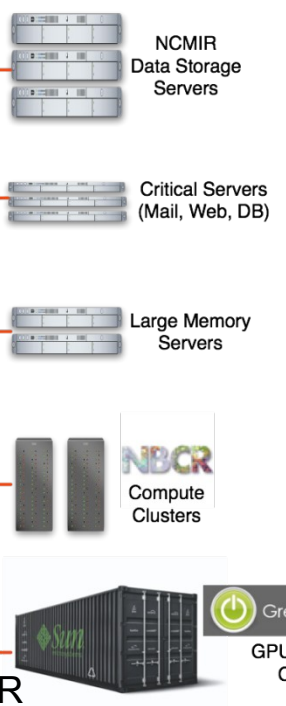


Scientific Instruments





End User Workstations



Local SOM Infrastructure

SDSC

Source: Steve Peltier, Mark Ellisman, NCMIR

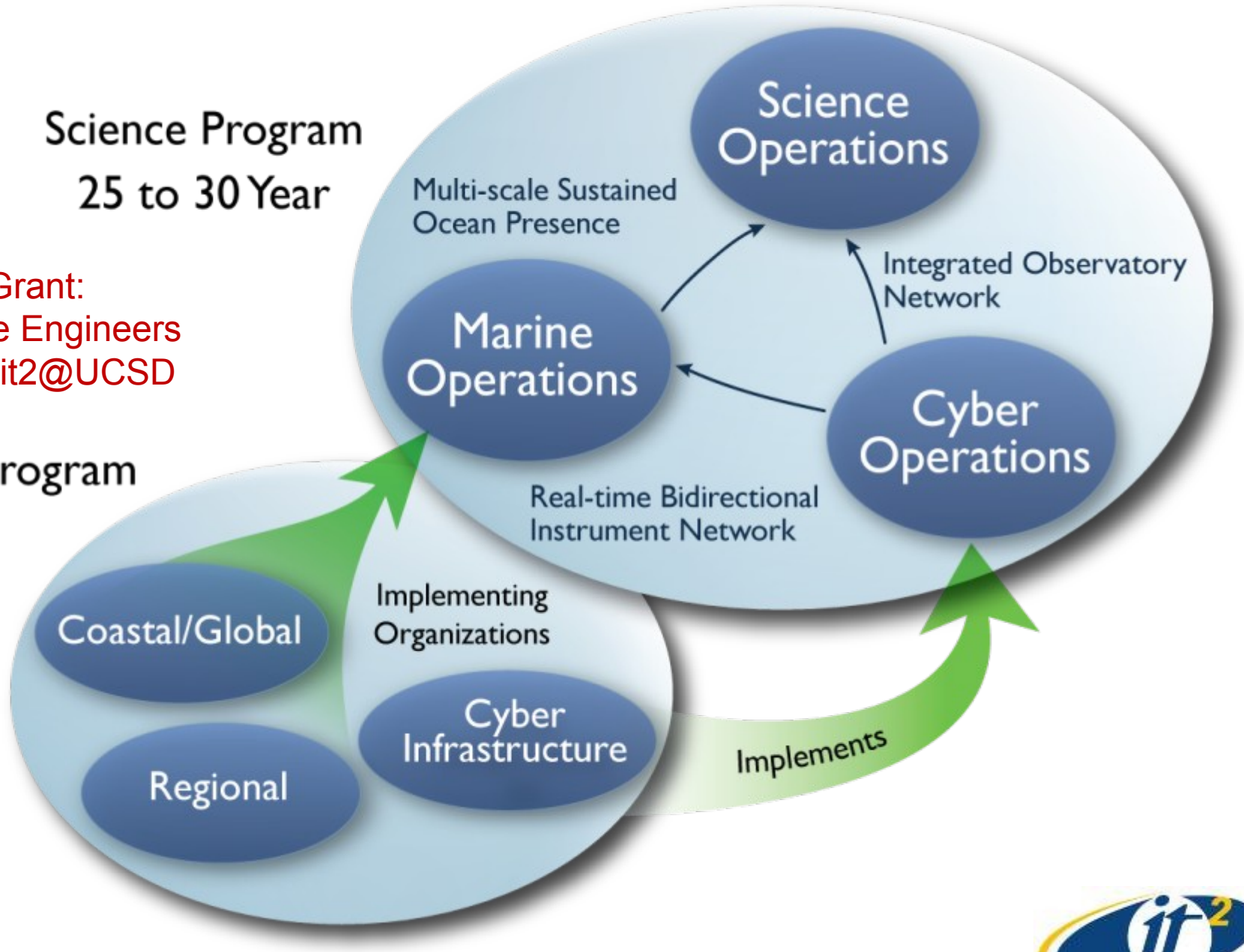
NSF's Ocean Observatory Initiative Has the Largest Funded NSF CI Grant

OCEAN OBSERVATORIES INITIATIVE

Science Program
25 to 30 Year

OOI CI Grant:
30-40 Software Engineers
Housed at Calit2@UCSD

Construction Program
5 Year

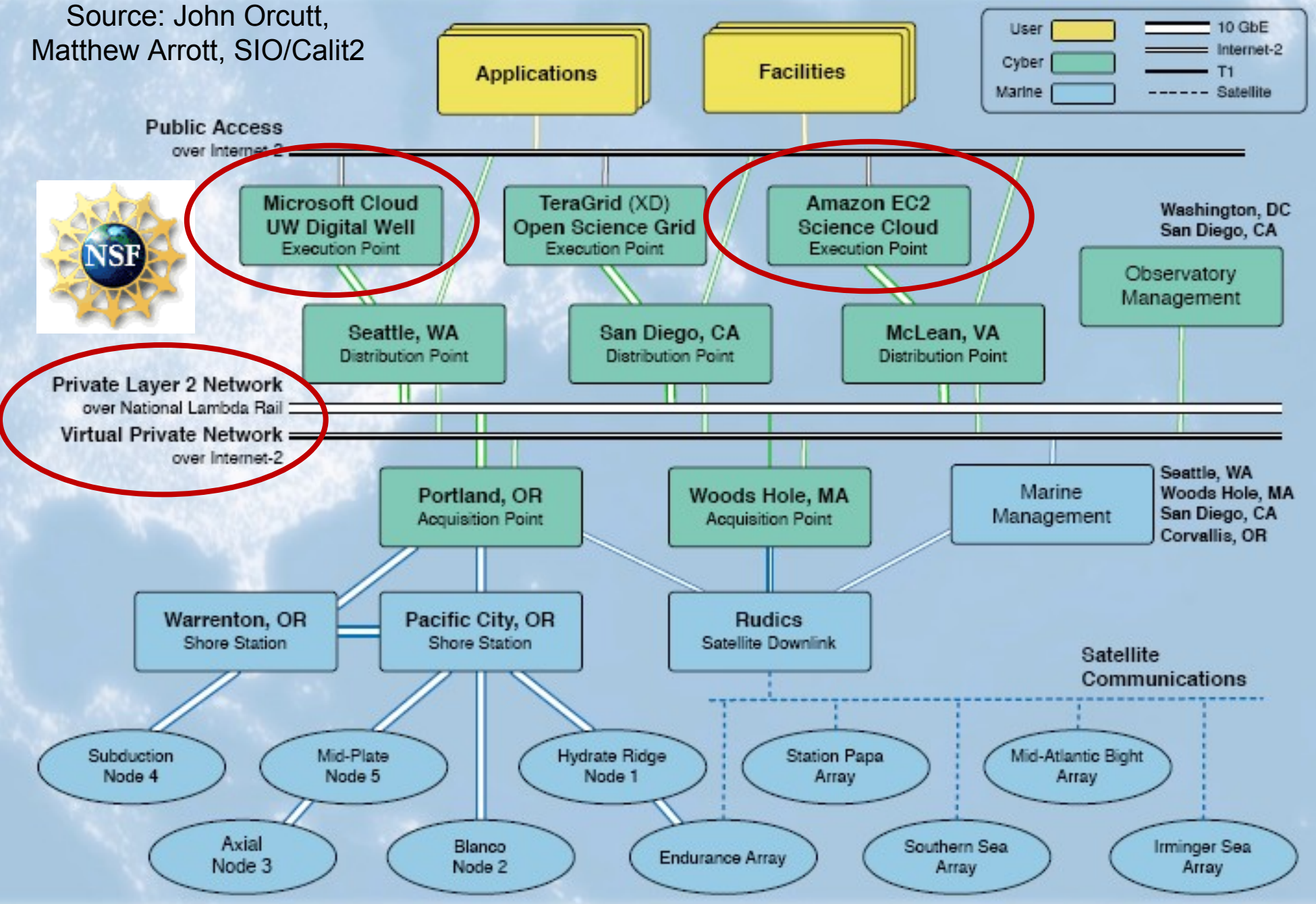


Source: Matthew Arrott, Calit2 Program Manager for OOI CI

OOI CI is Built on Dedicated Optical Infrastructure Using Clouds

OCEAN OBSERVATORIES INITIATIVE

Source: John Orcutt, Matthew Arrott, SIO/Calit2

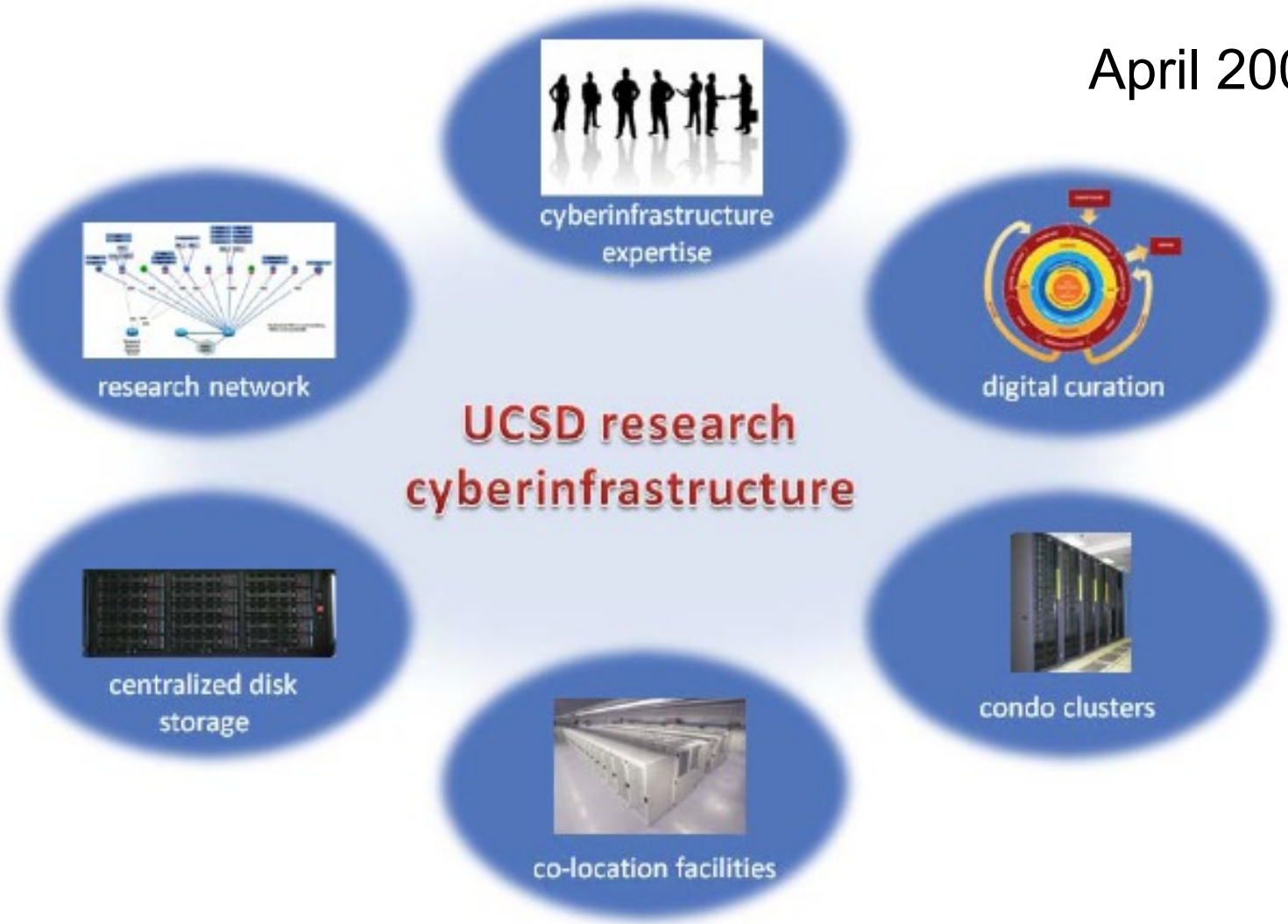


"Blueprint for the Digital University"--Report of the UCSD Research Cyberinfrastructure Design Team

- **A Five Year Process Began Pilot Deployment Last Year**

April 2009

No Data Bottlenecks--Design for Gigabit/s Data Flows

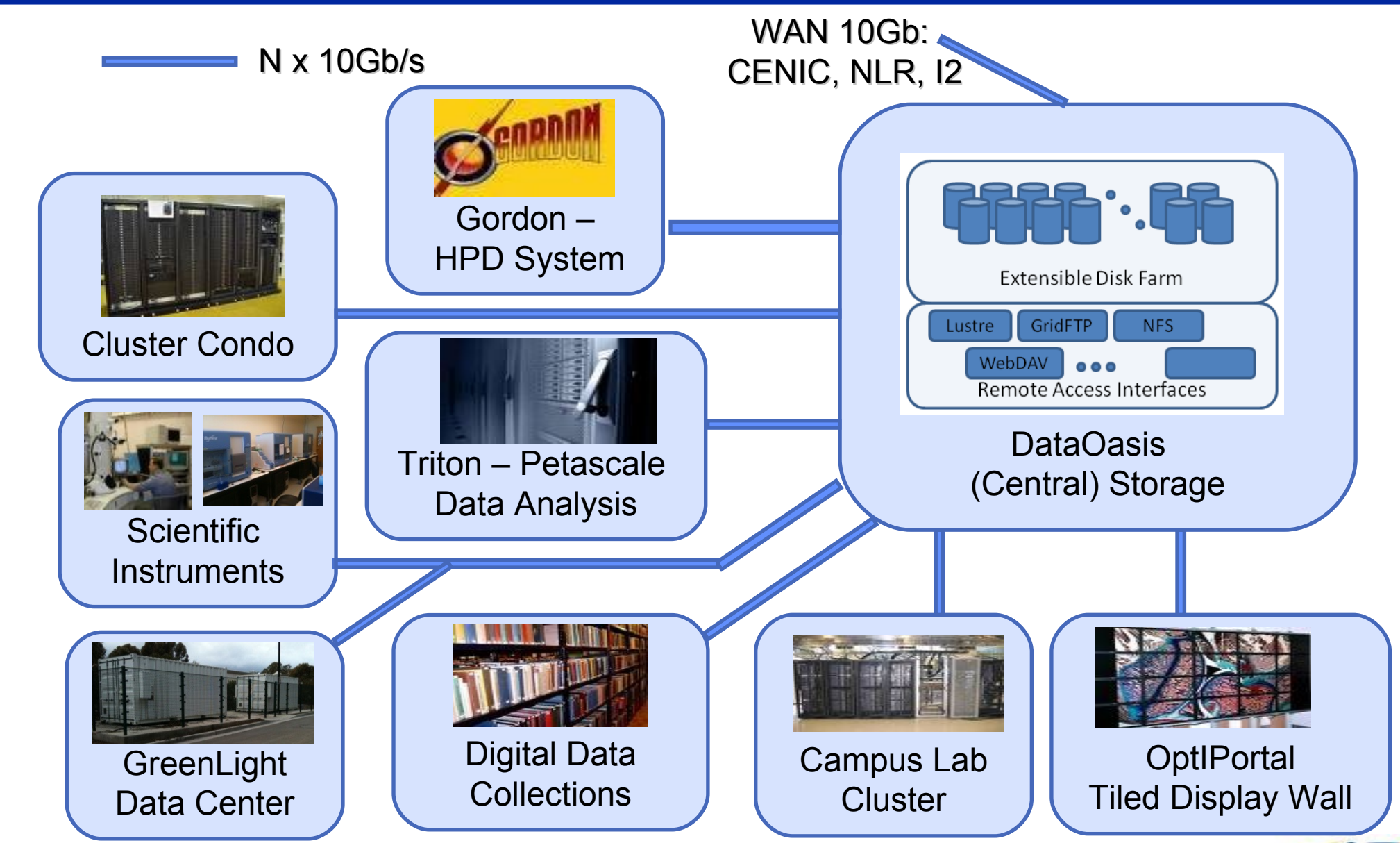


SDSC

http://rci.ucsd.edu

it²

UCSD Campus Investment in Fiber Enables Consolidation of Energy Efficient Computing & Storage



SDSC

Source: Philip Papadopoulos, SDSC, UCSD

Calit2 Sunlight OptIPuter Exchange Connects 60 Campus Sites Each Dedicated at 10Gbps



Maxine Brown, EVL, UIC OptIPuter Project Manager

SDSC

it²

NSF Funds a Big Data Supercomputer: SDSC's Gordon-Dedicated Dec. 5, 2011

- **Data-Intensive Supercomputer Based on SSD Flash Memory and Virtual Shared Memory SW**
  - *Emphasizes MEM and IOPS over FLOPS*
  - *Supernode has Virtual Shared Memory:*
    - *2 TB RAM Aggregate*
    - *8 TB SSD Aggregate*
  - *Total Machine = 32 Supernodes*
    - *4 PB Disk Parallel File System >100 GB/s I/O*



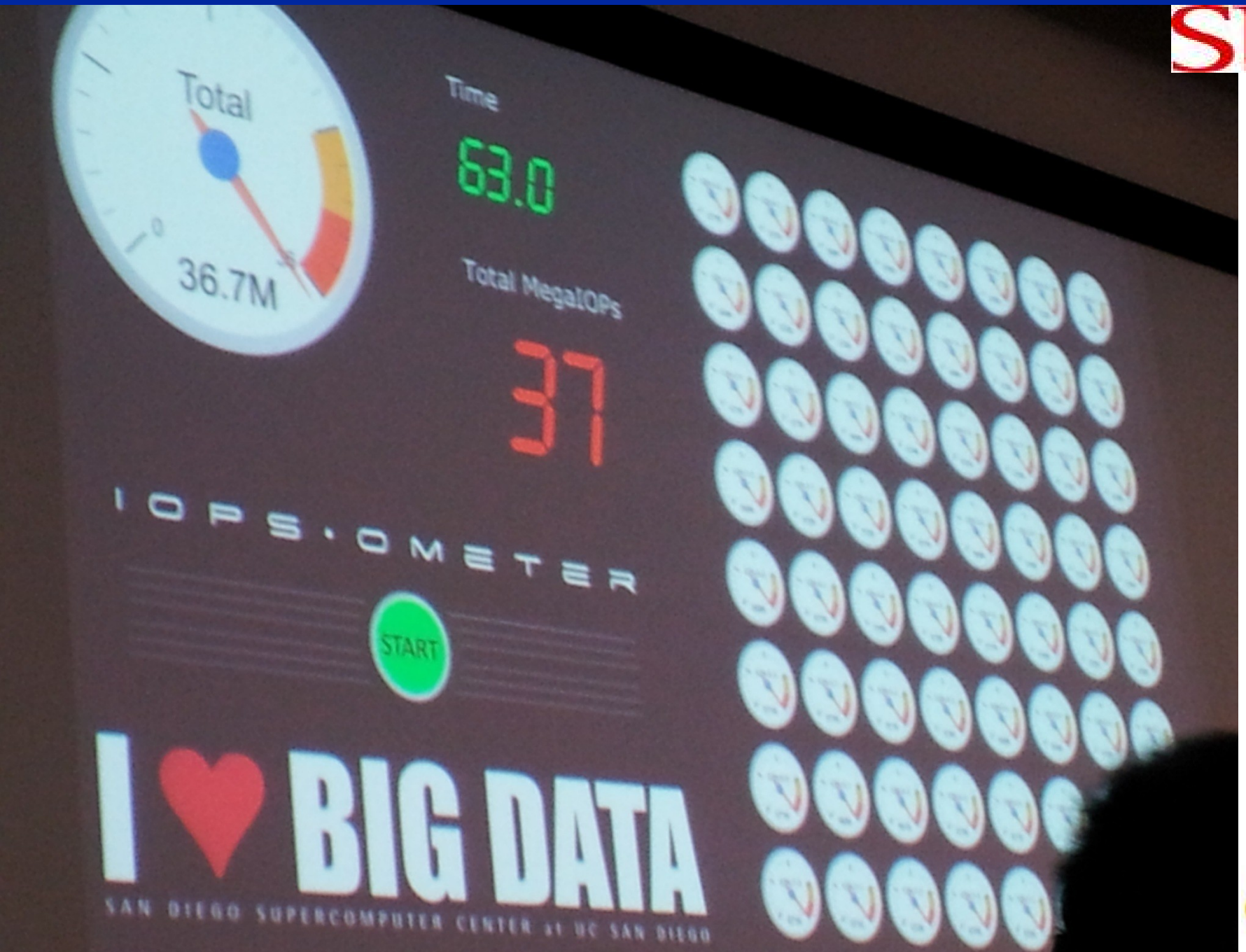
- **System Designed to Accelerate Access to Massive Datasets being Generated in Many Fields of Science, Engineering, Medicine, and Social Science**

SDSC

Source: Mike Norman, Allan Snavely SDSC

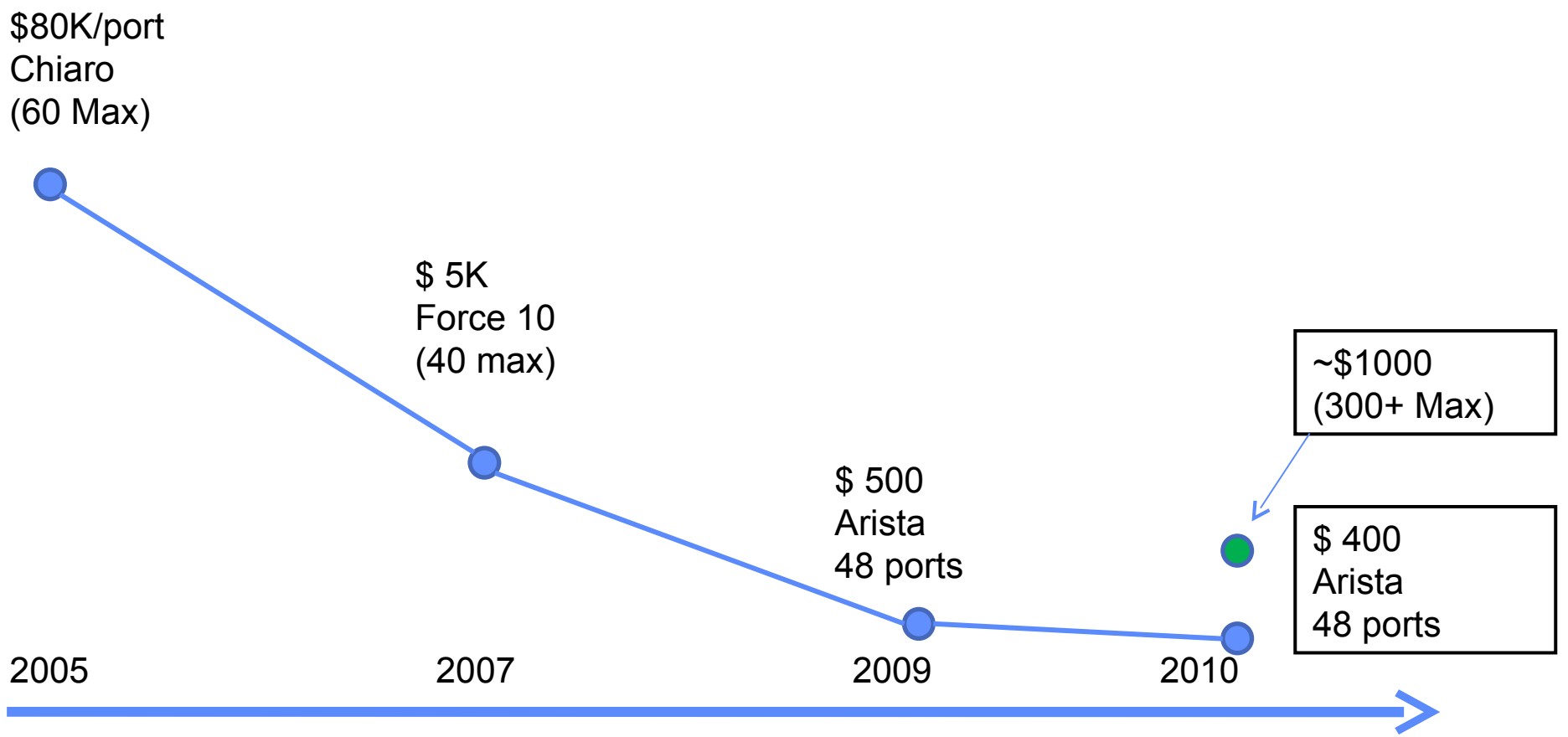
it²

Gordon Bests Previous Mega I/O per Second by 25x



Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

- Port Pricing is Falling
- Density is Rising – Dramatically
- Cost of 10GbE Approaching Cluster HPC Interconnects

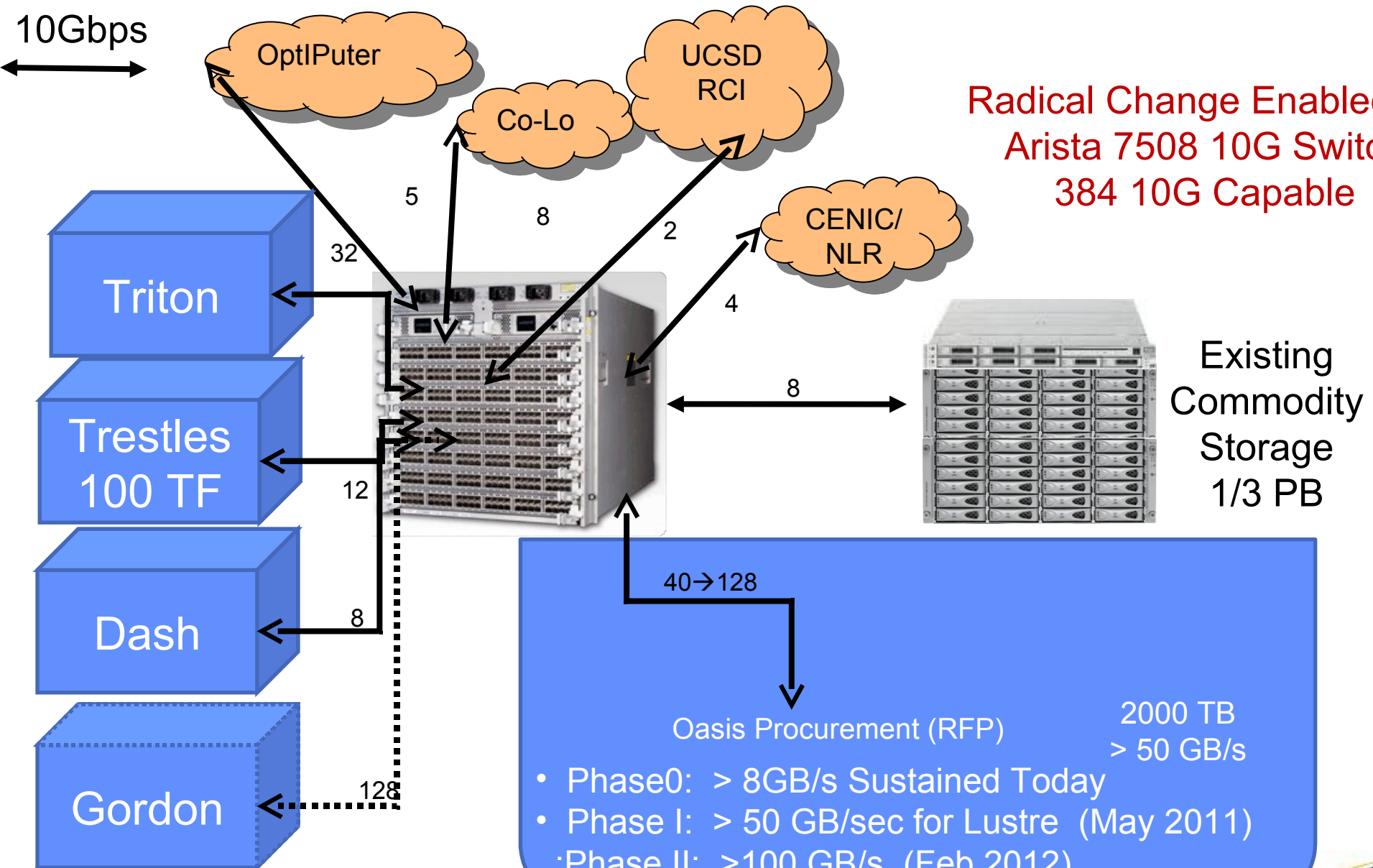


SDSC

Source: Philip Papadopoulos, SDSC/Calit2

it²

Arista Enables SDSC's Massive Parallel 10G Switched Data Analysis Resource



Radical Change Enabled by Arista 7508 10G Switch 384 10G Capable

Existing Commodity Storage 1/3 PB

SDSC

Source: Philip Papadopoulos, SDSC/Calit2

it²

The Next Step for Data-Intensive Science: Pioneering the HPC Cloud



HPC Cloud Computing

Self Service and Dynamically Scalable

High Performance Computing (HPC) on









