"High Performance Cyberinfrastructure Enables Data-Driven Science in the Globally Networked World"

Invited Speaker

Grand Challenges in Data-Intensive Discovery Conference

San Diego Supercomputer Center, UC San Diego

La Jolla, CA

October 28, 2010

Dr. Larry Smarr

Director, California Institute for Telecommunications and Information Technology

Harry E. Gruber Professor, Dept. of Computer Science and Engineering

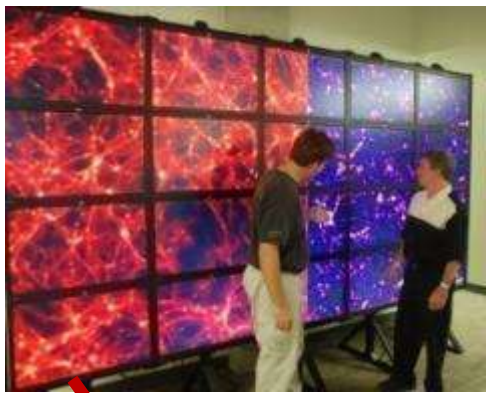
Jacobs School of Engineering, UCSD

Follow me on Twitter: lsmarr

Abstract

Today we are living in a data-dominated world where distributed scientific instruments, as well as supercomputers, generate terabytes to petabytes of data. It was in response to this challenge that the NSF funded the OptIPuter project to research how user-controlled 10Gbps dedicated lightpaths (or "lambdas") could provide direct access to global data repositories, scientific instruments, and computational resources from "OptIPortals," PC clusters which provide scalable visualization, computing, and storage in the user's campus laboratory. The use of dedicated lightpaths over fiber optic cables enables individual researchers to experience "clear channel" 10,000 megabits/sec, 100-1000 times faster than over today's shared Internet—a critical capability for data-intensive science. The seven-year OptIPuter computer science research project is now over, but it stimulated a national and global build-out of dedicated fiber optic networks. U.S. universities now have access to high bandwidth lambdas through the National LambdaRail, Internet2's WaveCo, and the Global Lambda Integrated Facility. A few pioneering campuses are now building on-campus lightpaths to connect the data-intensive researchers, data generators, and vast storage systems to each other on campus, as well as to the national network campus gateways. I will give examples of the application use of this emerging high performance cyberinfrastructure in genomics, ocean observatories, radio astronomy, and cosmology.

Academic Research "OptIPlatform" Cyberinfrastructure: A 10Gbps "End-to-End" Lightpath Cloud



End User OptIPortal



Instruments



HD/4k Telepresence



HD/4k Video Cams

HPC



10G Lightpaths

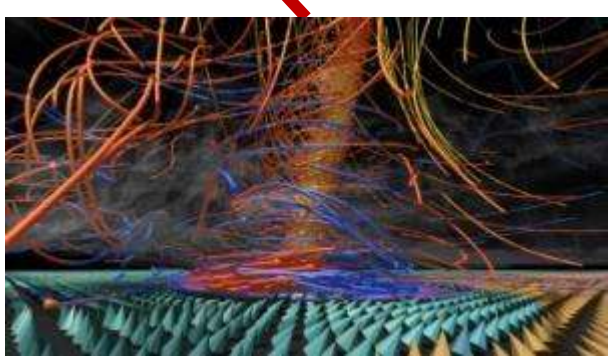


TeraGrid

Campus Optical Switch

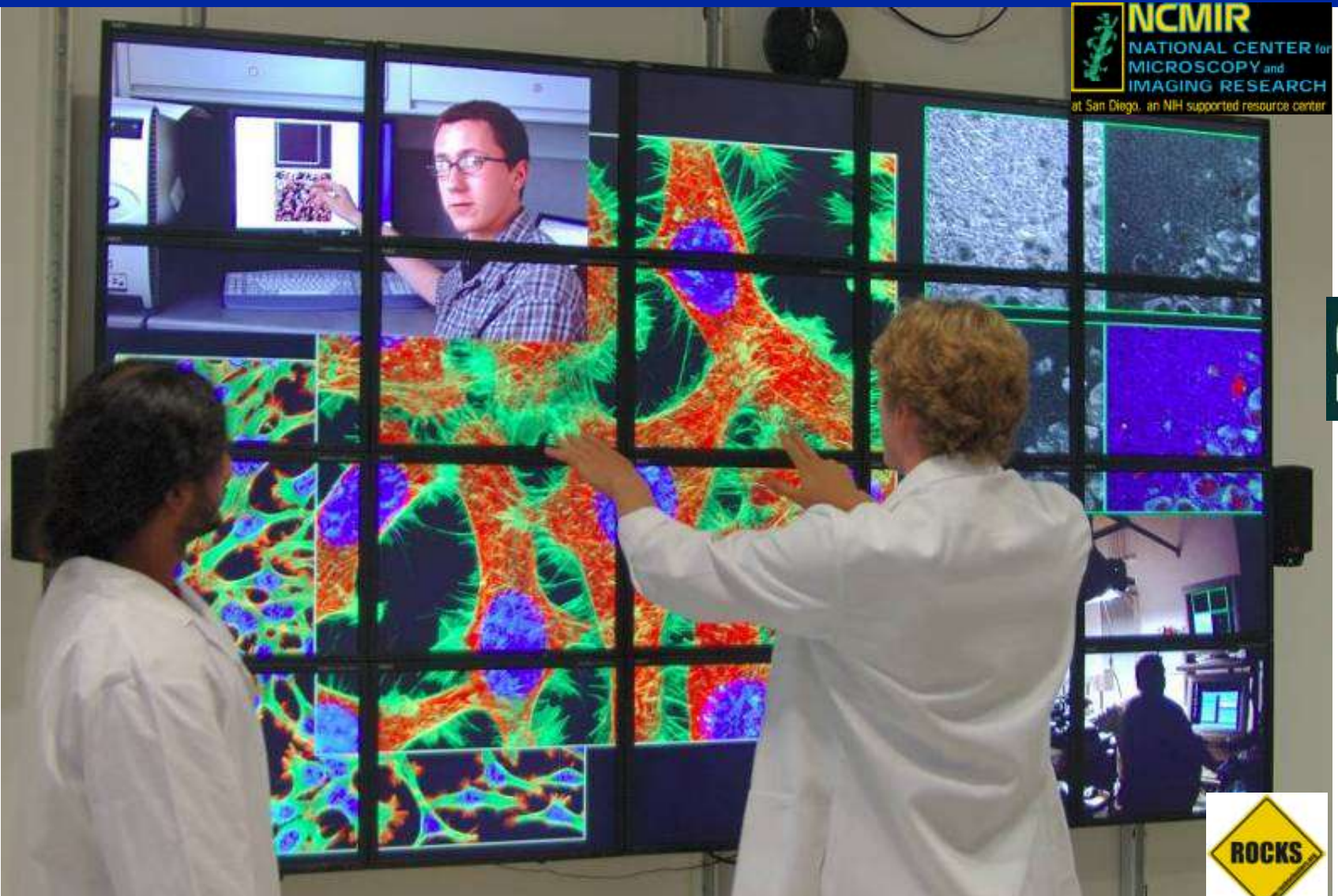


Data Repositories & Clusters



HD/4k Video Images

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

Picture Source: Mark Ellisman, David Lee, Jason Leigh

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

On-Line Resources
Help You Build Your Own OptIPortal

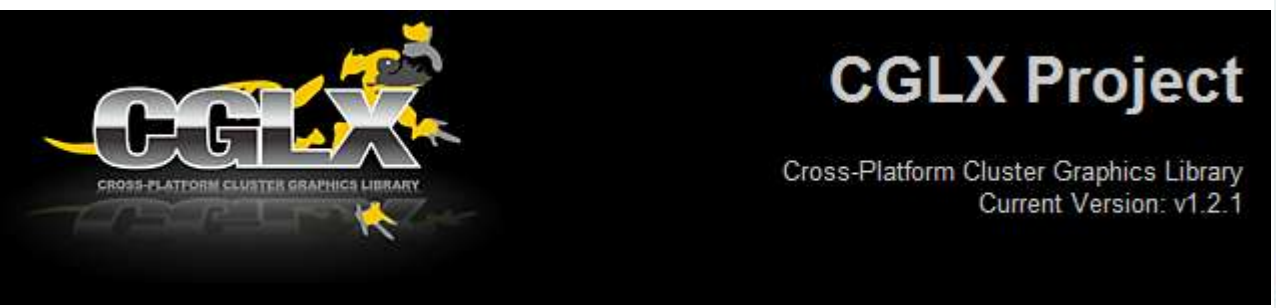
www.optiputer.net

http://wiki.optiputer.net/optiportal



www.evl.uic.edu/cavern/sage/





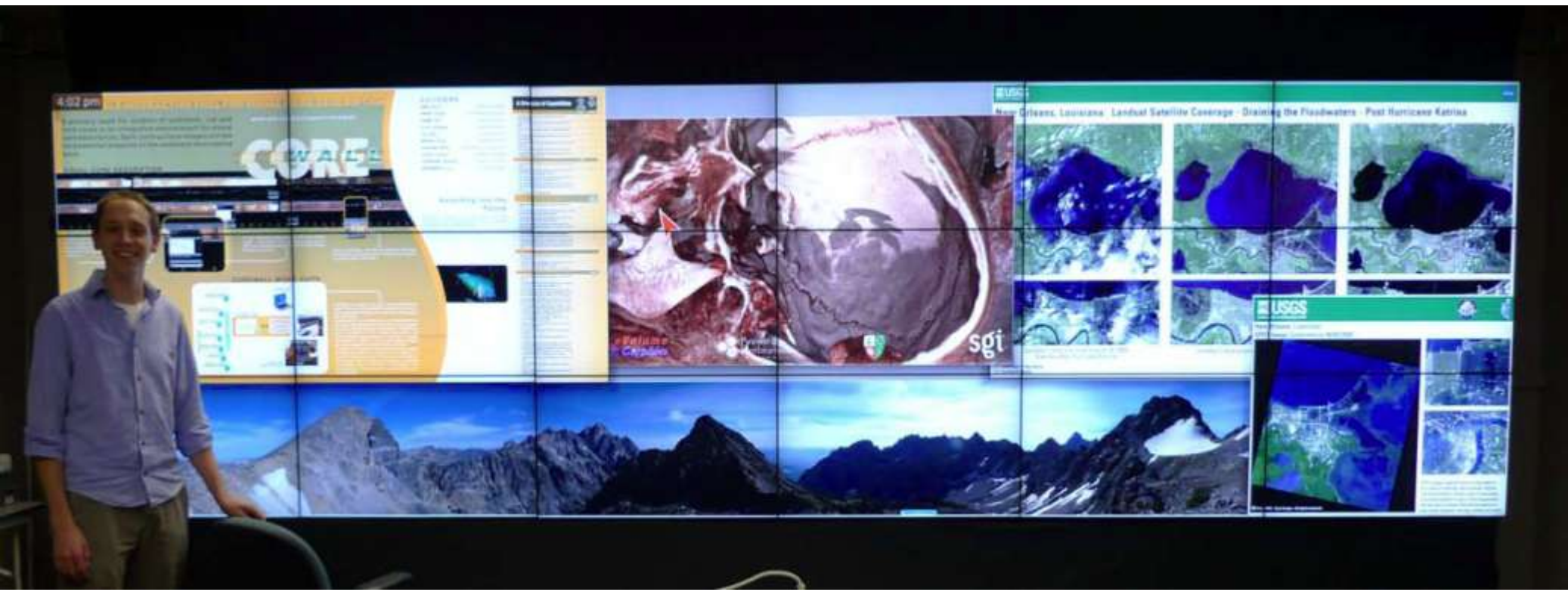
http://vis.ucsd.edu/~cglx/



OptIPortals Are Built
From Commodity PC Clusters and LCDs
To Create a 10Gbps Scalable Termination Device

Nearly Seamless AESOP OptIPortal

46” NEC Ultra-Narrow Bezel 720p LCD Monitors



Source: Tom DeFanti, Calit2@UCSD;

3D Stereo Head Tracked OptIPortal: NexCAVE



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

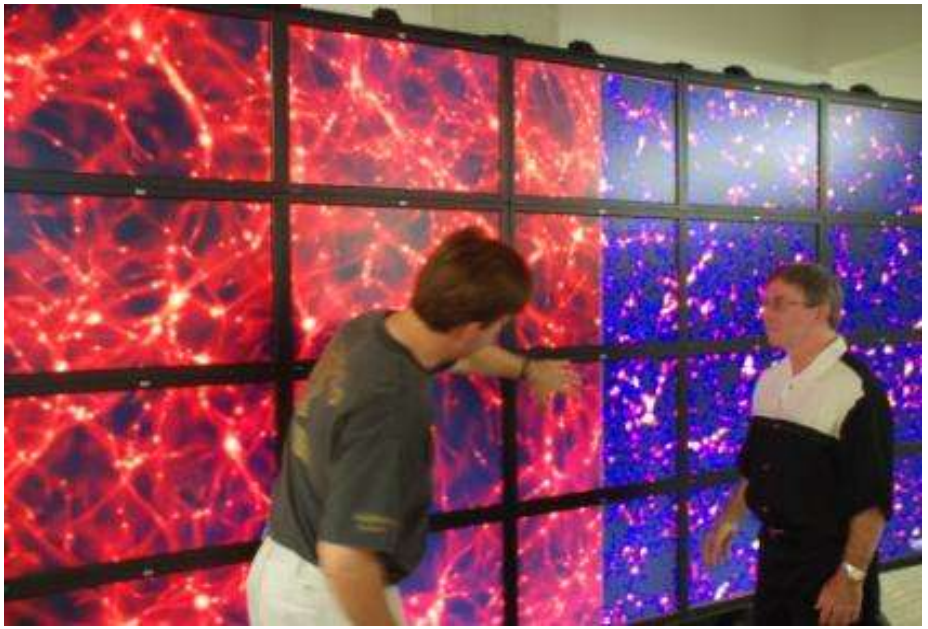
www.calit2.net/newsroom/article.php?id=1584

Source: Tom DeFanti, Calit2@UCSD

Project StarGate Goals: Combining Supercomputers and Supernetworks

- Create an "End-to-End" 10Gbps Workflow
- Explore Use of OptIPortals as Petascale Supercomputer "Scalable Workstations"
- Exploit Dynamic 10Gbps Circuits on ESnet
- Connect Hardware Resources at ORNL, ANL, SDSC
- Show that Data Need Not be Trapped by the Network "Event Horizon"

OptIPortal@SDSC



Rick Wagner Mike Norman

Source: Michael Norman, SDSC, UCSD

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

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



SDSC

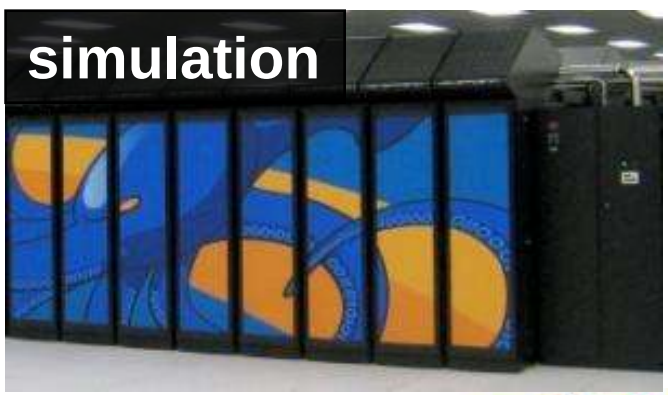


ESnet
10 Gb/s fiber optic network

NICS ORNL

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

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



SDSC

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

it²

National-Scale Interactive Remote Rendering of Large Datasets

SDSC



OptIPortal (40M pixels LCDs)
10 NVIDIA FX 4600 Cards
10 Gb/s Network Throughout

ESnet

Science Data Network (SDN)
> 10 Gb/s Fiber Optic Network
Dynamic VLANs Configured
Using OSCARS

ALCF



Eureka
100 Dual Quad Core Xeon Servers
200 NVIDIA FX GPUs
3.2 TB RAM

Interactive Remote Rendering

Real-Time Volume Rendering Streamed from ANL to SDSC

Last Year

High-Resolution (4K+, 15+ FPS)—But:
- Command-Line Driven
- Fixed Color Maps, Transfer Functions
- Slow Exploration of Data

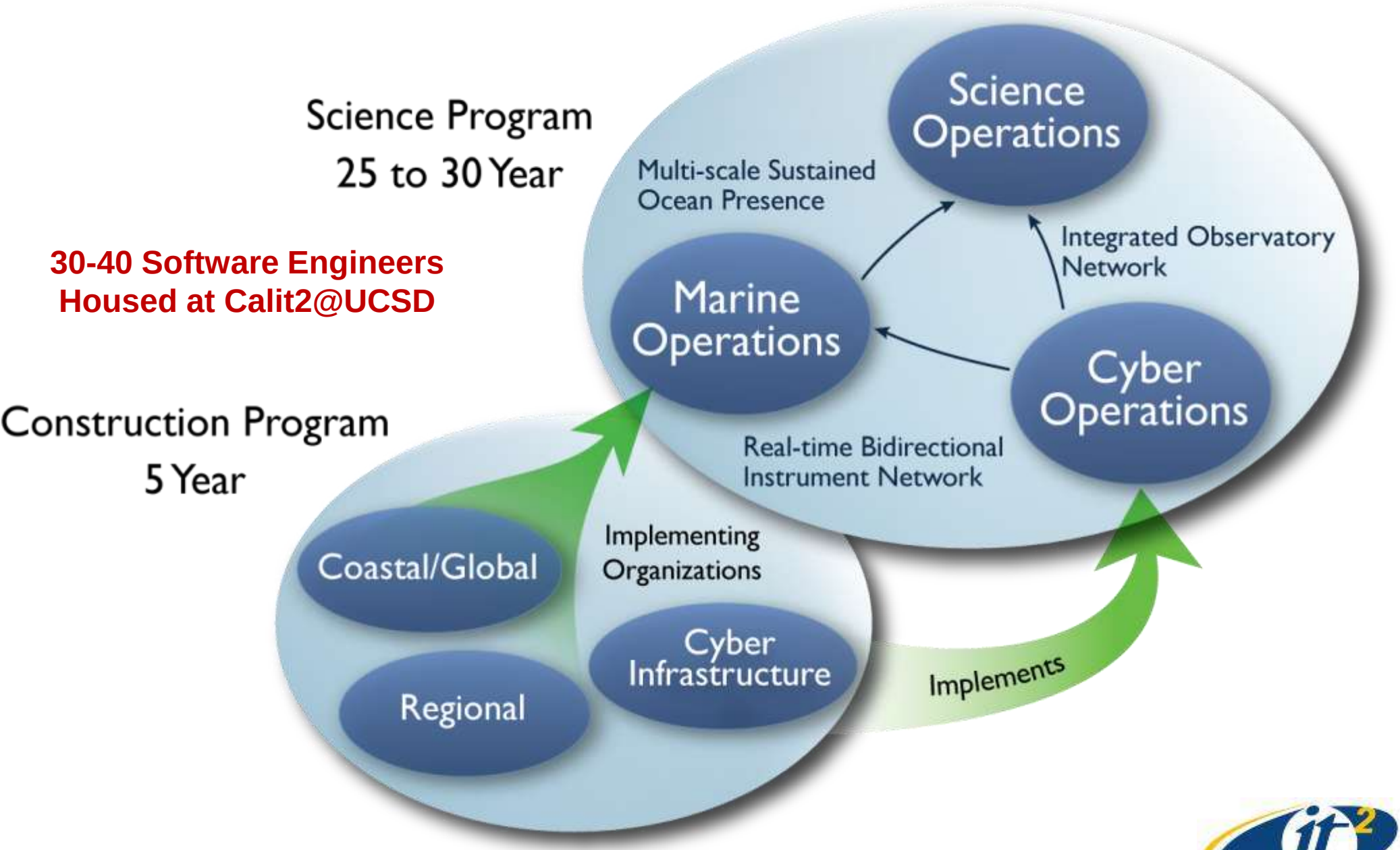
Last Week

Now Driven by a Simple Web GUI
- Rotate, Pan, Zoom
- GUI Works from Most Browsers
- Manipulate Colors and Opacity
- Fast Renderer Response Time

Source: Rick Wagner, SDSC

NSF OOI is a $400M Program -OOI CI is $34M Part of This

OCEAN OBSERVATORIES INITIATIVE



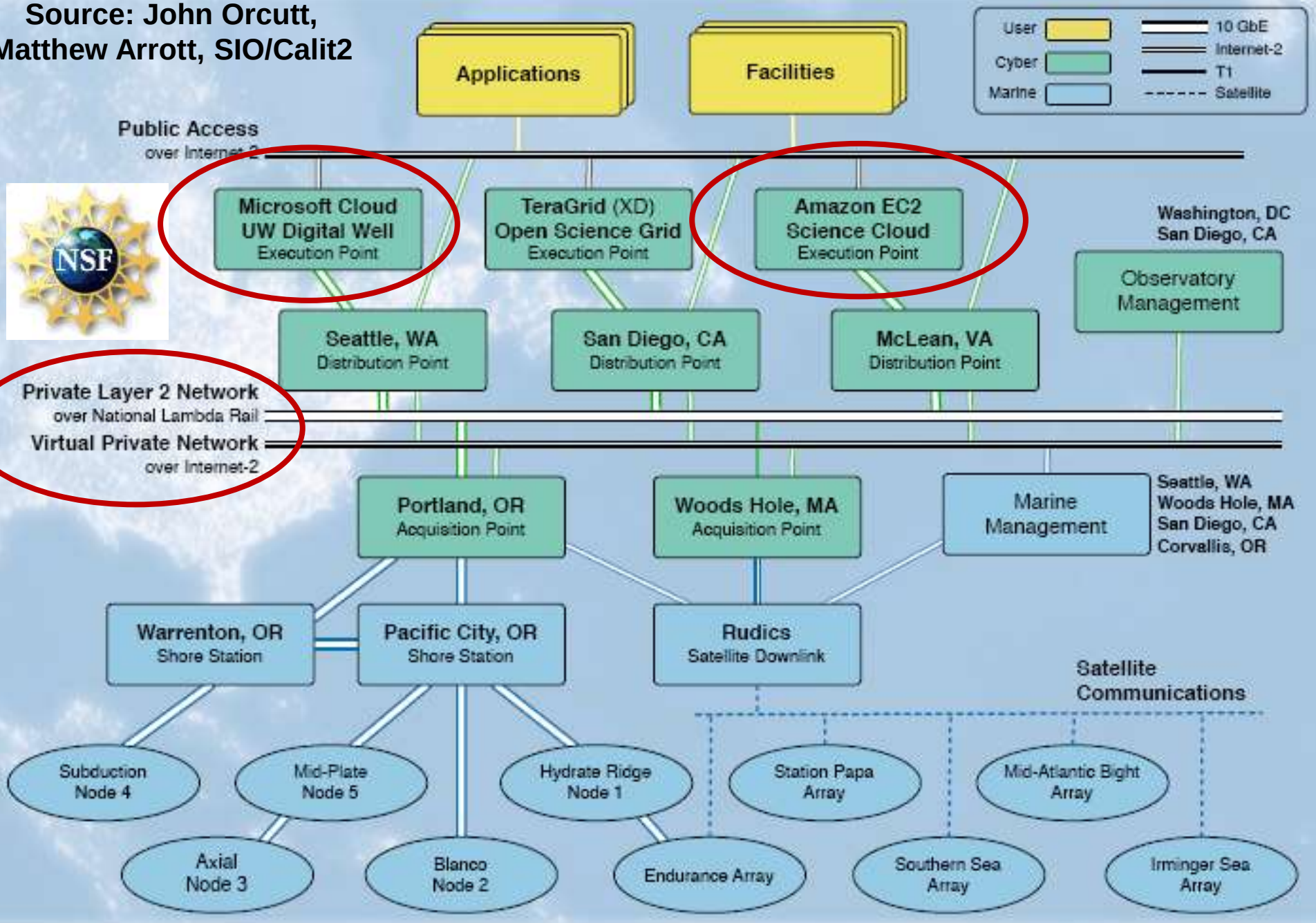
30-40 Software Engineers Housed at Calit2@UCSD

Source: Matthew Arrott, Calit2 Program Manager for OOI CI

OOI CI
is Built on NLR/I2 Optical Infrastructure

OCEAN OBSERVATORIES INITIATIVE

Source: John Orcutt, Matthew Arrott, SIO/Calit2



California and Washington Universities Are Testing a 10Gbps Connected Commercial Data Cloud

- **Amazon Experiment for Big Data**
  - **Only Available Through CENIC & Pacific NW GigaPOP**
    - **Private 10Gbps Peering Paths**
  - **Includes Amazon EC2 Computing & S3 Storage Services**
- **Early Experiments Underway**
  - **Robert Grossman, Open Cloud Consortium**
  - **Phil Papadopoulos, Calit2/SDSC Rocks**

amazon web services™

cenic

PACIFIC NORTHWEST GIGAPOP

it²

Open Cloud OptIPuter Testbed--Manage and Compute Large Datasets Over 10Gbps Lambdas



- 9 Racks
- 500 Nodes
- 1000+ Cores
- 10+ Gb/s Now
- Upgrading Portions to 100 Gb/s in 2010/2011

Calit2 — CENIC

NLR C-Wave

Dragon

STARLIGHT SM
The Optical STAR TAP SM

MREN

UIC

Open Source SW

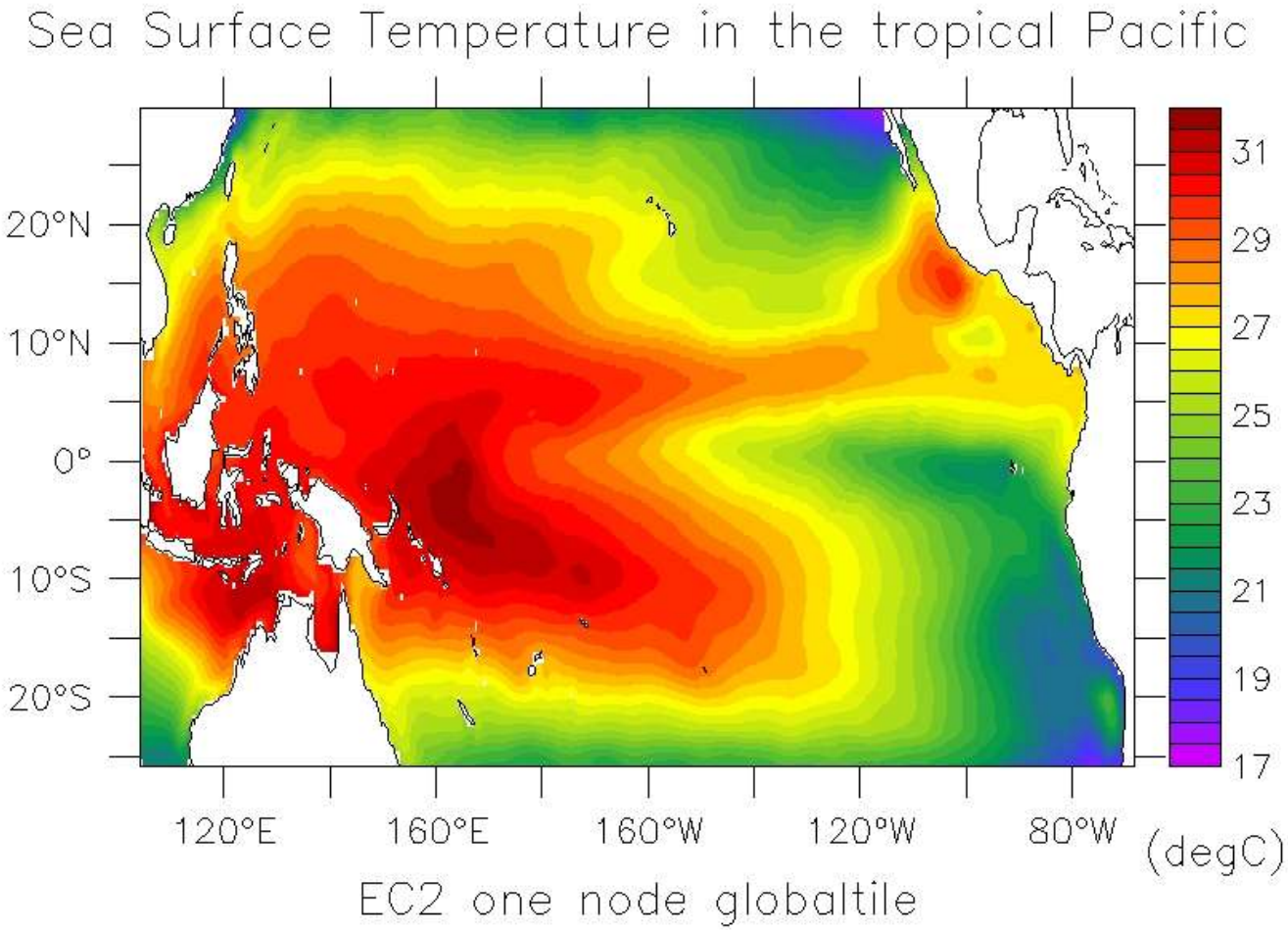
- Hadoop
- Sector/Sphere
- Nebula
- Thrift, GPB
- Eucalyptus
- Benchmarks

Open Cloud Consortium

Source: Robert Grossman, UChicago

Ocean Modeling HPC In the Cloud: Tropical Pacific SST (2 Month Ave 2002)

MIT GCM 1/3 Degree Horizontal Resolution, 51 Levels, Forced by NCEP2. Grid is 564x168x51, Model State is T,S,U,V,W and Sea Surface Height



Run on EC2 HPC Instance. In Collaboration with OOI CI/Calit2

Source: B. Cornuelle, N. Martinez, C.Papadopoulos COMPAS, SIO

Run Timings of Tropical Pacific: Local SIO ATLAS Cluster and Amazon EC2 Cloud

	ATLAS Ethernet NFS	ATLAS Myrinet, NFS	ATLAS Myrinet Local Disk	EC2 HPC Ethernet 1 Node	EC2 HPC Ethernet Local Disk
Wall Time*	4711	2986	2983	14428	2379
User Time*	3833	2953	2933	1909	1590
System Time*	798	17	19	2764	750

*All times in Seconds

Atlas: 128 Node Cluster @ SIO COMPAS. Myrinet 10G, 8GB/node, ~3yrs old
EC2: HPC Computing Instance, 2.93GHz Nehalem, 24GB/Node, 10GbE

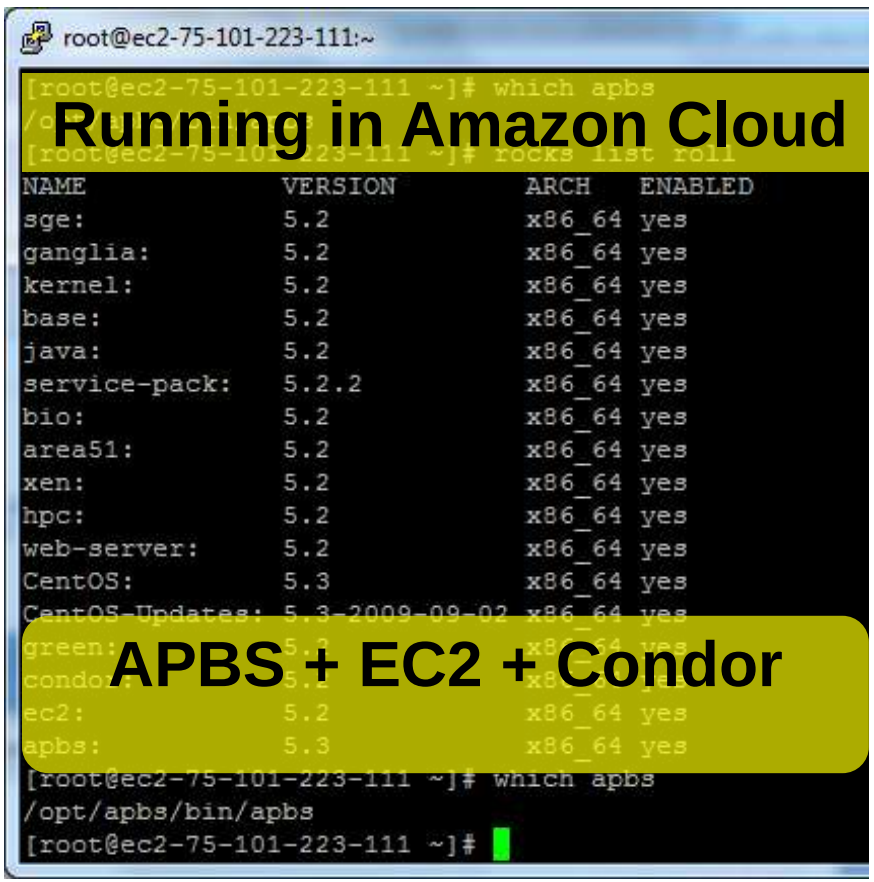
Compilers: Ethernet – GNU FORTRAN with OpenMPI
Myrinet – PGI FORTRAN with MPICH1

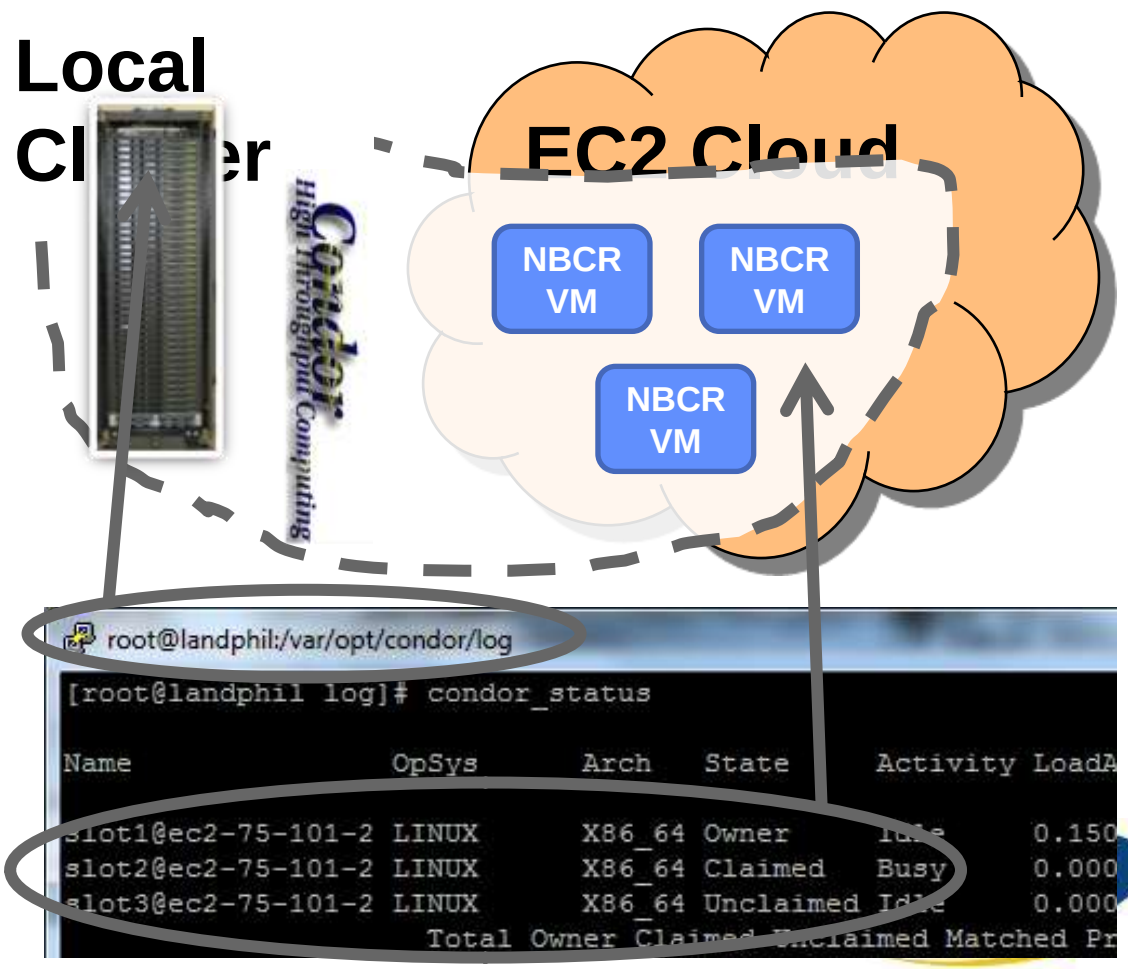
Single Node EC2 was Oversubscribed, 48 Process. All Other Parallel Instances used 6 Physical Nodes, 8 Cores/Node. Model Code has been Ported to Run on ATLAS, Triton (@SDSC) and in EC2.

Source: B. Cornuelle, N. Martinez, C.Papadopoulos COMPAS, SIO

Using Condor and Amazon EC2 on Adaptive Poisson-Boltzmann Solver (APBS)

- APBS Rocks Roll (NBCR) + EC2 Roll + Condor Roll = Amazon VM
- Cluster extension into Amazon using Condor





SDSC

Source: Phil Papadopoulos, SDSC/Calit2

Moving into the Clouds: Rocks and EC2

- **We Can Build Physical Hosting Clusters & Multiple, Isolated Virtual Clusters:**
  - **Can I Use Rocks to Author "Images" Compatible with EC2? (We Use Xen, They Use Xen)**
  - **Can I Automatically Integrate EC2 Virtual Machines into My Local Cluster (Cluster Extension)**
    - **Submit Locally**
    - **My Own Private + Public Cloud**
- **What This Will Mean**
  - **All your Existing Software Runs Seamlessly Among Local and Remote Nodes**
  - **User Home Directories Can Be Mounted**
  - **Queue Systems Work**
  - **Unmodified MPI Works**

SDSC

Source: Phil Papadopoulos, SDSC/Calit2

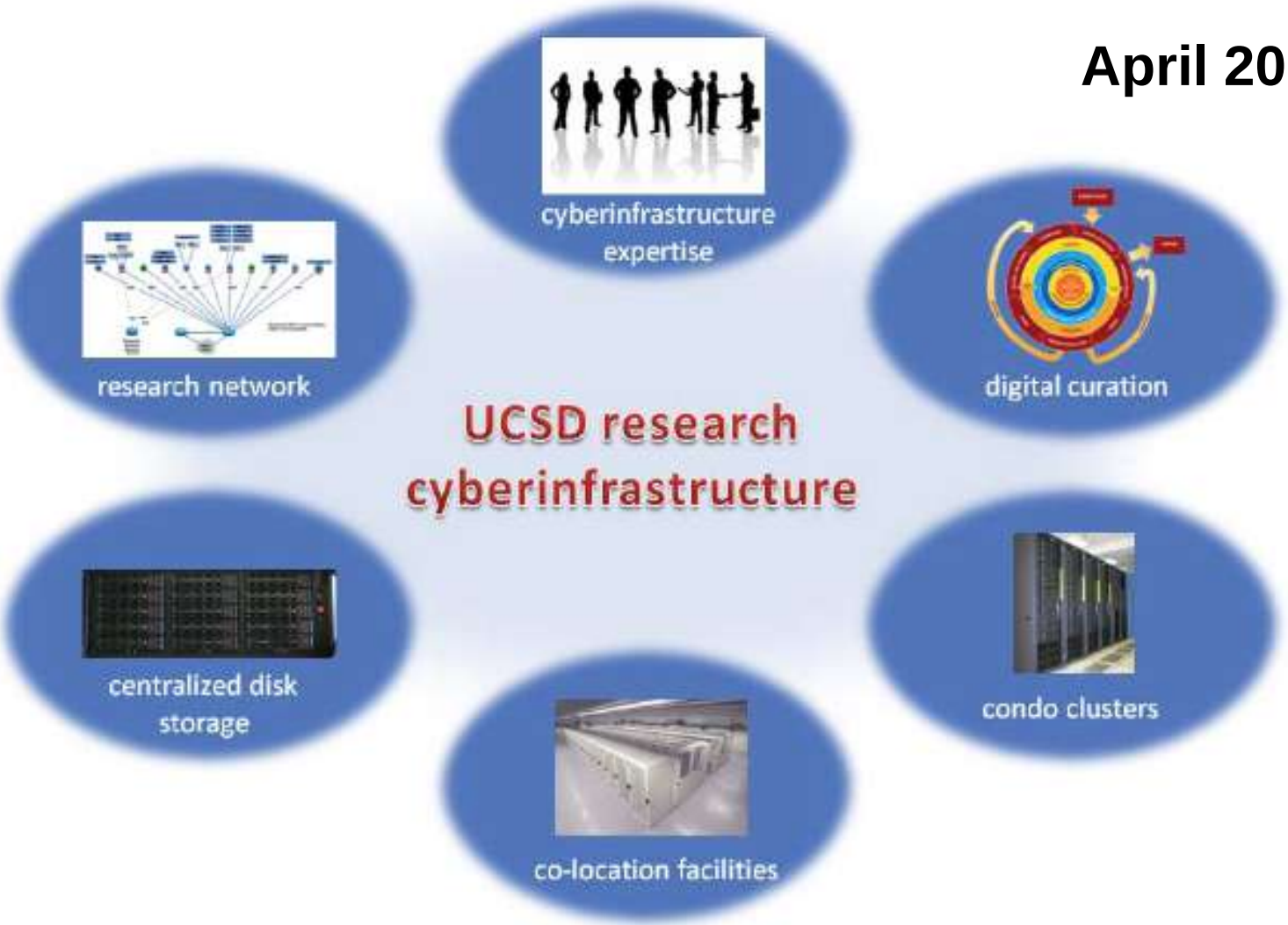
it²

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

- **Focus on Data-Intensive Cyberinfrastructure**

April 2009

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

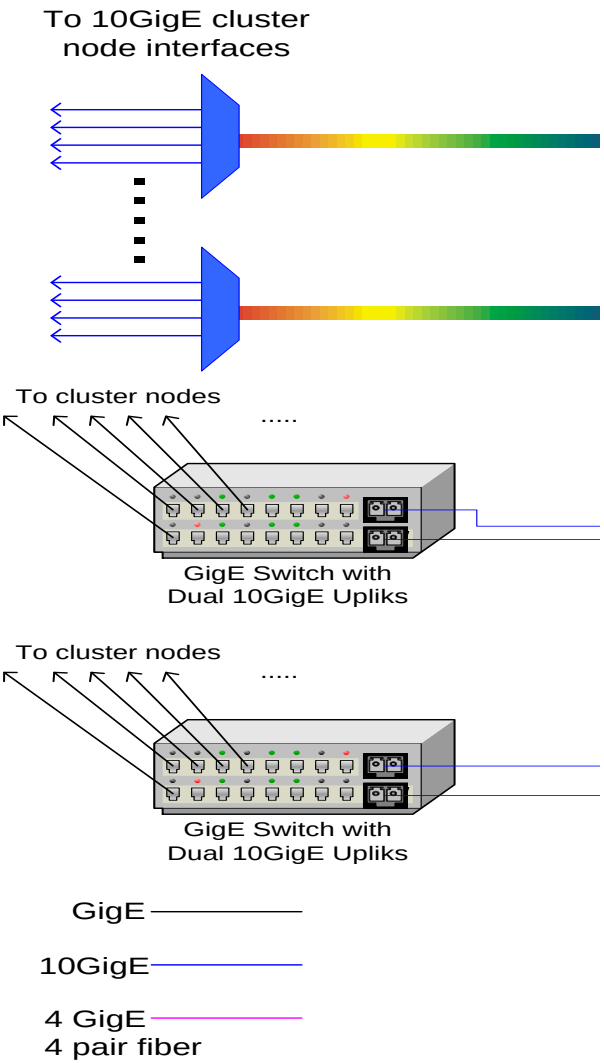


http://research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf

SDSC

it²

Current UCSD Optical Core: Bridging End-Users to CENIC L1, L2, L3 Services



Enpoints:

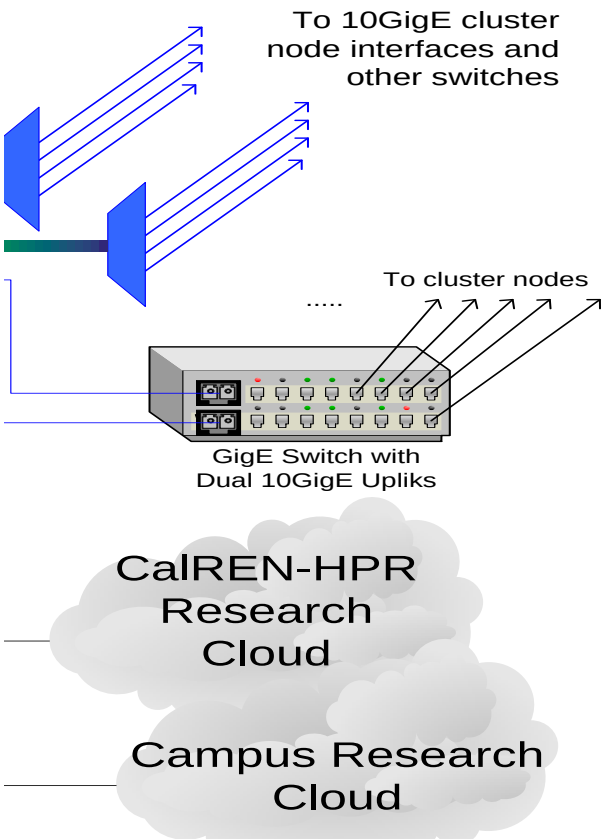
>= 60 endpoints at 10 GigE

>= 32 Packet switched

>= 32 Switched wavelengths

>= 300 Connected endpoints

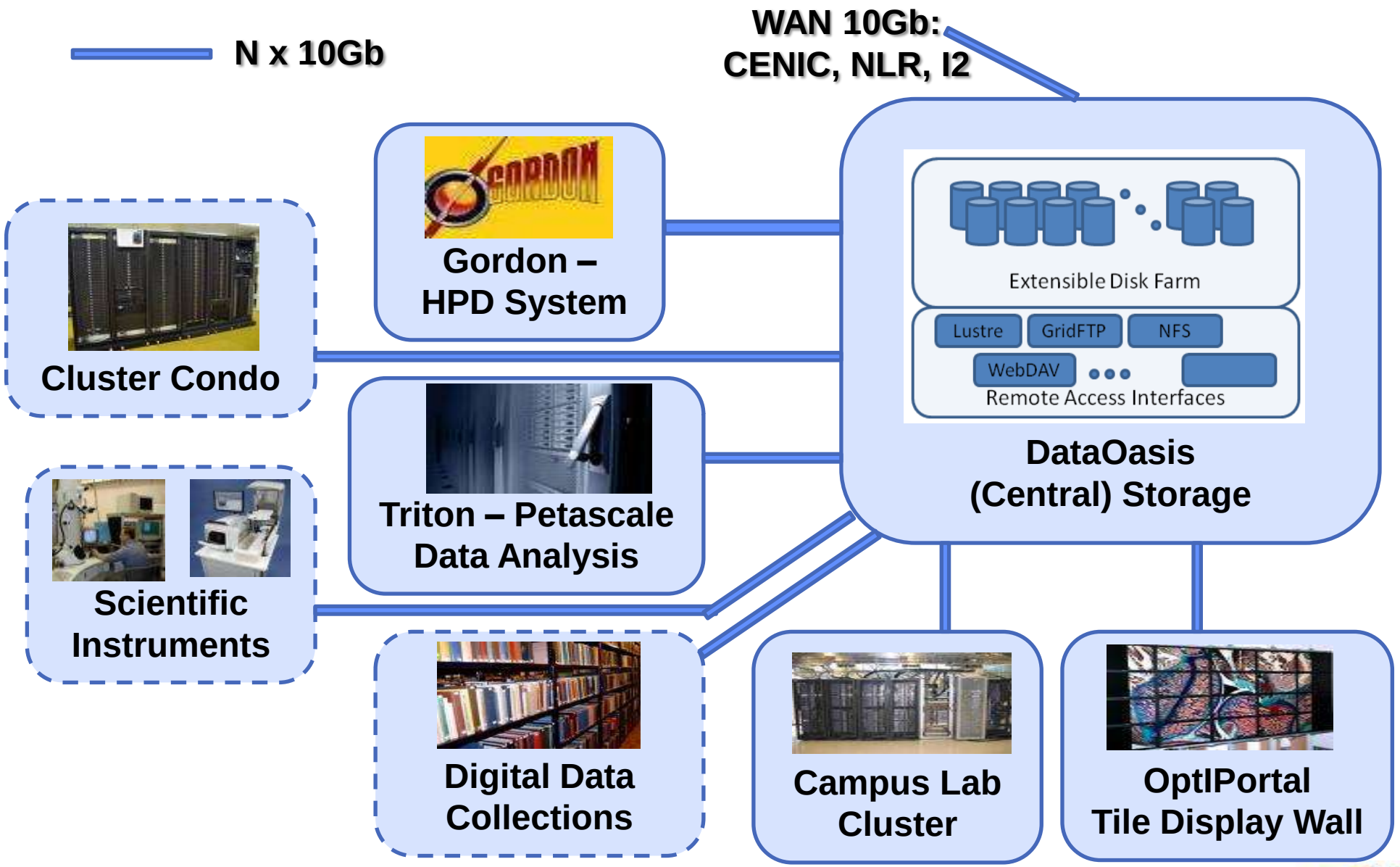
Approximately 0.5 TBit/s Arrive at the "Optical" Center of Campus. Switching is a Hybrid of: Packet, Lambda, Circuit -- OOO and Packet Switches



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

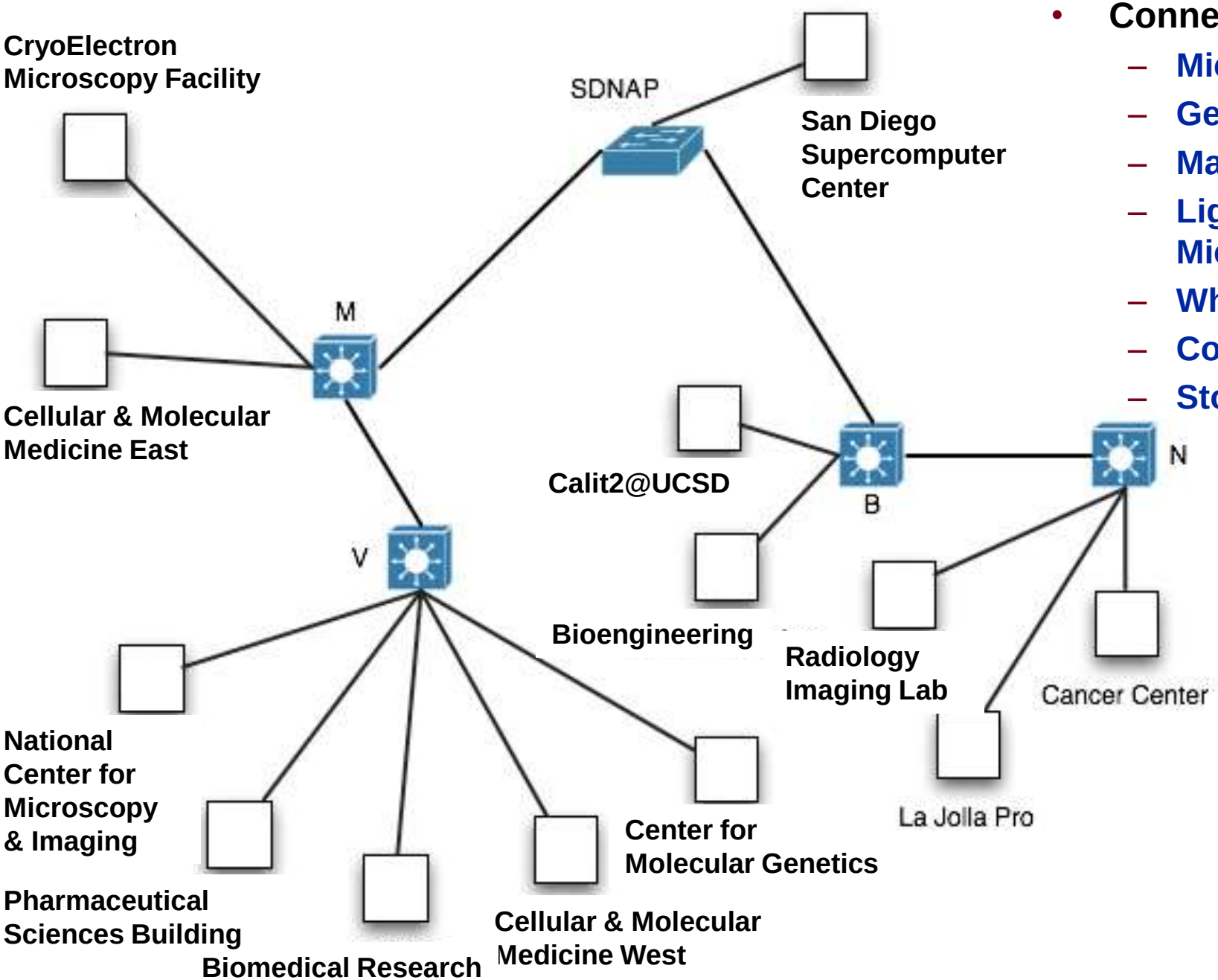
Quartzite Network MRI #CNS-0421555; OptIPuter #ANI-0225642

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



Source: Philip Papadopoulos, SDSC/Calit2

UCSD Planned Optical Networked Biomedical Researchers and Instruments



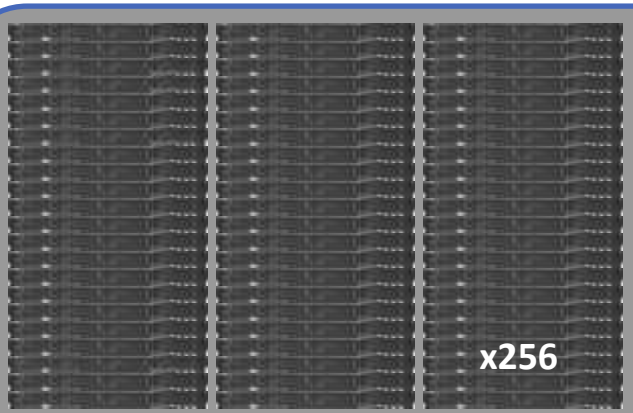
- **Connects at 10 Gbps :**
  - **Microarrays**
  - **Genome Sequencers**
  - **Mass Spectrometry**
  - **Light and Electron Microscopes**
  - **Whole Body Imagers**
  - **Computing**
  - **Storage**

Moving to a Shared Campus Data Storage and Analysis Resource: Triton Resource @ SDSC

Triton Resource

Large Memory PSDAF
- 256/512 GB/sys
- 9TB Total
- 128 GB/sec
- ~ 9 TF





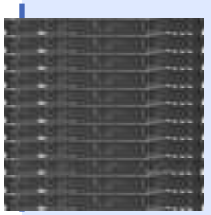
Shared Resource Cluster
- 24 GB/Node
- 6TB Total
- 256 GB/sec
- ~ 20 TF



Large Scale Storage
- 2 PB
- 40 – 80 GB/sec
- 3000 – 6000 disks
- Phase 0: 1/3 TB, 8GB/s



UCSD Research Labs







Campus Research Network

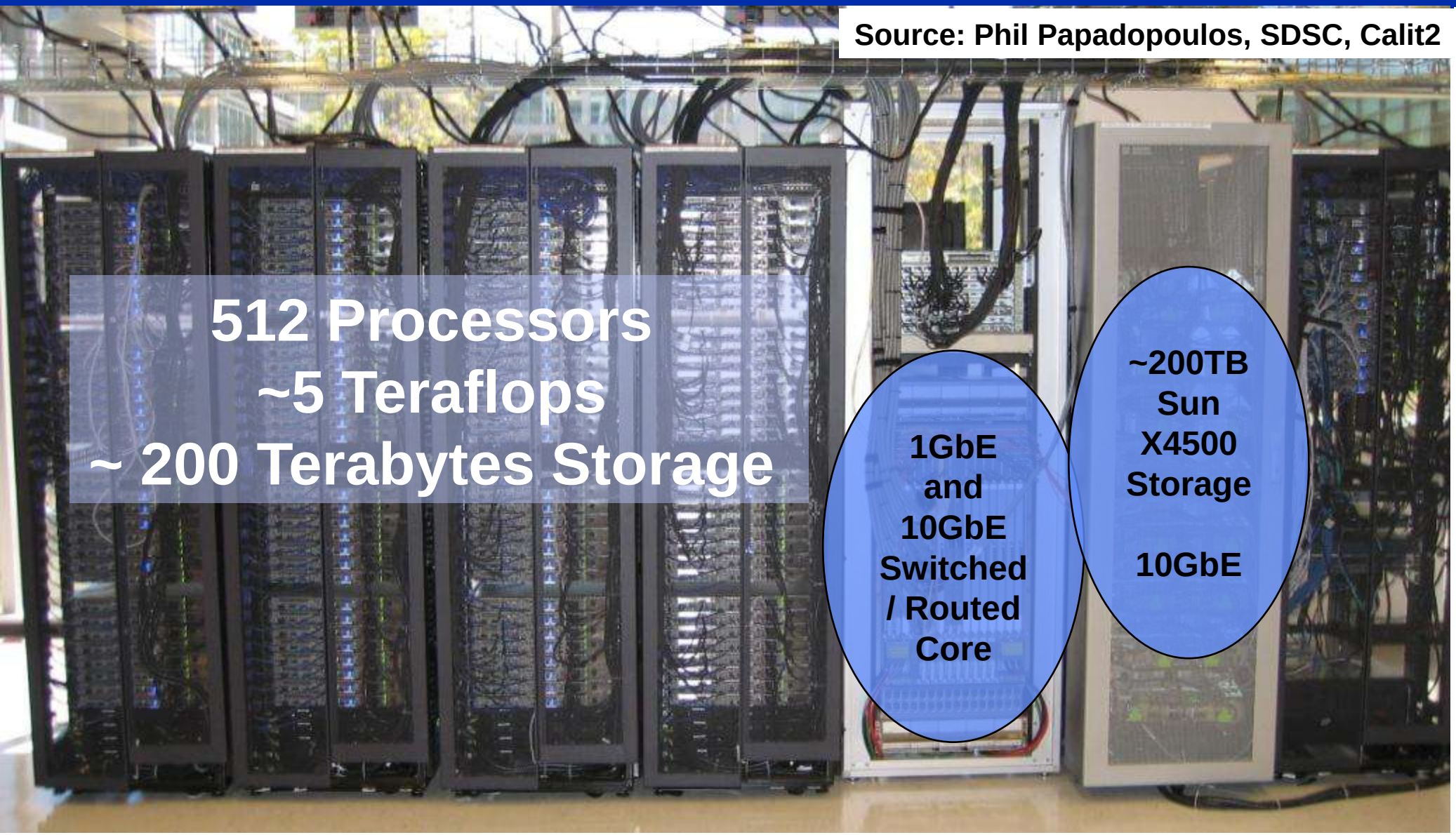
SDSC

Source: Philip Papadopoulos, SDSC/Calit2

it²

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

Source: Phil Papadopoulos, SDSC, Calit2

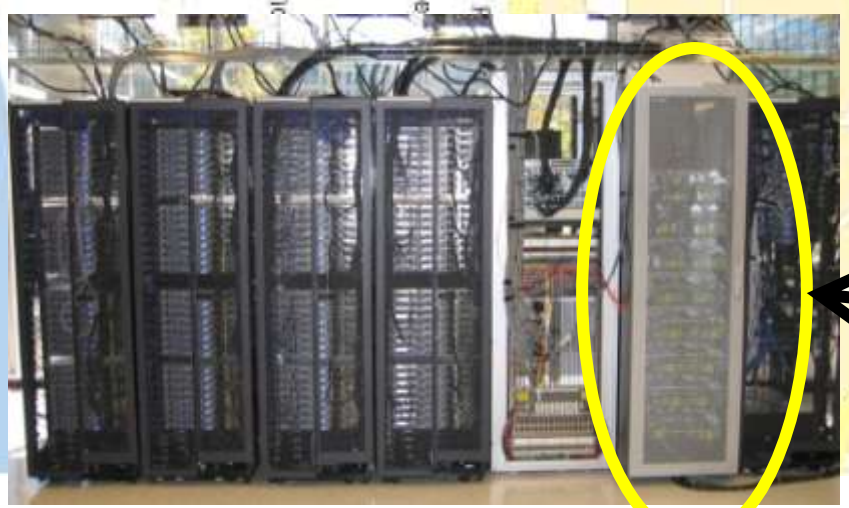


Gordon and Betty MOORE FOUNDATION ROCKS SDSC J. Craig Venter INSTITUTE it²

Calit2 CAMERA Automatic Overflows into SDSC Triton



@ CALIT2



@ SDSC

Triton Resource

CAMERA - Managed Job Submit Portal (VM)

Transparently Sends Jobs to Submit Portal on Triton

10Gbps

CAMERA DATA

Direct Mount == No Data Staging

SDSC

ROCKS

it²

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



OptIPuter

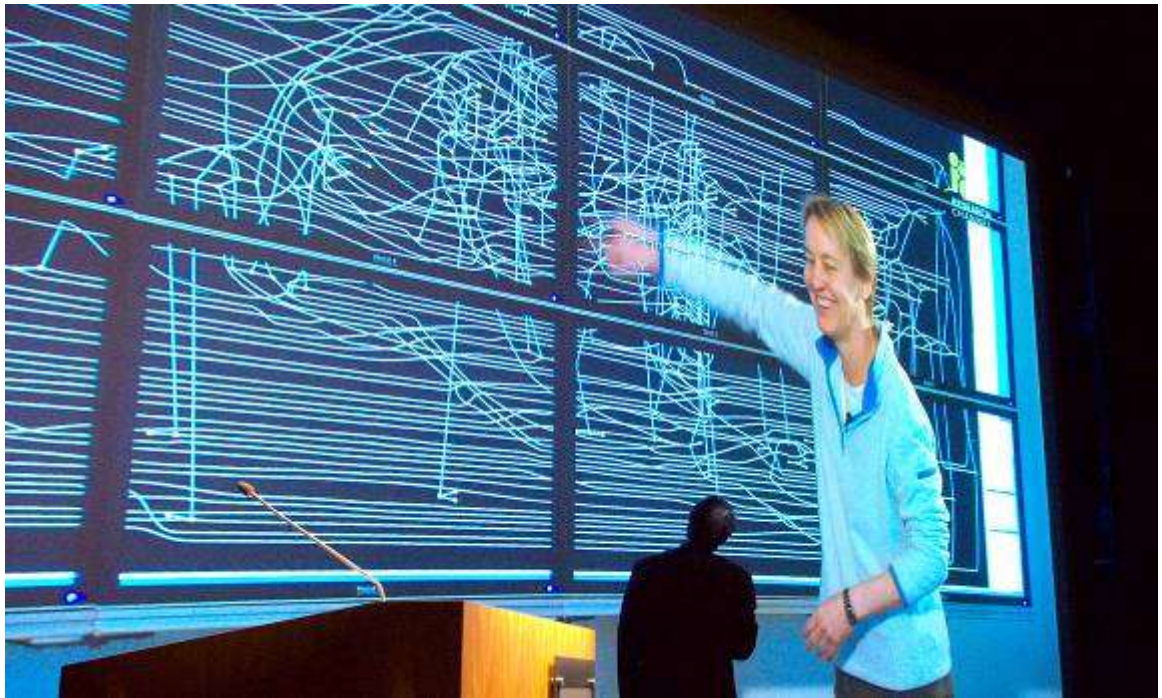
Ginger Armbrust's Diatoms: Micrographs, Chromosomes, Genetic Assembly

ROCKS



NSF

OptIPuter

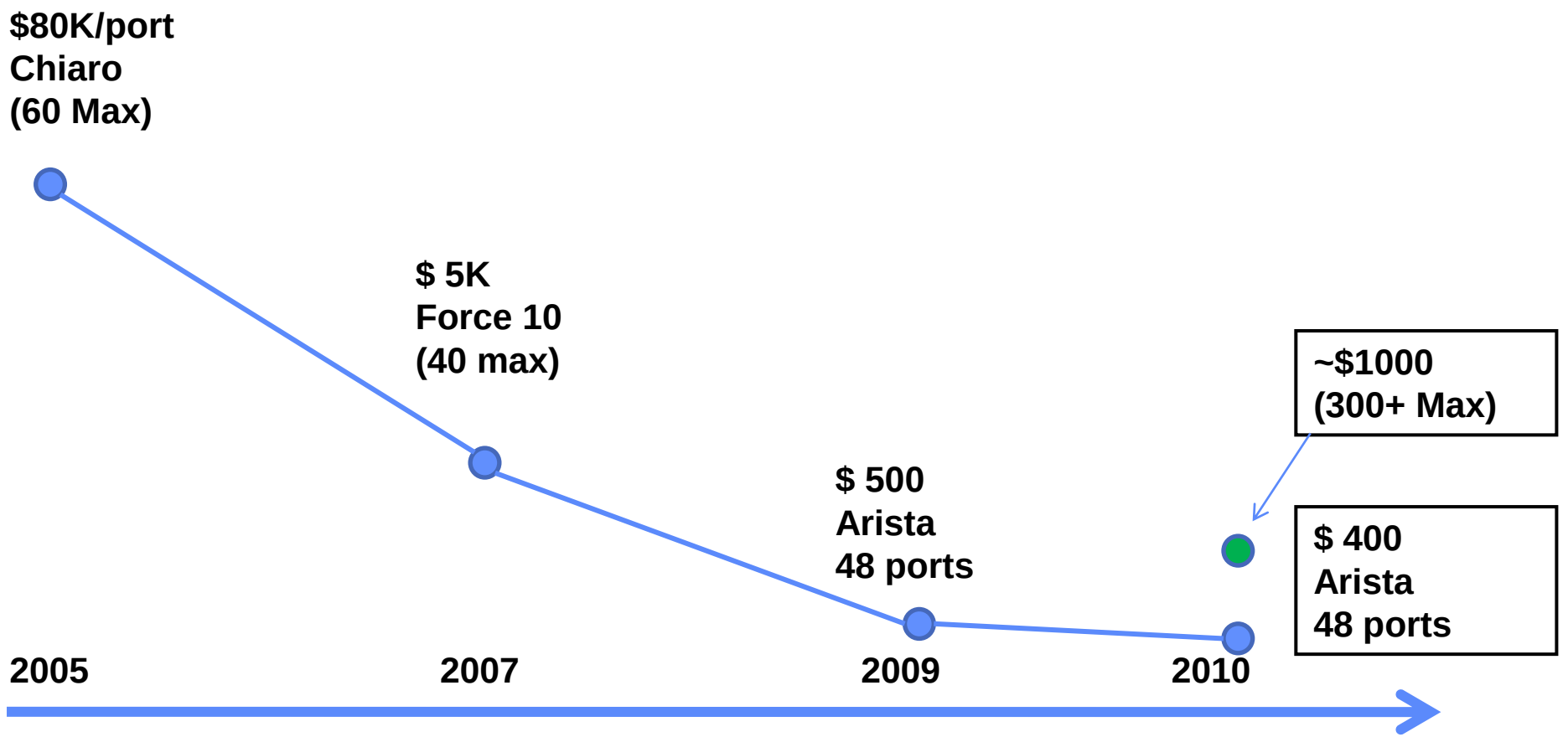


GORDON AND BETTY MOORE FOUNDATION

it²

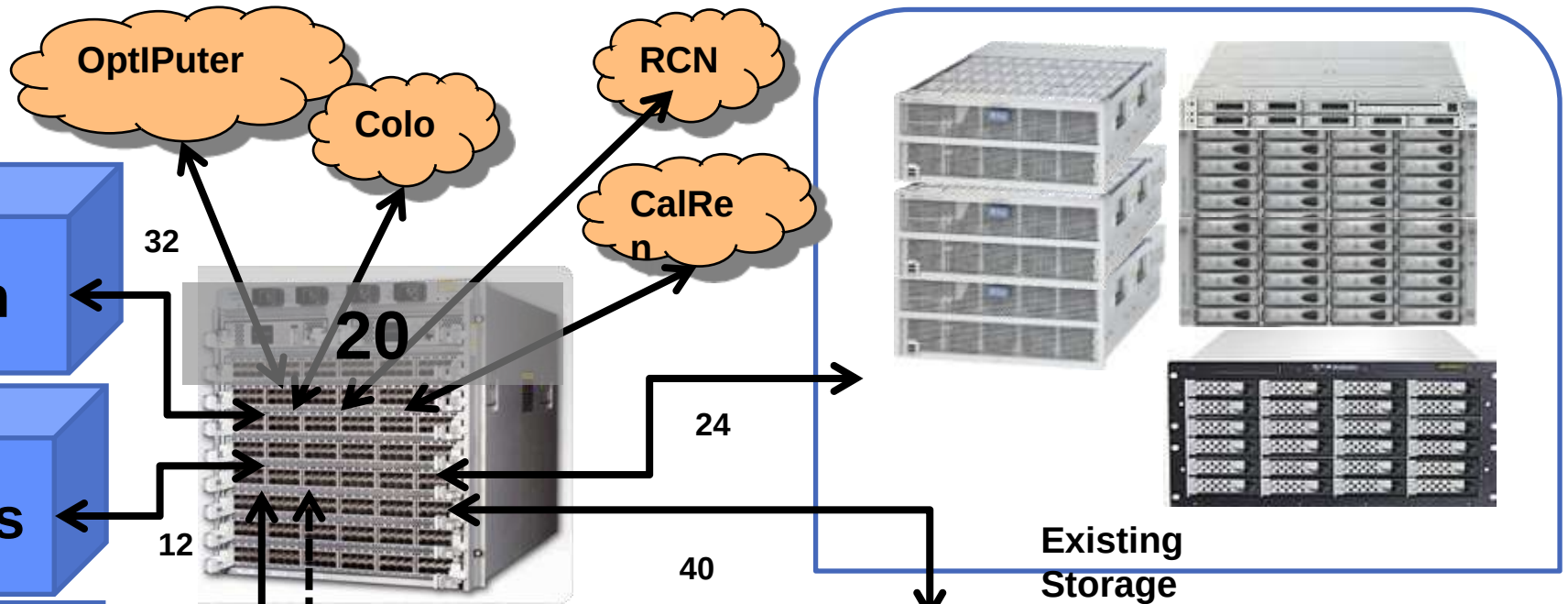
Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

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



Source: Philip Papadopoulos, SDSC/Calit2

10G Switched Data Analysis Resource: Data Oasis (RFP Responses Due 10/29/2010)



Triton

Trestles

Dash

Gordon

8

100

Oasis Procurement (RFP)

- Phase0: > 8GB/s sustained, today
- RFP for Phase1: > 40 GB/sec for Lustre
- Nodes must be able to function as Lustre OSS (Linux) or NFS (Solaris)
- Connectivity to Network is 2 x 10GbE/Node
- Likely Reserve dollars for inexpensive replica servers

1500 – 2000 TB > 40 GB/s

SDSC

Source: Philip Papadopoulos, SDSC/Calit2

it²

You Can Download This Presentation at lsmarr.calit2.net



ABOUT LARRY SMARR

Larry Smarr is the founding Director of the California Institute for Telecommunications and Information Technology (Calit2), a UC San Diego/UC Irvine partnership, and holds the Harry E. Gruber professorship in Computer Science and Engineering (CSE) at UCSD's Jacobs School. At Calit2, Smarr has

WHAT I'M READING

Complete List of Larry's Recommendations

Homo Britannicus: The Incredible Story of Human Life in Britain (2006)

American College and University

RECENT TWITTER POSTS

Twitter? Click To Follow Me

Ismarr: SDSC Founding Director Sid Karin and I are making a few comments on the history of the NSF Supercomputing Centers at