High Performance Cyberinfrastructure Enabling Data-Driven Science in the Biomedical Sciences

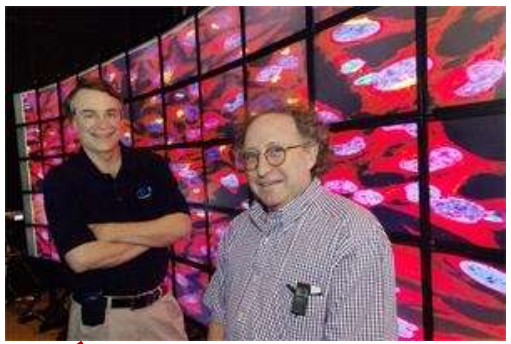
Joint Presentation

UCSD School of Medicine Research Council

Larry Smarr, Calit2 & Phil Papadopoulos, SDSC/Calit2

April 6, 2011

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



End User OptIPortal





Local or Remote Instruments



HD/4k Live Video

HPC

TeraGrid

National LambdaRail



10G Lightpaths



Campus Optical Switch

amazon web services



Data Repositories & Clusters



HD/4k Video Repositories

SDSC

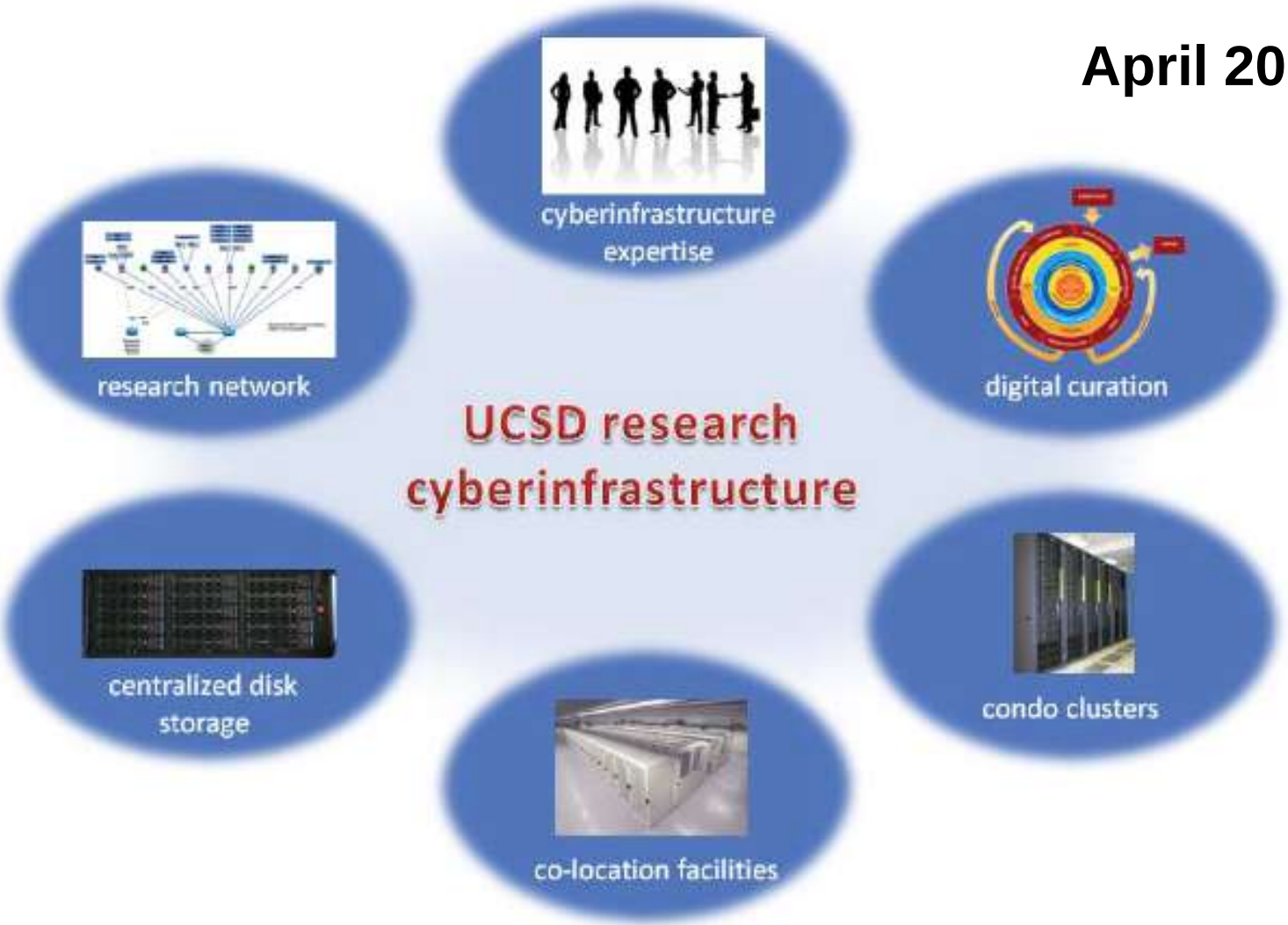
it²

"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



SDSC research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf

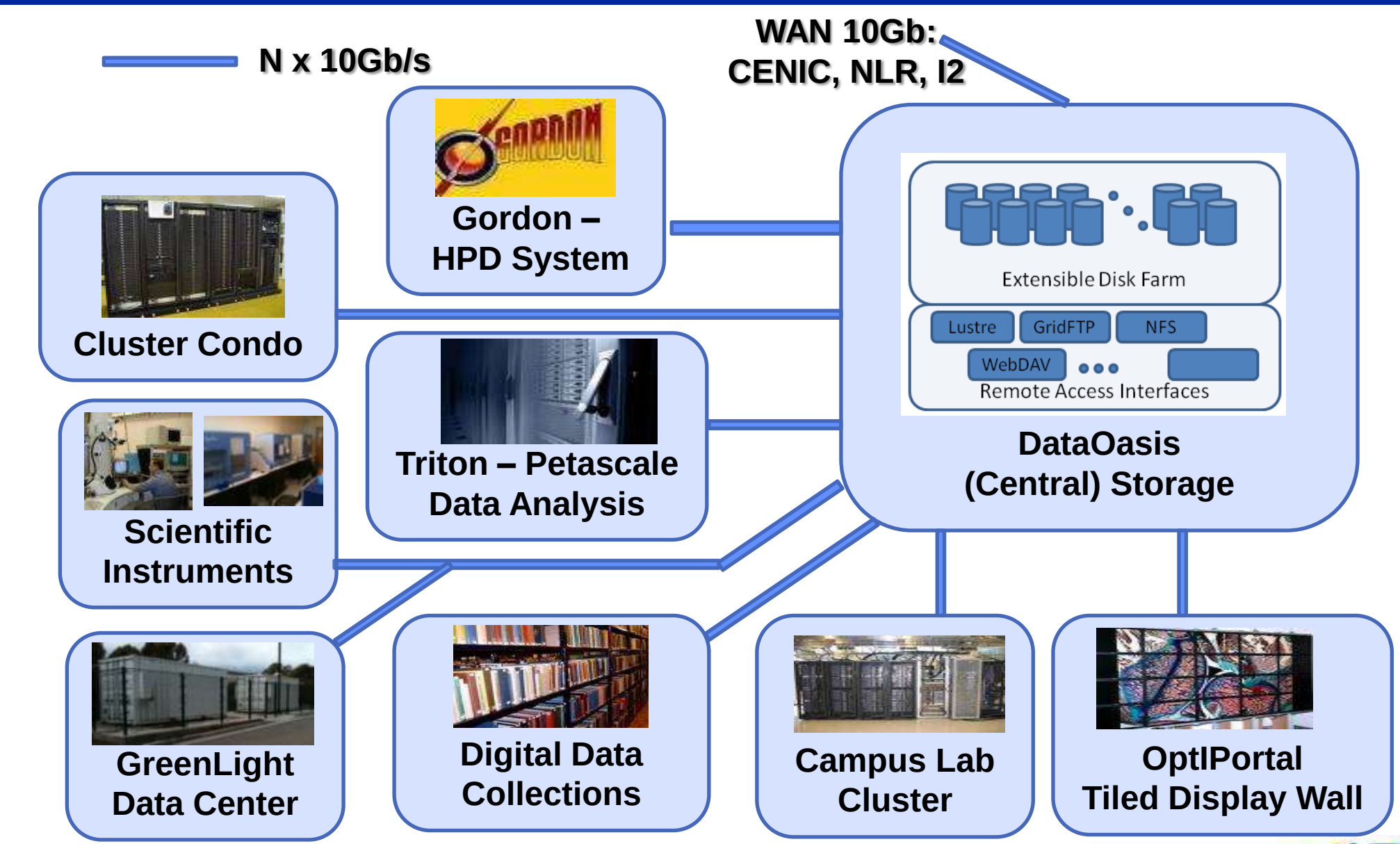
"Digital Shelter"

- **21st Century Science is Dependent on High-Quality Digital Data**
  - **It Needs to be:**
    - **Stored Reliably**
    - **Discoverable for Scientific Publication and Re-use**
- **The RCI Design Team Centered its Architecture on Digital Data**
- **The Fundamental Questions/Observations:**
  - **Large-Scale Data Storage is Hard!**
    - **It's "Expensive" to do it WELL**
      - Performance AND Reliable Storage
      - People are Expensive
  - **What Happens to ANY Digital Data Product at the End of a Grant?**
    - **Who Should be Fundamentally Responsible?**

SDSC

it²

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

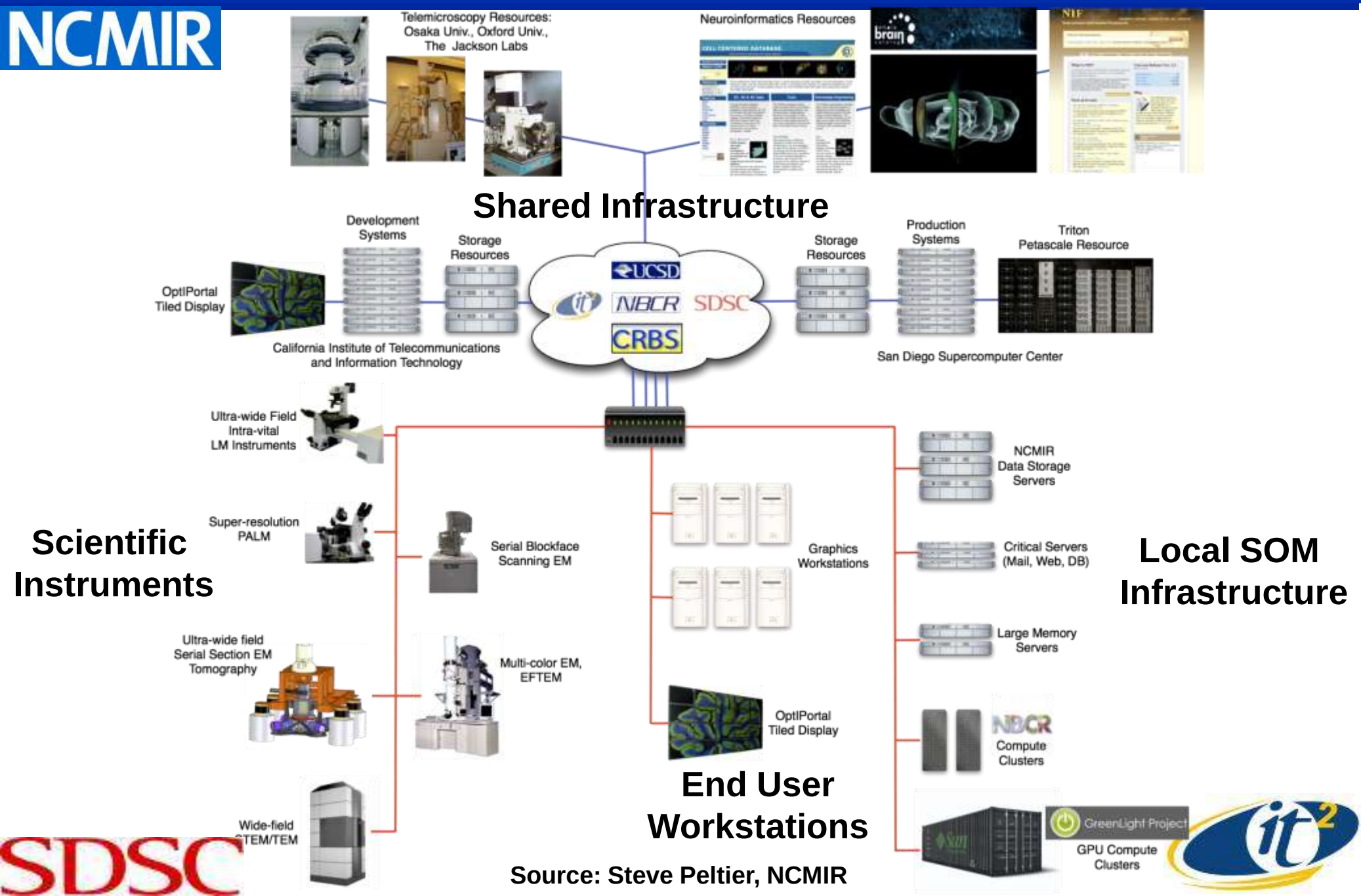


Source: Philip Papadopoulos, SDSC, UCSD

Applications Built on RCI:
Example #1 NCMIR Microscopes

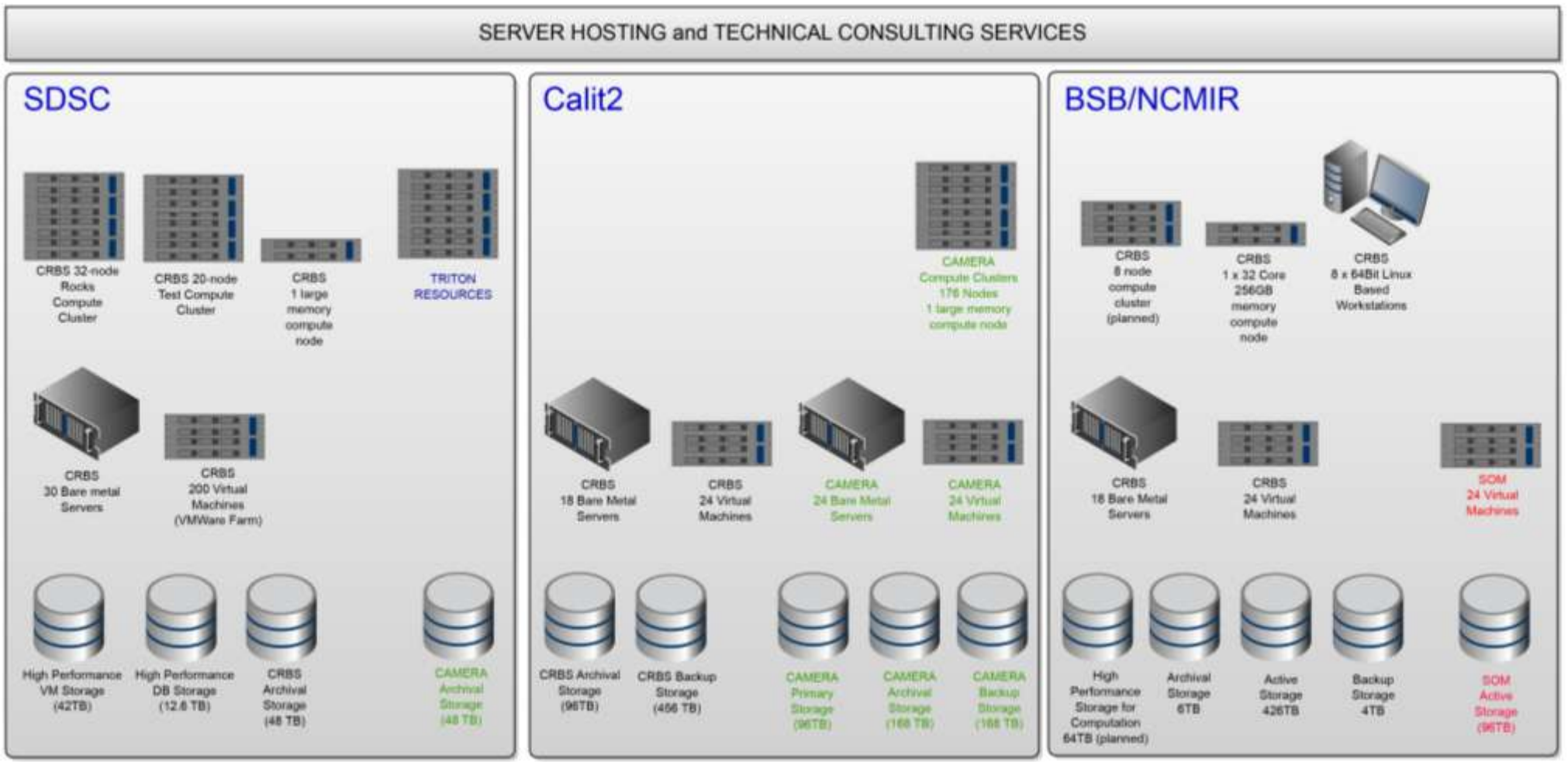
SDSC NCMIR it²

NCMIR's Integrated Infrastructure of Shared Resources



Source: Steve Peltier, NCMIR

Detailed Map of CRBS/SOM Computation and Data Resources



System Wide Upgrade to 10Gb Underway

SDSC NCMIR it²

Applications Built on RCI:
Example #2 Next Gen Sequencers

SDSC

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

Next Generation Genome Sequencers Produce Large Data Sets

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	

Source: Chris Misleh, SOM

The Growing Sequencing Data Load Runs over RCI Connecting GreenLight and Triton

- **Data from the Sequencers Stored in GreenLight SOM Data Center**
  - **Data Center Contains Cisco Catalyst 6509-connected to Campus RCI at 2 x 10Gb.**
  - **Attached to the Cisco Catalyst is a 48 x 1Gb switch and an Arista 7148 switch which has 48 x 10Gb ports.**
  - **The two Sun Disks connect directly to the Arista switch for 10Gb connectivity.**
- **With our current configuration of two Illumina GAIIx, one GAII, and one HiSeq 2000, we can produce a maximum of 3TB of data per week.**
- **Processing uses a combination of local compute nodes and the Triton resource at SDSC.**
  - **Triton comes in particularly handy when we need to run 30 seqmap/blat/blast jobs. On a standard desktop computer this analysis could take several weeks. On Triton, we have the ability submit these jobs in parallel and complete computation in a fraction of the time. Typically within a day.**
- **In the coming months we will be transitioning another lab to the 10Gbit Arista switch. In total we will have 6 Sun Disks connected at 10Gbit speed, and mounted via NFS directly on the Triton resource..**
- **The new PacBio RS is scheduled to arrive in May, which will also utilize the Campus RCI in Leichtag and the SOM GreenLight Data Center.**

SDSC

Source: Chris Misleh, SOM

it²

Applications Built on RCI:
Example #3 Microbial Metagenomic Services

SDSC

camera
Marine Microbial Ecology

it²

Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis

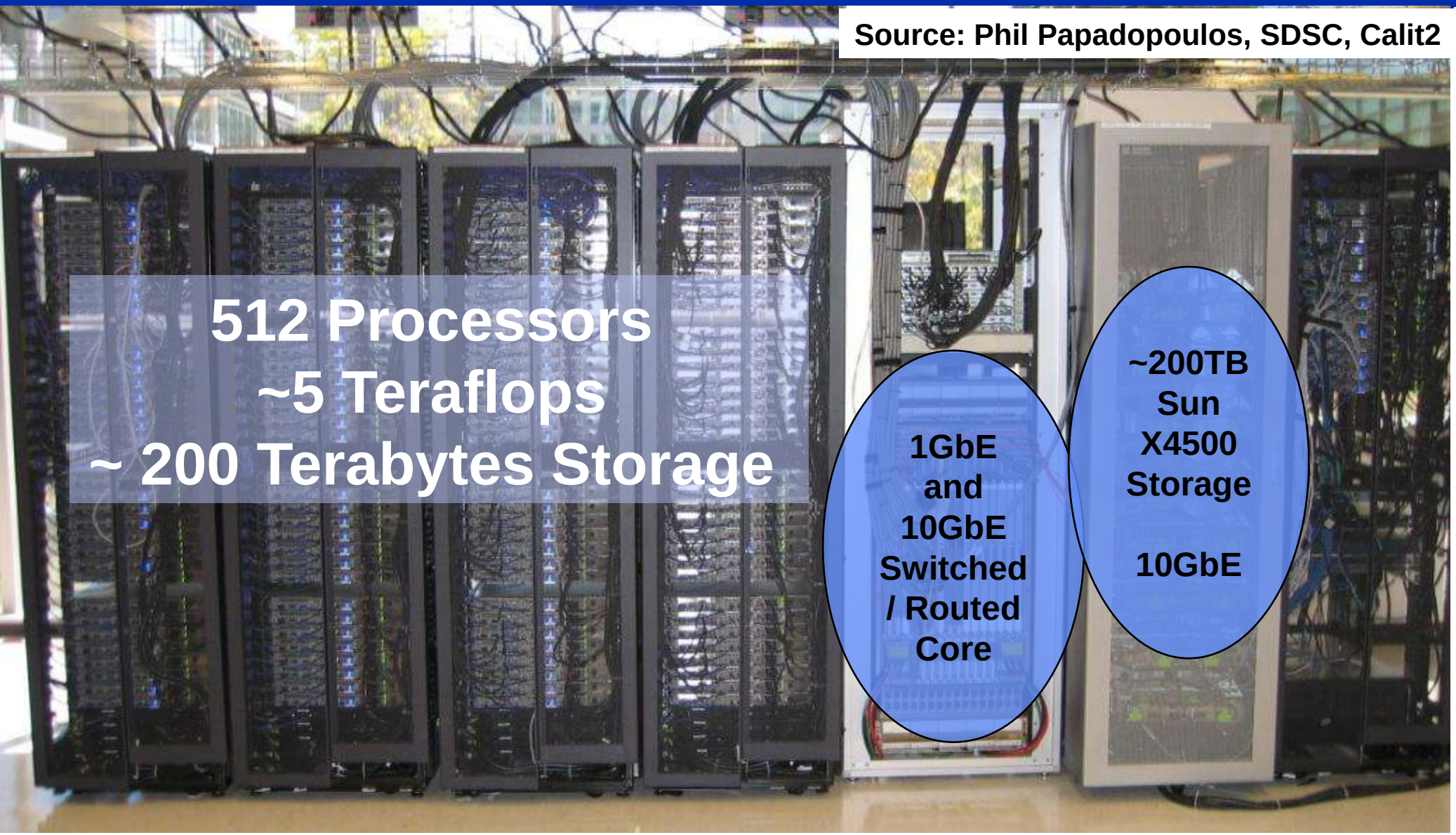


SDSC

Gordon and Betty MOORE FOUNDATION

it²

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



Source: Phil Papadopoulos, SDSC, Calit2

4000 Users
From 90 Countries

Gordon and Betty MOORE FOUNDATION

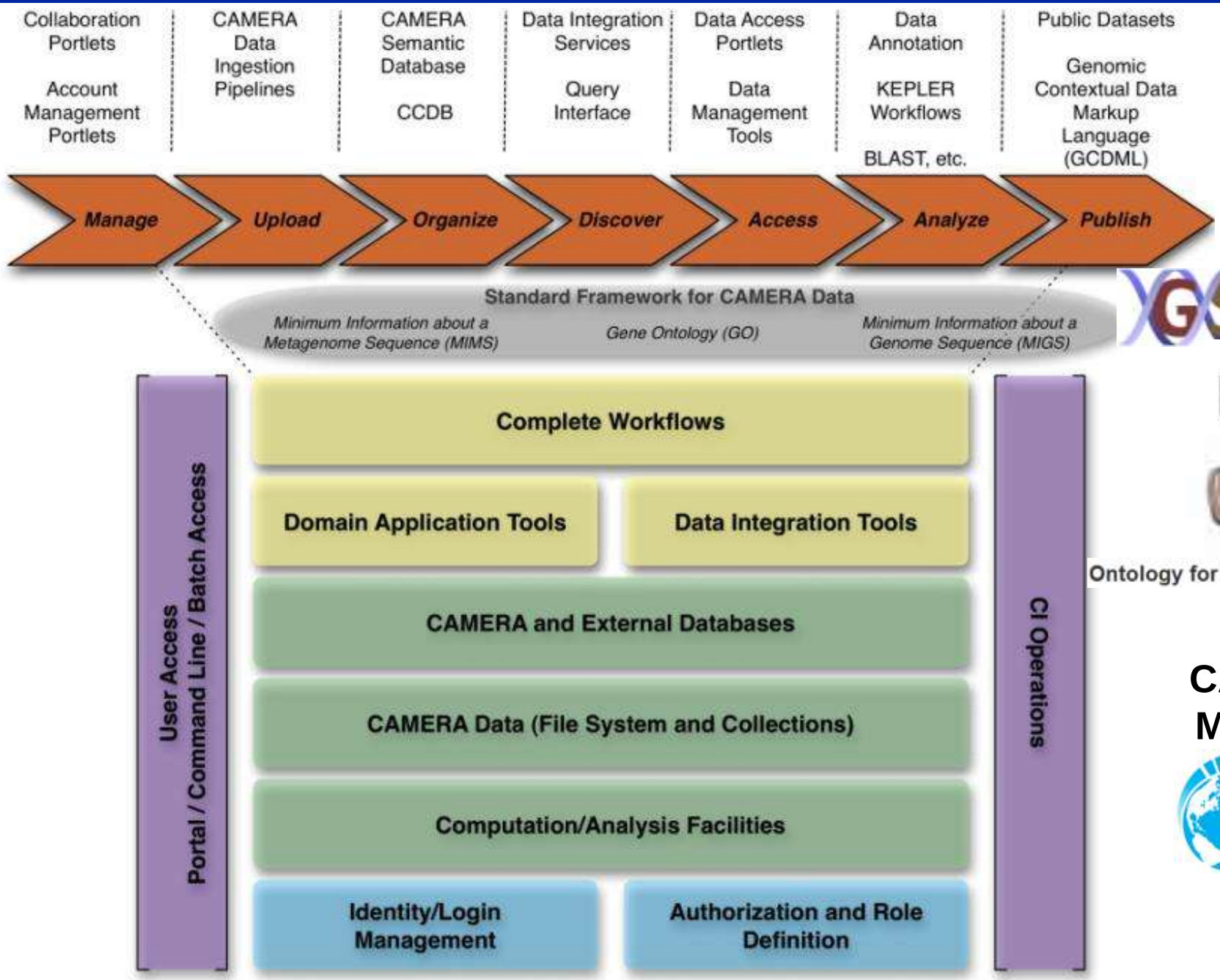
ROCKS

SDSC

J. Craig Venter INSTITUTE

it²

Creating CAMERA 2.0 -
Advanced Cyberinfrastructure Service Oriented Architecture



GSC Genomic Standards Consortium

Env

OBI

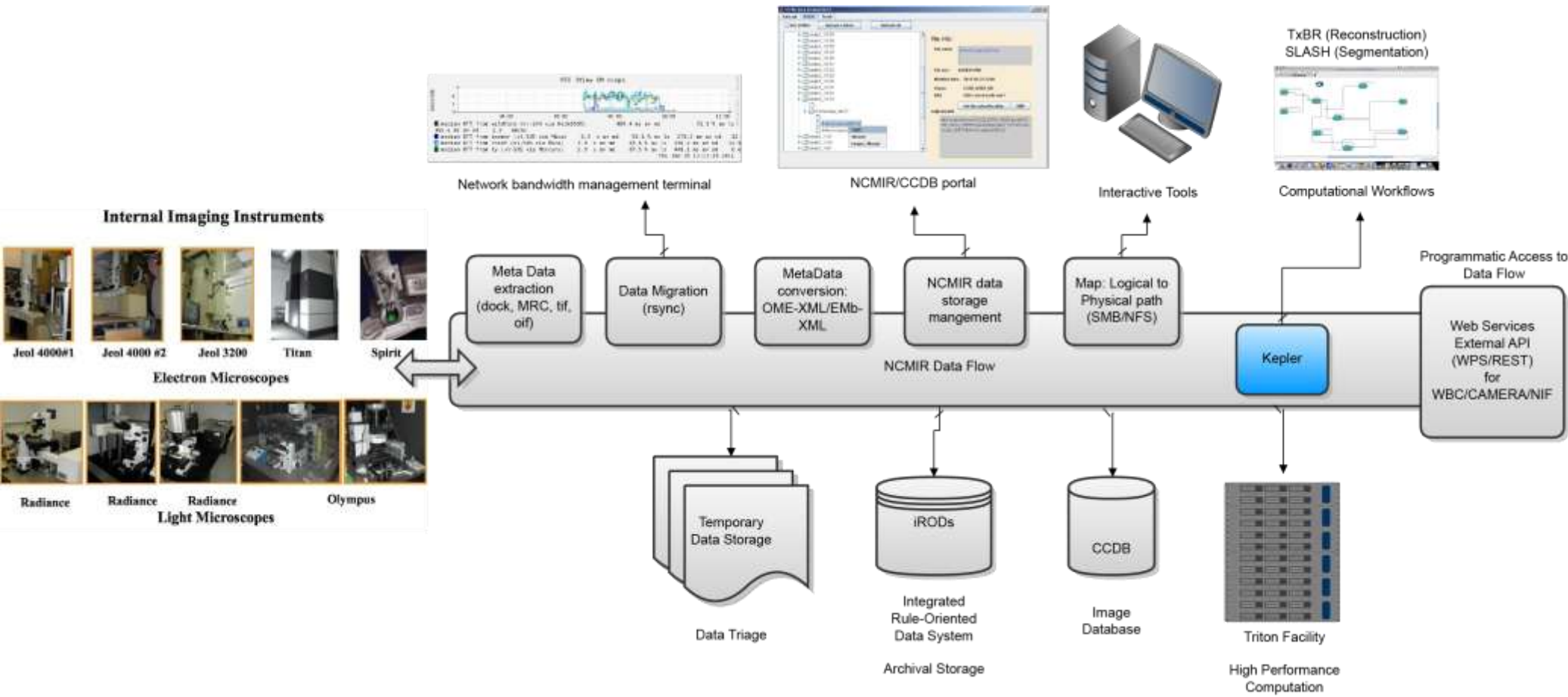
Ontology for Biomedical Investigations

Source:
CAMERA CTO
Mark Ellisman

camera
Marine Microbial Ecology

it2

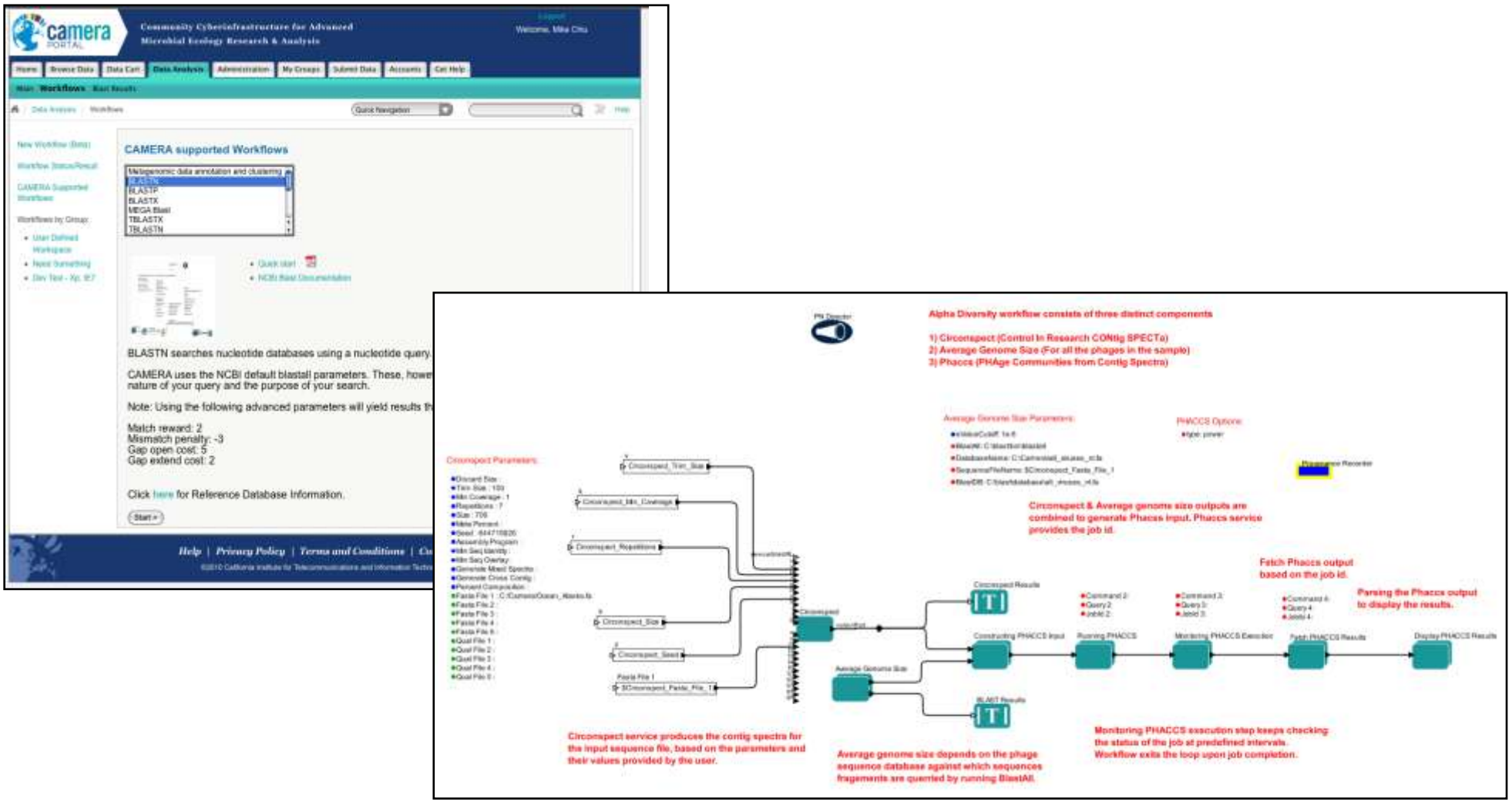
Fully Integrated UCSD CI Manages the End-to-End Lifecycle of Massive Data from Instruments to Analysis to Archival



UCSD CI Features Kepler Workflow Technologies

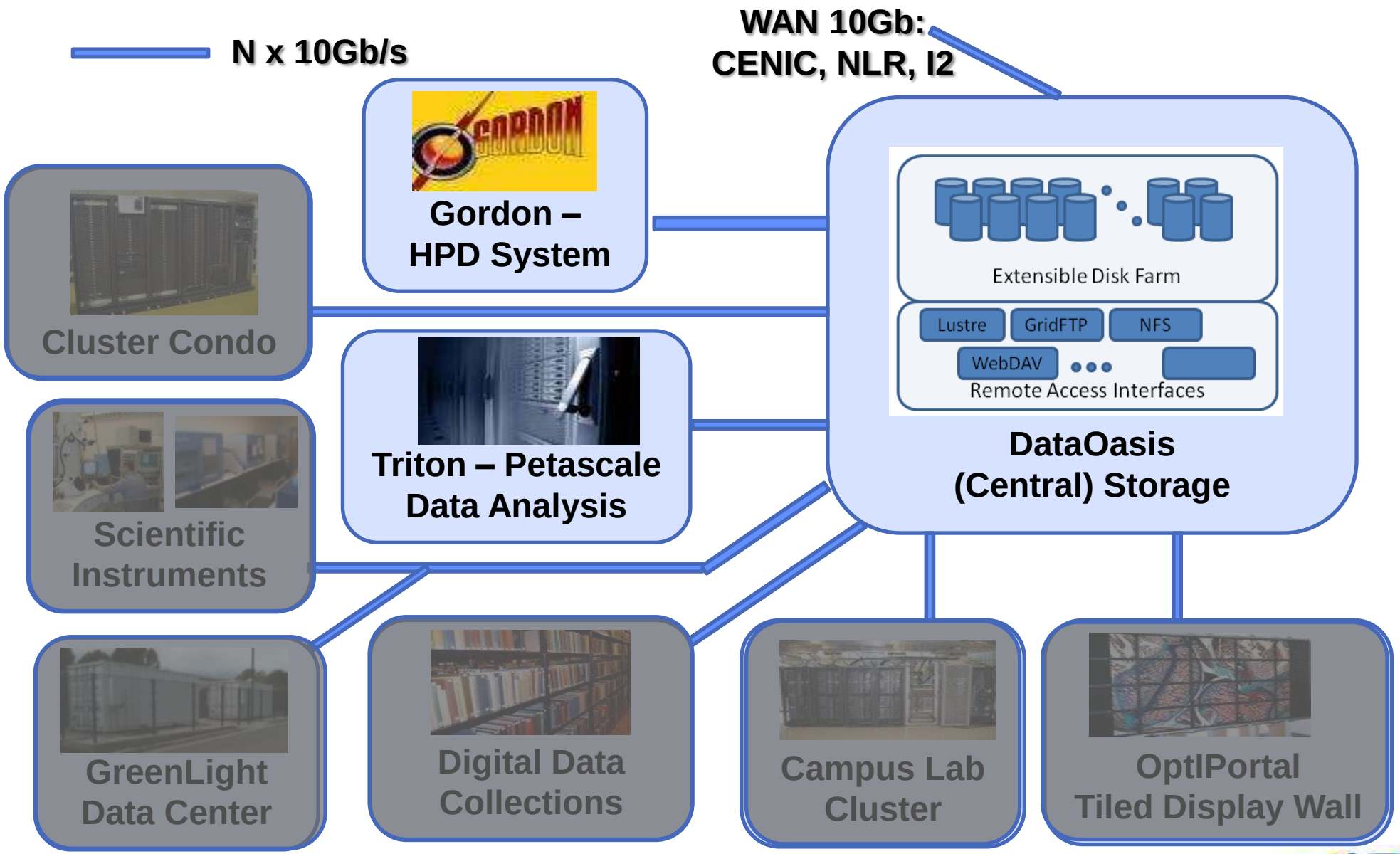
SDSC NCMIR it²

UCSD CI and Kepler Workflows Power CAMERA 2.0 Community Portal (4000+ users)



SDSC NCMIR it²

SDSC Investments in the CI Design Team Architecture



Source: Philip Papadopoulos, SDSC, UCSD

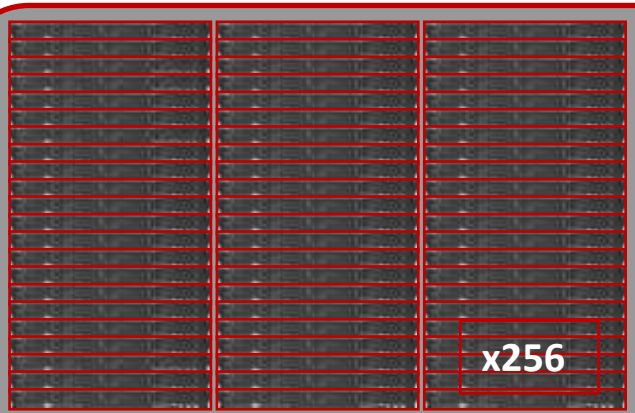
Moving to Shared Enterprise Data Storage & Analysis Resources: SDSC Triton Resource & Calit2 GreenLight

http://tritonresource.sdsc.edu

Source: Philip Papadopoulos, SDSC, UCSD

SDSC Large Memory Nodes
- 256/512 GB/sys
- 8TB Total
- 128 GB/sec
- ~ 9 TF





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

ROCKS



SDSC Data Oasis Large Scale Storage
- 2 PB
- 50 GB/sec
- 3000 – 6000 disks
- Phase 0: 1/3 PB, 8GB/s



UCSD Research Labs





N x 10Gb/s



Campus Research Network

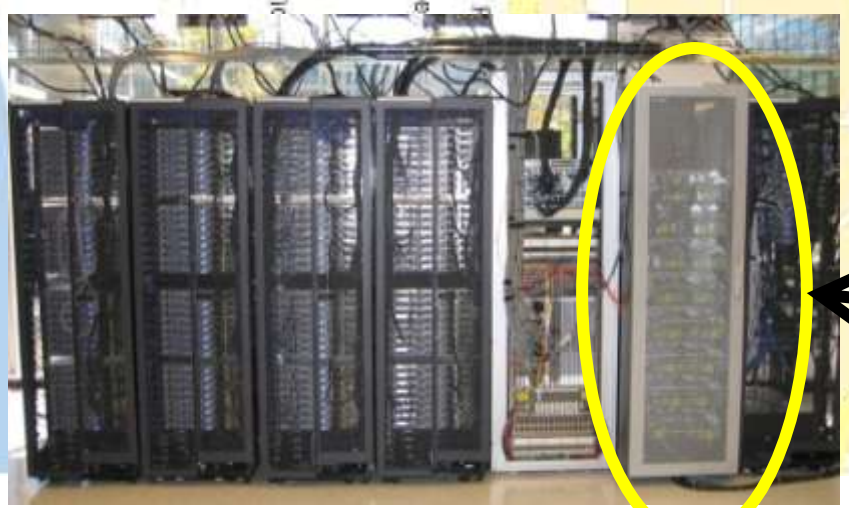
SDSC

it²

Calit2 CAMERA Automatic Overflows Use Triton as a Computing "Peripheral"



@ CALIT2



@ SDSC

Triton Resource

Transparently Sends Jobs to Submit Portal on Triton

CAMERA - Managed Job Submit Portal (VM)

10Gbps

CAMERA DATA

Direct Mount == No Data Staging

SDSC

ROCKS

it²

NSF Funds a Data-Intensive Track 2 Supercomputer: SDSC's Gordon-Coming Summer 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 Data Bases being Generated in Many Fields of Science, Engineering, Medicine, and Social Science**

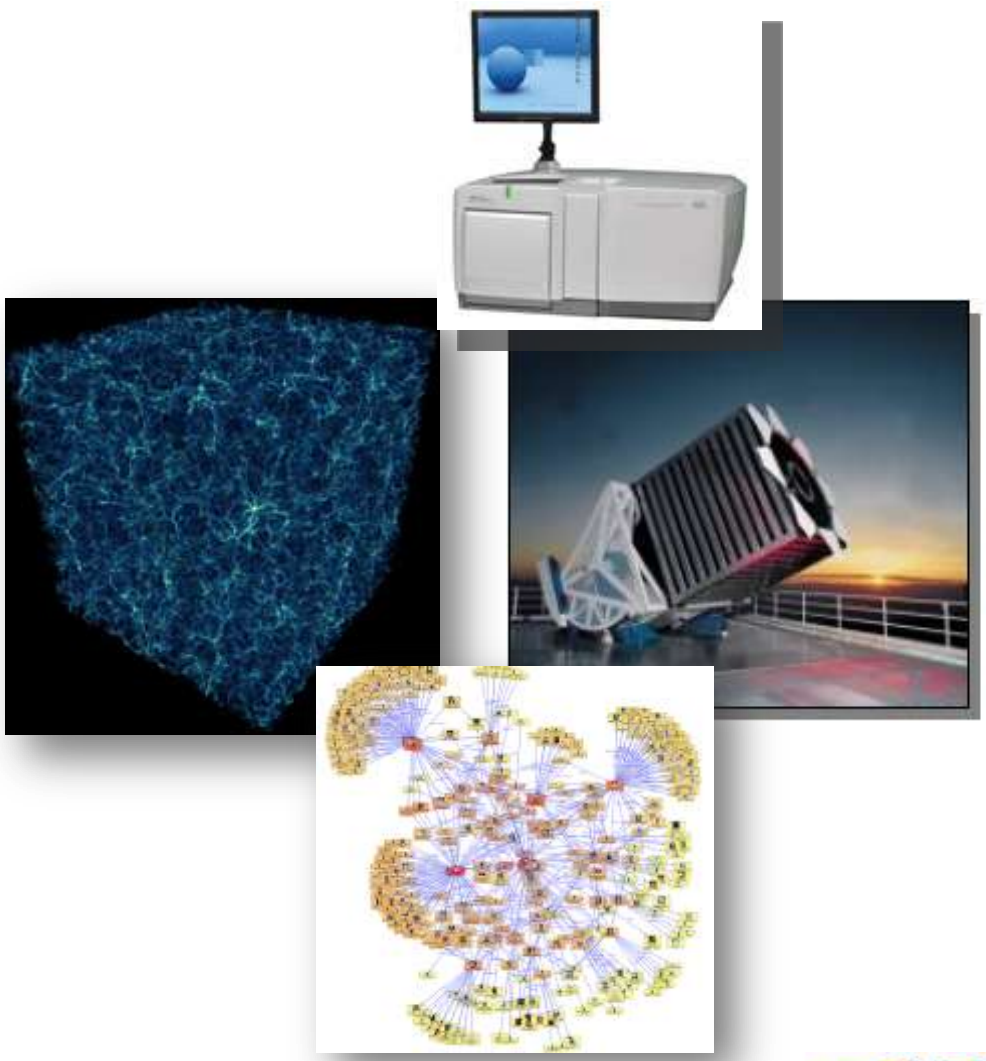
SDSC

Source: Mike Norman, Allan Snavely SDSC

it²

Data Mining Applications will Benefit from Gordon

- ***De Novo* Genome Assembly from Sequencer Reads & Analysis of Galaxies from Cosmological Simulations & Observations**
  - *Will Benefit from Large Shared Memory*
- **Federations of Databases & Interaction Network Analysis for Drug Discovery, Social Science, Biology, Epidemiology, Etc.**
  - *Will Benefit from Low Latency I/O from Flash*



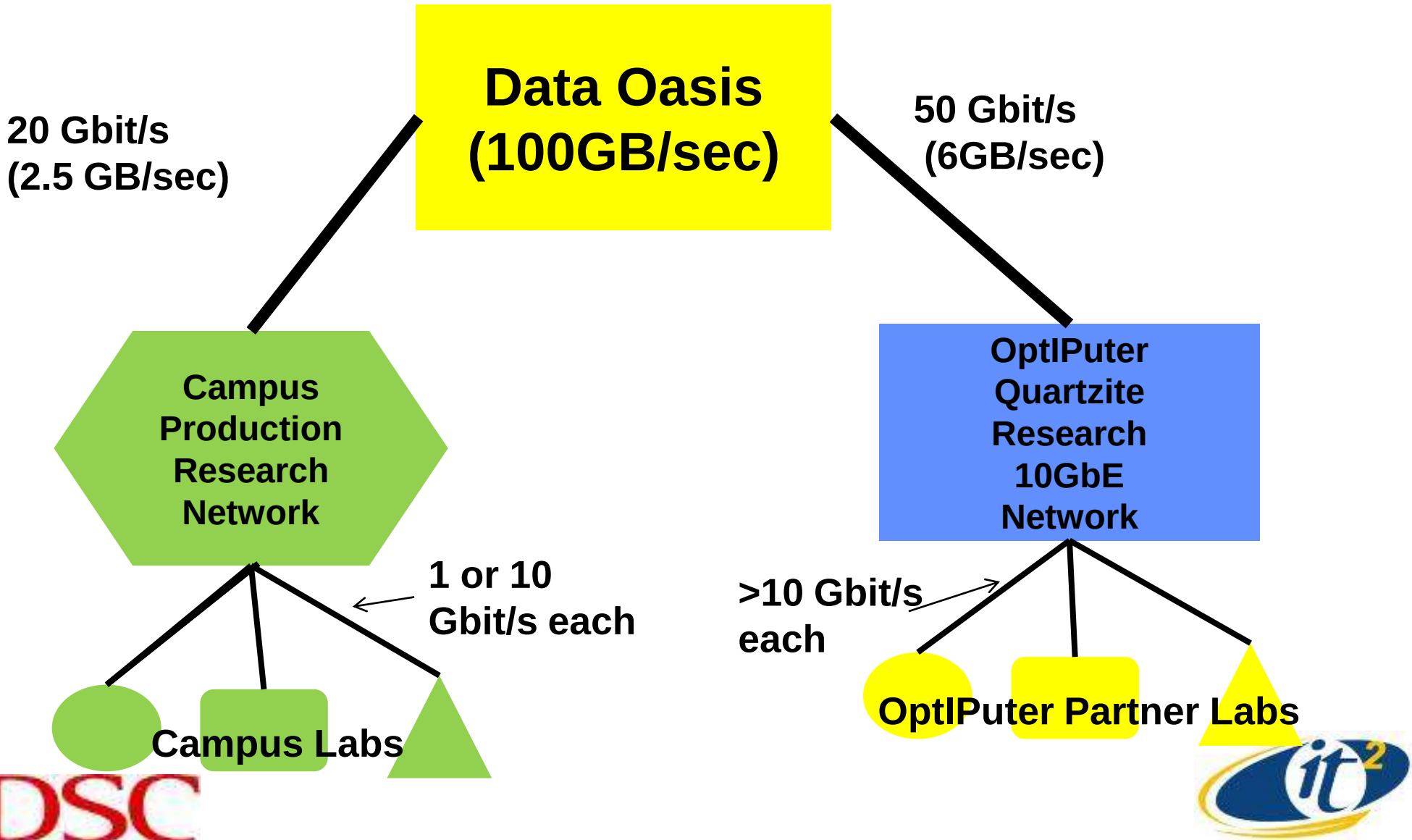
SDSC

Source: Mike Norman, SDSC

it²

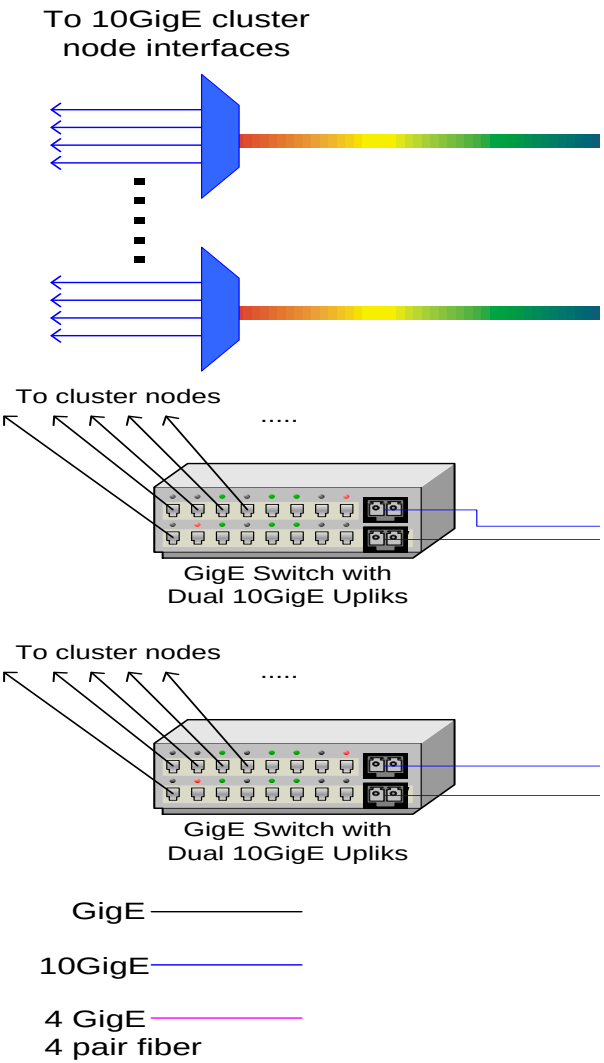
IF Your Data is Remote, Your Network Better be "Fat"

1TB @ 10 Gbit/sec = ~20 Minutes
1TB @ 1 Gbit/sec = 3.3 Hours



SDSC

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



Enpoints:

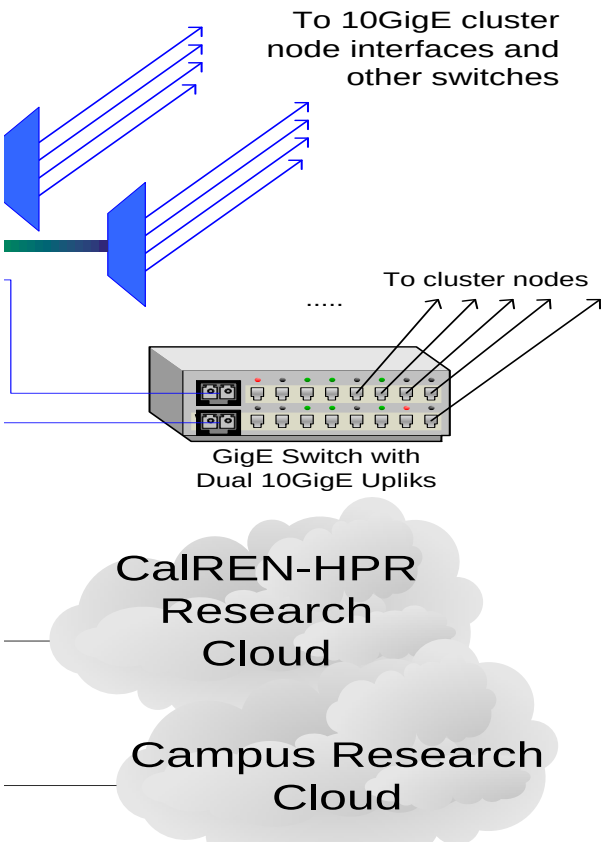
>= 60 endpoints at 10 GigE

>= 32 Packet switched

>= 32 Switched wavelengths

>= 300 Connected endpoints

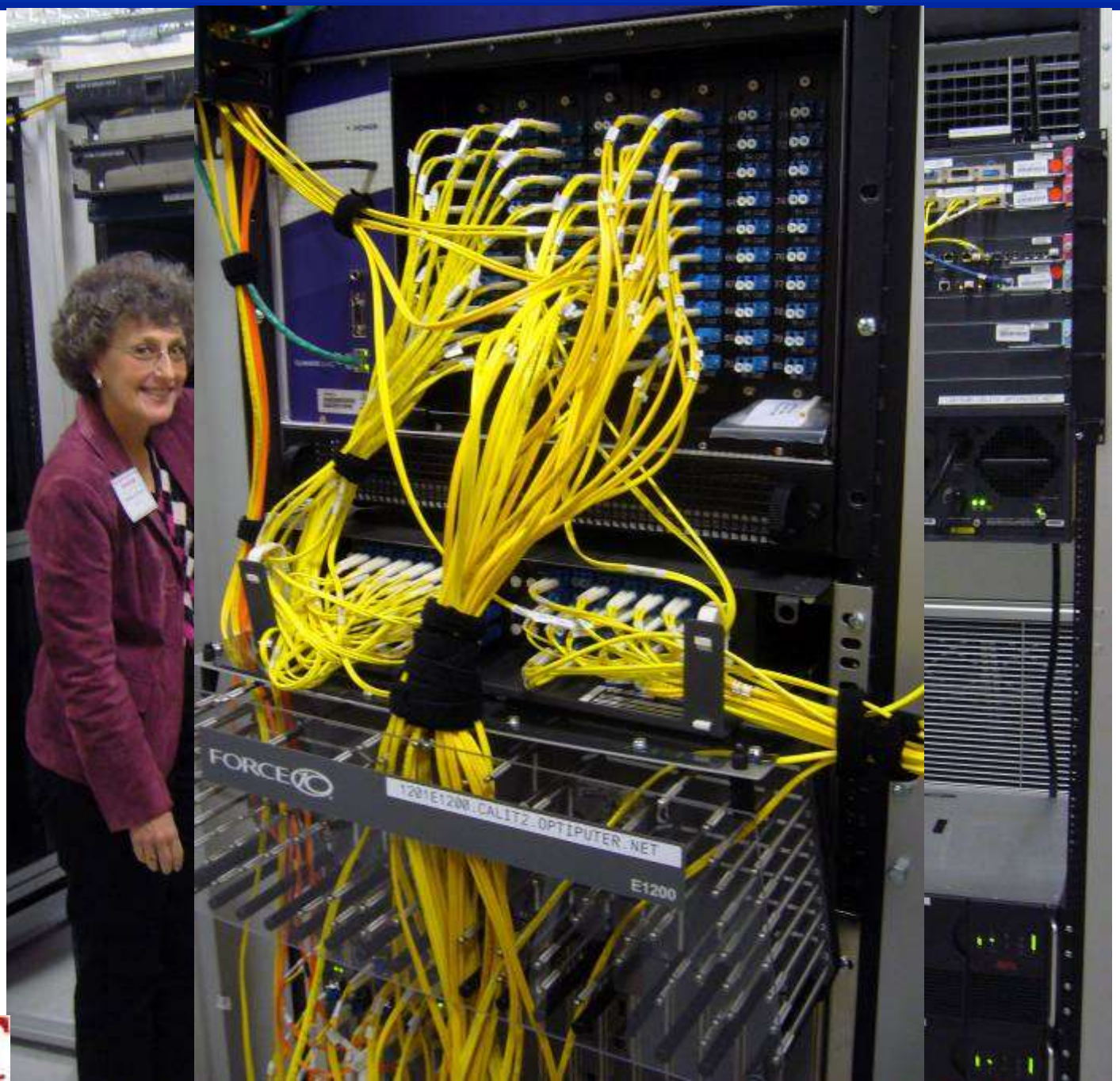
Approximately 0.5 TBit/s Arrive at the "Optical" Center of Campus. Switching is a Hybrid of: Packet, Lambda, Circuit -- OOO and Packet Switches



Source: Phil Papadopoulos, SDSC/Calit2 (Quartzite PI, OptIPuter co-PI)
Quartzite Network MRI #CNS-0421555; OptIPuter #ANI-0225642

SDSC

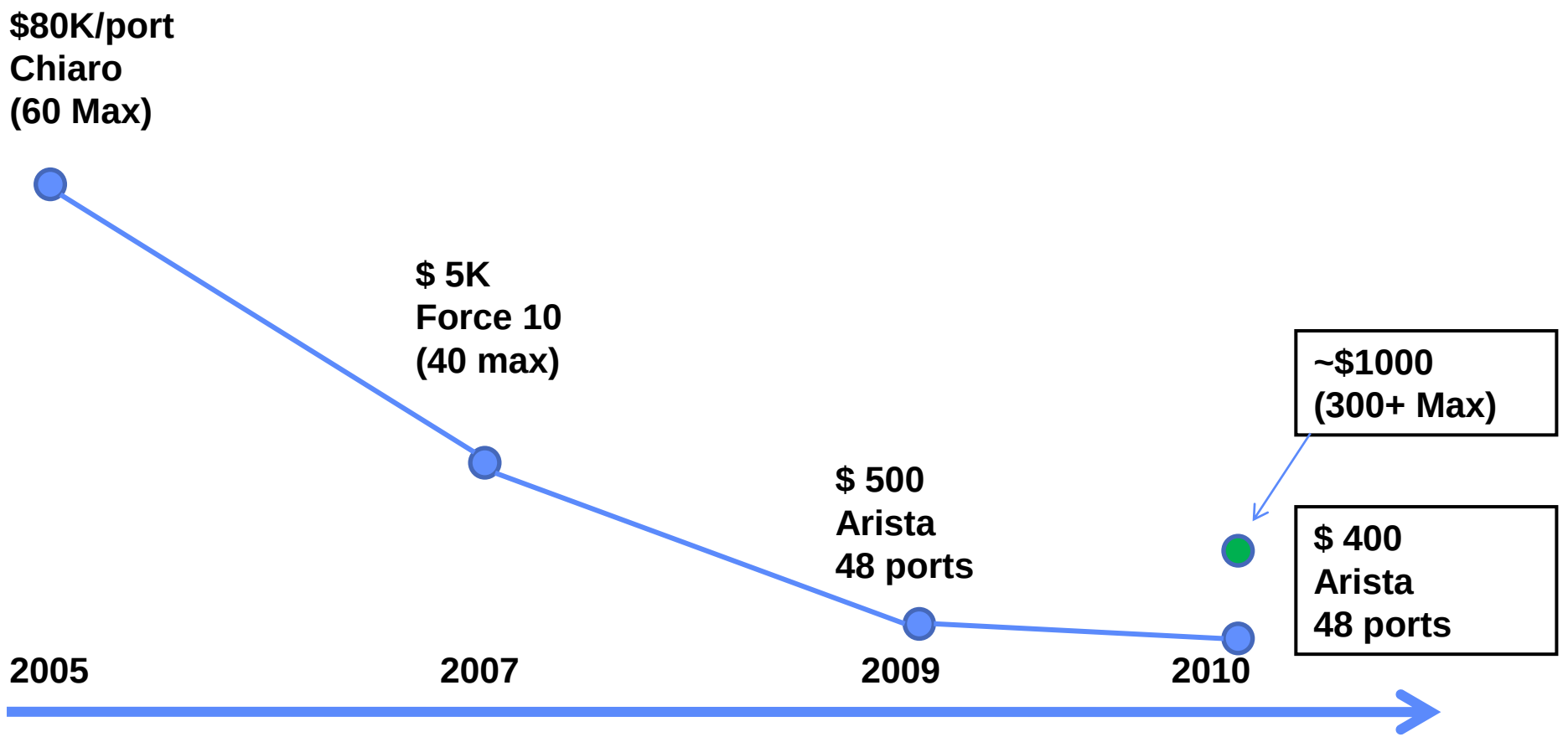
Calit2 Sunlight OptIPuter Exchange Contains Quartzite



Maxine Brown, EVL, UIC OptIPuter Project Manager

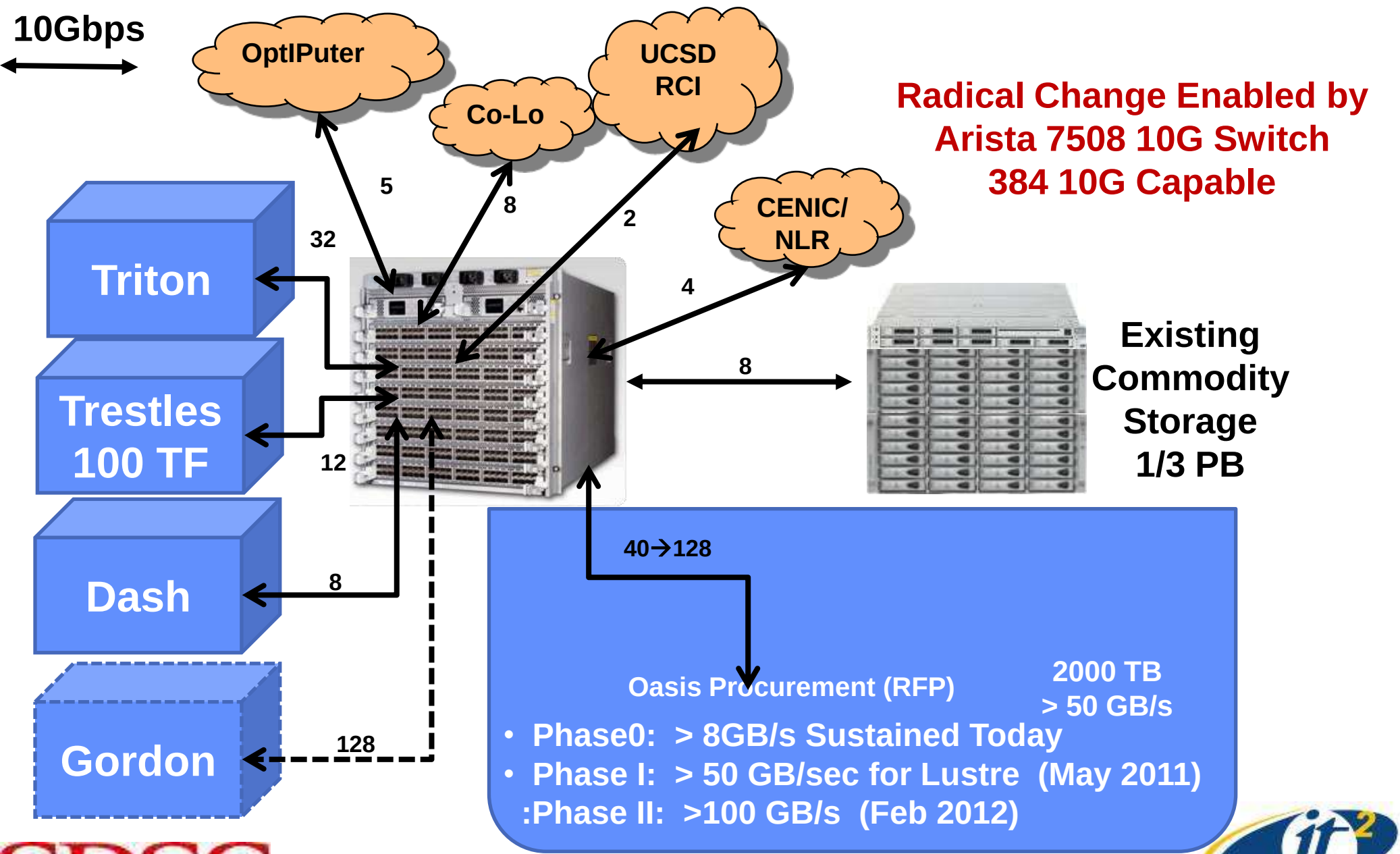
Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

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



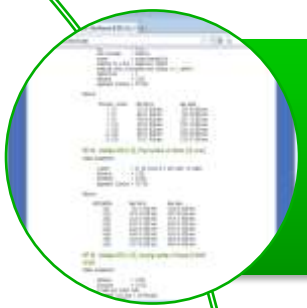
Source: Philip Papadopoulos, SDSC/Calit2

10G Switched Data Analysis Resource: SDSC's Data Oasis – Scaled Performance



Source: Philip Papadopoulos, SDSC/Calit2

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

Campus Now Starting RCI Pilot (http://rci.ucsd.edu)



SDSC

it²

UCSD Research Cyberinfrastructure (RCI) Stages

- **RCI Design Team (RCIDT)**
  - **Norman, Papadopoulos Co-Chairs**
  - **Report Completed in 2009--Report to VCR**
- **RCI Planning and Operations Committee**
  - **Ellis, Subramani Co-Chairs**
  - **Report to Chancellor**
  - **Recommended Pilot Phase--Completed 2010**
- **RCI Oversight Committee –**
  - **Norman, Gilson Co-Chairs. Started 2011**
  - **Subsidy to Campus Researchers for Co-Location & Electricity**
  - **Storage & Curation Pilot**
    - **Will be a Call for "Participation" and/or "Input" Soon**
    - **SDSC Mostly Likely Place for Physical Storage**
      - Could Add onto Data Oasis
    - **UCSD Libraries Leading the Curation Pilot**

SDSC

it²