"A Campus-Scale High Performance Cyberinfrastructure is Required for Data-Intensive Research"

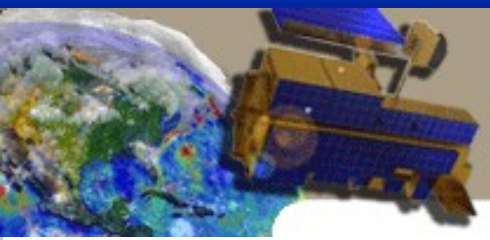
Seminar Presentation
Princeton Institute for Computational Science and Engineering (PICSciE)
Princeton University
Princeton, NJ
December 12, 2011

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

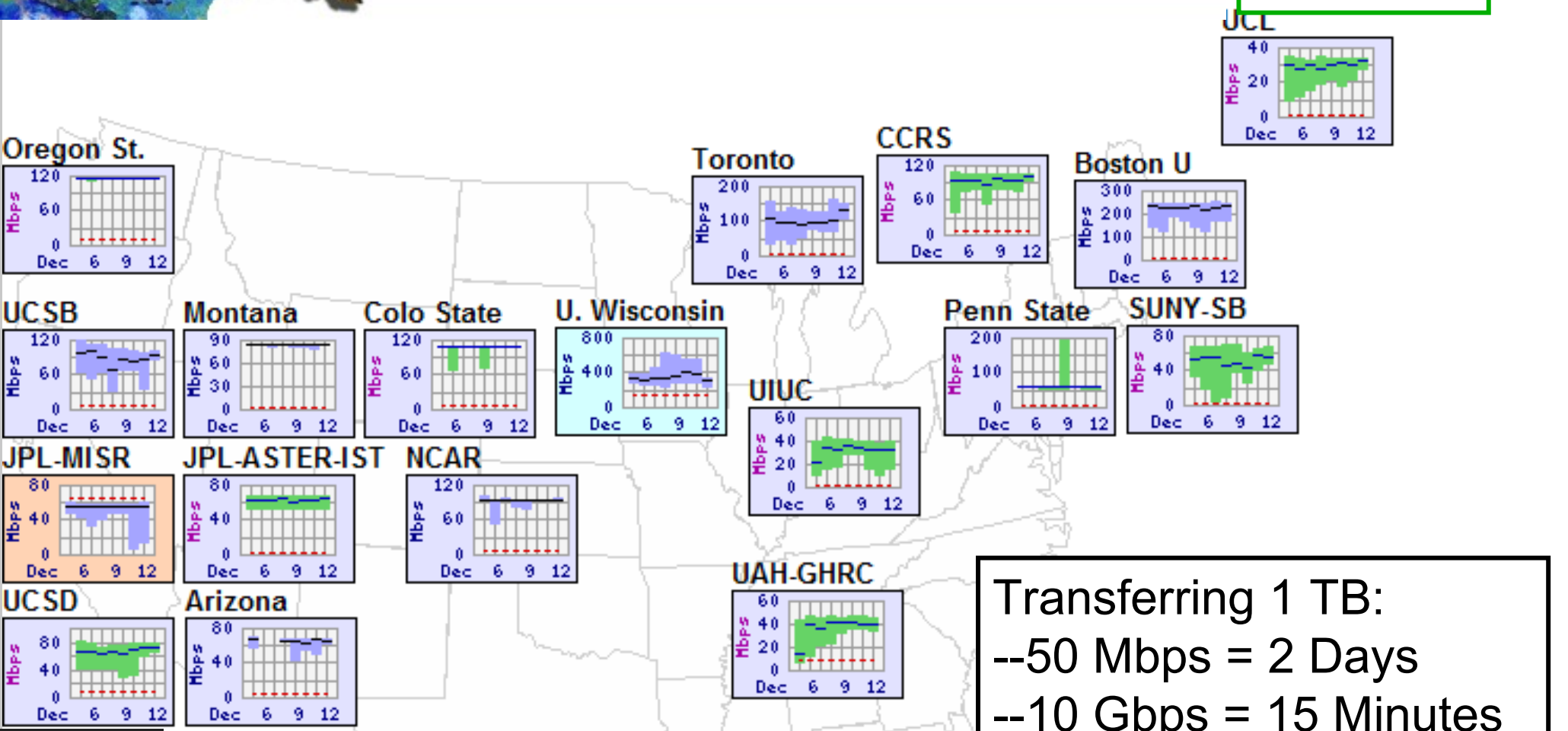
Campuses are experiencing an enormous increase in the quantity of data generated by scientific instruments and computational clusters and stored in massive data repositories. The shared Internet, engineered to enable interaction with megabyte-sized data objects is not capable of dealing with the typical gigabytes to terabytes of modern scientific data. Instead, a high performance cyberinfrastructure is emerging to support data-intensive research. Fortunately, multi-channel optical fiber can support both the traditional internet and this new data utility. I will give examples of early prototypes which integrate data generation, transmission, storage, analysis, visualization, curation, and sharing, driven by applications as diverse as genomics, ocean observatories, and cosmology.

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

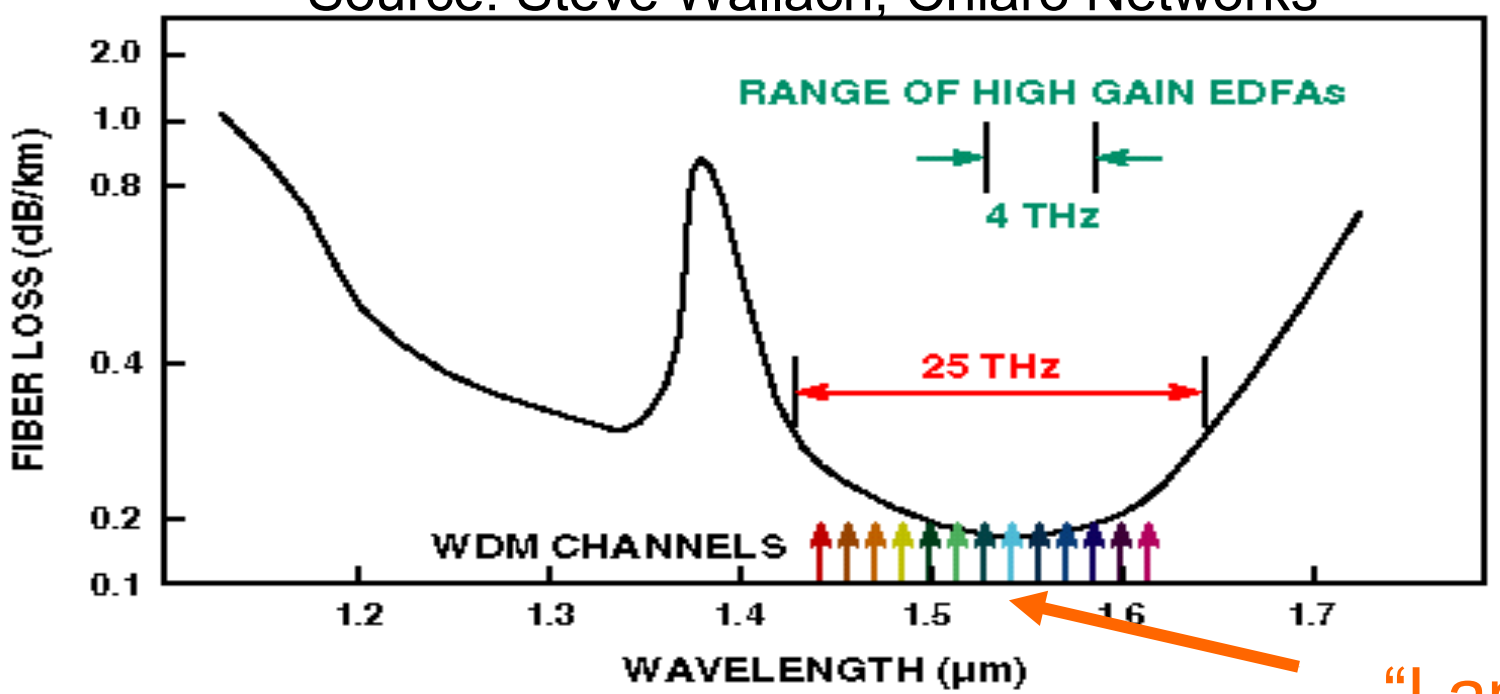
OptIPuter Solution:
Give Dedicated Optical Channels to Data-Intensive Users

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

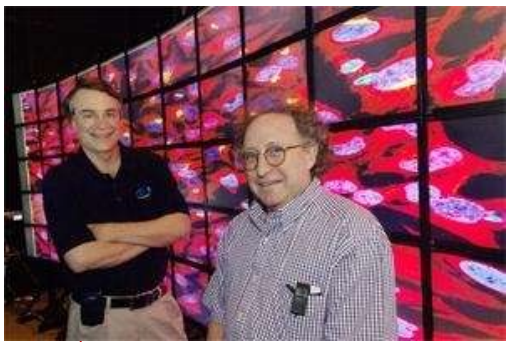
The Global Lambda Integrated Facility--
Creating a Planetary-Scale High Bandwidth Collaboratory

Research Innovation Labs Linked by 10G Dedicated Lambdas



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

Academic Research OptIPlanet Collaboratory: A 10Gbps "End-to-End" Lightpath Cloud

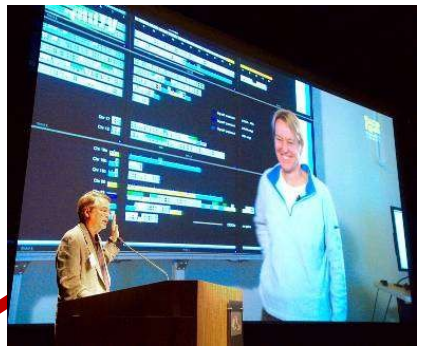


End User OptIPortal





Local or Remote Instruments



HD/4k Live Video

HPC

TeraGrid





10G Lightpaths

Campus Optical Switch

amazon web services



Data Repositories & Clusters

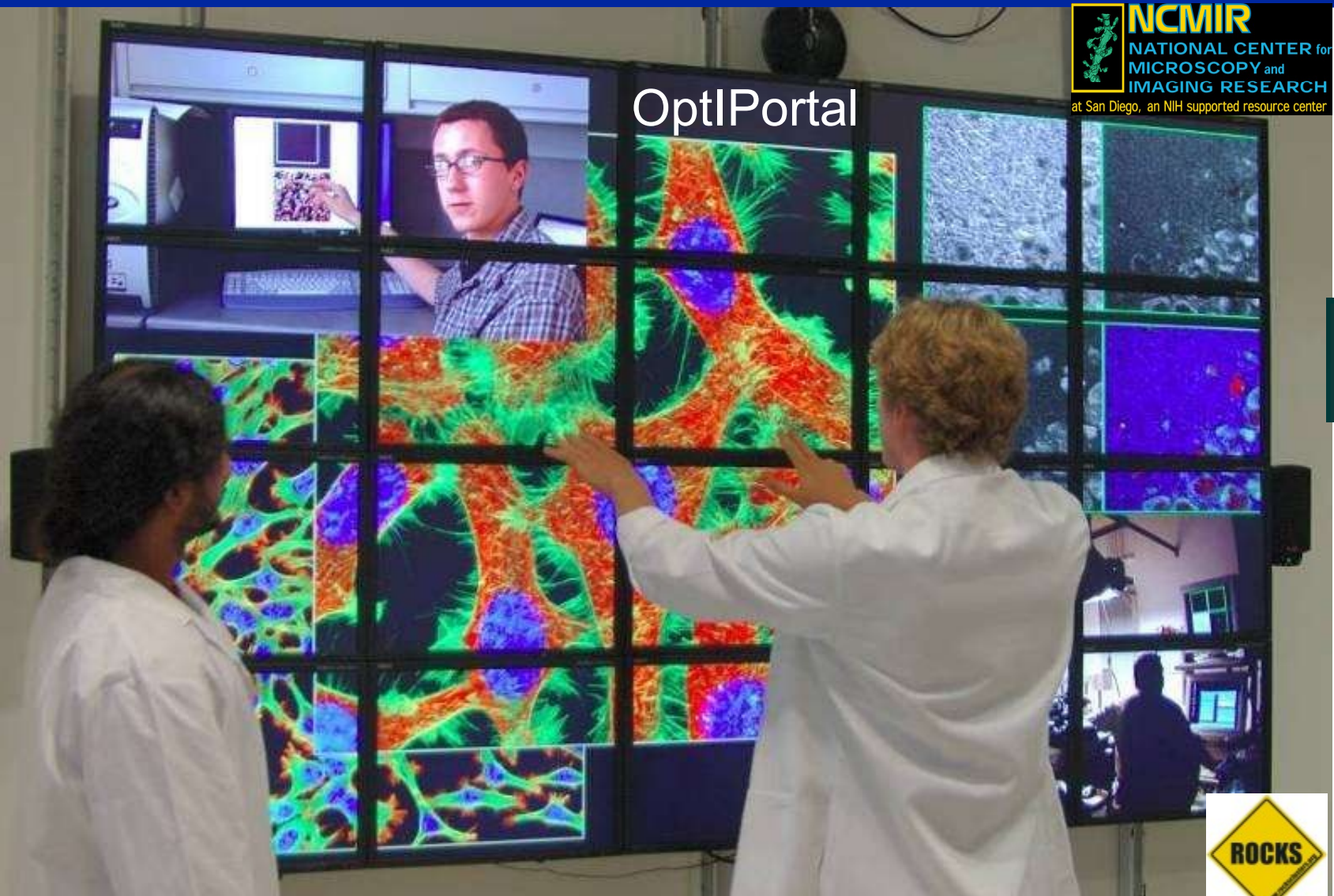


HD/4k Video Repositories

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

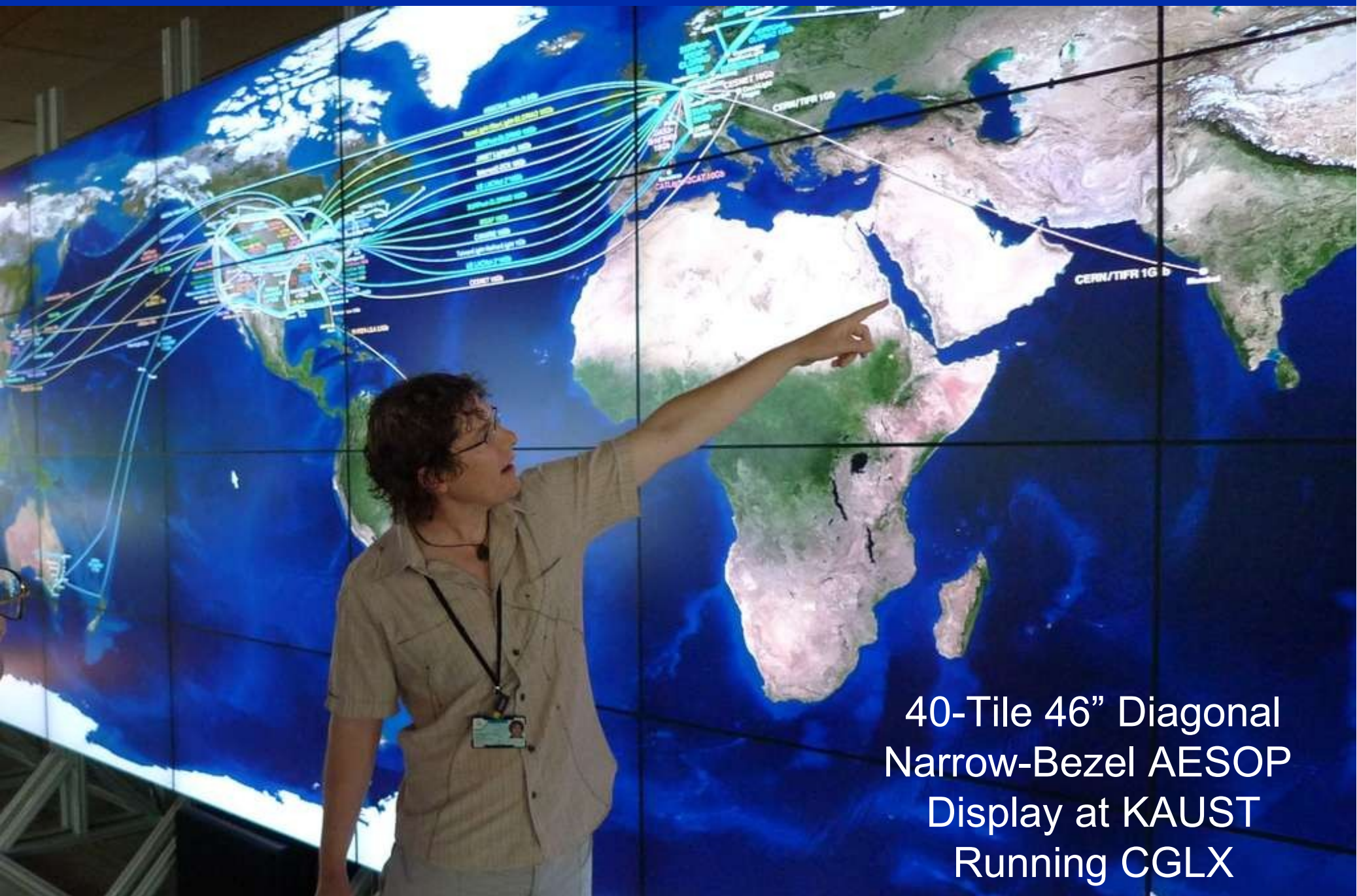
MIT's Ed DeLong and Darwin Project Team Using OptIPortal to Analyze 10km Ocean Microbial Simulation



GORDON AND BETTY MOORE FOUNDATION

Cross-Disciplinary Research at MIT, Connecting Systems Biology, Microbial Ecology, Global Biogeochemical Cycles and Climate

AESOP Display built by Calit2 for KAUST-- King Abdullah University of Science & Technology



40-Tile 46” Diagonal Narrow-Bezel AESOP Display at KAUST Running CGLX

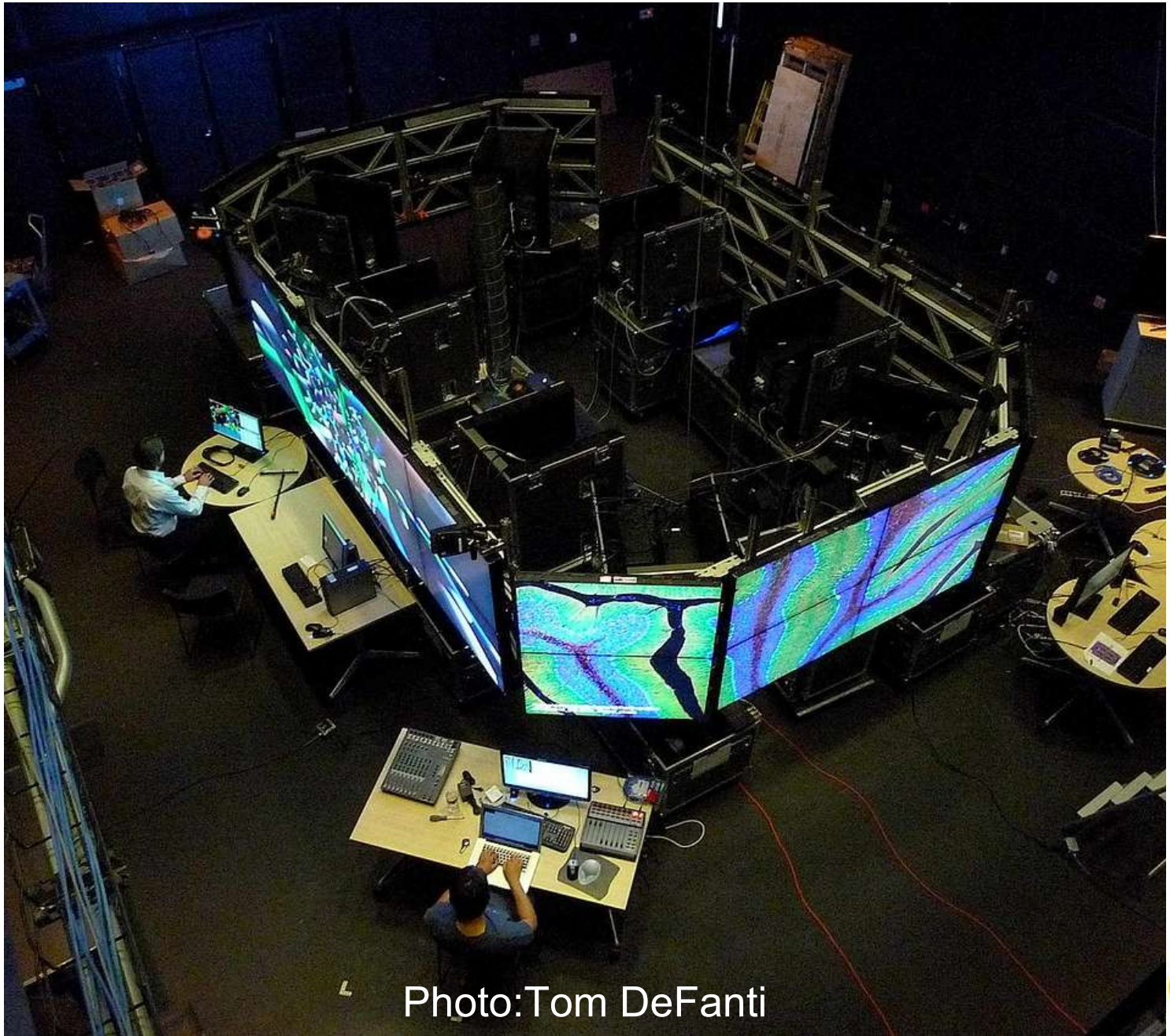
The Latest OptIPuter Innovation: Quickly Deployable Nearly Seamless OptIPortables

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



3D Stereo Head Tracked OptIPortal: NexCAVE

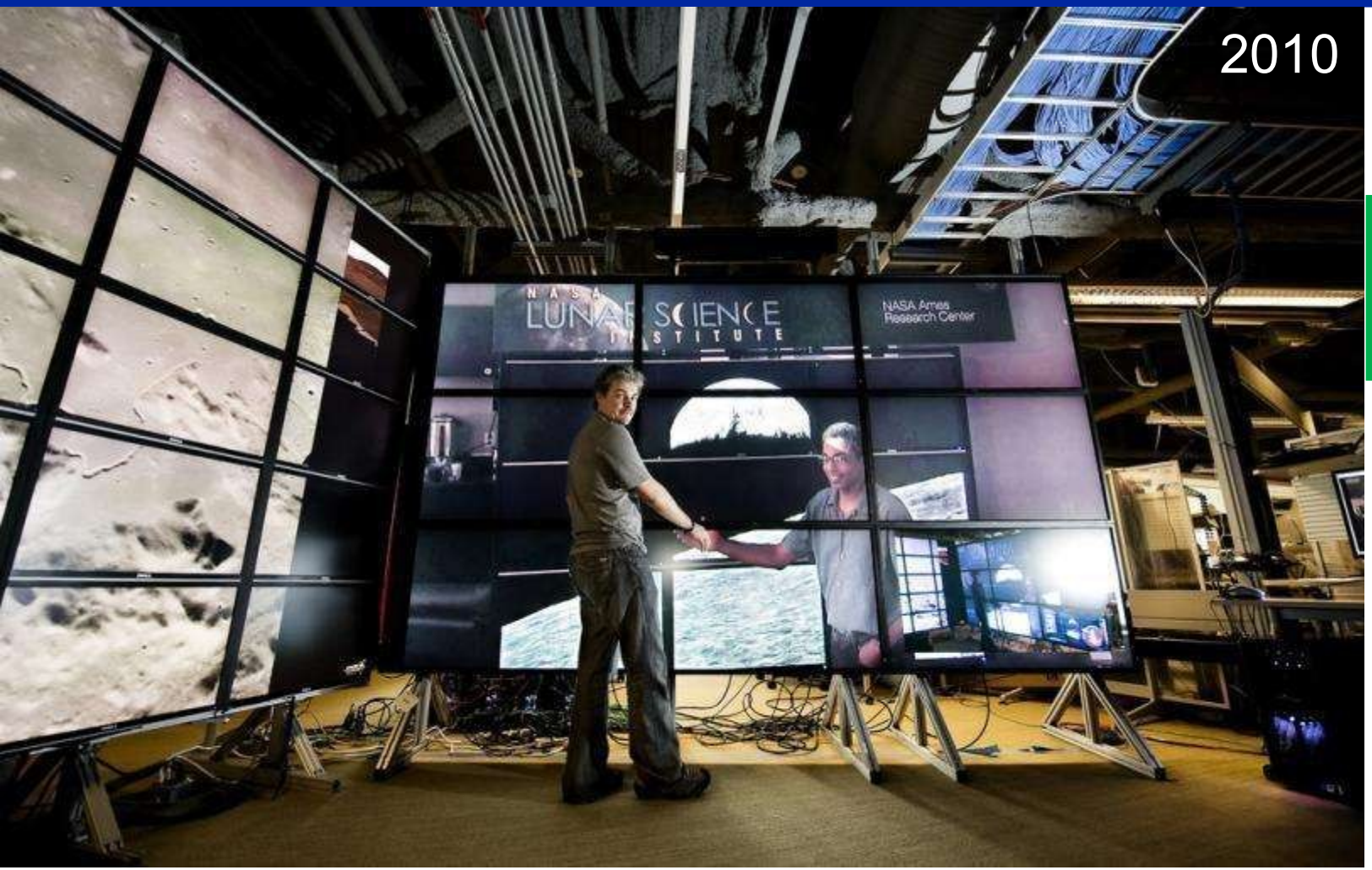


Array of JVC HDTV 3D LCD Screens
KAUST NexCAVE = 22.5MPixels

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

Source: Tom DeFanti, Calit2@UCSD

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

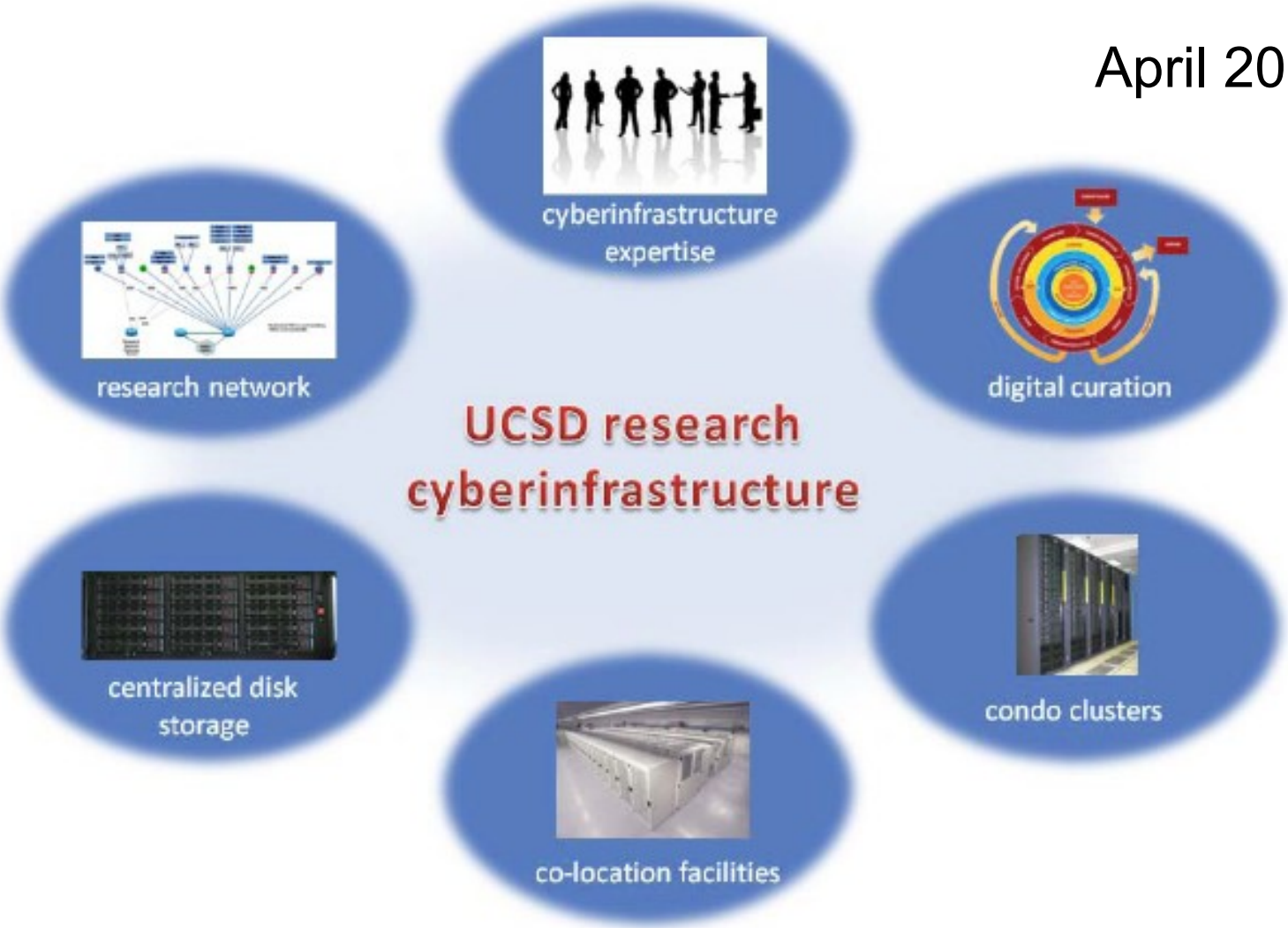
it²

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

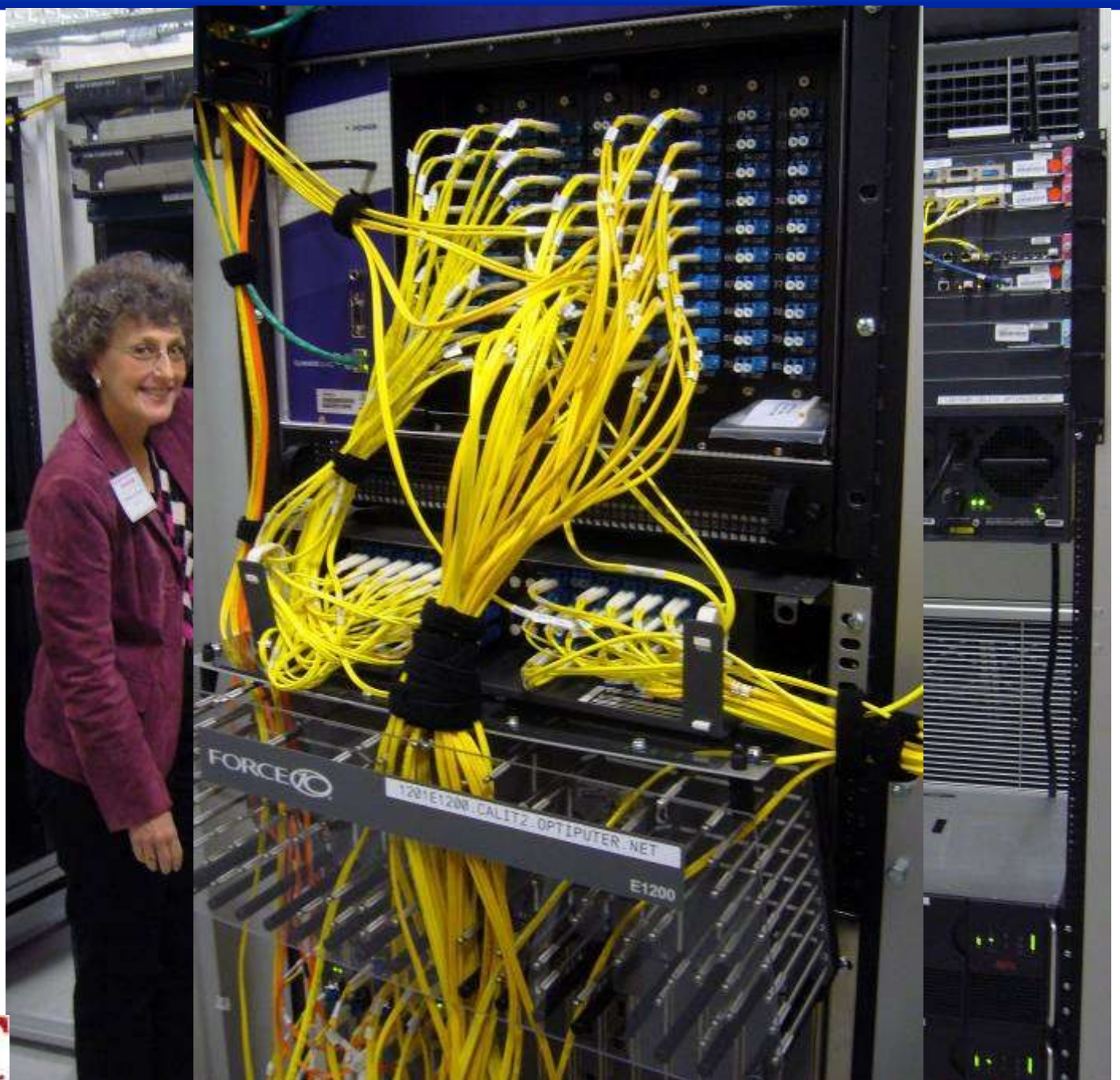
- **A Five Year Process Begins Pilot Deployment This Year**

April 2009

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



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

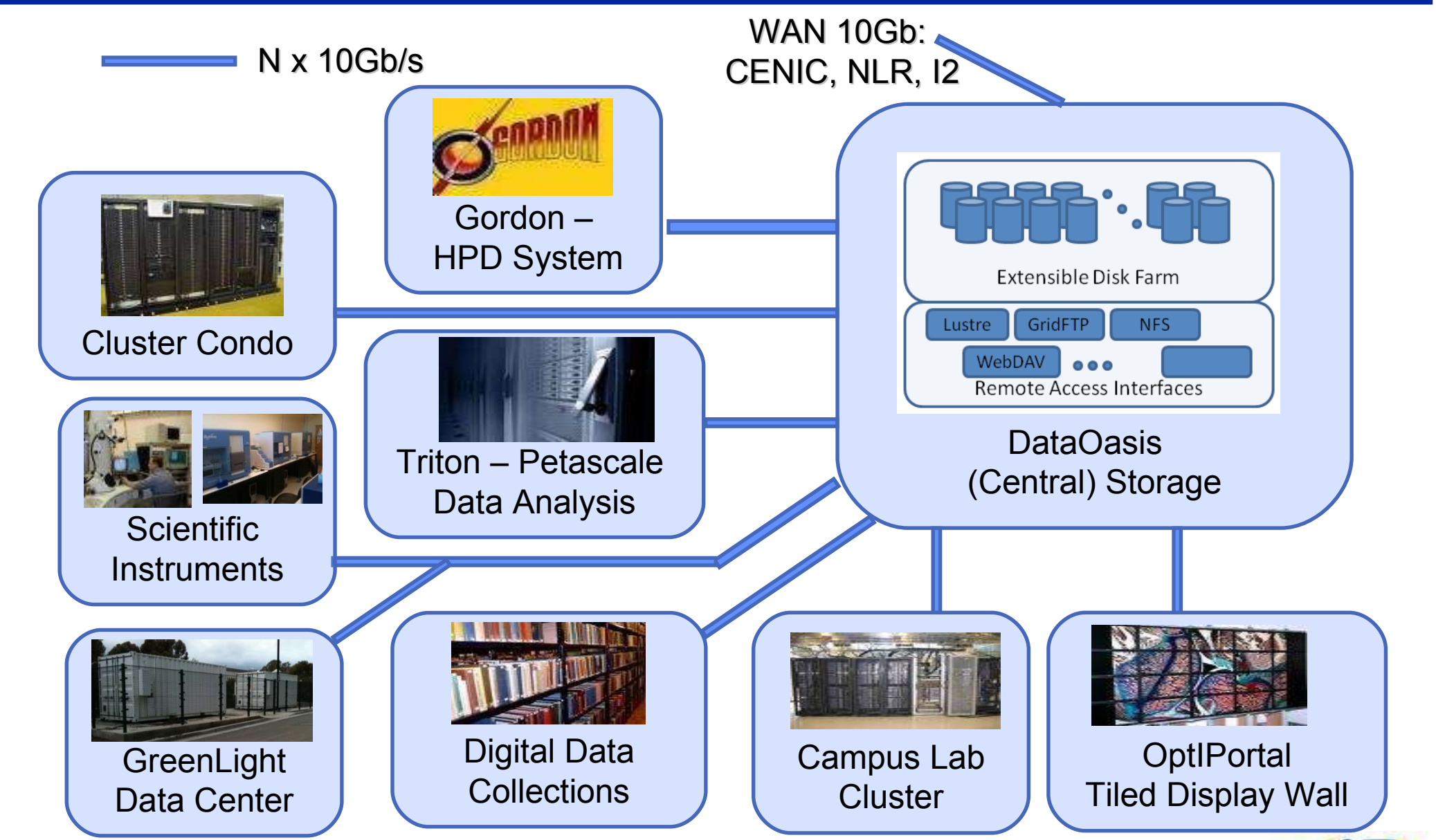


Maxine Brown, EVL, UIC OptIPuter Project Manager

SDSC

it²

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



Source: Philip Papadopoulos, SDSC, UCSD

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*



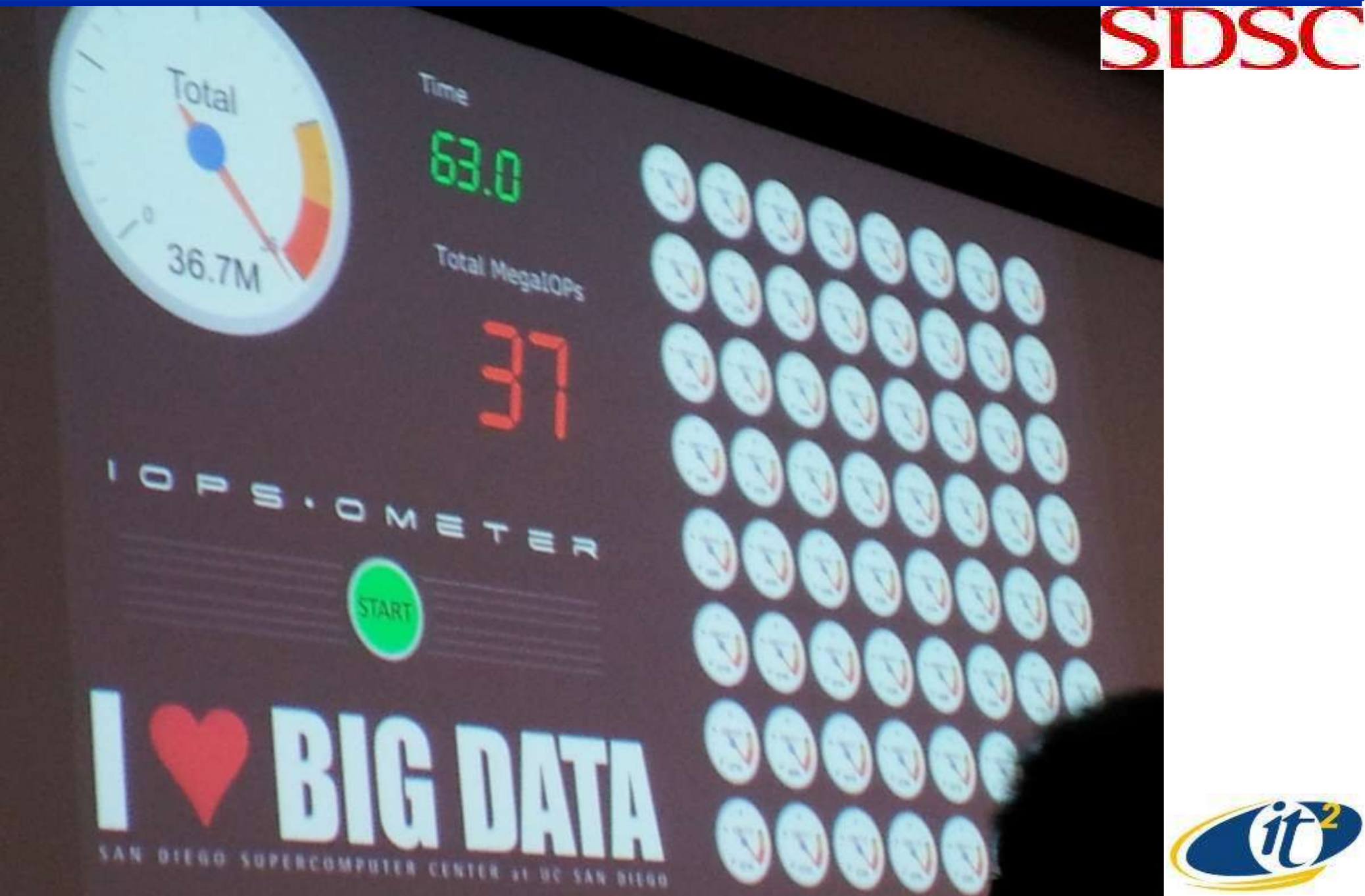
- **System Designed to Accelerate Access to Massive Datasets being Generated in Many Fields of Science, Engineering, Medicine, and Social Science**

SDSC

Source: Mike Norman, Allan Snavely SDSC

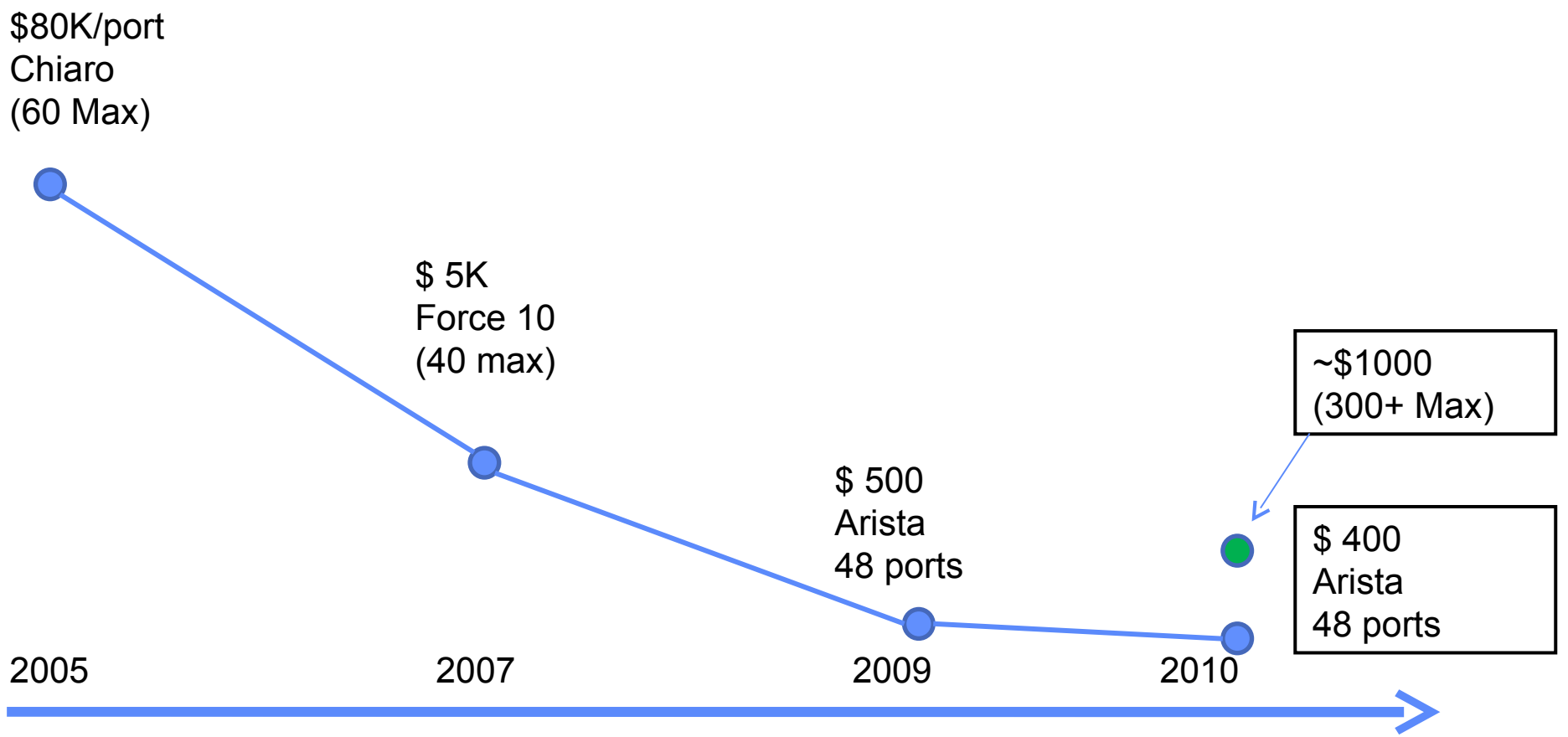
it²

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



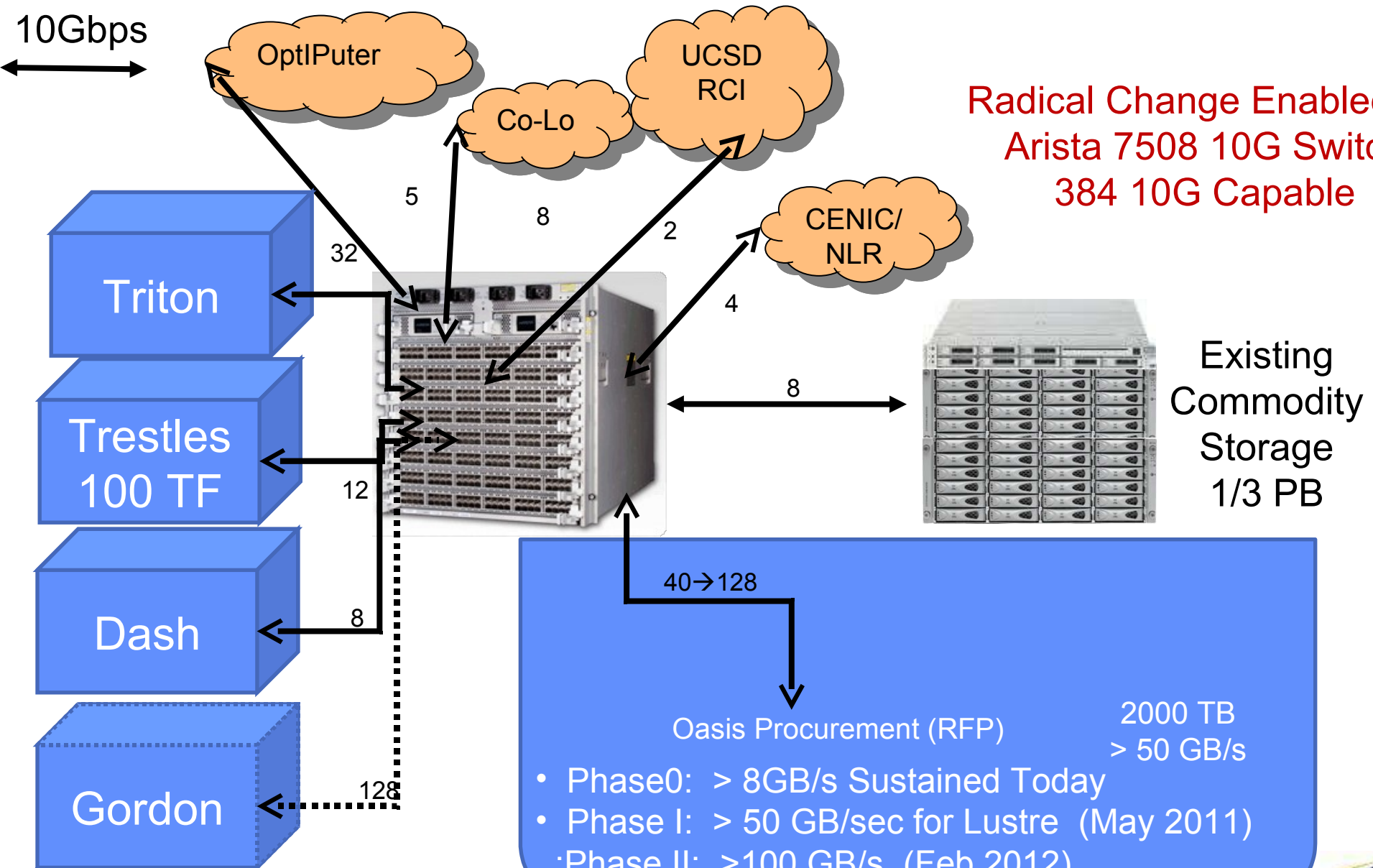
Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

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



Source: Philip Papadopoulos, SDSC/Calit2

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

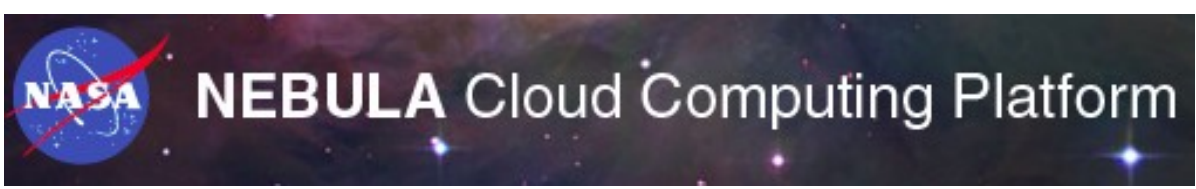
Self Service and Dynamically Scalable

High Performance Computing (HPC) on











Data Oasis – 3 Different Types of Storage



HPC Storage (Lustre-Based PFS)

- Purpose: Transient Storage to Support HPC, HPD, and Visualization
- Access Mechanisms: Lustre Parallel File System Client



Project (Traditional File Server) Storage

- Purpose: Typical Project / User Storage Needs
- Access Mechanisms: NFS/CIFS "Network Drives"



Cloud Storage

- Purpose: Long-Term Storage of Data that will be Infrequently Accessed
- Access Mechanisms: S3 interfaces, DropBox-esq web interface, CommVault

SDSC

it²

Examples of Applications Built on UCSD RCI

- DOE Remote Use of Petascale HPC
- Moore Foundation Microbial Metagenomics Server
- NSF GreenLight Instrumented Data Center
- NIH Next Generation Gene Sequencers
- NIH Shared Scientific Instruments

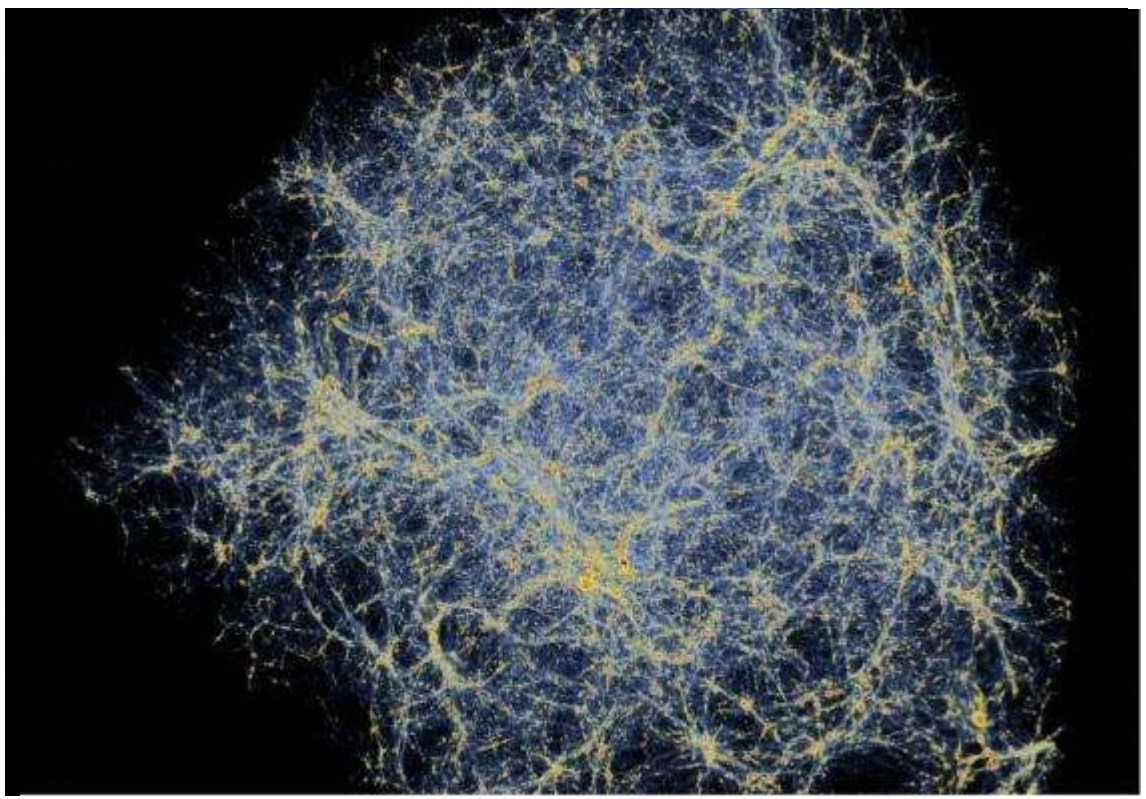
SDSC

it²

Exploring Cosmology With Supercomputers, Supernetworks, and Supervisualization

Source: Mike Norman, SDSC

Intergalactic Medium on 2 GLyr Scale

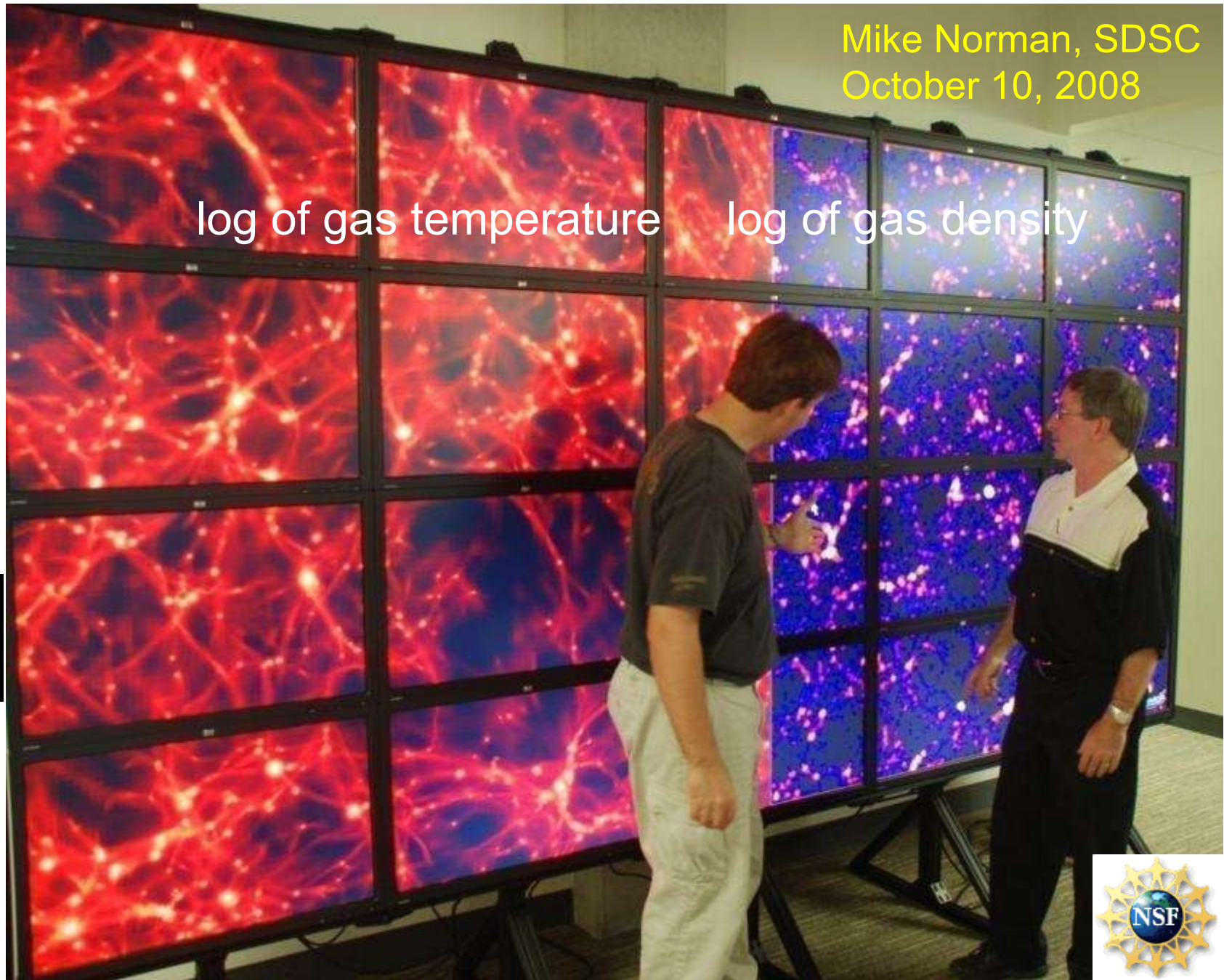


Science: Norman, Harkness,Paschos SDSC
Visualization: Insley, ANL; Wagner SDSC

- **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

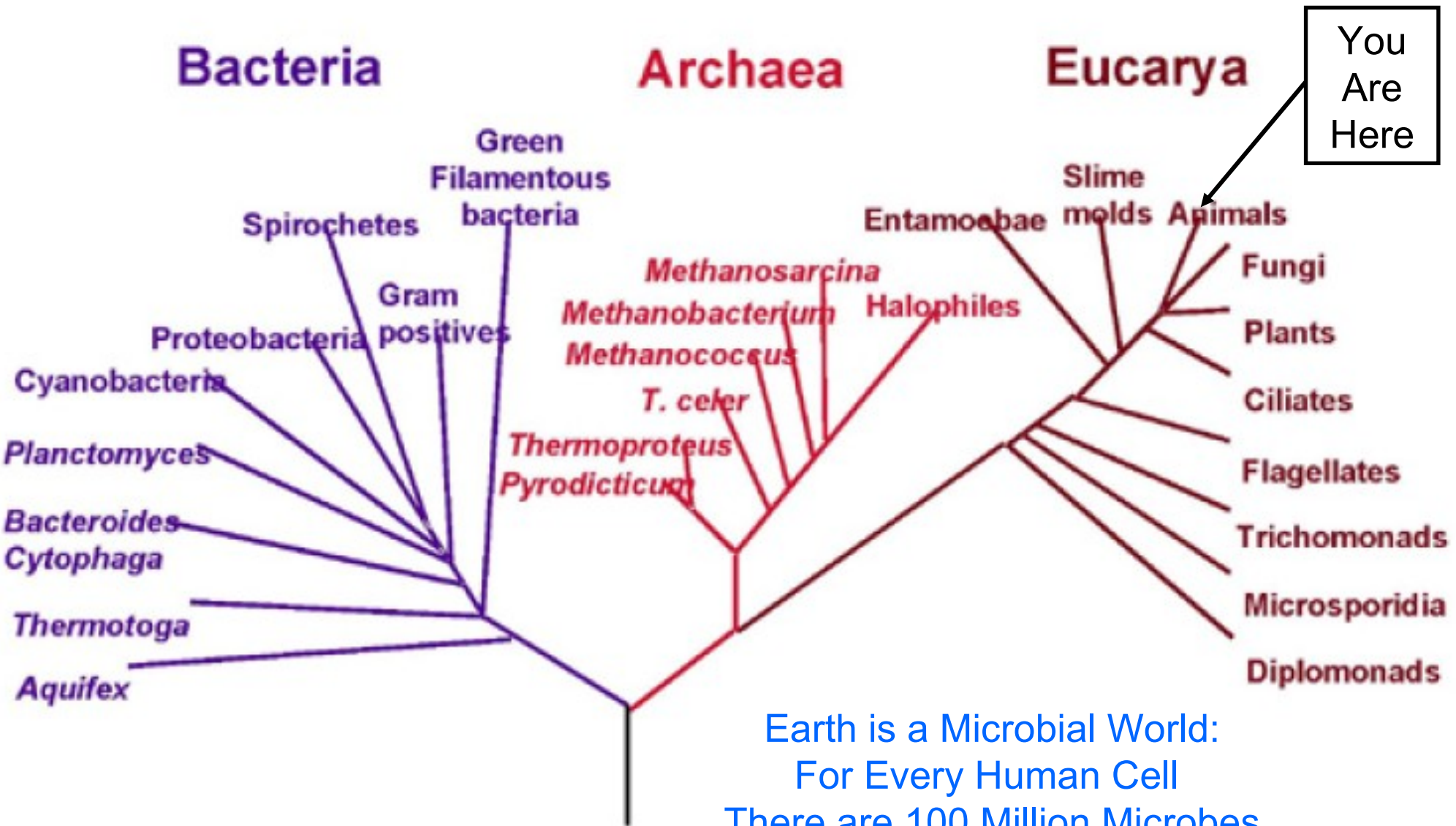
NSF TeraGrid Kraken
Cray XT5
8,256 Compute Nodes
99,072 Compute Cores
129 TB RAM



SDSC

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

Most of Evolutionary Time Was in the Microbial World



Earth is a Microbial World:
For Every Human Cell
There are 100 Million Microbes

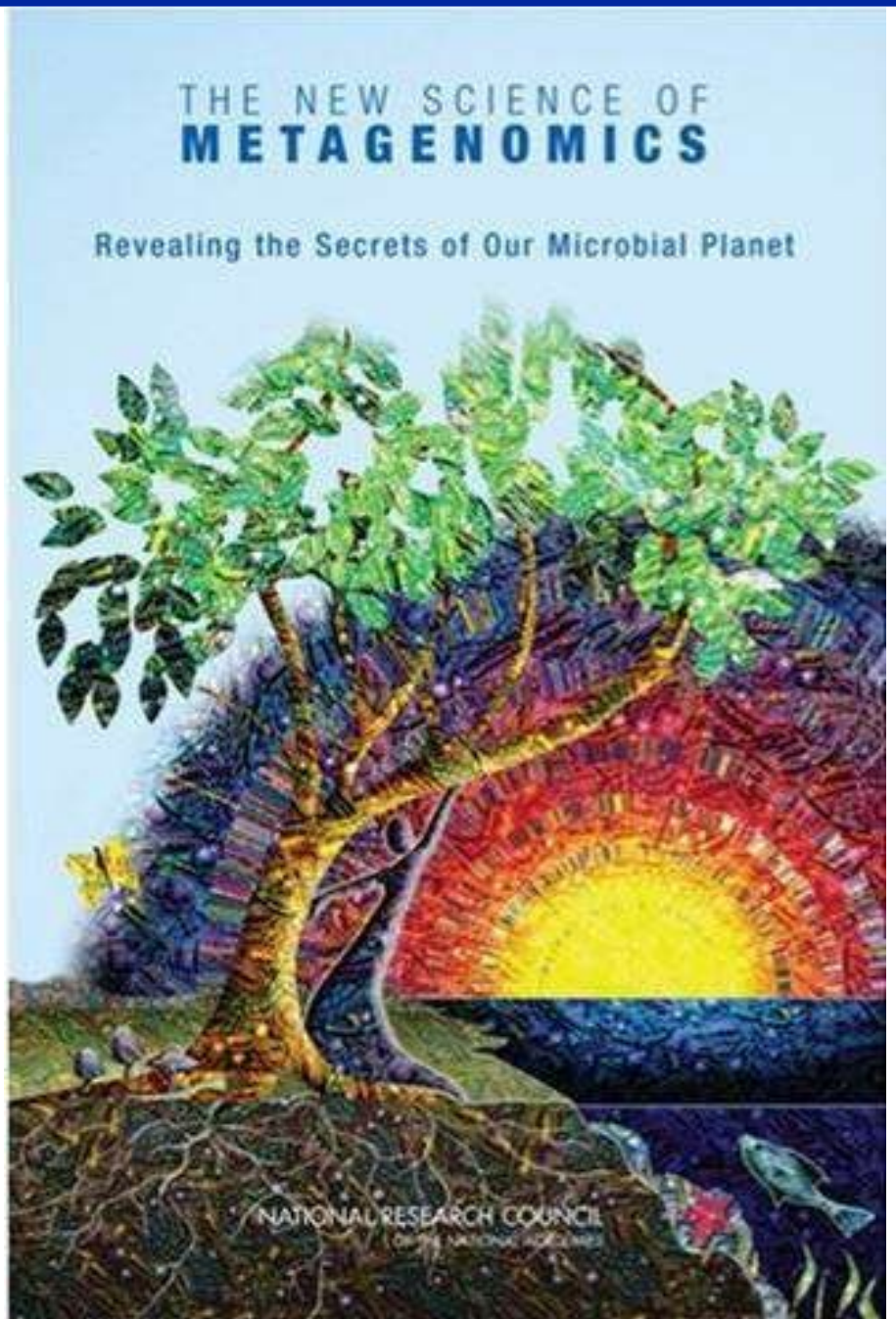
Tree of Life Derived from 16S rRNA Sequences

Source: Carl Woese, et al

The New Science of Microbial Metagenomics

NRC Report:

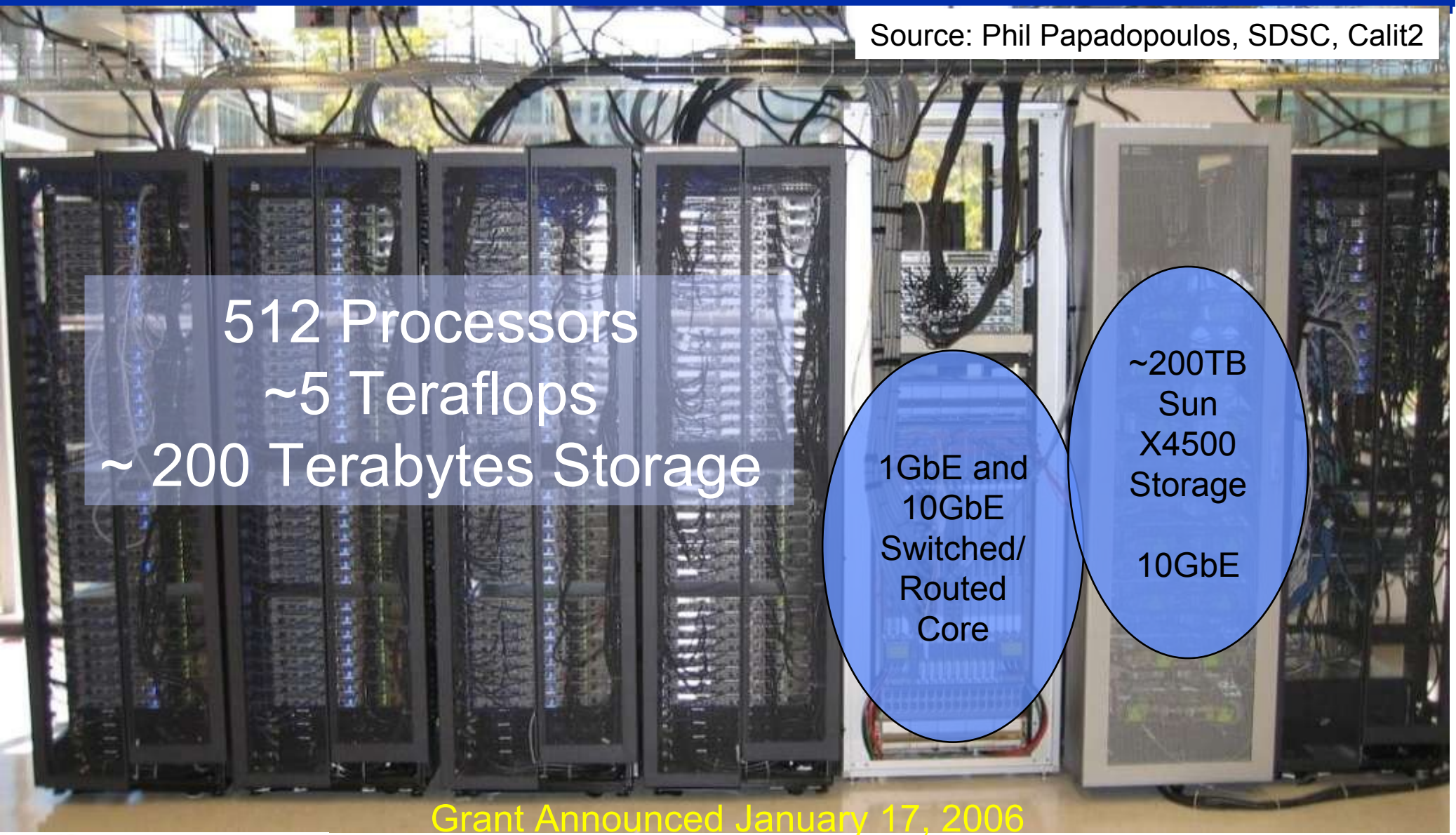
Metagenomic data should be made publicly available in international archives as rapidly as possible.



"The emerging field of metagenomics, where the DNA of entire communities of microbes is studied simultaneously, presents the greatest opportunity – perhaps since the invention of the microscope – to revolutionize understanding of the microbial world." –

National Research Council
March 27, 2007

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



Grant Announced January 17, 2006

camera Marine Microbial Ecology

Gordon and Betty MOORE FOUNDATION

ROCKS

SDSC

J. Craig Venter INSTITUTE

it²

Calit2 CAMERA:
Over 4000 Registered Users From Over 80 Countries

Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis

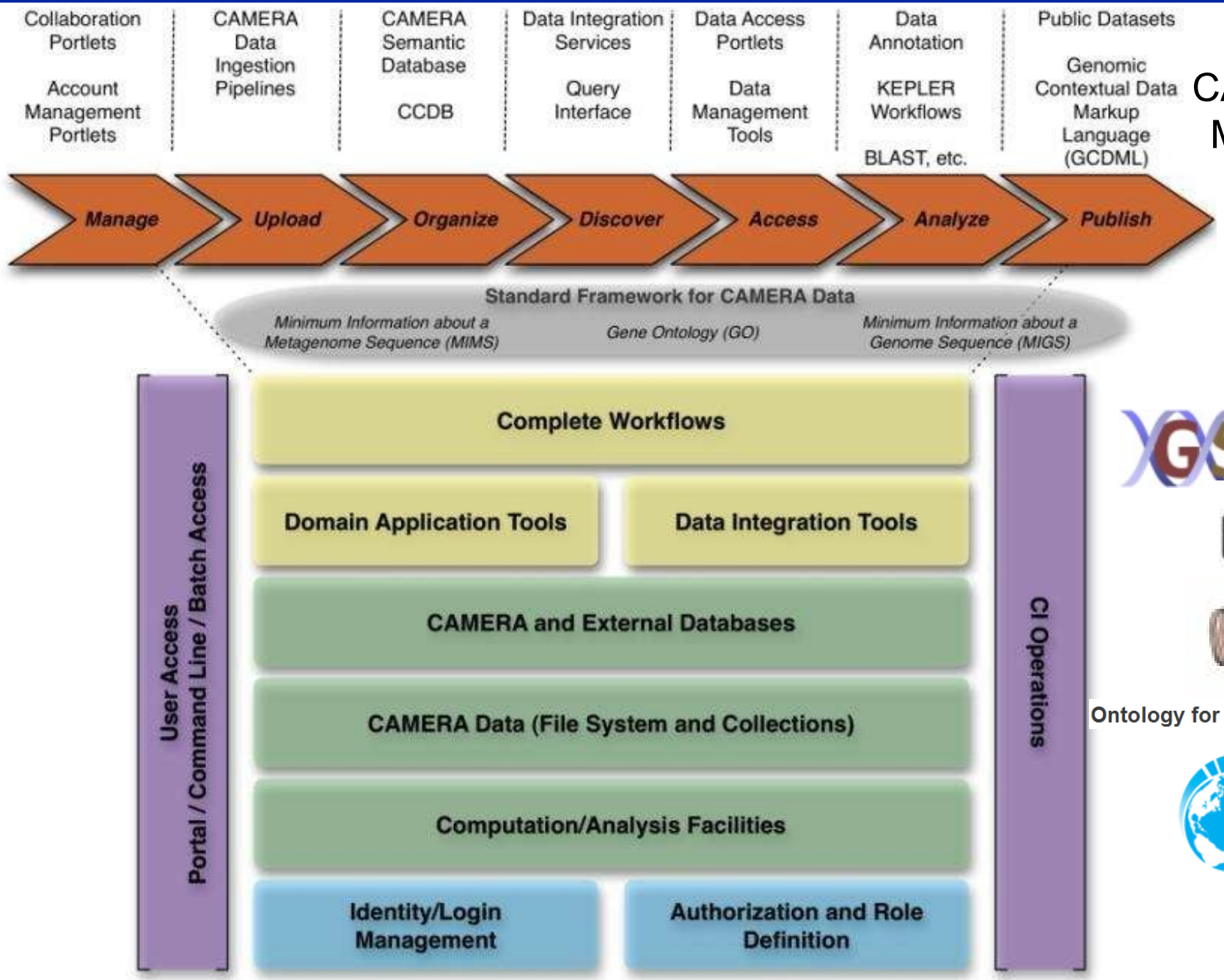


Gordon and Betty MOORE FOUNDATION

camera
Marine Microbial Ecology

http://camera.calit2.net/

Creating CAMERA 2.0 -
Advanced Cyberinfrastructure Service Oriented Architecture



Source:
CAMERA CTO
Mark Ellisman

GSC Genomic Standards Consortium

Envo

OBI
Ontology for Biomedical Investigations

camera
Marine Microbial Ecology

it2

The GreenLight Project: Instrumenting the Energy Cost of Computational Science

- **Focus on 5 Communities with At-Scale Computing Needs:**
  - Metagenomics
  - Ocean Observing
  - Microscopy
  - Bioinformatics
  - Digital Media
- **Measure, Monitor, & Web Publish Real-Time Sensor Outputs**
  - Via Service-oriented Architectures
  - Allow Researchers Anywhere To Study Computing Energy Cost
  - Enable Scientists To Explore Tactics For Maximizing Work/Watt
- **Develop Middleware that Automates Optimal Choice of Compute/RAM Power Strategies for Desired Greenness**
- **Data Center for School of Medicine Illumina Next Gen Sequencer Storage and Processing**





SDSC

Source: Tom DeFanti, Calit2; GreenLight PI

NSF

it²

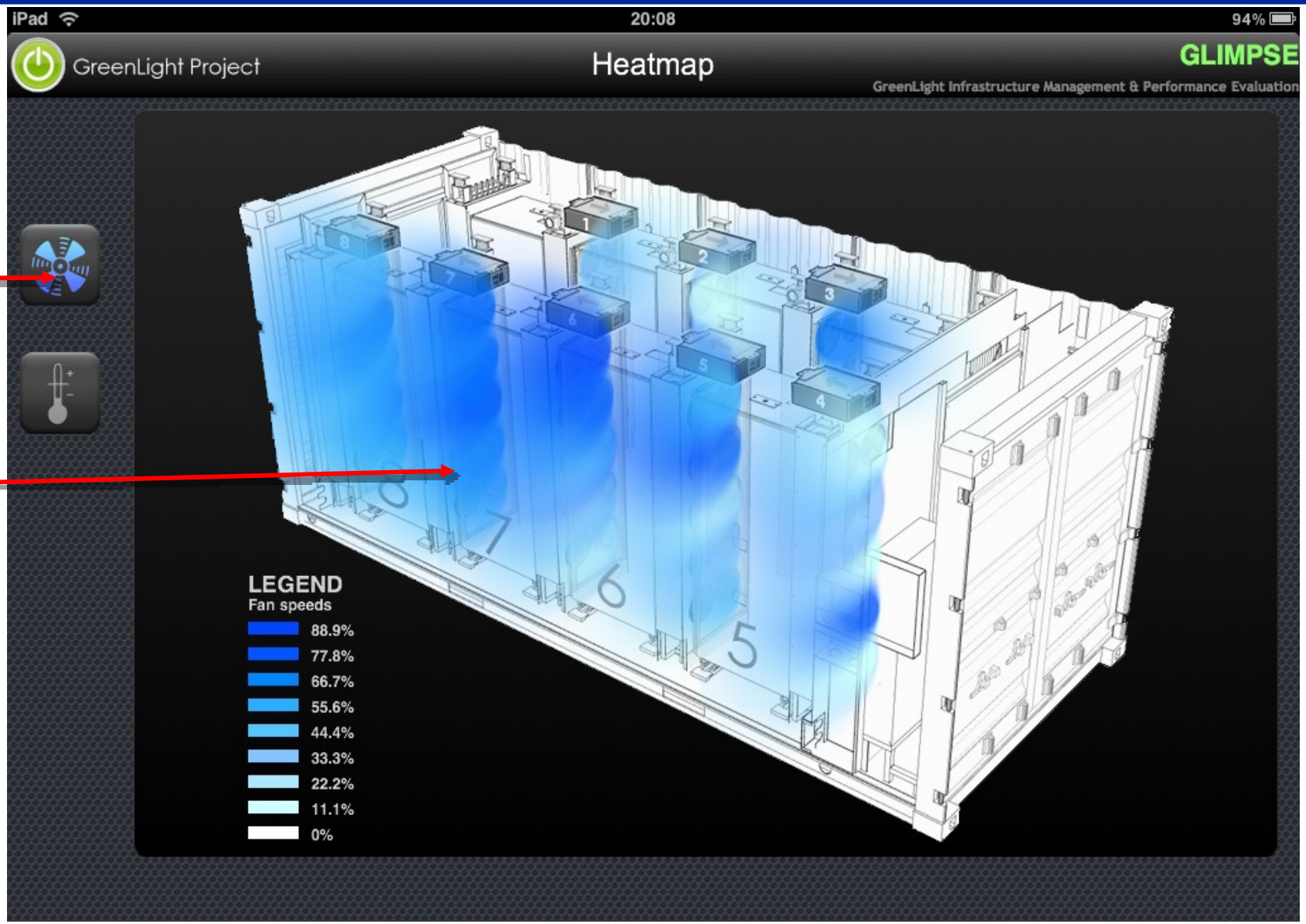
GreenLight Project: Remote Visualization of Data Center



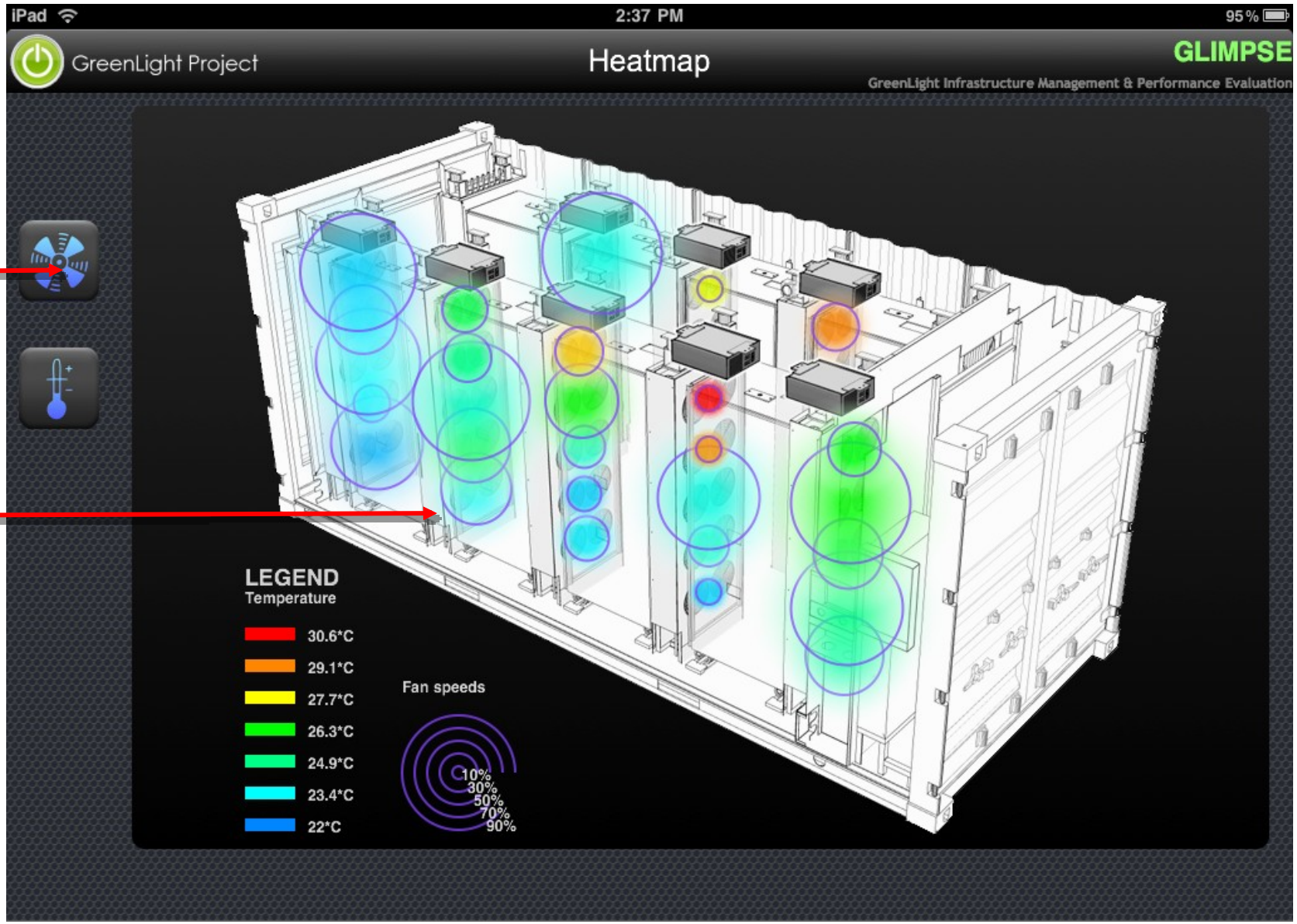
GreenLight Projects
Airflow dynamics

Live fan speeds

Airflow dynamics



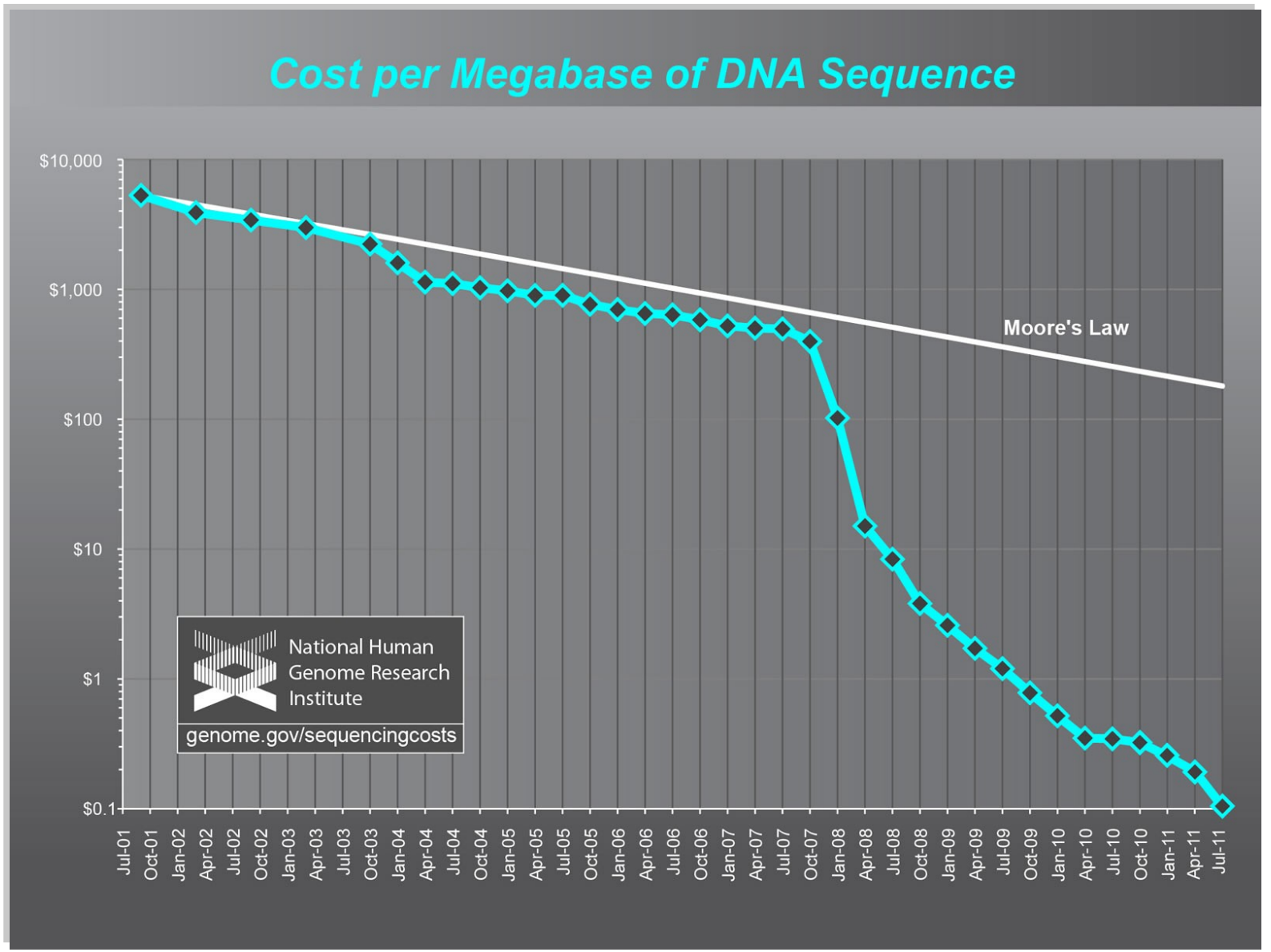
GreenLight Project
Heat Distribution



Combined heat + fans

Realistic correlation

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.!

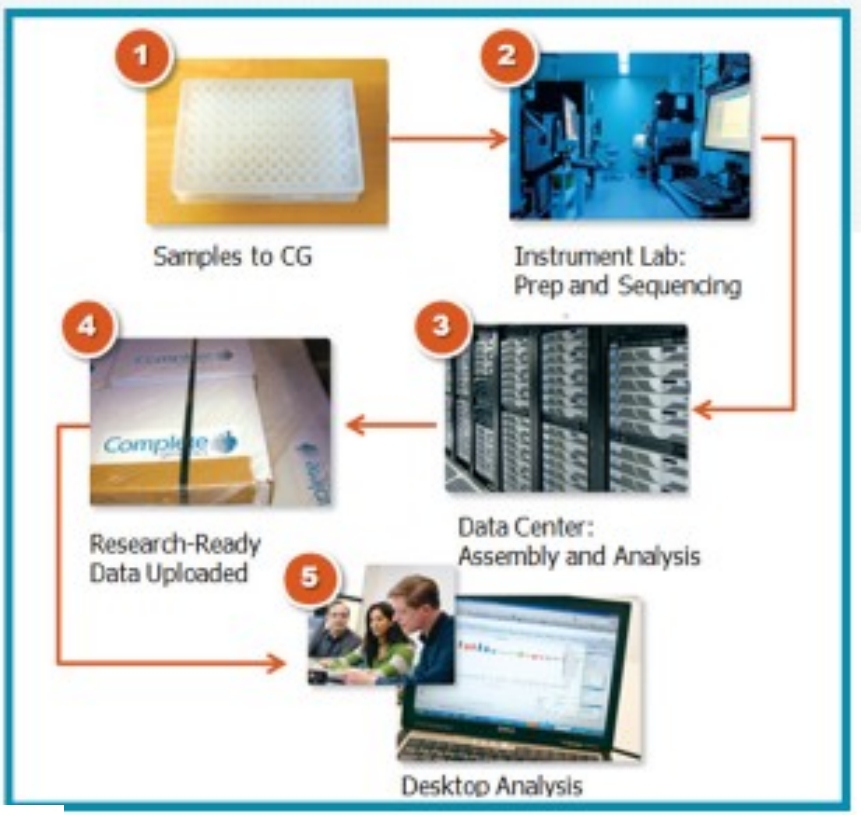
Proprietary Sequencing Technology

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

Calit2 Brings Together Computer Science and Bioinformatics





National Biomedical Computation Resource
an NIH supported resource center













GreenLight Project Allows for Testing of Novel Architectures on Bioinformatics Algorithms



Convey Computer™ Corporation Ships to First Customer: University of California, San Diego

New Convey HC-1™ System to Boost Performance of Bioinformatics Application

June 23, 2009

"Our version of MS-Alignment [a proteomics algorithm] is more than 115x faster than a single core of an Intel Nehalem processor, is more than 15x faster than an eight-core version, and reduces the runtime for a few samples from 24 hours to just a few hours."

—From "Computational Mass Spectrometry in a Reconfigurable Coherent Co-processing Architecture," *IEEE Design & Test of Computers*, Yalamarthy (ECE), Coburn (CSE), Gupta (CSE), Edwards (Convey), and Kelly (Convey) (2011)

http://research.microsoft.com/en-us/um/cambridge/events/date2011/msalignment_dateposter_2011.pdf



Using UCSD RCI to Store and Analyze Next Gen Sequencer Datasets

Data Volume	Total	Final	Comment
HiSeq 2000 200G run			
Image Data	32 TB	0	
Intensity Data	2 TB	0	Optionally transferred
Base Call / Quality Score Data	250 GB	250 GB	1 byte/base (raw) assuming qseq generation offline
Alignment Output	6 TB (3 TB)	1.2 TB	Remove intermediate files
GA_{IIx} 50G run			
Image Data	6.9	0	Optionally transferred
Intensity Data	0.93	0.93	
Base Call / Quality Score Data	0.17	0.17	
Alignment Output	1.2TB	1.2 TB	

Stream Data from Genomics Lab to GreenLight Storage,
NFS Mount Over 10Gbps to Triton Compute Cluster

Source: Chris Misleh, SOM/Calit2 UCSD

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

NCMIR

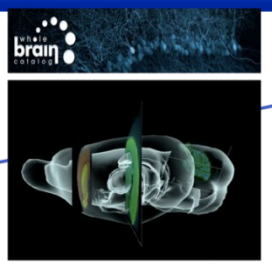
Telemicroscopy Resources:
Osaka Univ., Oxford Univ.,
The Jackson Labs





Neuroinformatics Resources



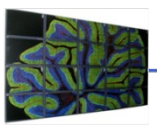



Shared Infrastructure

Development Systems

Storage Resources

OptIPortal Tiled Display









California Institute of Telecommunications and Information Technology

Storage Resources

Production Systems

Triton Petascale Resource







San Diego Supercomputer Center



Scientific Instruments

Ultra-wide Field Intra-vital LM Instruments



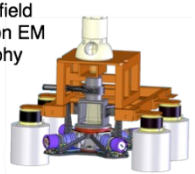
Super-resolution PALM





Serial Blockface Scanning EM

Ultra-wide field Serial Section EM Tomography





Multi-color EM, EFTEM

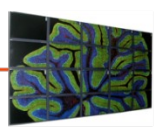
Wide-field TEM/TEM



End User Workstations



Graphics Workstations



OptIPortal Tiled Display

Local SOM Infrastructure



NCMIR Data Storage Servers



Critical Servers (Mail, Web, DB)



Large Memory Servers



NBCR Compute Clusters



GreenLight Project

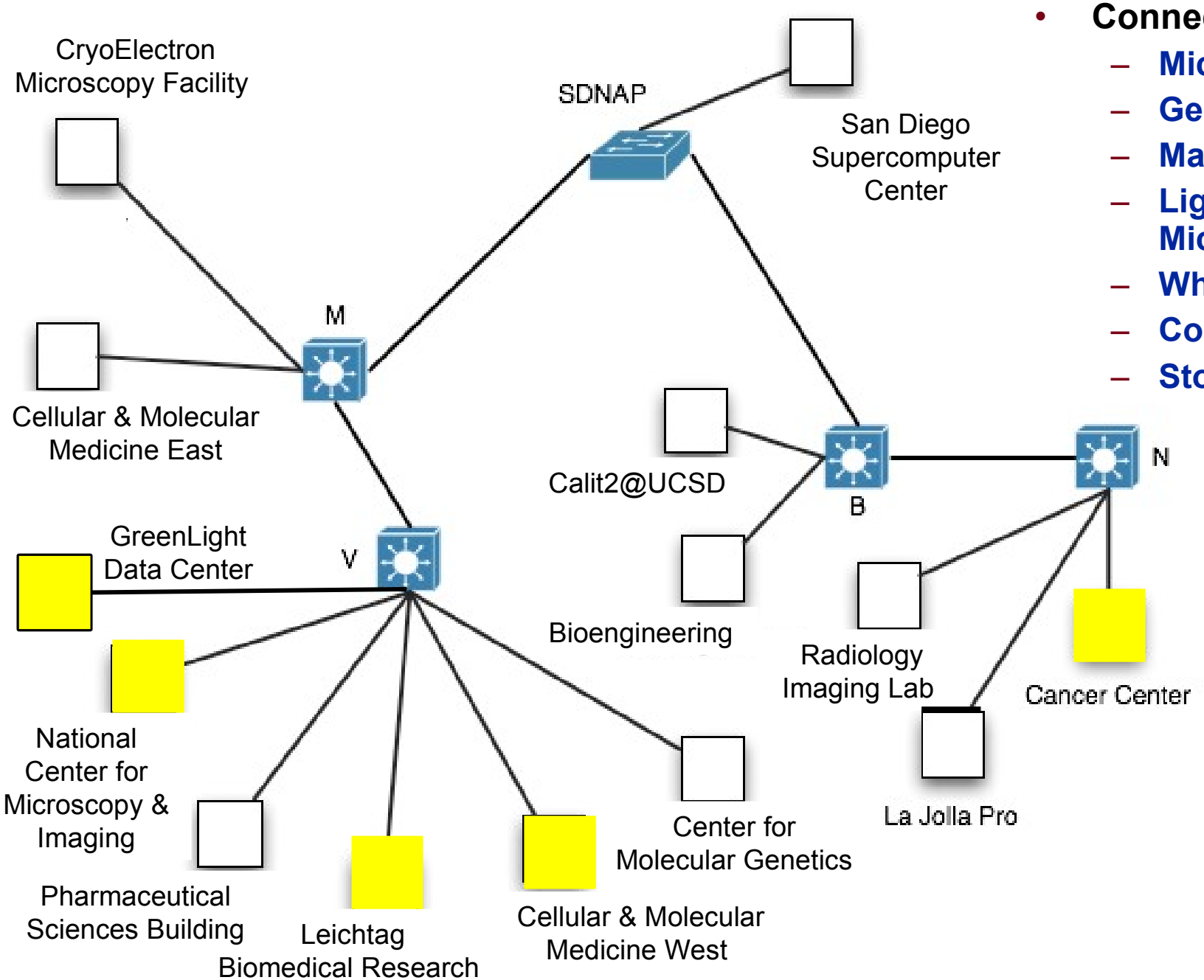
GPU Compute Clusters

SDSC

it²

Source: Steve Peltier, Mark Ellisman, NCMIR

UCSD Planned Optical Networked Biomedical Researchers and Instruments



- **Connects at 10 Gbps :**
  - **Microarrays**
  - **Genome Sequencers**
  - **Mass Spectrometry**
  - **Light and Electron Microscopes**
  - **Whole Body Imagers**
  - **Computing**
  - **Storage**