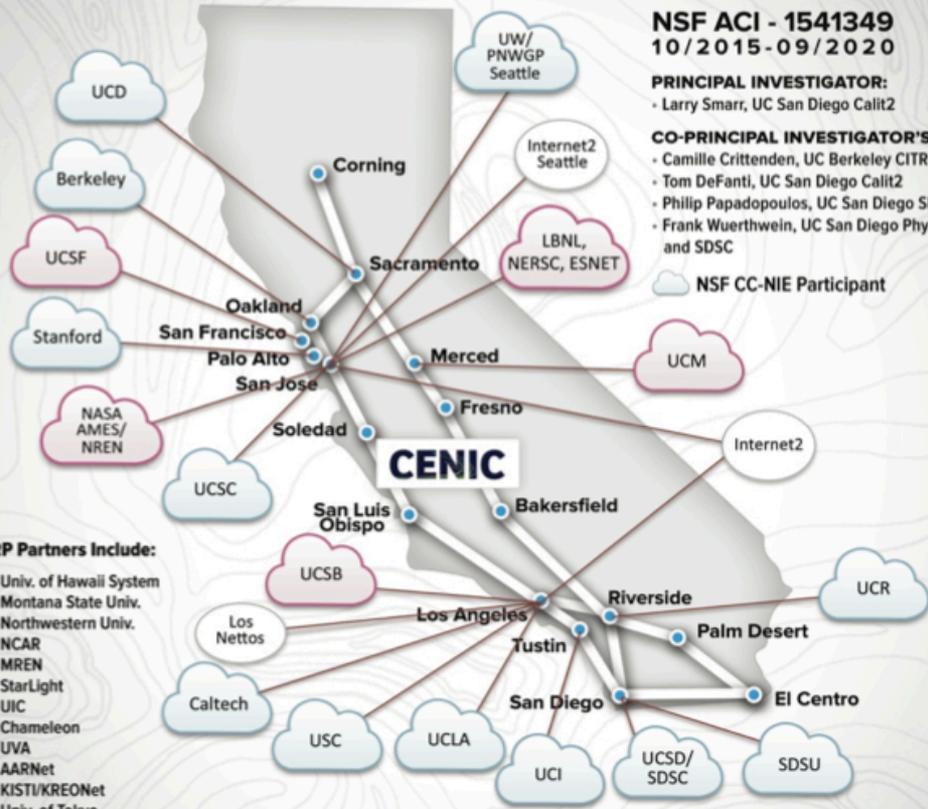


# Pacific Research Platform: Partial Status & UC @ LHC

XRootd @ Tokyo University, 11/8/2016

Matevž Tadel, UCSD

# PACIFIC RESEARCH PLATFORM — P A R T N E R —



**NSF ACI - 1541349**  
10/2015-09/2020

**PRINCIPAL INVESTIGATOR:**  
- Larry Smarr, UC San Diego Calit2

**CO-PRINCIPAL INVESTIGATOR'S:**  
- Camille Crittenden, UC Berkeley CITRIS  
- Tom DeFanti, UC San Diego Calit2  
- Philip Papadopoulos, UC San Diego SDSC  
- Frank Wuerthwein, UC San Diego Physics and SDSC

☁ NSF CC-NIE Participant

**PRP Partners Include:**

- Univ. of Hawaii System
- Montana State Univ.
- Northwestern Univ.
- NCAR
- MREN
- StarLight
- UIC
- Chameleon
- UVA
- AARNet
- KISTI/KREONet
- Univ. of Tokyo
- NCSA
- Clemson Univ.

NOTE: This diagram represents a subset of sites and connections.

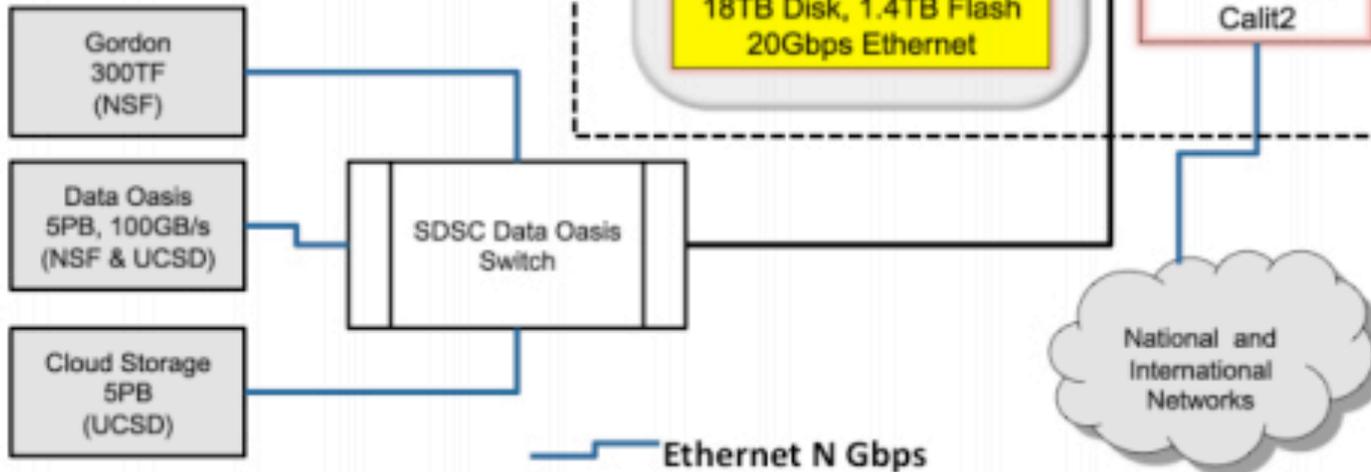
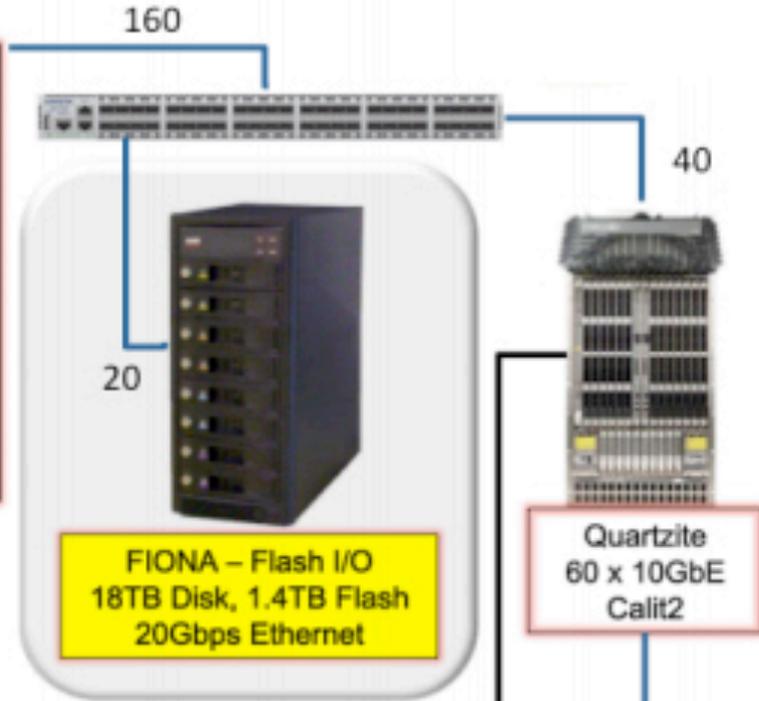
SOURCE: John Hess, CENIC

# What is Pacific Research Platform

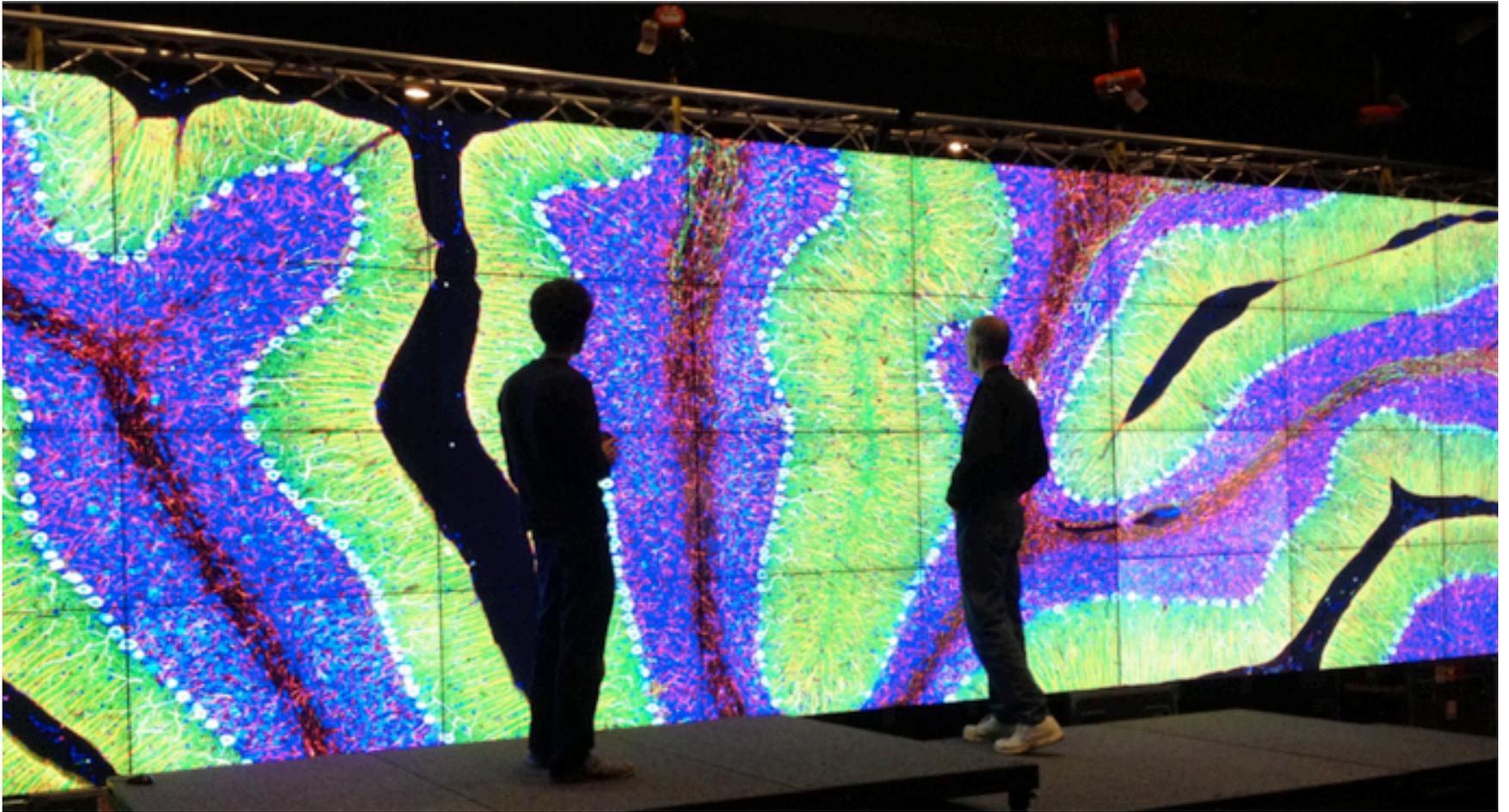
- PRP manifesto:
  - Satisfy high-speed networking needs of collaborative, big-data science.
  - Extend the campus Science DMZ model to a regional model for data-intensive networking.
  - Enable researchers to quickly and easily move data between collaborator labs, supercomputer centers, and data repositories, creating a big-data freeway that allows the data to traverse multiple, heterogeneous networks without performance degradation.
  - The PRP's data sharing architecture, with end-to-end 10–100Gb/s connections, will enable region-wide virtual co-location of data with computing.
- Principal Investigator: Larry Smarr, Director of CalIT<sup>2</sup>
- \$ 5M over three years ... 3 FTEs only ... money goes for end-point demonstrator hardware.
  - People are expected to chip in for production use.
  - More than 20 institutions, several out of California.

# Flash Input Output Network Appliance

## FIONAs: Phil Papadopoulos



# Visualization: Tom DeFanti



# HEP: Frank Würthwein

- Driving the Science part of PRP with CMS and ATLAS as examples
  - Have a couple of his slides from PRP kick-off meeting held Oct 12-14 at UCSD.
  - Interesting as it includes input from all kinds of Sciences!
- Flagship: LHC @ UC Project
  - Have CHEP slides by Jeff Dost
- Brings in XRootd as one of technologies that will be used in PRP!

**Selection from:**

# PRP End-to-End Technical Requirements From Science Applications

Frank Würthwein moderator

16 October 2015

# What is the PRP?

## Ecosystem and networking

- PRP is the answer to all research IT problems?
  - PRP is the answer to all networking problems?
- PRP is 3 FTEs to change the world
- Requirements we heard focus on much more than just networking.
- The entire gambit of compute, storage, and networking issues were present in the wish lists.
- And whatever we do must scale out faster than technology/\$\$, and be sustainable beyond the project.

# Quotes paraphrased without attribution

- Bringing home results produced at large compute centers is a **long tedious and painful process**.
- Ability to use cached data and/or **elastically scaling out** processing local data **is a game changer**.
- I want to **process any data, anytime, anywhere**.
  - processing on my laptop and supercomputer without recompiling or human data handling.
    - **transparent computing** without me having to think about where to run.
    - **from a single interface**, rather than a dozen different accounts and environments ... **but fully controlled by me where I run what**.
- **KISS** – don't give me layers of complicated middleware that are impenetrable for me.
- Whatever tools you give me must require **minimal maintenance on my part**.
- Want **to make my data accessible for download & analysis by anybody** at their home institution because I can't support everybody to compute at my institution.
- Want the same network connectivity to external resources as I have to internal resources: **Science irrespective of Geography!**
- **No matter what network bandwidth you give me, I will want more, low latency streaming bandwidth from anywhere to anywhere.**

**Large mismatch between what's needed and what can realistically be done with 3 FTE ?!?**

Leverage, Leverage, Leverage

Partner with other projects as much as we can!

Pick low hanging fruit and exemplarize them.

# Connectivity beyond PRP

- Many sciences want to connect at 10Gbps to XSEDE resources & US National labs
  - feeding data to large compute resources is a widely shared requirement.
- LHC/HEP locations want to route across LHCOne
- Connectivity to international GCM archives at “PRP quality”
- Connectivity to ALMA at “PRP quality”
- Connectivity to AWS at “PRP quality”
  - feeding data to large compute resources is a widely shared requirement.

# “Use Cases”

- Bring my data to a large compute resource for processing, and bring the output back to me when done.
- Make my data available for others to download and process with their local compute resources.
- Bring my friend’s data to me for processing with my local resources.
  - It’s probably true that nobody cares to manage these transfers, but would rather use caches that make transfer management superfluous.
- Support incorporating all of the above into orchestrated pipelines and workflows
- Support this all to happen in (quasi-)realtime, within human attention spans, and in “batched” mode “overnight”.



*Selection from:*



# Site in a Box

## Improving the Tier 3 Experience

University of California as a prototype

across [UC Davis](#), [UC Santa Cruz](#), [UC Santa Barbara](#), [UC Irvine](#), [UC Riverside](#), [UC San Diego](#)

Jeff Dost (UCSD)



# Tier 3 Challenges

## System Administration

- Software maintenance
  - WN software stack
  - CA Certs / CRLs
  - Security patches
  - Experiment software maintenance
- Network debugging
- Performance Tuning
- Non-dedicated staff (Campus IT not HEP specialized)

## Workflow Administration

- Transporting data in / out of site
- Submitting to external resources
  - Q: Don't we have WMS (CRAB / PANDA) for that?
  - A: Not all workflows are WMS appropriate
  - Want to submit seamlessly between local resources and external (common interface)
- Letting external users in

For more details, see [Lightweight sites \(#165\)](#), M. Litmaath

# Solution

- Provide a physical host to the Tier 3 that is centrally maintained by dedicated Tier 2 staff
  - Software is installed, configured, and maintained remotely from a central Puppet instance
- Host provides necessary software to handle data transfers and send users out to the grid
  - Should integrate seamlessly with Tier 3 cluster
- No LHC or Grid knowledge required from local IT

# Hardware Specifications

(aka the “brick”)



## Hardware:

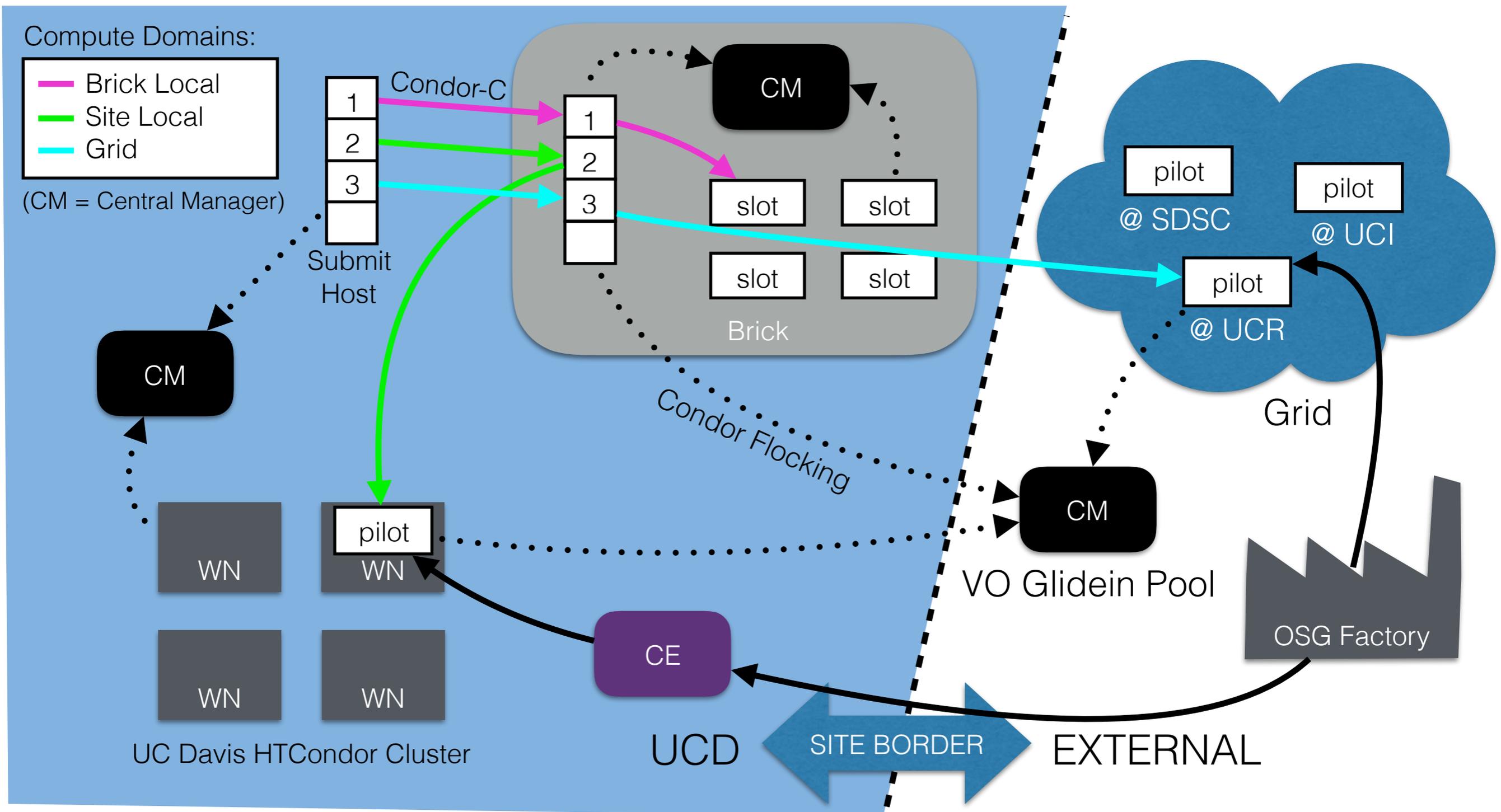
- 40 cores
- 12 x 4TB data disks (RAID6, XFS)
- 128 GB ram
- 2 x 10 gbit network interface

## Software:

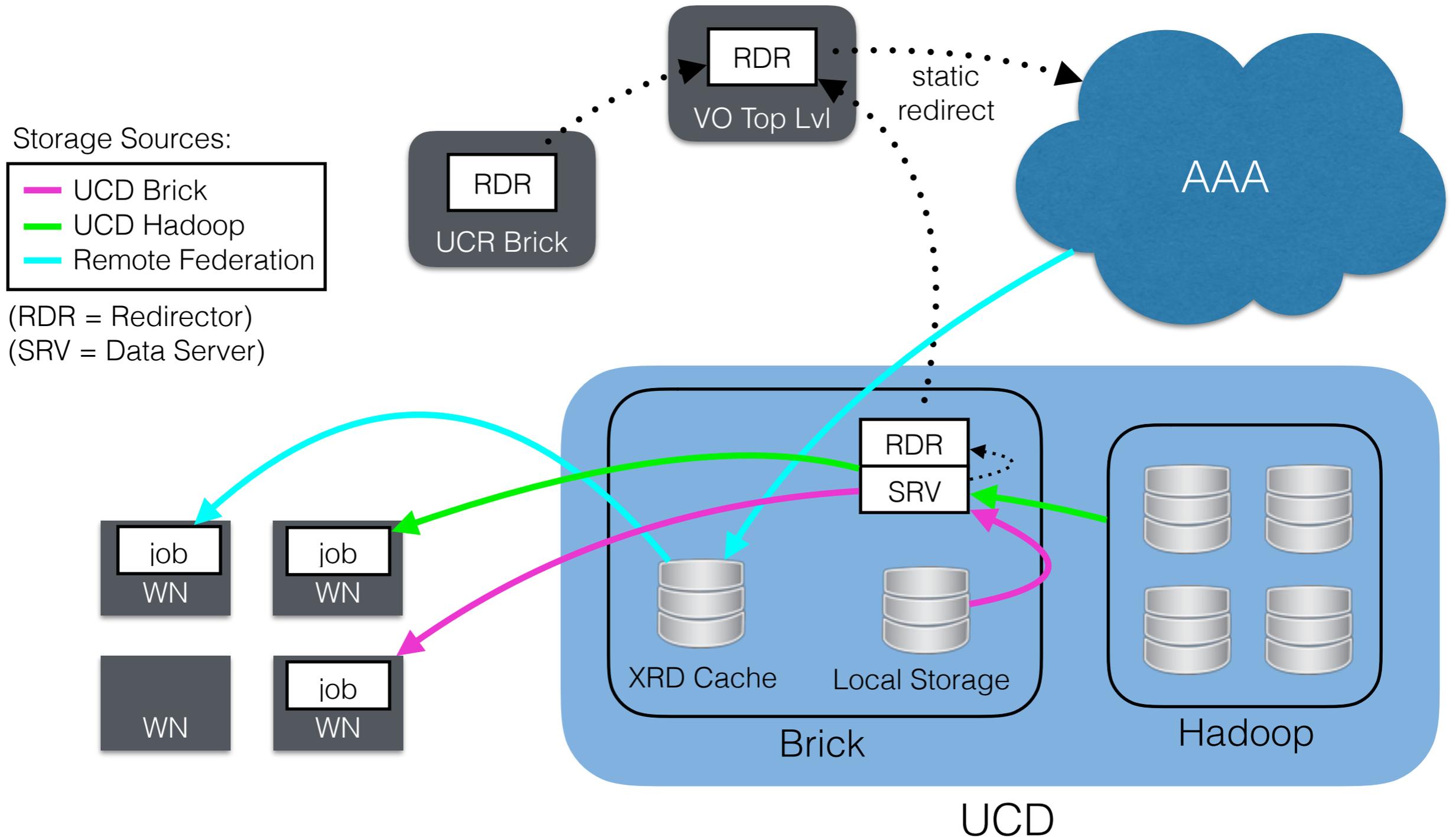
- Full HTCondor pool
- XRootD server, redirector, and proxy cache
- cvmfs w/ optional Squid

**The brick is effectively a site in a box for \$10k**

# Job Submission (UC Davis)



# XRootD Data Access



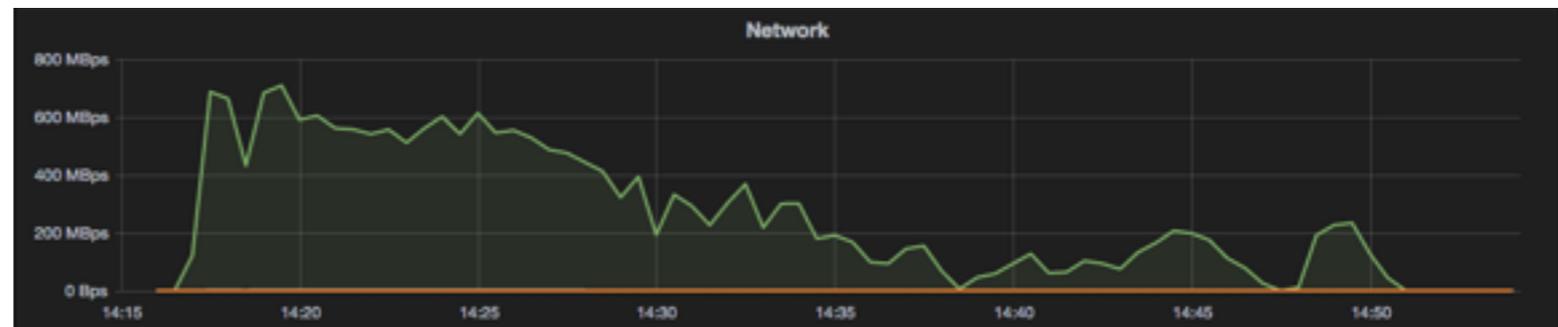
# Case Study: LHC @ UC

- Participating UCs:
  - ATLAS
  - CMS
- Science DMZs interconnected over 100gbps Pacific Research Platform network
- Resources currently provided:
  - Each UC
  - 50k core Comet cluster at SDSC
- Eventually:
  - Any other non-UC participating PRP site
  - Any OSG site beyond PRP
  - Other NSF XSEDE and DOE super computing centers
  - Cloud resources



# XRD Cache Scale Test

- 100 files ~1TB total stored at UCSC
- 10 clients running on UCSD brick fetch files at random simultaneously
- As cache fills, inbound network IO decreases and files begin to be read from disk cache
- Caching IO maxes at 5 Gbps (~640 MBytes/sec) during simultaneous disk reads and writes



# Conclusion

- The **site in a box** enhances Tier 3 sites by providing:
  - A unified way to submit locally and compute globally
  - The ability to decouple data placement from where the jobs run
- The central management of the services by dedicated Tier 2 admins allows the local Tier 3 users to worry less about infrastructure maintenance and focus more on getting science done