Accelerating Science with the Open Science Grid

Ian Stokes-Rees, PhD
Harvard Medical School, Boston, USA

http://portal.sbgrid.org
ijstokes@hkl.hms.harvard.edu

Slides and Contact

ijstokes@hkl.hms.harvard.edu

http://linkedin.com/in/ijstokes
http://slidesha.re/ijstokes-cs50

http://www.sbgrid.org
http://portal.sbgrid.org
http://www.opensciencegrid.org
http://www.xsede.org

Abstract

In the mid-1990s, the high-energy physics community (think FermiLab and CERN) started planning for the Large Hadron Collider. Managing the petabytes of data that would be generated by the facility and sharing it with the globally distributed community of over 10,000 researchers would be a major infrastructure and technology problem. This same community that brought us the web has now developed standards, software, and infrastructure for grid computing. In this seminar I'll present some of the exciting science that is being done on the Open Science Grid, the US national cyberinfrastructure linking 60 institutions (Harvard included) into a massive distributed computing and data processing system.


CERN


Earth System Grid


LIGO


IceCube
SOUTH POLE NEUTRINO DETECTOR


LSST

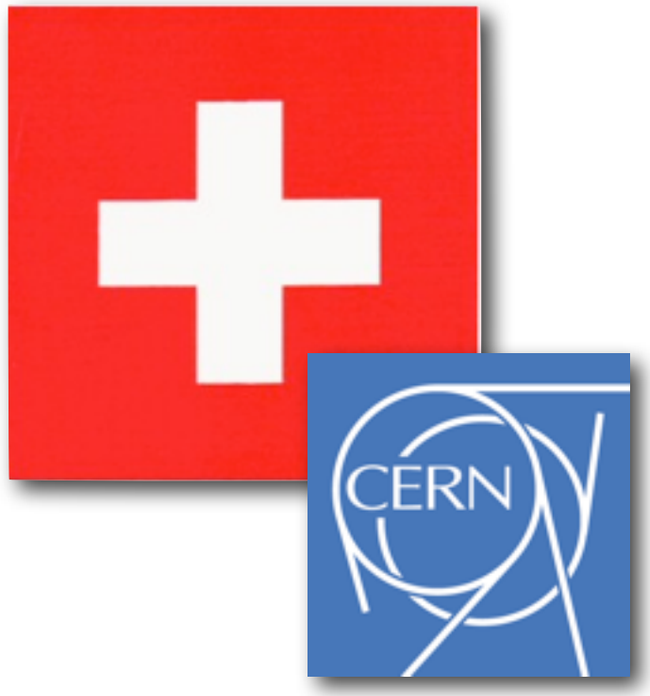

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH


NASA
JPL
Jet Propulsion Laboratory
California Institute of Technology


RCSB
PDB
PROTEIN DATA BANK

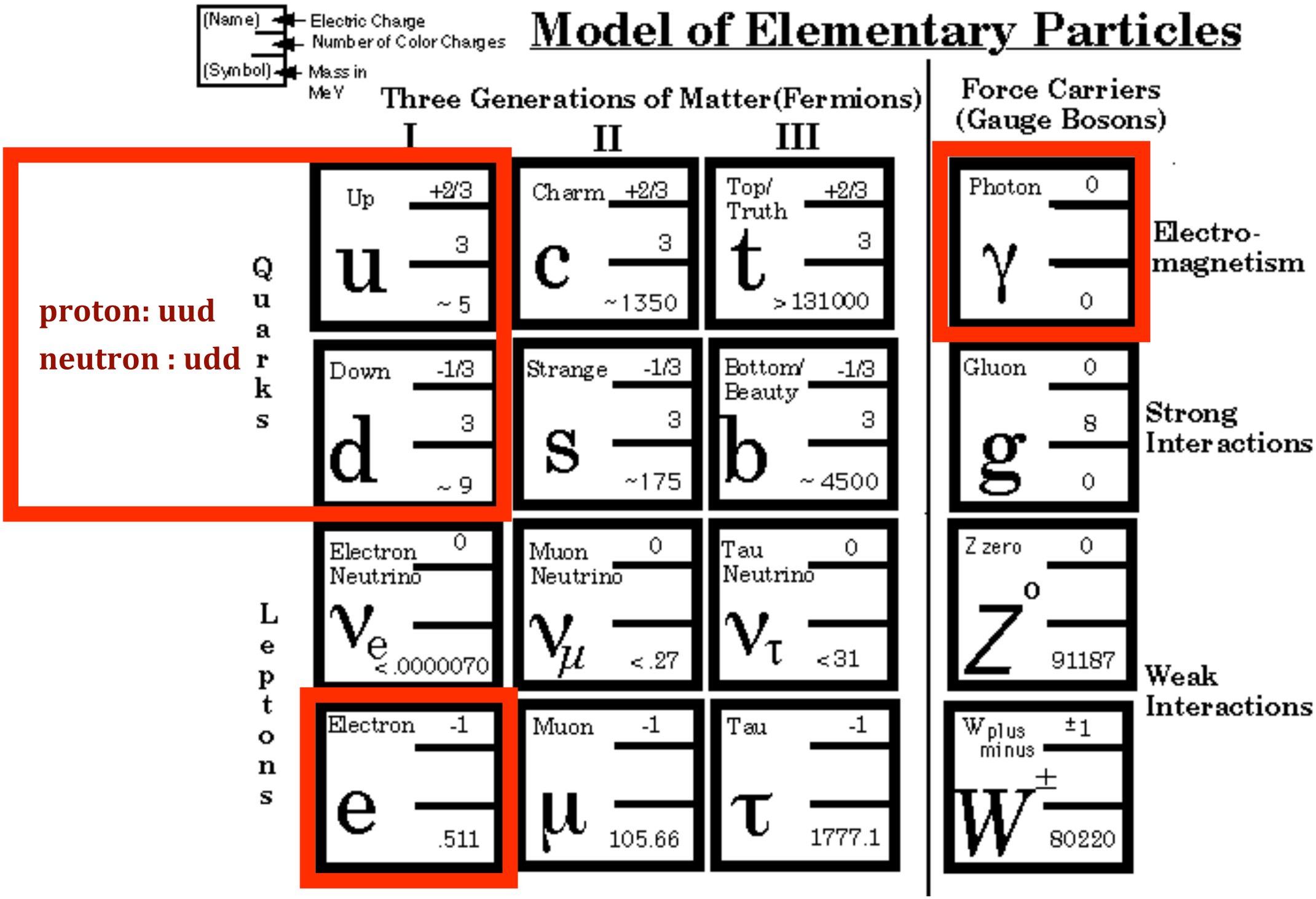

NCBI

About Me



ijstokes@hkl.hms.harvard.edu

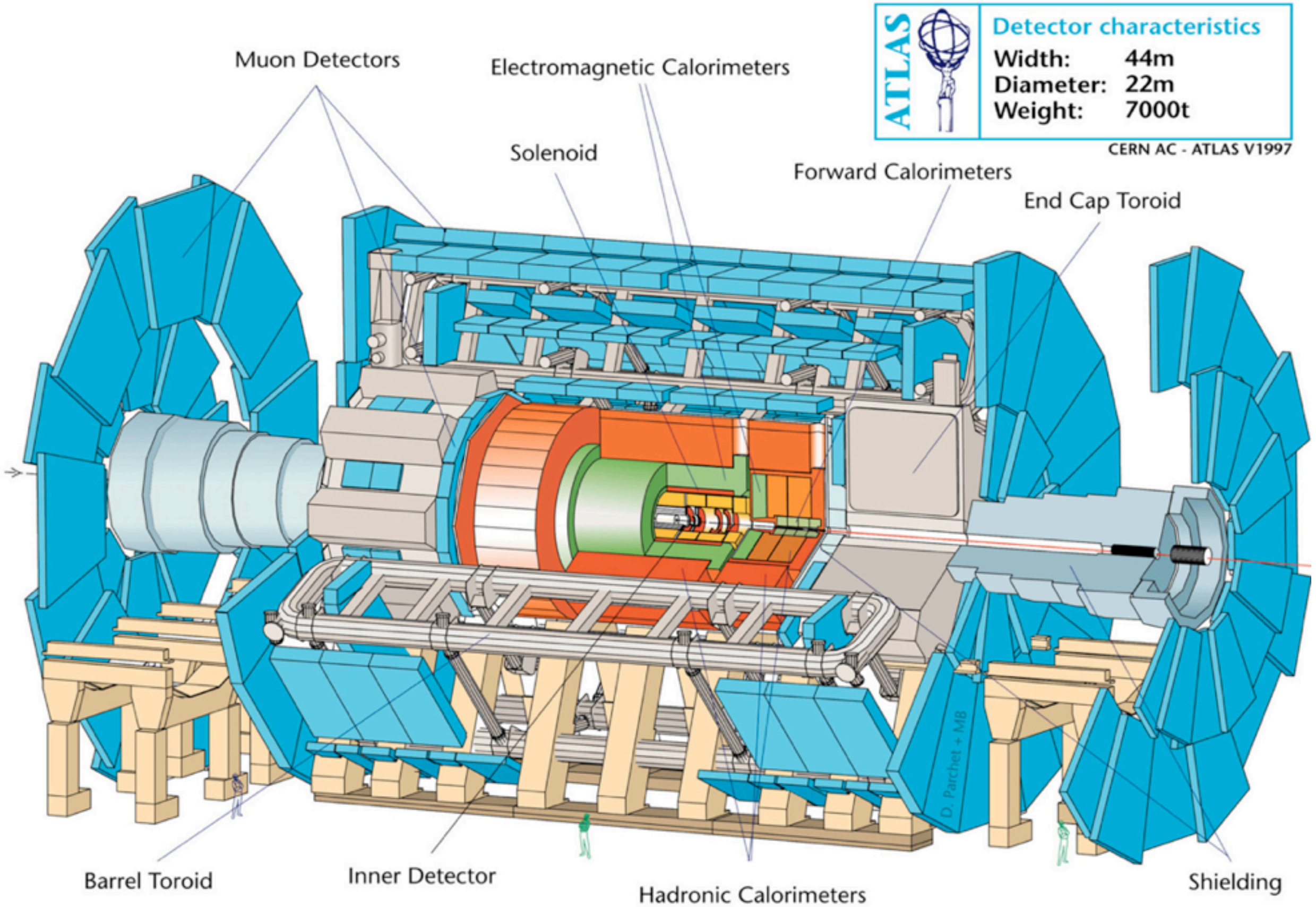
Particle Physics Standard Model

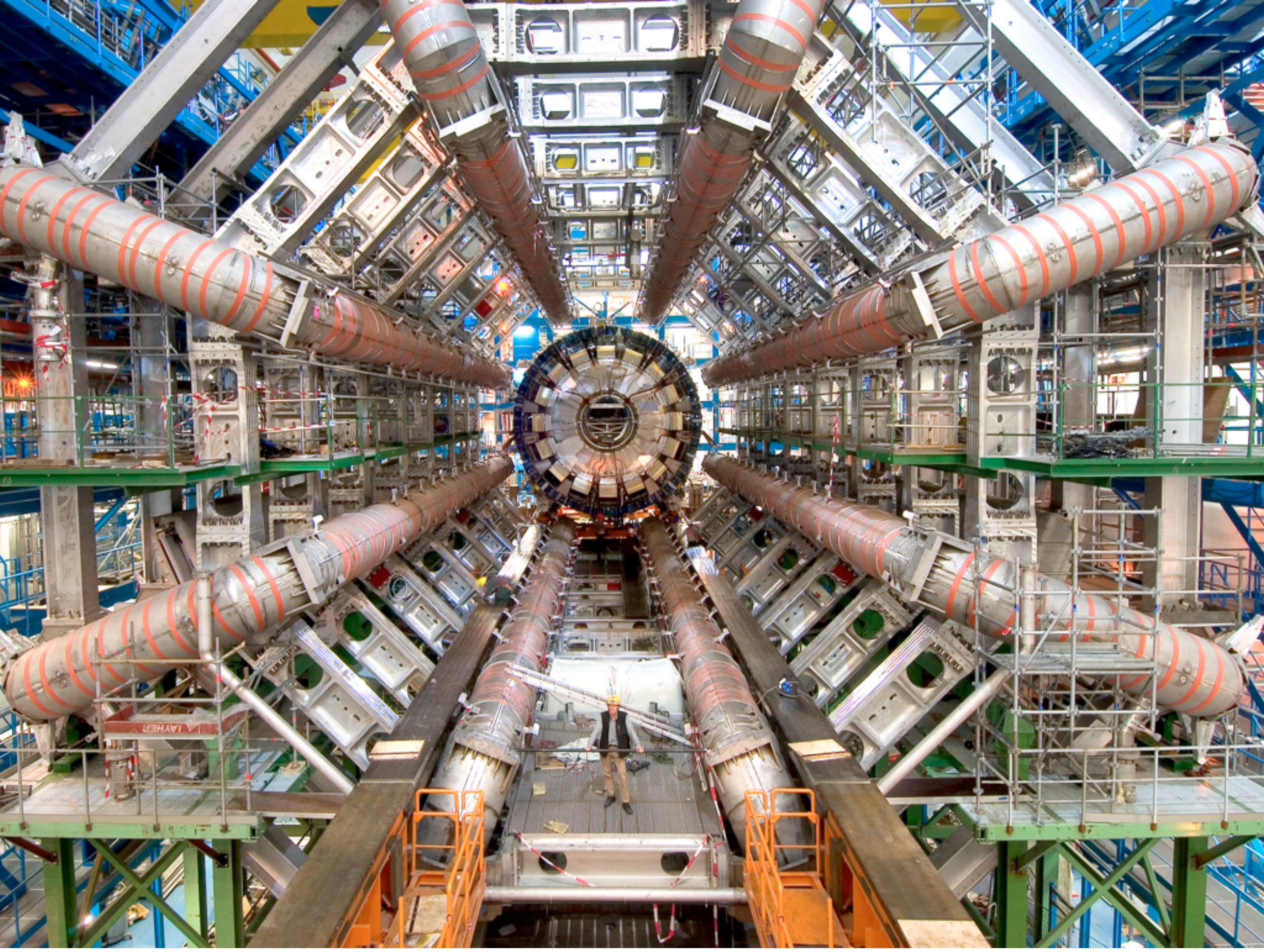


proton: uud

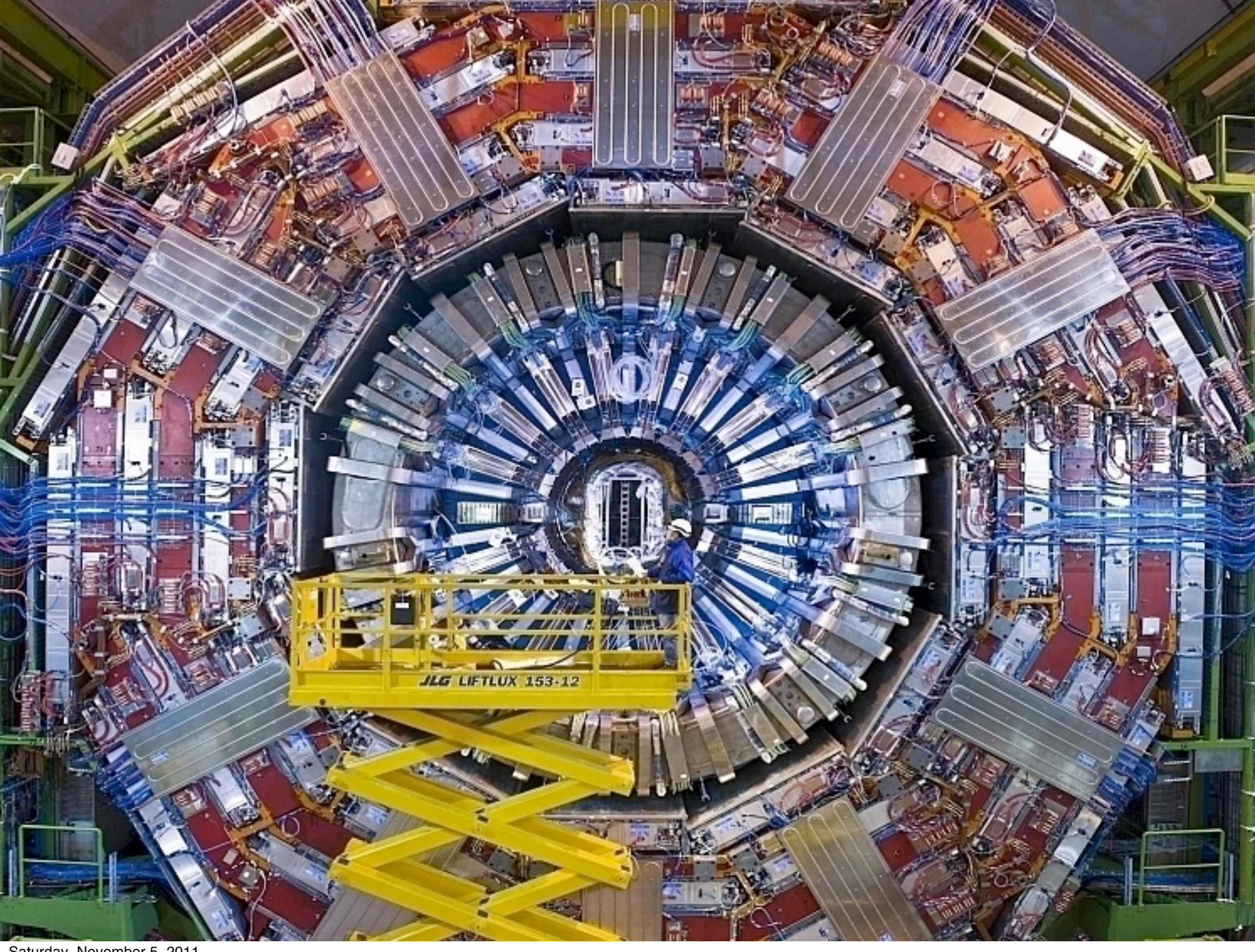
neutron : udd

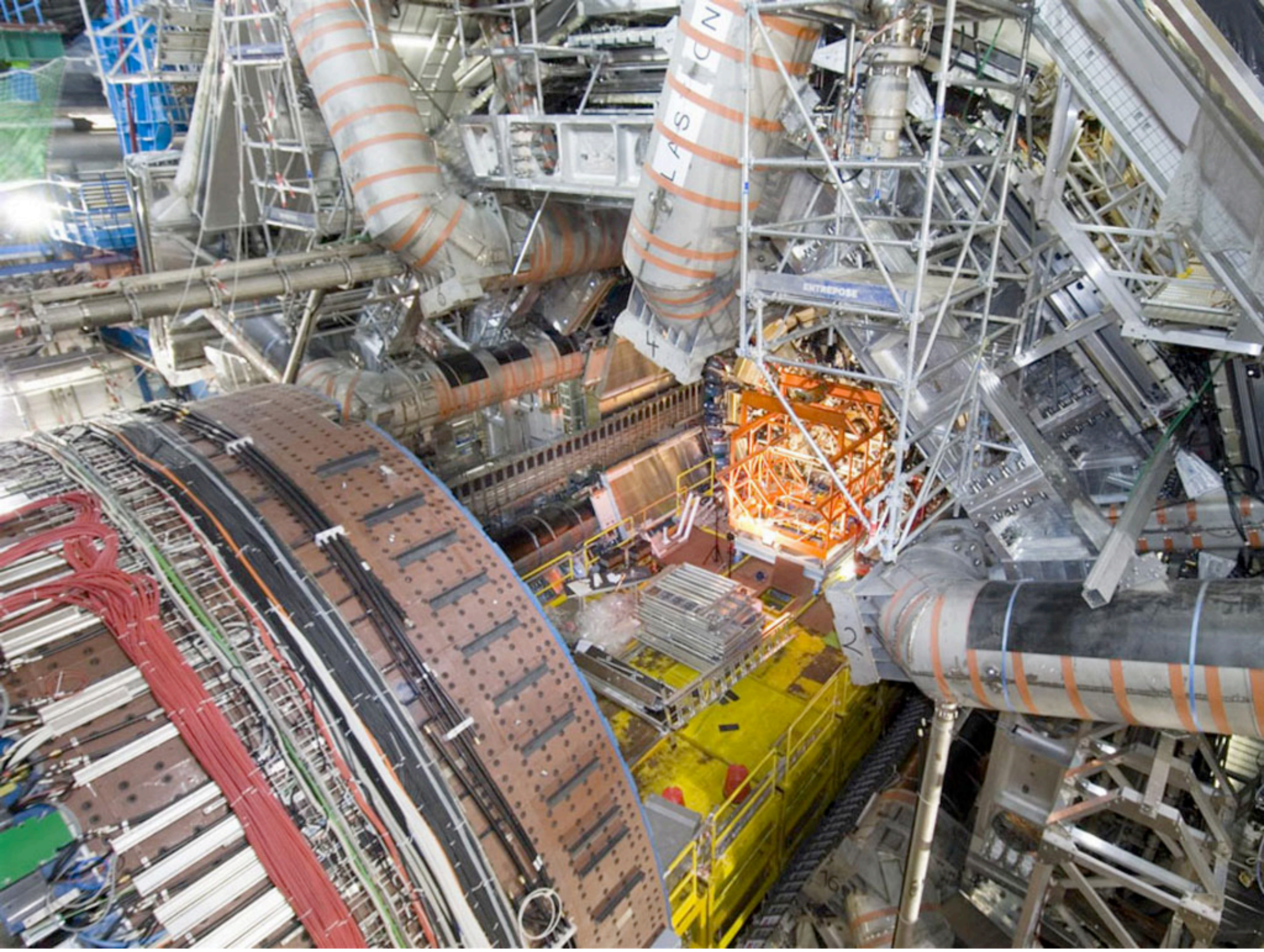
Saturday, November 5, 2011


ATLAS
Detector characteristics
Width: 44m
Diameter: 22m
Weight: 7000t
CERN AC - ATLAS V1997
Muon Detectors
Electromagnetic Calorimeters
Solenoid
Forward Calorimeters
End Cap Toroid
Barrel Toroid
Inner Detector
Hadronic Calorimeters
Shielding
D. Parchet + MB

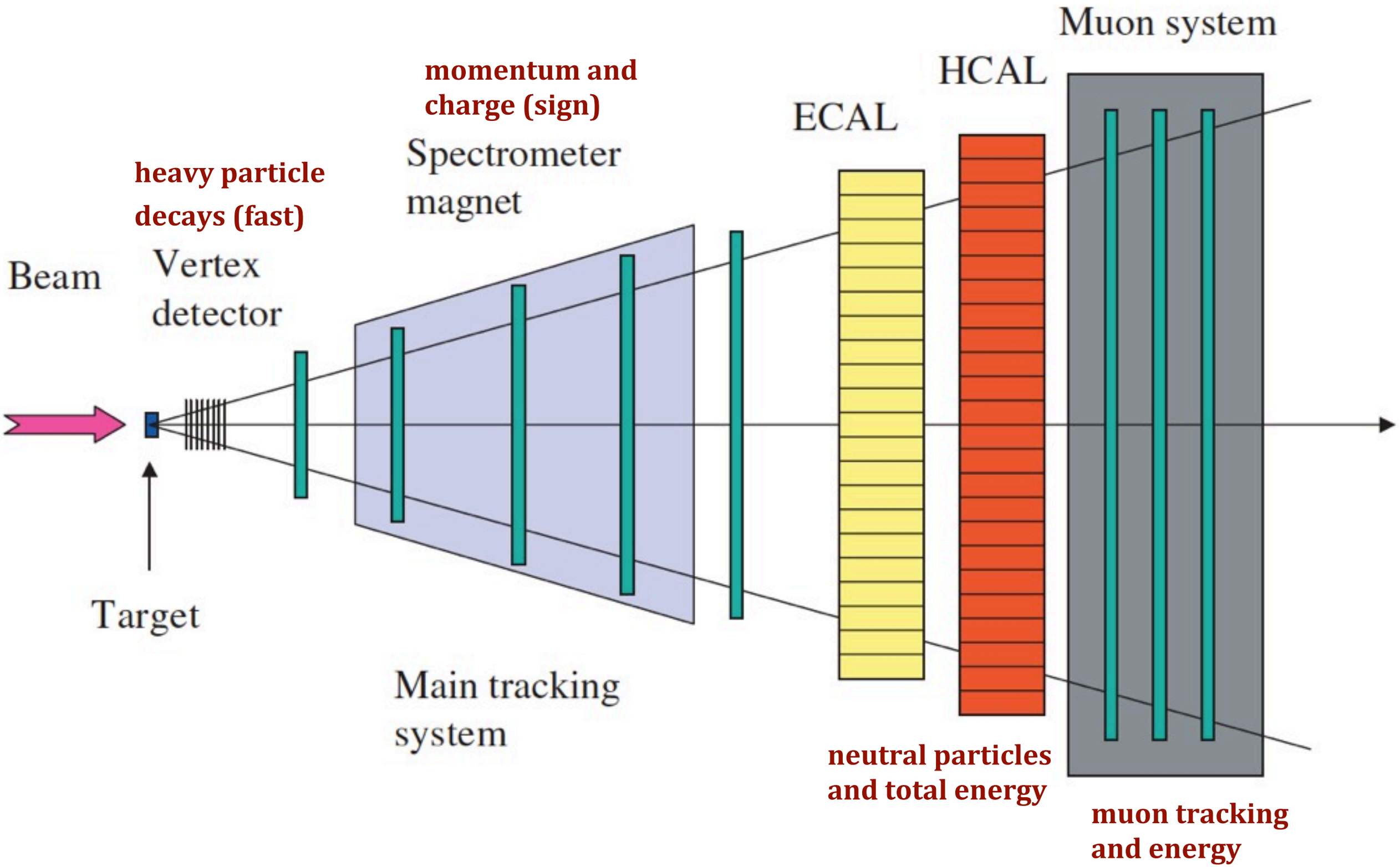


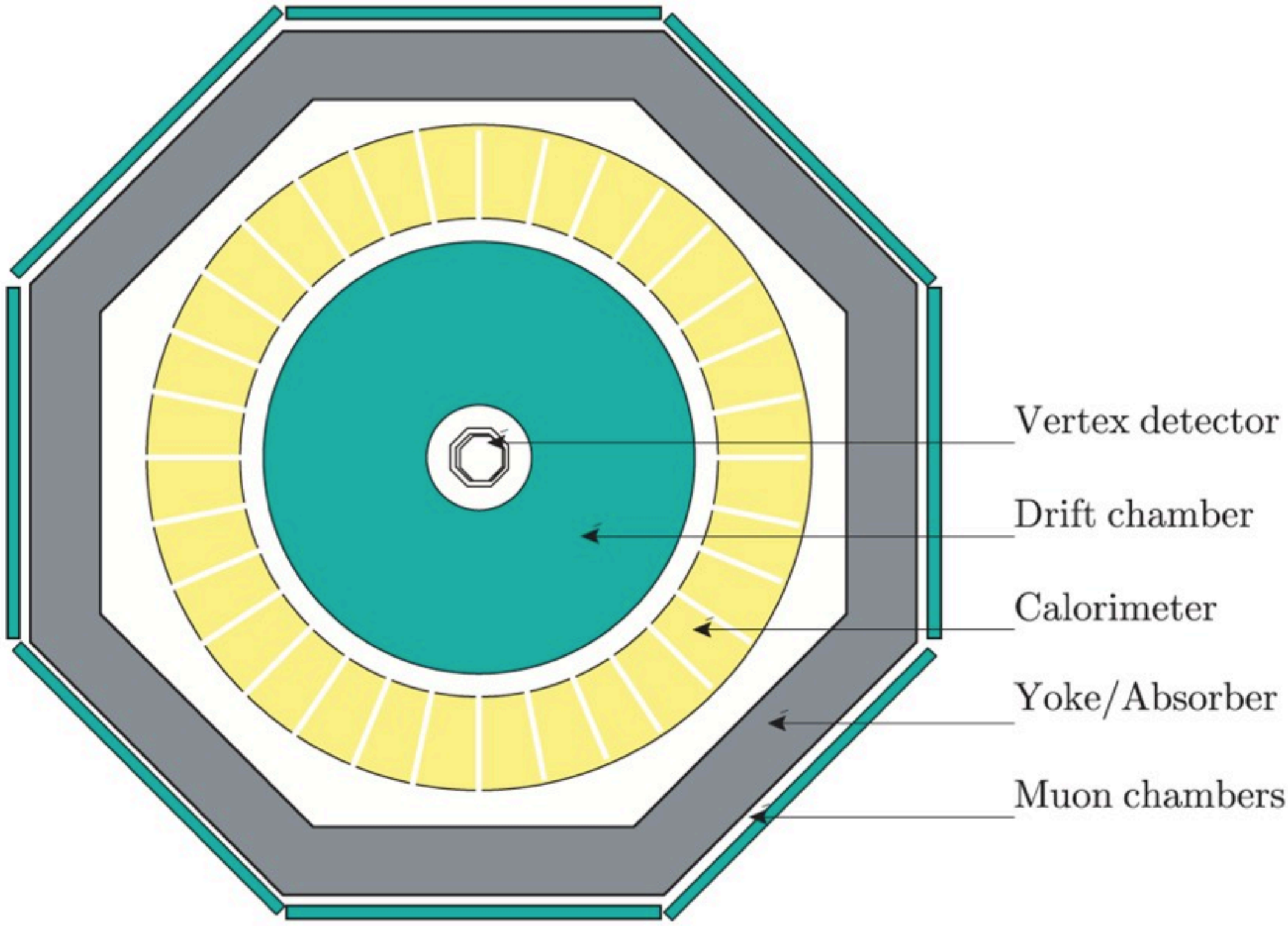
Saturday, November 5, 2011

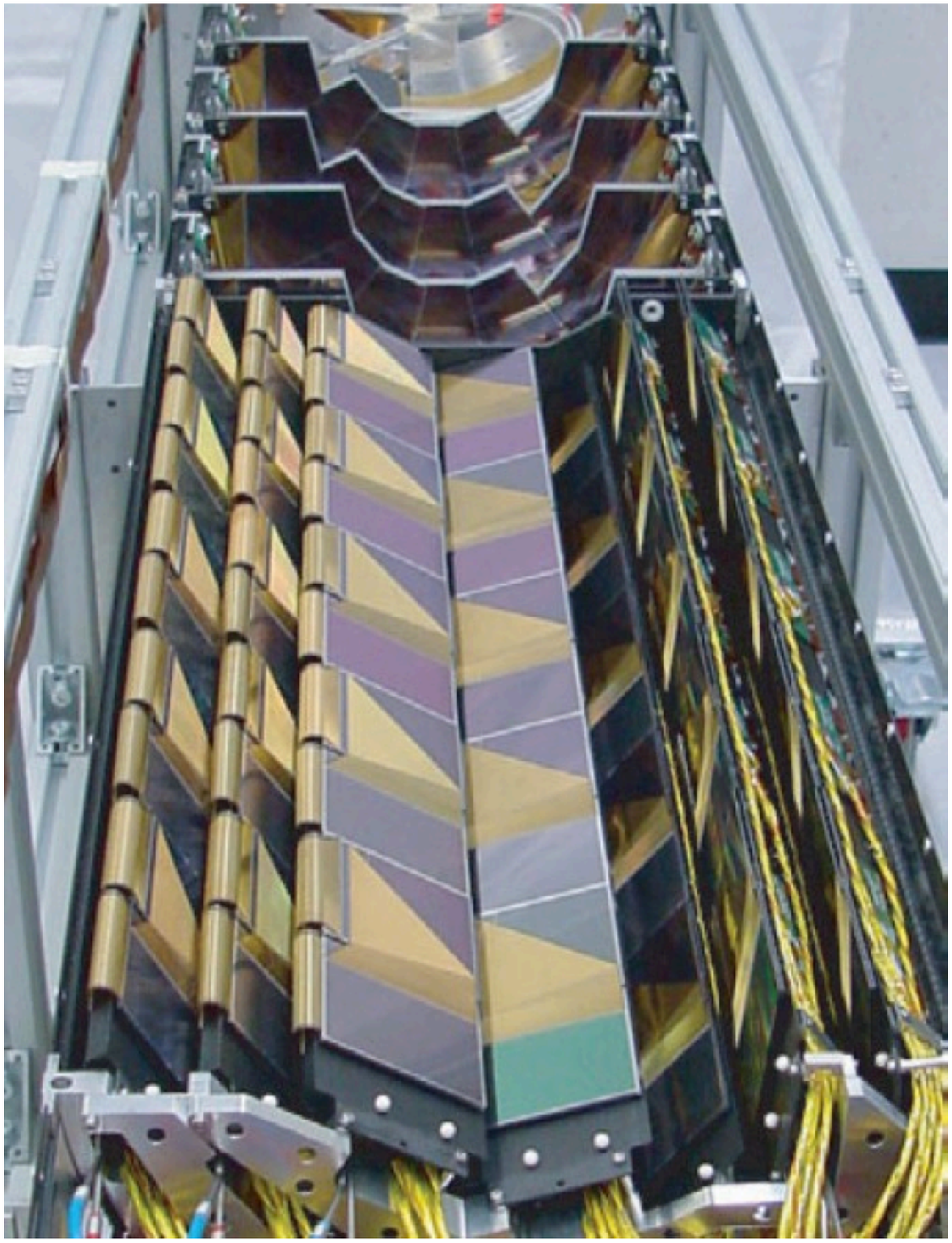

JLG LIFTLUX 153-12


ATLAS

animation 1 - ATLAS collision




Vertex detector
Drift chamber
Calorimeter
Yoke/Absorber
Muon chambers



animation 2 -
tracking and energy

40 MHz bunch crossing rate

10 million data channels

1 KHz level 1 event recording rate

1-10 MB per event

14 hours per day, 7+ months / year

4 detectors

6 PB of data / year

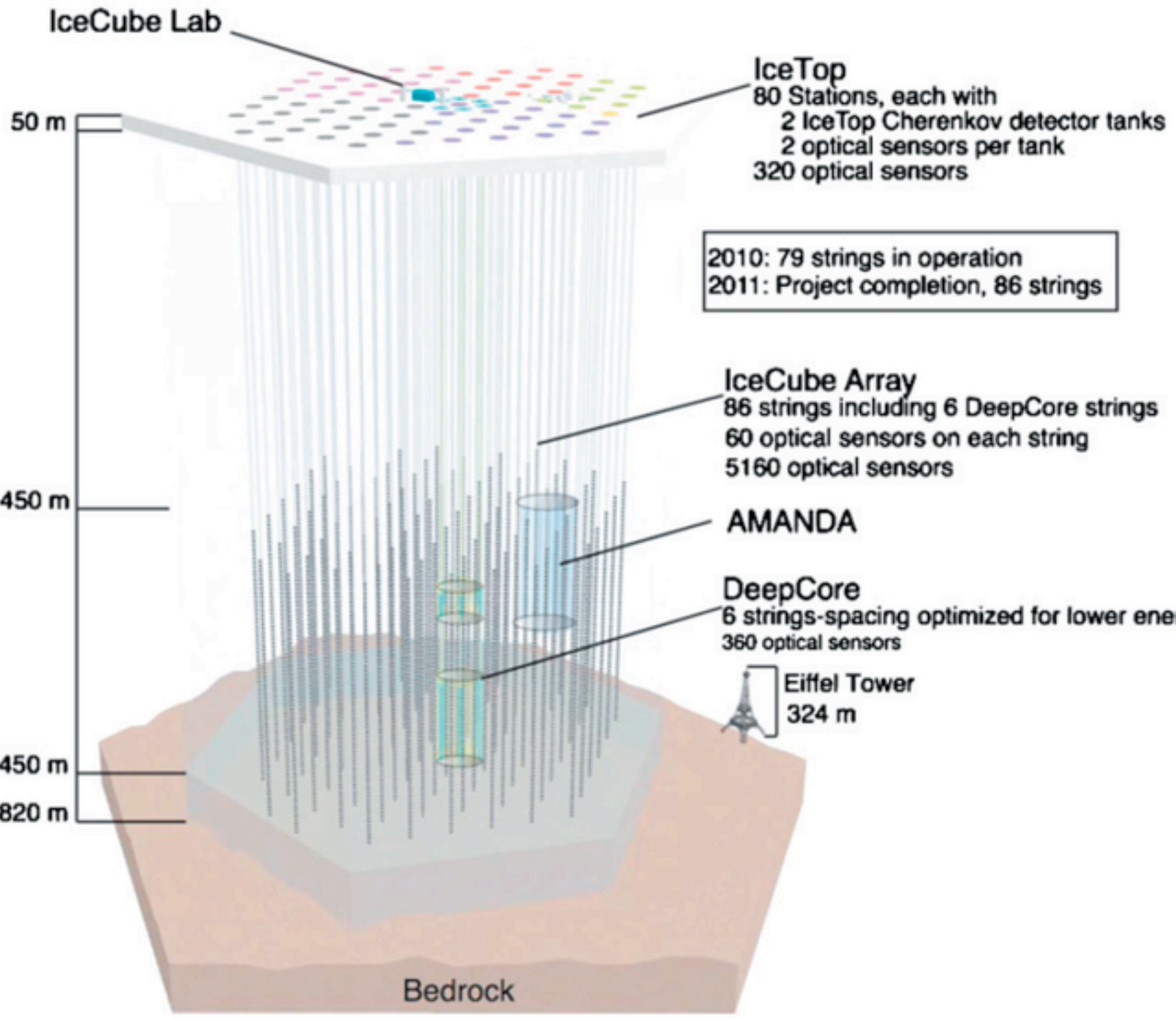
globally distribute data for analysis (x2)

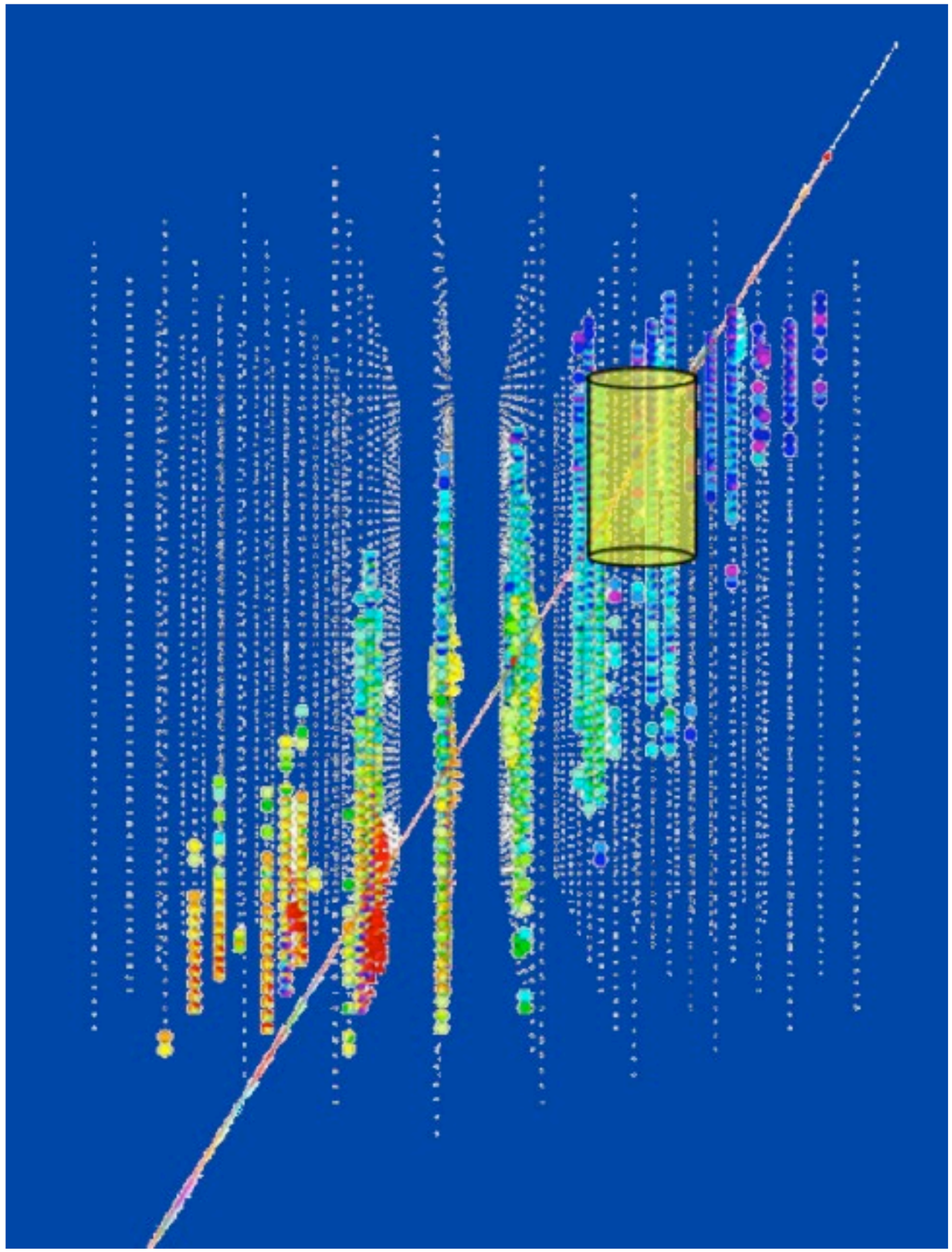
What are the computing challenges?

- Track reconstruction
  - Data distributed globally, processed, and re-processed around the clock
- Search for new physics
  - Modify physics model, see if better predictions are produced
- Data infrastructure
  - A lot of data to move around
- Computing infrastructure
  - A lot of data to process
- Global collaboration
  - thousands of users around the world: how to securely share systems, data, computations, results
  - federated scientific computing infrastructure

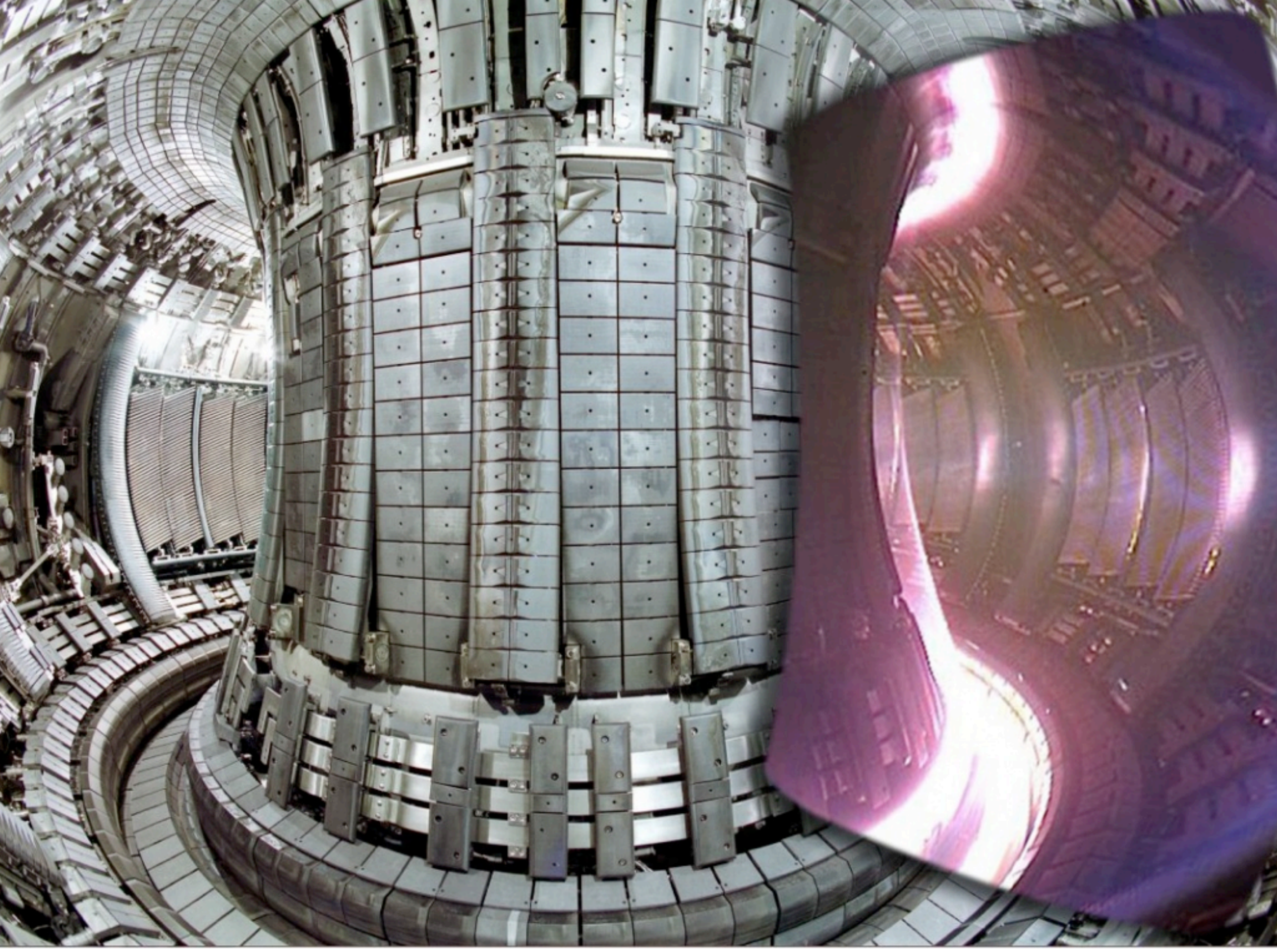
IceCube

SOUTH POLE NEUTRINO DETECTOR

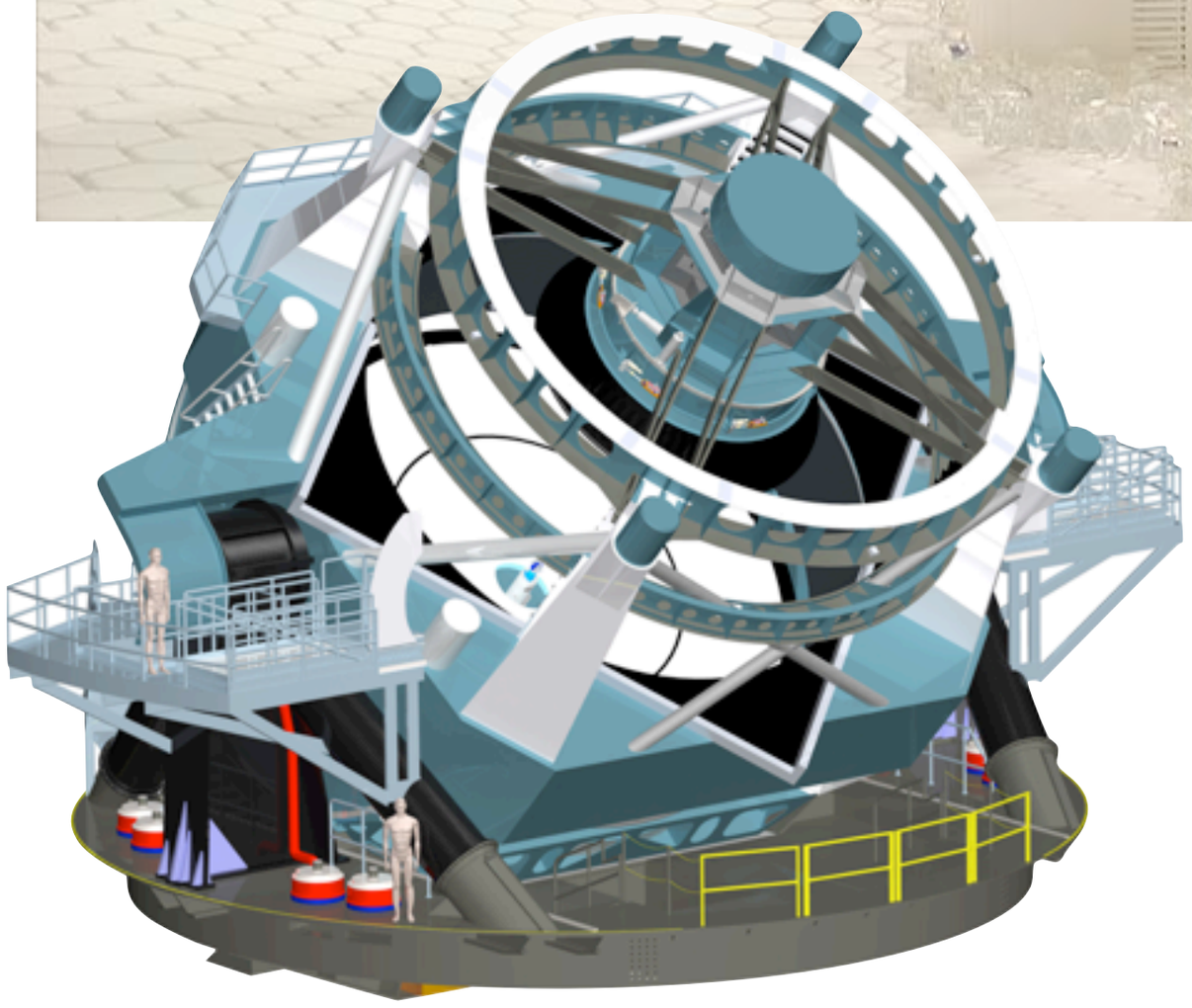





SCIENCE photoLIBRARY



Saturday, November 5, 2011

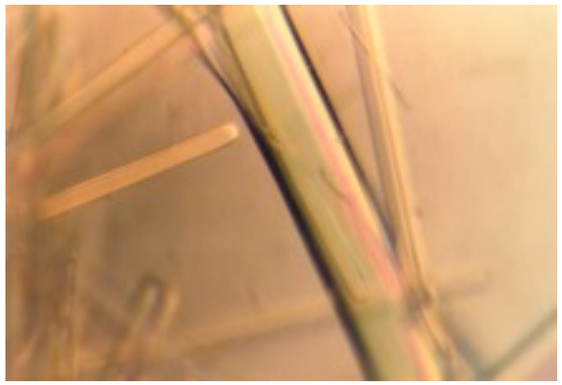


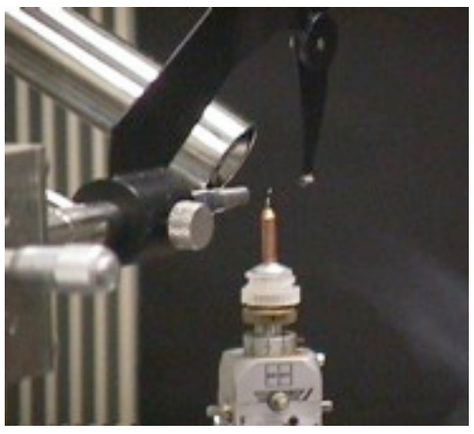
3 billion pixels

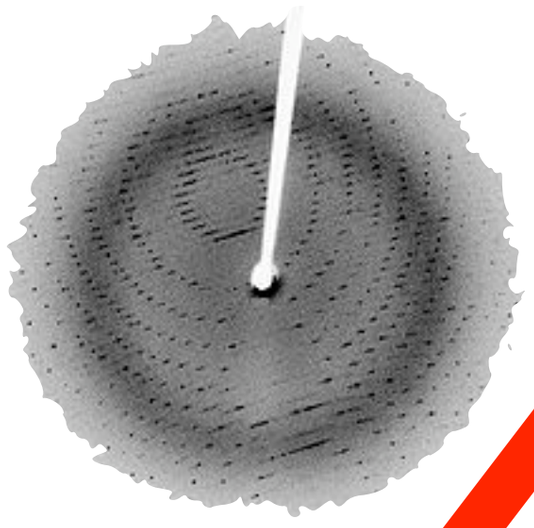
55k x 55k equivaltent

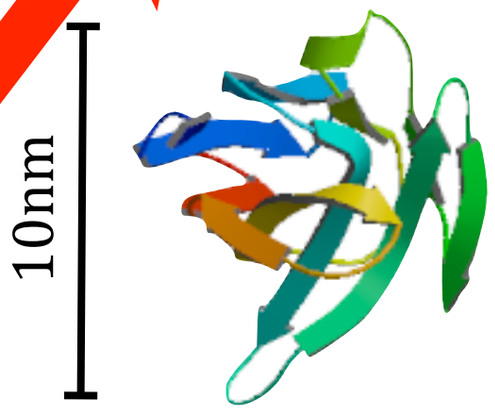
ijstokes@hkl.hms.harvard.edu

Study of Protein Structure and Function



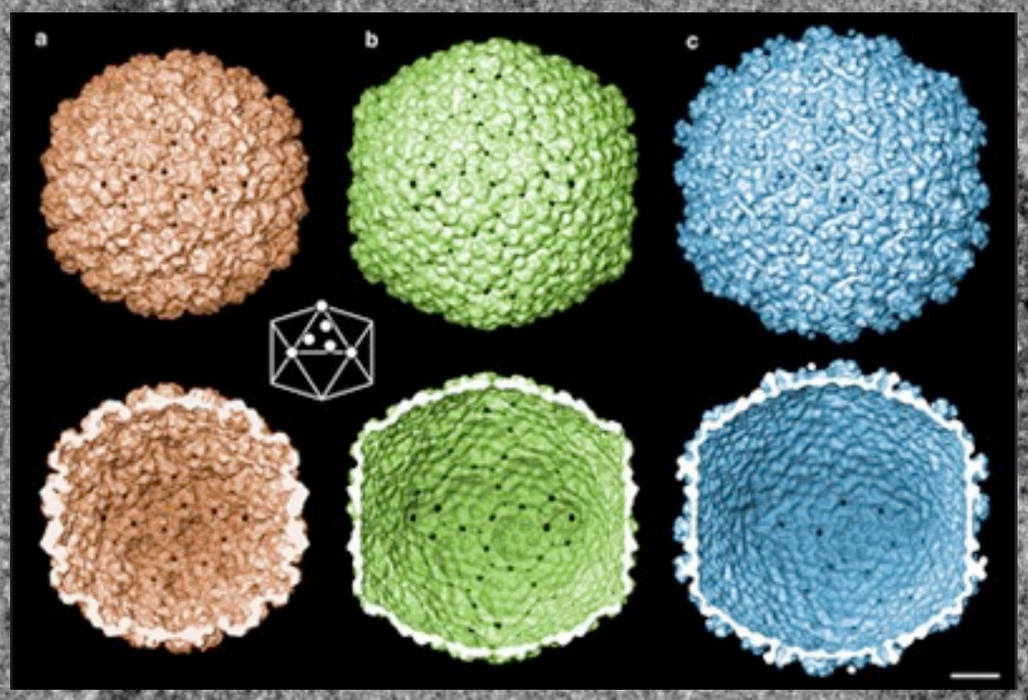


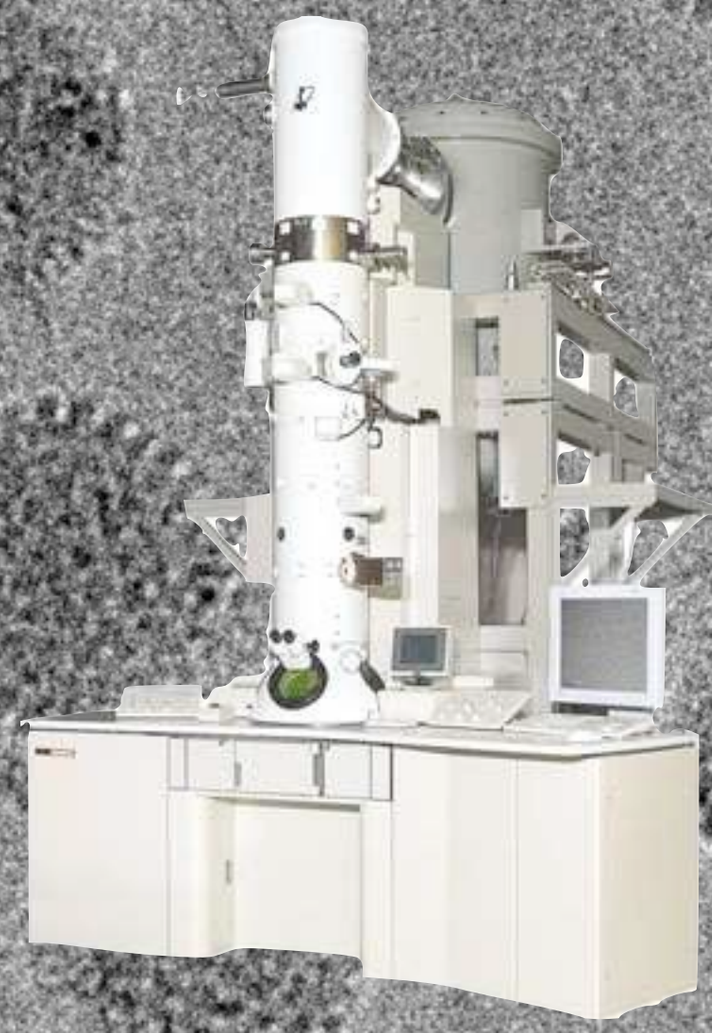




- Shared scientific data collection facility
- Data intensive (10-100 GB/day)

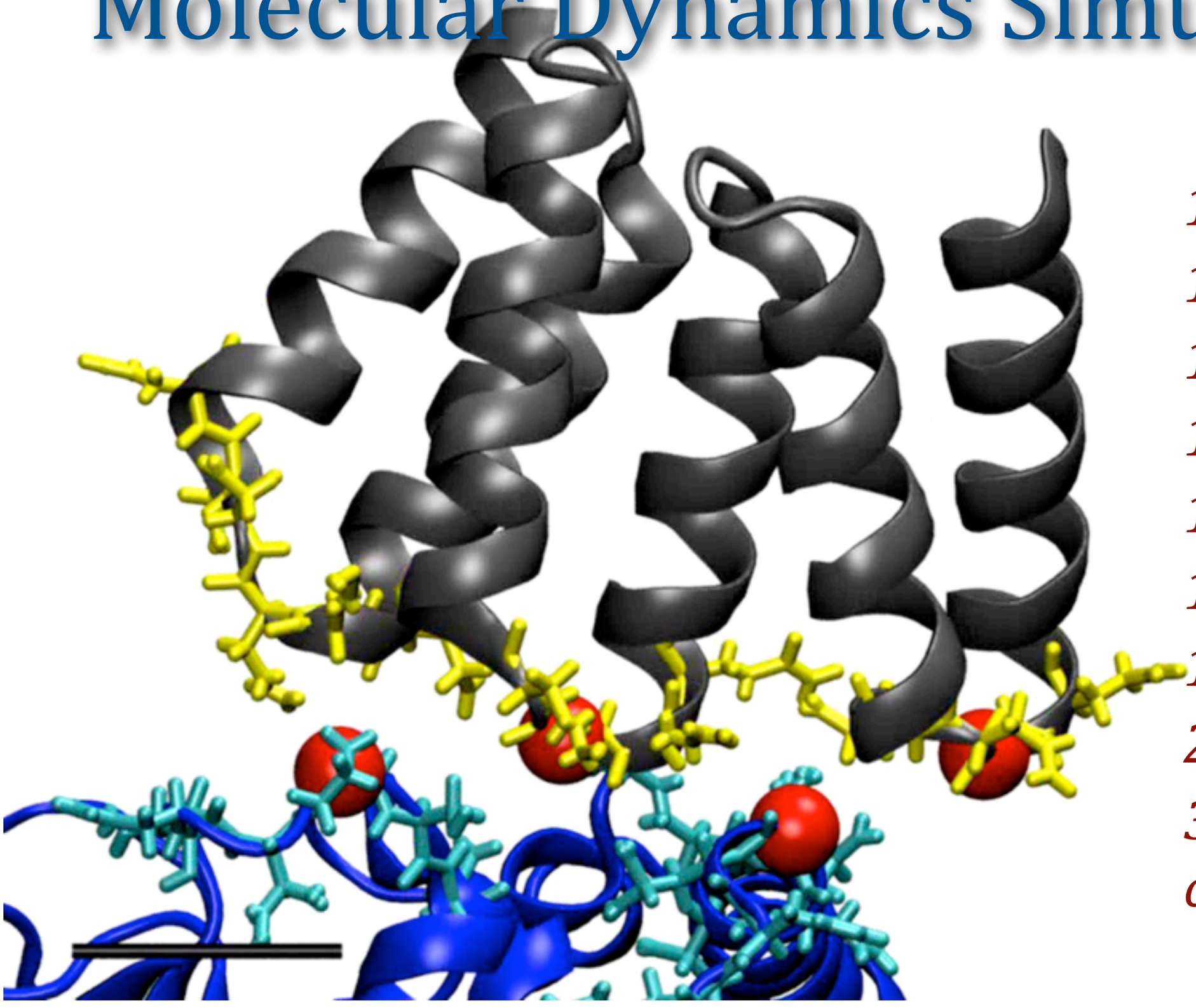
Cryo Electron Microscopy





- *Previously, 1-10,000 images, managed by hand*
- *Now, robotic systems collect millions of hi-res images*
- *estimate 250,000 CPU-hours to reconstruct model*

Molecular Dynamics Simulations



1 fs time step

1ns snapshot

1 us simulation

1e6 steps

1000 frames

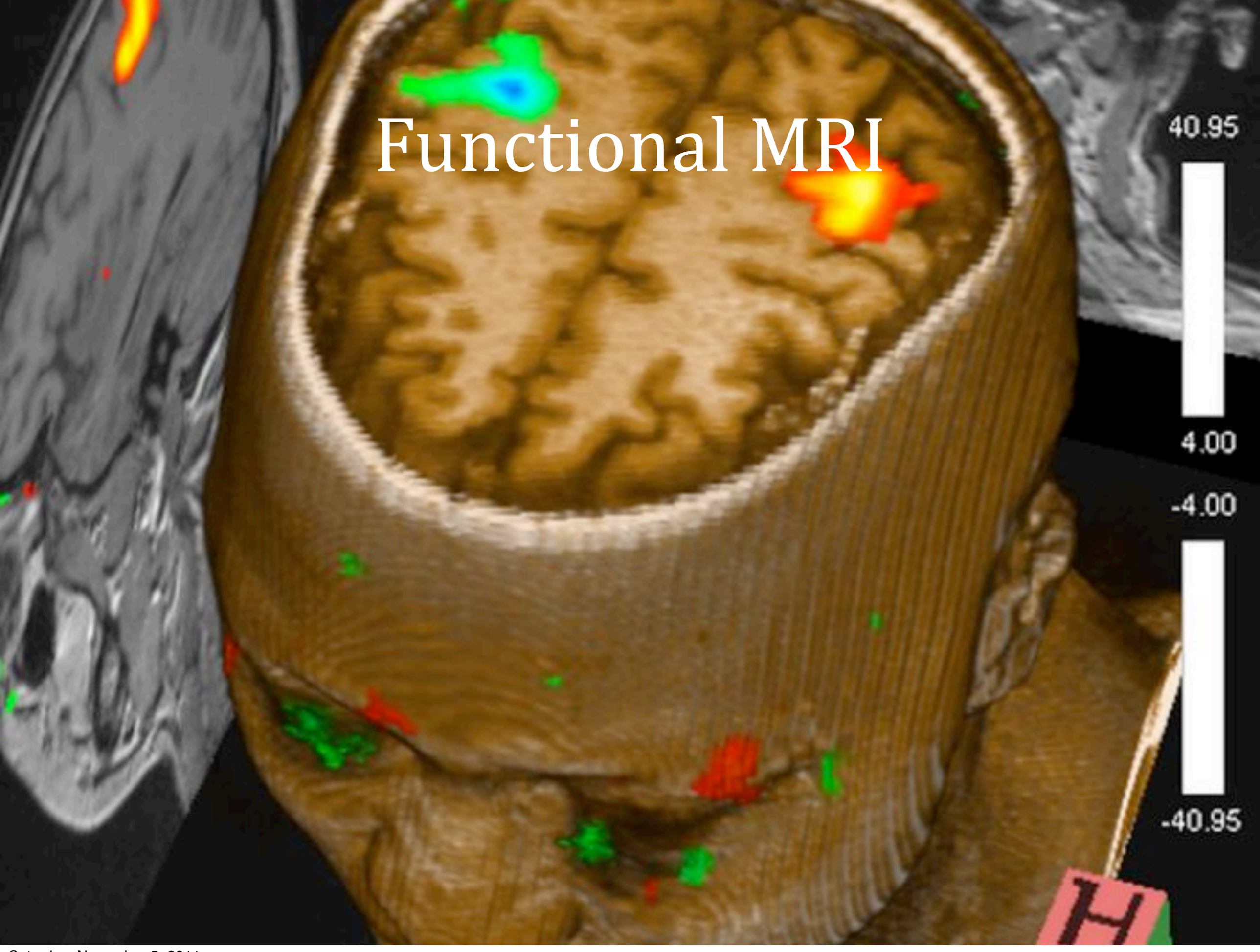
10 MB / frame

10 GB / sim

20 CPU-years

3 months (wall-clock)

Functional MRI



Next Generation Sequencing

Saturday, November 5, 2011

This is the edge of science ....
.... and computing is everywhere

- computational model of system
- simulation
- analysis
- visualization
- data management infrastructure
- computational infrastructure
- collaboration, identity mgmt, access control

Boston Life Sciences Hub

- Biomedical researchers
- Government agencies
- Life sciences
- Universities
- Hospitals

Scientific Research Today

- International collaborations
  - IT becomes embedded into research process: data, results, analysis, visualization
  - Crossing institutional and national boundaries

- Computational techniques increasingly important
  - **...** and computationally **intensive** techniques as well
  - requires use of high performance computing systems

- Data volumes are growing fast
  - hard to share
  - hard to manage

- Scientific software often difficult to use
  - or to use properly

- Web based tools increasingly important
  - but often lack disconnect from persisted and shared results

Required:

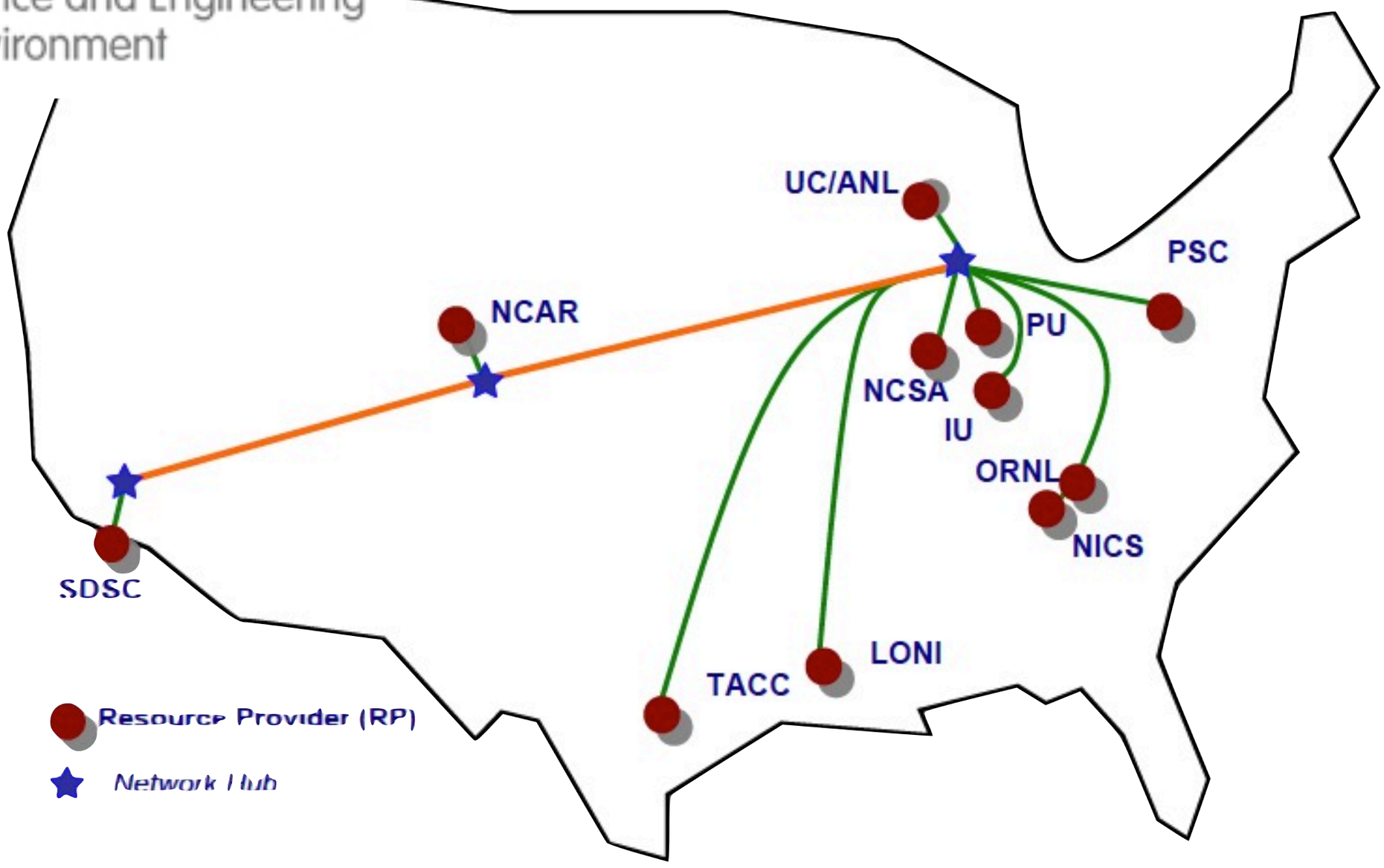
Collaborative environment for compute and data intensive science

XSEDE

Extreme Science and Engineering Discovery Environment

http://www.xsede.org

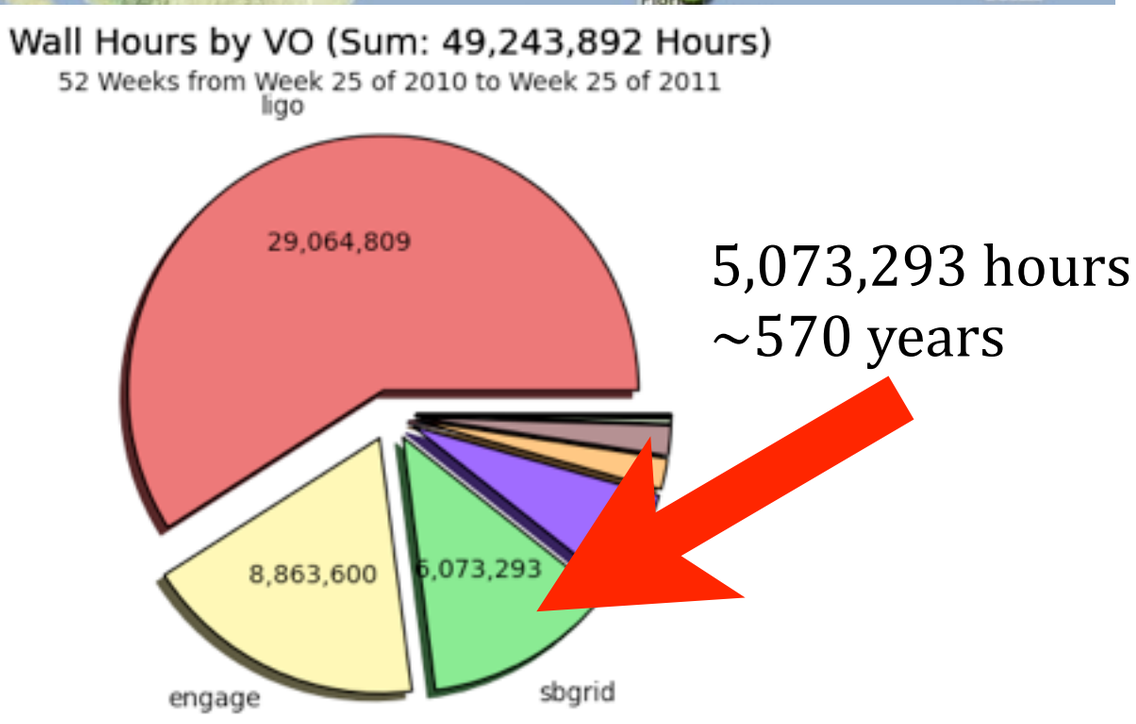
TeraGrid™



Open Science Grid

http://opensciencegrid.org

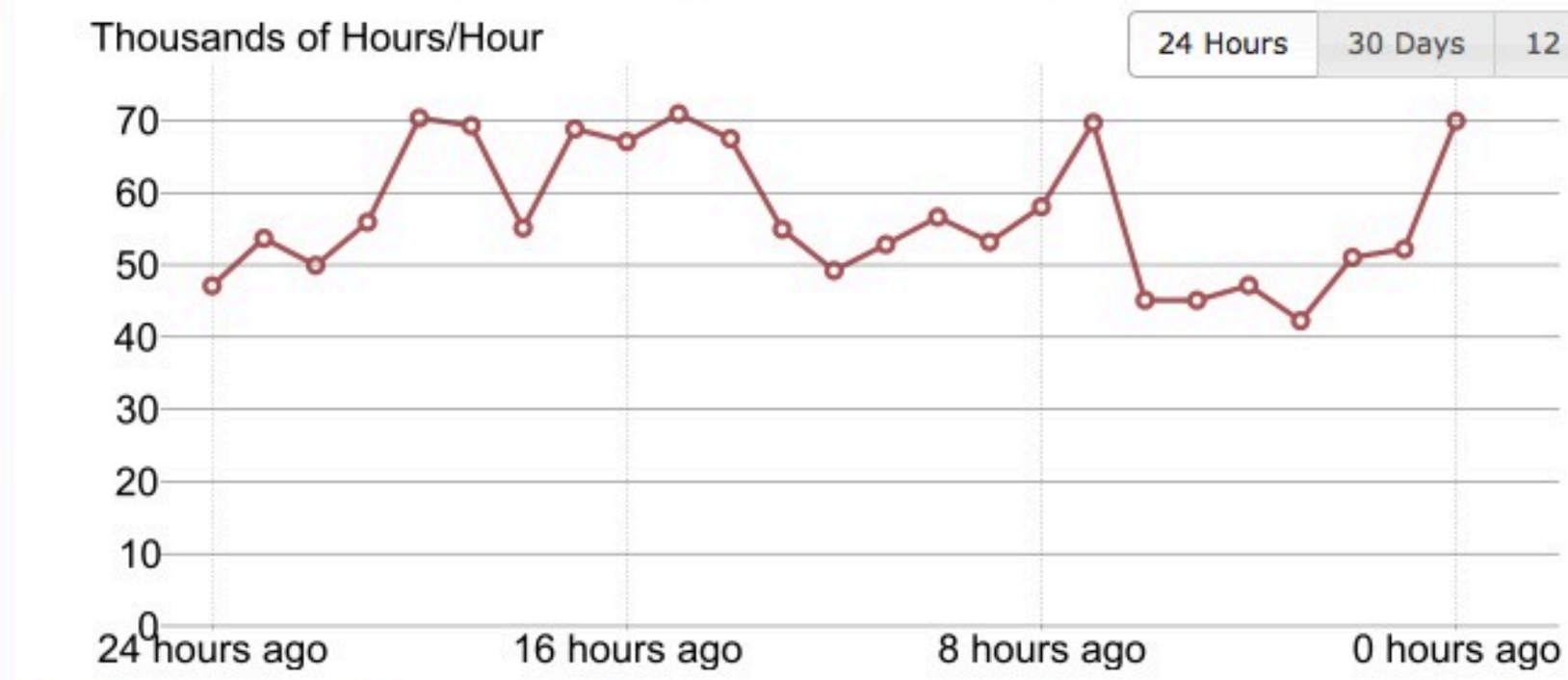
- US National Cyberinfrastructure
- Primarily used for high energy physics computing
- 80 sites
- 100,000 job slots
- 1,500,000 hours per day
- PB scale aggregate storage
- 1 PB transferred each day
- Virtual Organization-based



Open Science Grid

A national, distributed computing partnership for data-intensive research

Jobs | CPU Hours | Transfers | TB Transferred | Status Map



CPU hours spend on an OSG resource is reported to the central accounting system. The above graph shows the number of CPU hours per hour. A total of 1,423,000 CPU hours were spent.

OSG delivered across 95 sites

In the last 24 Hours

817,000	Jobs
1,423,000	CPU Hours
2,020,000	Transfers
894	TB Transferred

In the last 30 Days

16,110,000	Jobs
48,719,000	CPU Hours
49,400,000	Transfers
20,595	TB Transferred

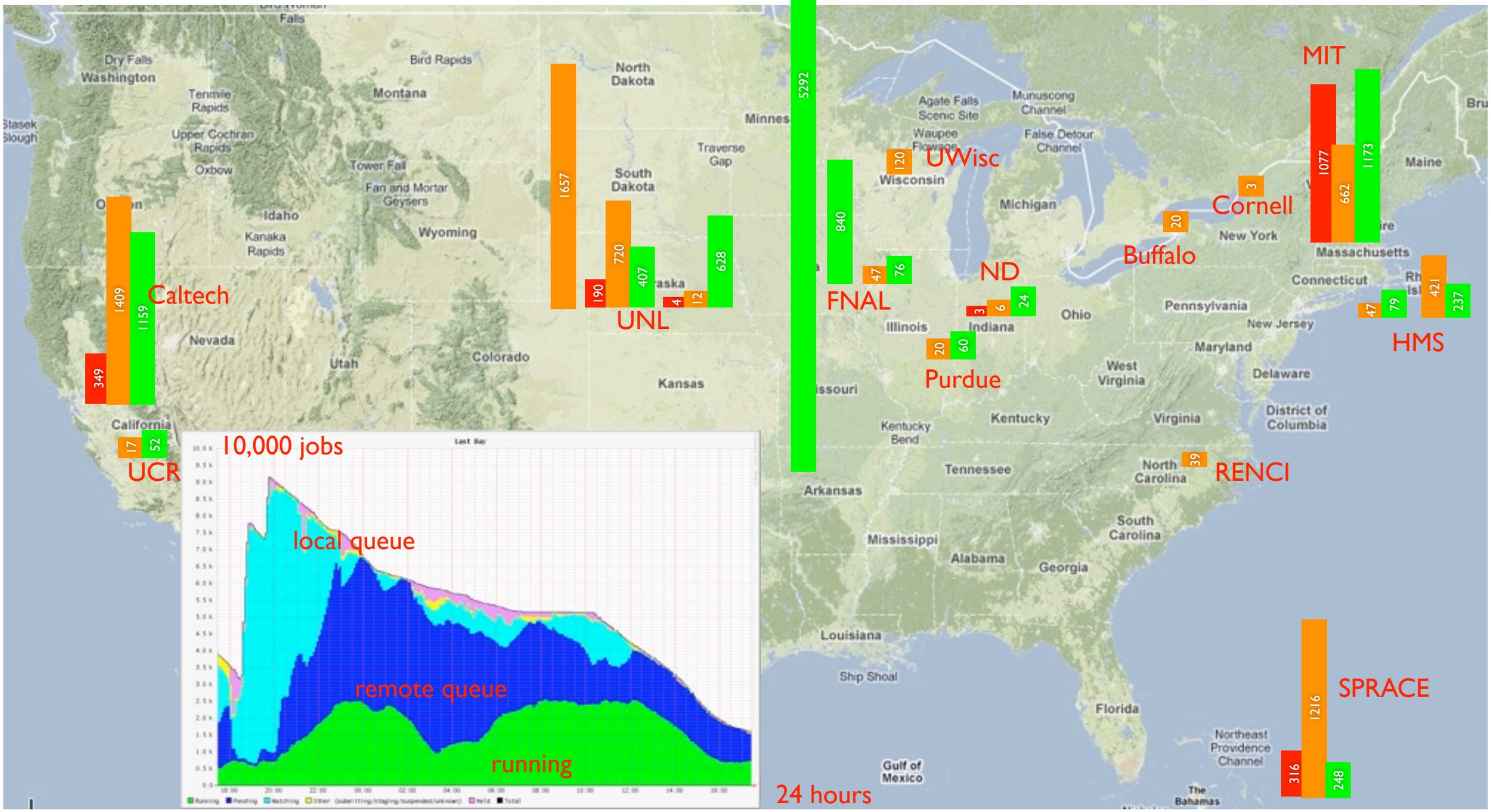
In the last Year

192,724,000	Jobs
453,369,000	CPU Hours
560,326,000	Transfers
285,161	TB Transferred

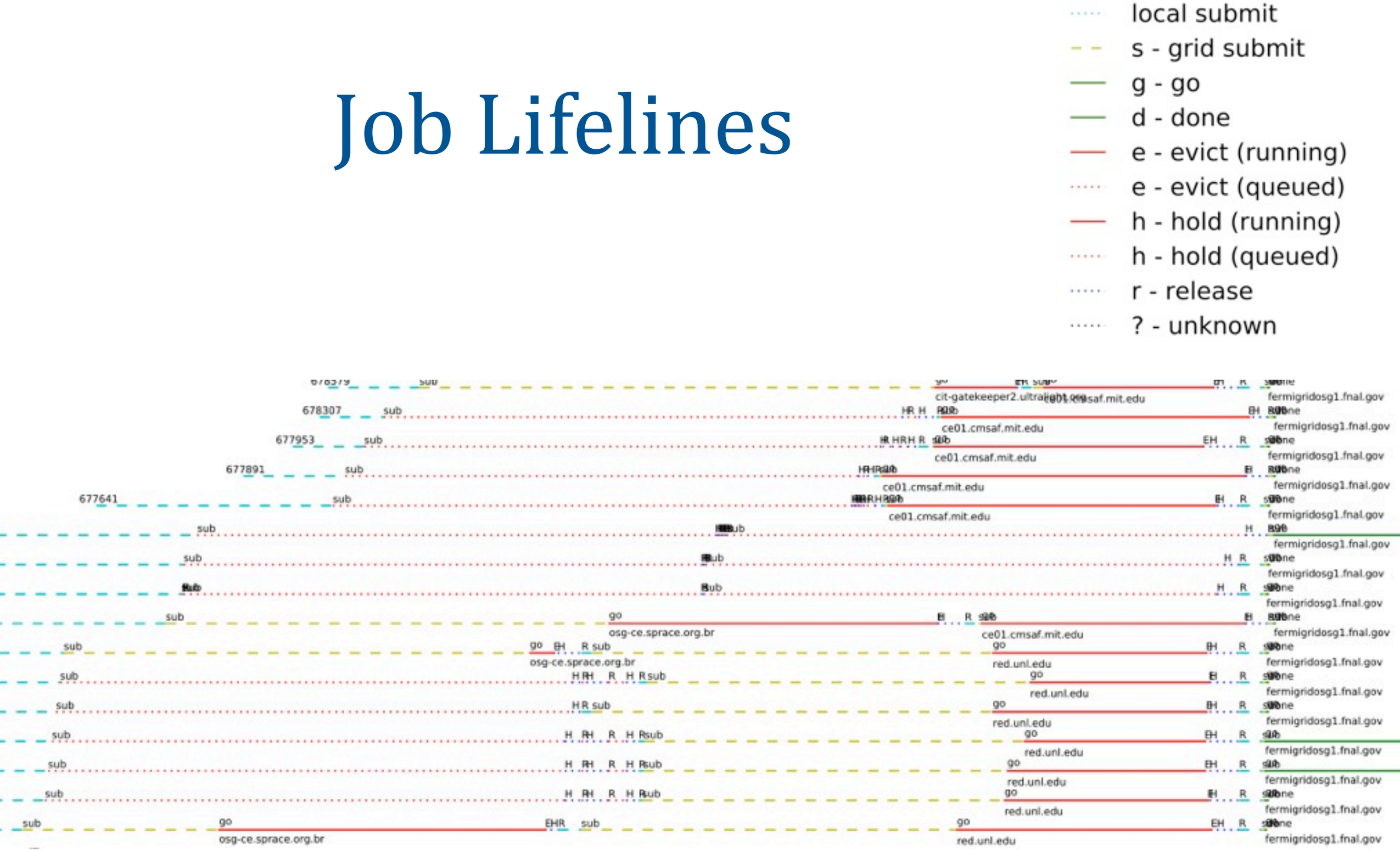
Example Job Set

10k grid jobs
approx 30k CPU hours
99.7% success rate
24 wall clock hours

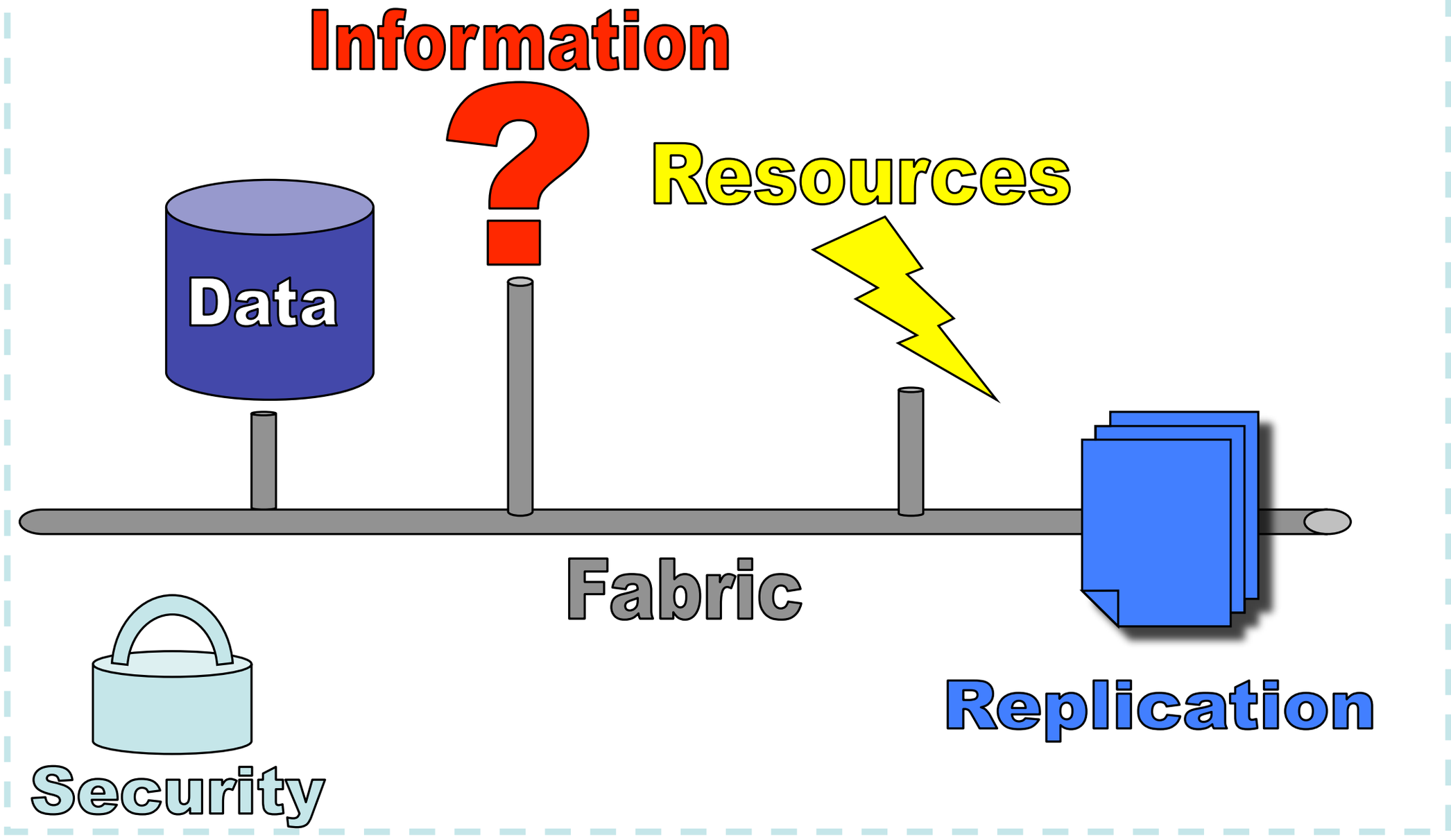
evicted - red
held - orange
completed - green



Job Lifelines



Simplified Grid Architecture



Grid Architectural Details

- Resources
  - Uniform compute clusters
  - Managed via batch queues
  - Local scratch disk
  - Sometimes high perf. network (e.g. InfiniBand)
  - Behind NAT and firewall
  - No shell access
- Data
  - Tape-backed mass storage
  - Disk arrays (100s TB to PB)
  - High bandwidth (multi-stream) transfer protocols
  - File catalogs
  - Meta-data
  - Replica management
- Information
  - LDAP based most common (not optimized for writes)
  - Domain specific layer
  - Open problem!
- Fabric
  - In most cases, assume functioning Internet
  - Some sites part of experimental private networks
- Security
  - Typically underpinned by X.509 Public Key Infrastructure
  - Same standards as SSL/TLS and “server certs” for “https”

OSG Components (I)

- Centralized
  - X.509 Certificate Authority: Energy Science Network CA @ LBL
  - Accounting: Gratia logging system to track usage (CPU, Network, Disk)
  - Status: LDAP directory with details of each participating system
  - Support: Central clearing house for support tickets
  - Software: distribution system, update testing, bug reporting and fixing
  - Communication: Wikis, docs, mailing lists, workshops, conferences, etc.
- Per Site
  - Compute Element/Gatekeeper (CE/GK): access point for external users, acts as frontend for any cluster. Globus GRAM + local batch system
  - Storage Element (SE): grid-accessible storage system, GridFTP-based + SRM
  - Worker Nodes (WN): cluster nodes with grid software stack
  - User Interface (UI): access point for local users to interact with remote grid
  - Access Control: GUMS + PRIMA for ACLs to local system by grid identities
  - Admin contact: need a local expert (or two!)

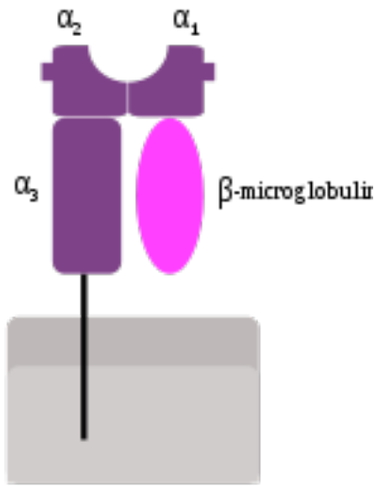
OSG Components (II)

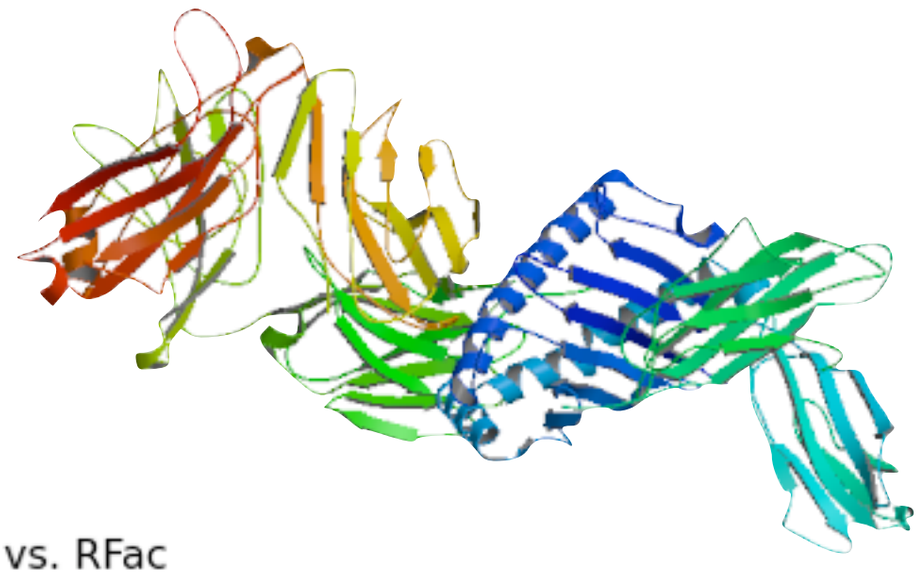
- Per Virtual Organization (user community)
  - VO Management System (VOMS): to organize and register users
  - Registration Authority (RA): to validate community users with X.509 issuer
  - User Interface system (UI): provide gateway to OSG for users
  - Support Contact: users are supported by their VO representatives
- Per User
  - X.509 user certificate (although I'd like to hide that part)
  - Induction: unless it is through a portal, grid computing is not shared file system batch computing! Many more failure modes and gotchas.

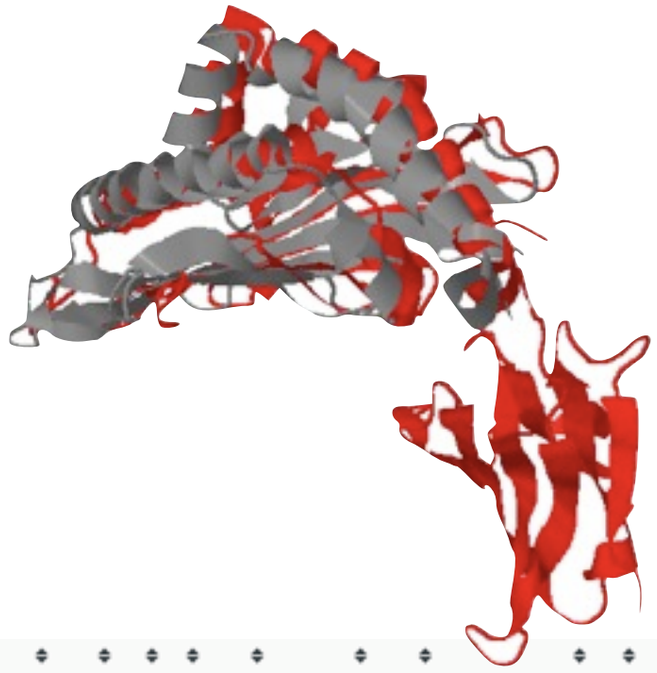
Grid Opportunities

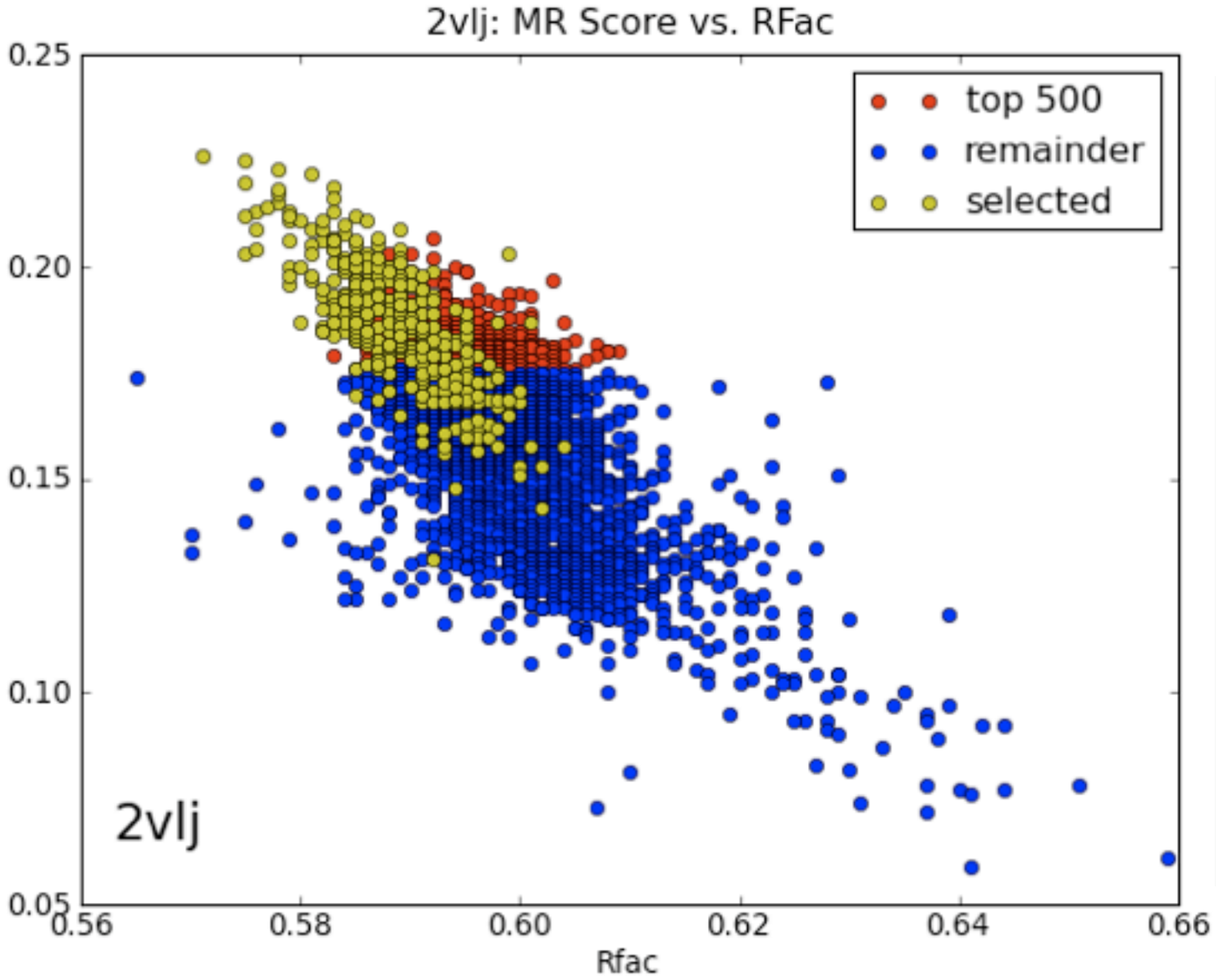
- New compute intensive workflows
  - think big: tens or hundreds of thousands of hours finished in 1-2 days
  - sharing resources for efficient and large scale utilization
- Data intensive problems
  - we mirror 20 GB of data to 30 computing centers
- Data movement, management, and archive
- Federated identity and user management
  - labs, collaborations or ad-hoc groups
  - role-based access control (RBAC) and IdM
- Collaborative environment
- Web-based access to applications

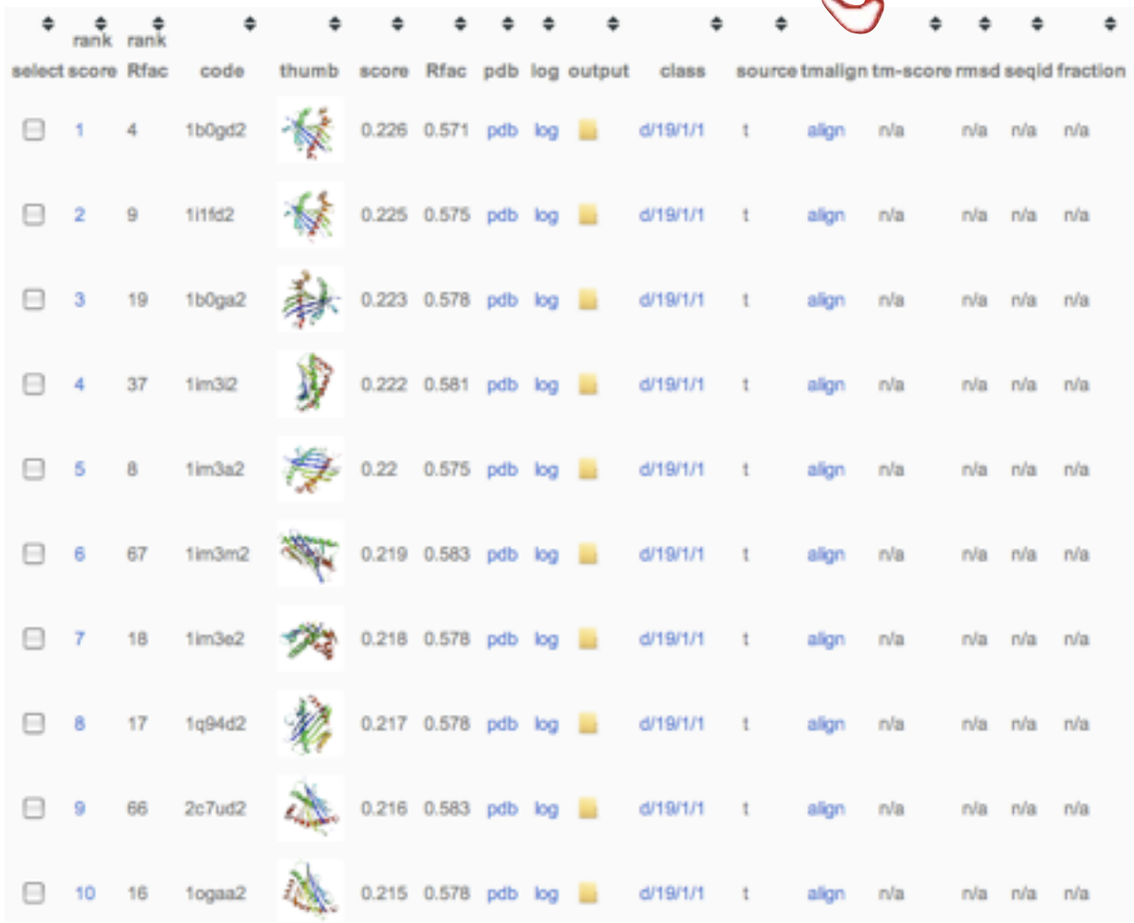
Protein Structure Determination





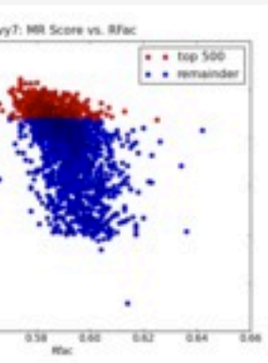


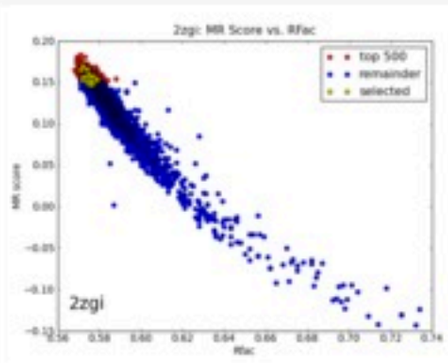


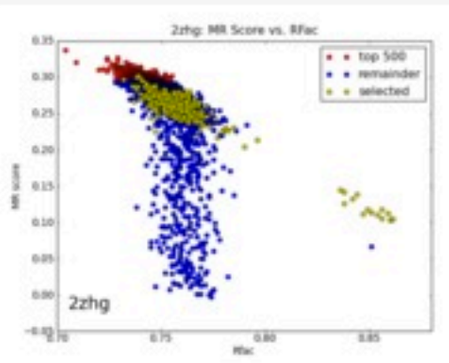


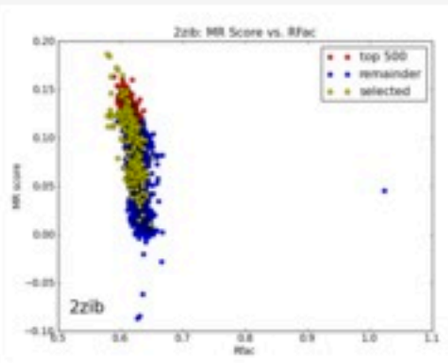
select	rank score	rank Rfac	code	thumb	score	Rfac	pdb	log	output	class	source	tmalign	tm-score	rmsd	seqid	fraction
☐	1	4	1b0gd2		0.226	0.571	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	2	9	1i1fd2		0.225	0.575	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	3	19	1b0ga2		0.223	0.578	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	4	37	1im3i2		0.222	0.581	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	5	8	1im3a2		0.22	0.575	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	6	67	1im3m2		0.219	0.583	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	7	18	1im3e2		0.218	0.578	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	8	17	1q94d2		0.217	0.578	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	9	66	2c7ud2		0.216	0.583	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a
☐	10	16	1ogaa2		0.215	0.578	pdb	log		d/19/1/1	t	align	n/a	n/a	n/a	n/a

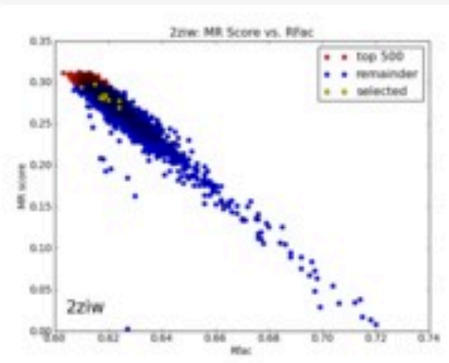
ijstokes@hkl.hms.harvard.edu

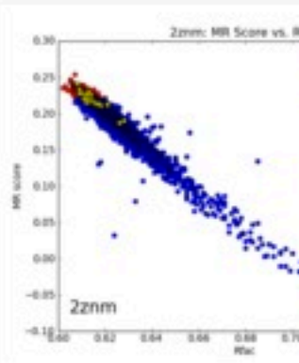

top 500
remainder


2zgi: MR Score vs. RFac
top 500
remainder
selected
2zgi
MR score
Rfac

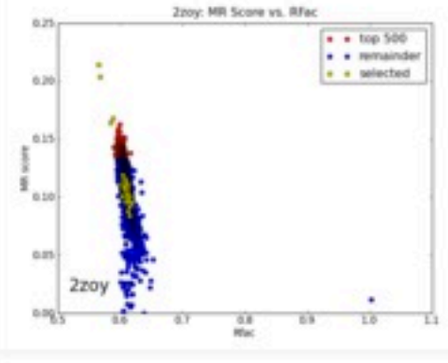

2zhg: MR Score vs. RFac
top 500
remainder
selected
2zhg
MR score
Rfac

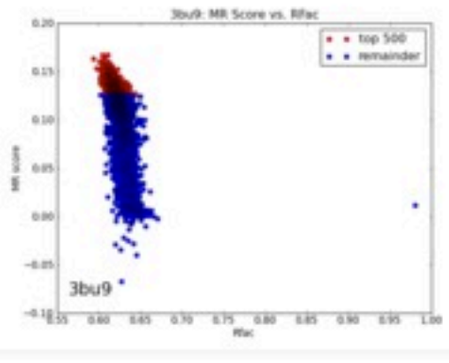

2zib: MR Score vs. RFac
top 500
remainder
selected
2zib
MR score
Rfac

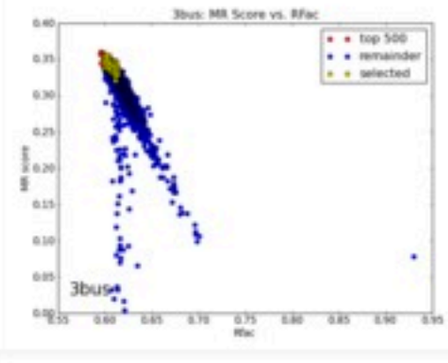

2ziw: MR Score vs. RFac
top 500
remainder
selected
2ziw
MR score
Rfac

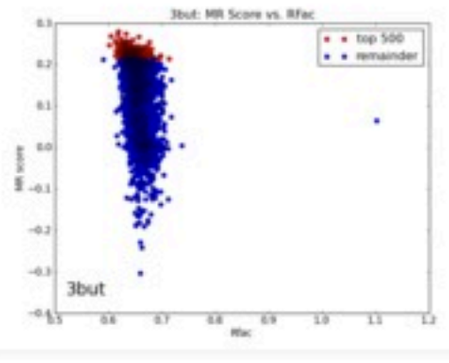

2znm: MR Score vs. R
2znm
MR score
Rfac

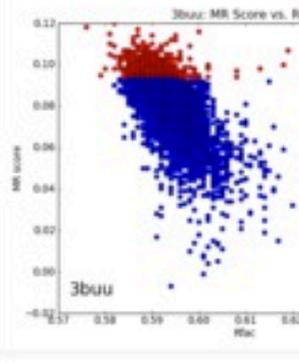

top 500
remainder
selected

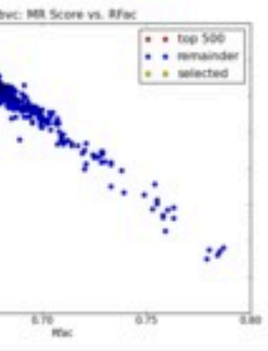

2zoy: MR Score vs. RFac
top 500
remainder
selected
2zoy
MR score
Rfac

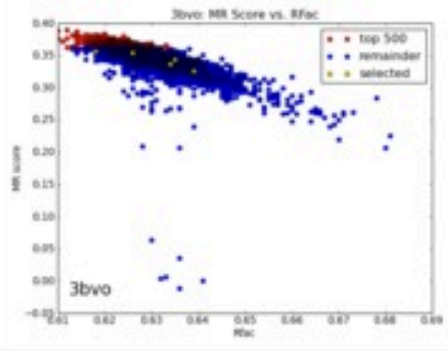

3bu9: MR Score vs. RFac
top 500
remainder
3bu9
MR score
Rfac

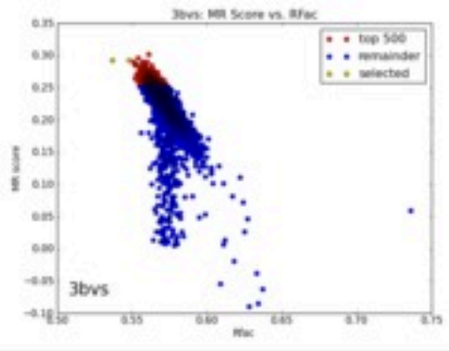

3bus: MR Score vs. RFac
top 500
remainder
selected
3bus
MR score
Rfac

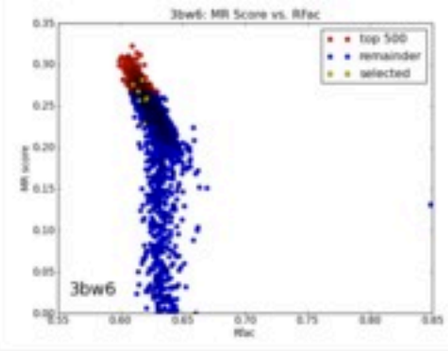

3but: MR Score vs. RFac
top 500
remainder
3but
MR score
Rfac

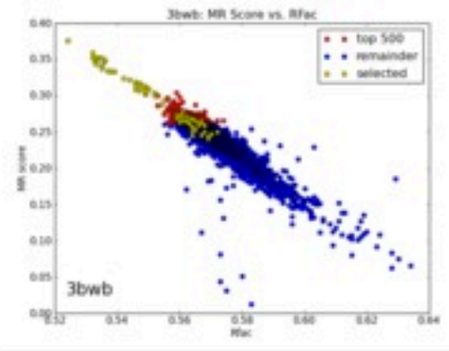

3buu: MR Score vs. R
3buu
MR score
Rfac

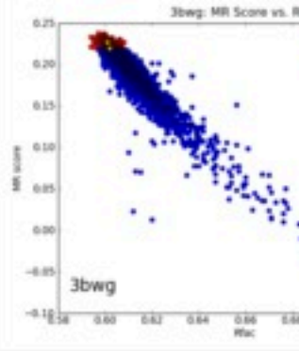

top 500
remainder
selected

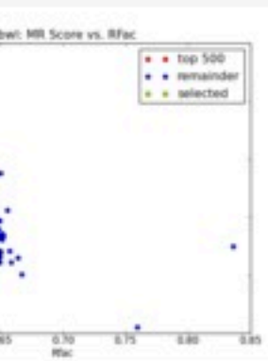

3bvo: MR Score vs. RFac
top 500
remainder
selected
3bvo
MR score
Rfac

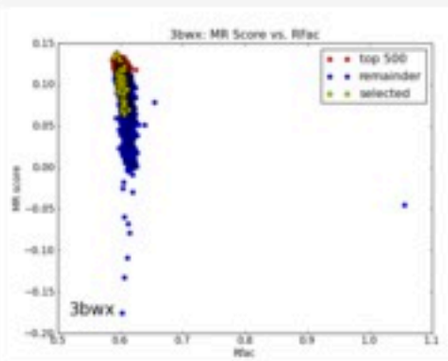

3bvs: MR Score vs. RFac
top 500
remainder
selected
3bvs
MR score
Rfac

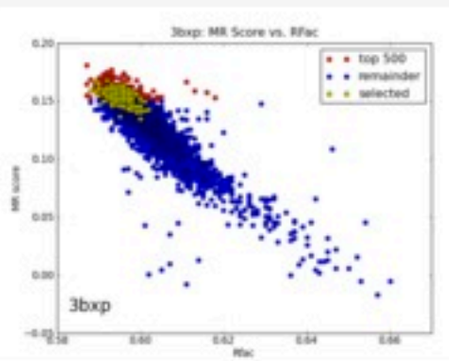

3bw6: MR Score vs. RFac
top 500
remainder
selected
3bw6
MR score
Rfac

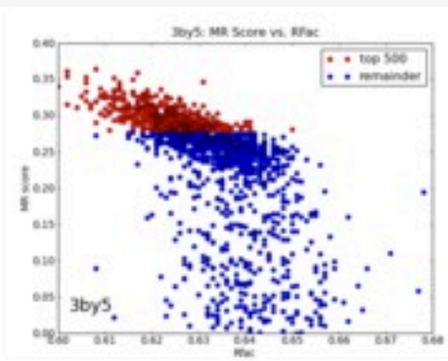

3bwb: MR Score vs. RFac
top 500
remainder
selected
3bwb
MR score
Rfac

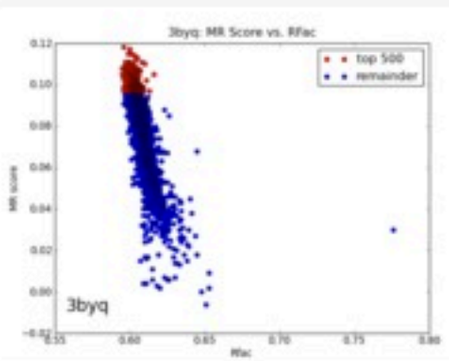

3bwg: MR Score vs. R
3bwg
MR score
Rfac

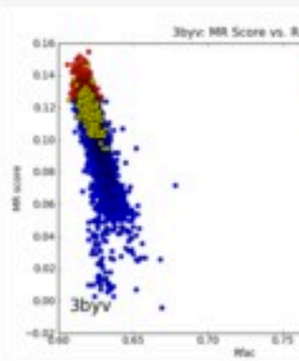

top 500
remainder
selected

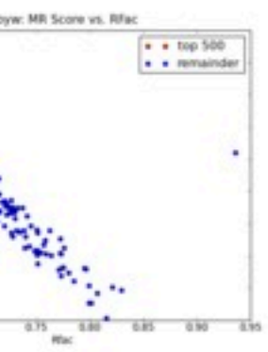

3bwx: MR Score vs. RFac
top 500
remainder
selected
3bwx
MR score
Rfac

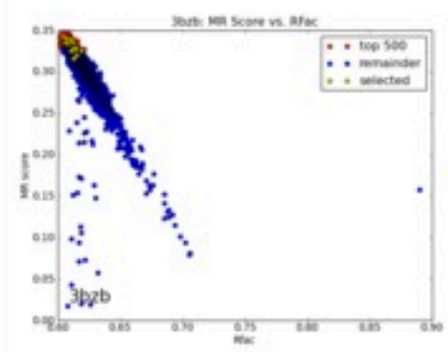

3bxp: MR Score vs. RFac
top 500
remainder
selected
3bxp
MR score
Rfac

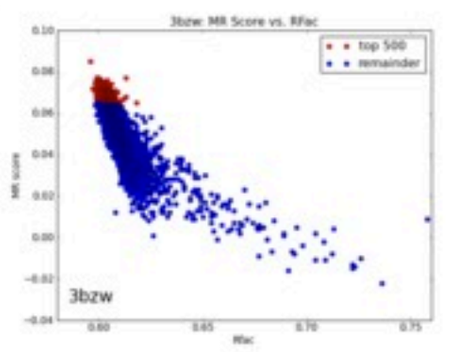

3by5: MR Score vs. RFac
top 500
remainder
3by5
MR score
Rfac

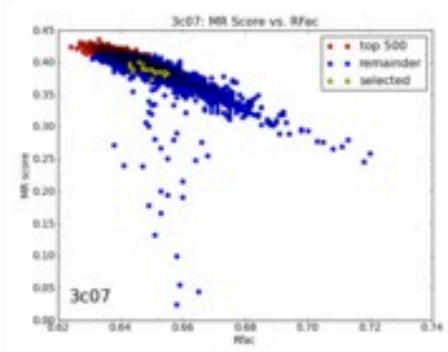

3byq: MR Score vs. RFac
top 500
remainder
3byq
MR score
Rfac

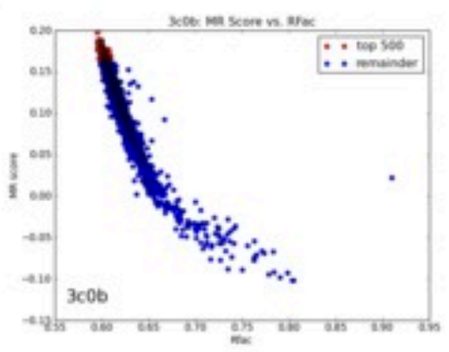

3byv: MR Score vs. R
3byv
MR score
Rfac

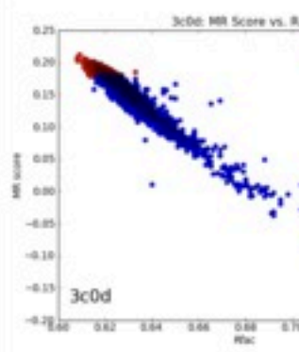

top 500
remainder


3bzb: MR Score vs. RFac
top 500
remainder
selected
3bzb
MR score
Rfac


3bzw: MR Score vs. RFac
top 500
remainder
3bzw
MR score
Rfac


3c07: MR Score vs. RFac
top 500
remainder
selected
3c07
MR score
Rfac


3c0b: MR Score vs. RFac
top 500
remainder
3c0b
MR score
Rfac


3c0d: MR Score vs. R
3c0d
MR score
Rfac

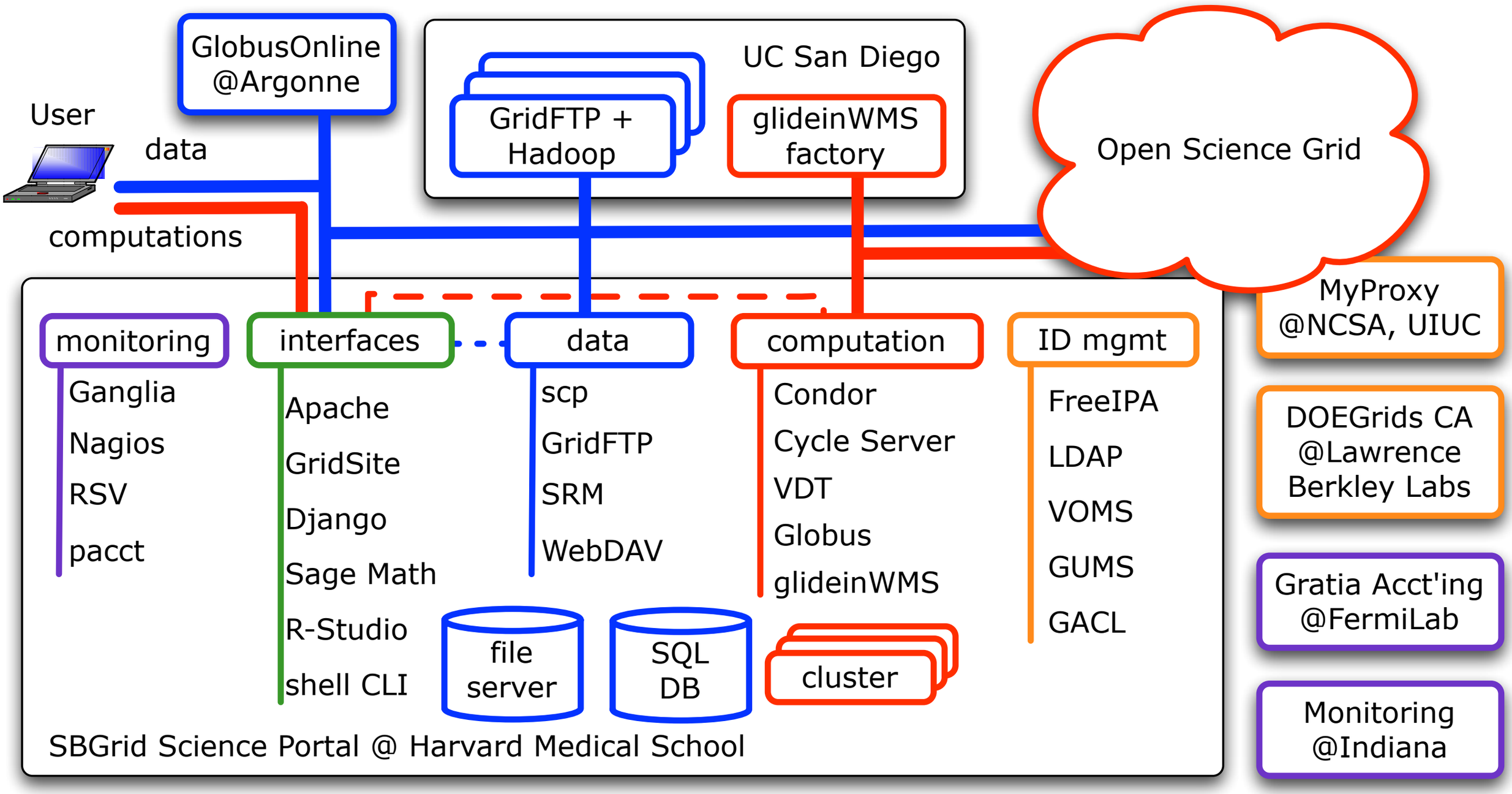
Typical Layered Environment

- Command line application (e.g. Fortran)
- Friendly application API wrapper
- Batch execution wrapper for N-iterations
- Results extraction and aggregation
- Grid job management wrapper
- Web interface
- forms, views, static HTML results
- GOAL eliminate shell scripts
- often found as “glue” language between layers

Map-Reduce

Architecture Diagrams

Service Architecture



internal

external

Grid

- Grid Monitoring Tools
- Grid prep and deployment
- Condor tools
- OSG middleware tools
- Universal grid wrapper

Portal

- Application "report" UI
- Application "job" UI
- Job mgmt UI and API
- User mgmt UI and API
- Django web framework
- Apache modules: python, gridsite, cernmeta
- Apache

Metadata

- metadata catalog
- xconfig
- cernmeta HTTP headers
- bespoke FS metadata

Application

- Report generation: tables, graphics, statistics
- Python modules: numpy, scipy, matplotlib, pylab
- Output post-processing and aggregation
- CLI
- API
- Application

database

filesystem

Presentation

- Django and Cheetah HTML templates
- CSS
- Bespoke JavaScript
- JavaScript libraries: jQuery, extJS, YUI, GWT

Security

- GridSecure
- PyGACL
- GACL
- VOMS
- Grid middleware
- POSIX security
- X.509 security

- **NEBioGrid Django Portal**

  Interactive dynamic web portal for workflow definition, submission, monitoring, and access control

- **NEBioGrid Web Portal**

  GridSite based web portal for file-system level access (raw job output), meta-data tagging, X.509 access control/sharing, CGI

- **PyCCP4**

  Python wrappers around CCP4 structural biology applications

- **PyCondor**

  Python wrappers around common Condor operations

  enhanced Condor log analysis

- **PyOSG**

  Python wrappers around common OSG operations

- **PyGACL**

  Python representation of GACL model and API to work with GACL files

- **osg_wrap**

  Swiss army knife OSG wrapper script to handle file staging, parameter sweep, DAG, results aggregation, monitoring

- **sbanalysis**

  data analysis and graphing tools for structural biology data sets

- **osg.monitoring**

  tools to enhance monitoring of job set and remote OSG site status

- **shex**

  Write bash scripts in Python: replicate commands, syntax, behavior

- **xconfig**

  Universal configuration

Web Portals for Collaborative, Multi-disciplinary Research...

... which leverage capabilities of federated grid computing environments

The Browser as the Universal Interface

- If it isn't already obvious to you
  - Any interactive application developed today should be web-based with a RESTful interface (if at all possible)
- A rich set of tools and techniques
  - AJAX, HTML4/5, CSS, and JavaScript
  - Dynamic content negotiation
  - HTTP headers, caching, security, sessions/cookies
- Scalable, replicable, centralized, multi-threaded, multi-user
- Alternatives
  - Command Line (CLI): great for scriptable jobs
  - GUI toolkits: necessary for applications with high graphics or I/O demands

What is a Science Portal?

- A web-based gateway to resources and data
  - simplified access
  - centralized access
  - unified access (CGI, Perl, Python, PHP, static HTML, static files, etc.)
- Attempt to provide uniform access to a range of services and resources
- Data access via HTTP
- Leverage brilliance of Apache HTTPD and associated modules

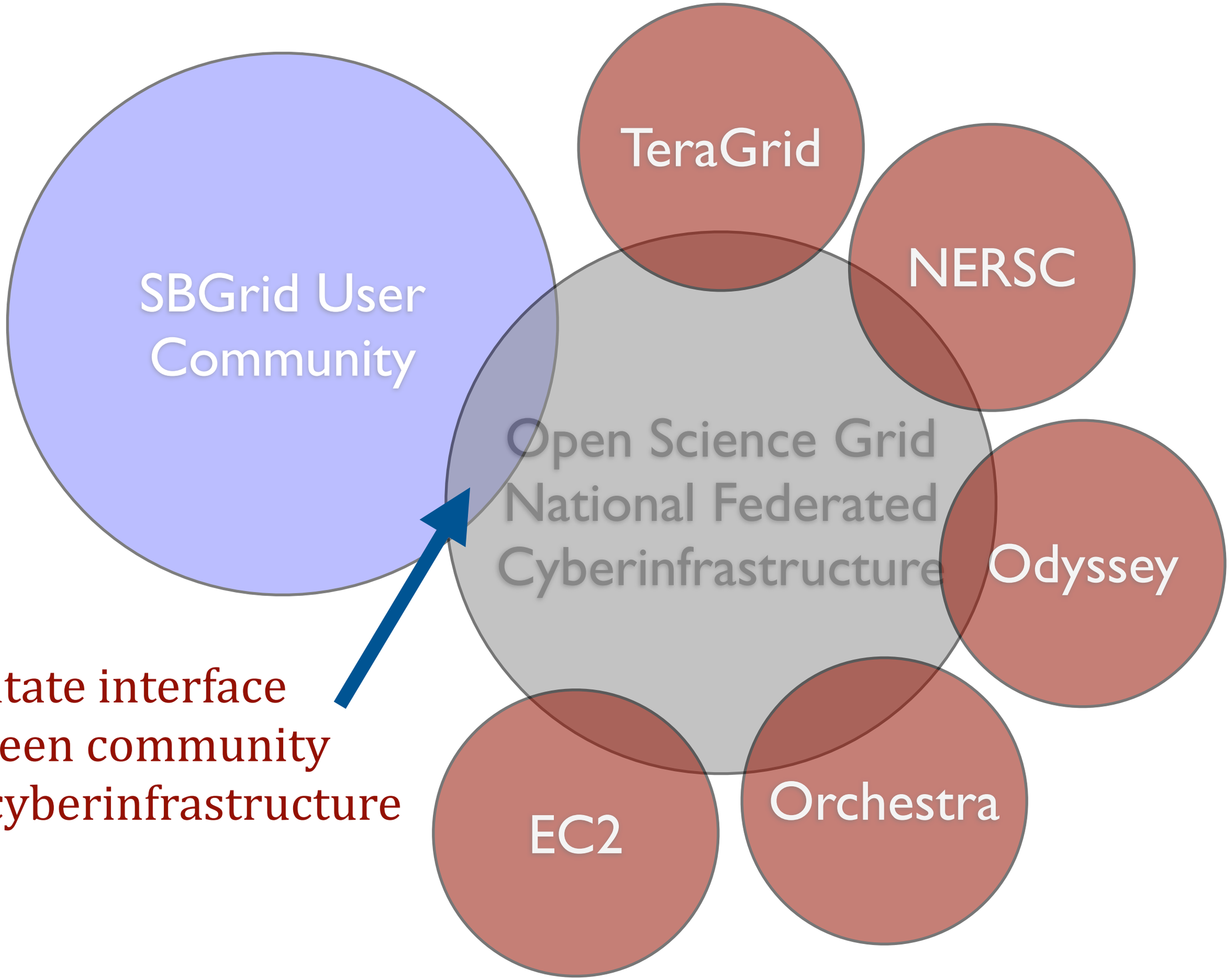
SBGrid Science Portal Objectives

A.

Extensible infrastructure to facilitate development and deployment of novel computational workflows

B.

Web-accessible environment for collaborative, compute and data intensive science



Facilitate interface between community and cyberinfrastructure

ijstokes@hkl.hms.harvard.edu

SBGrid Science Portal

https://portal.sbgrid.org/

Google

SBGrid Science Portal

Home | Applications | Resources | Support | About

Wide Search Molecular Replacement

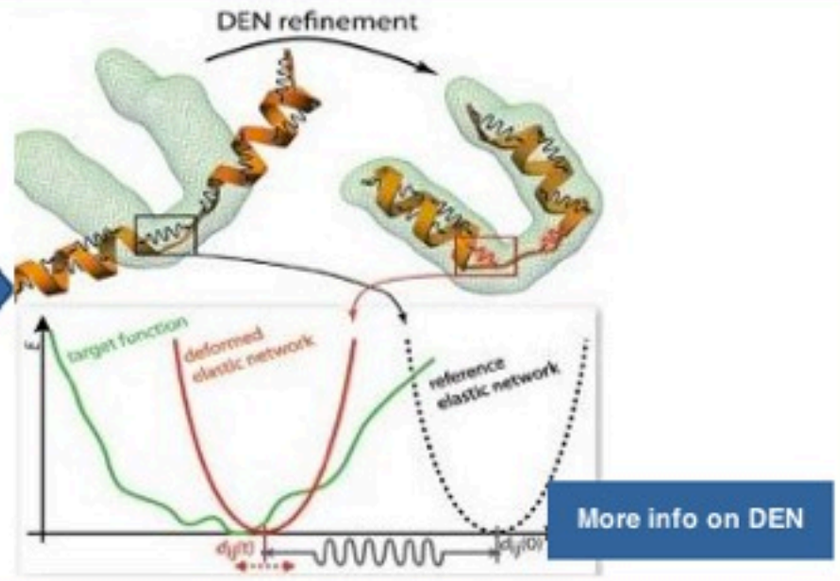
WSMR searches up to 100,000 domains for molecular replacement models given X-ray structure factors.

Deformable Elastic Network

DEN produces high resolution structures from low resolution X-Ray data by only refining known restraints and the use of homology models.

SBGrid Consortium

SBGrid provides a distribution of over 200 scientific software packages with monthly updates to member labs.



More info on DEN

The SBGrid Science Portal provides access to web-enabled structural biology applications, data sharing facilities, biological data sets, and other resources valuable to the computational structural biology community. This is in addition to the core SBGrid program of software distribution and scientific computing technical support.

Services

SBGrid has been funded by the NSF to encourage the adoption of advanced computational methods, grid computing facilities, and collaborative data-sharing systems within two communities:

- Structrual biology labs which are part of the SBGrid consortium
- Biomedical researchers in New England

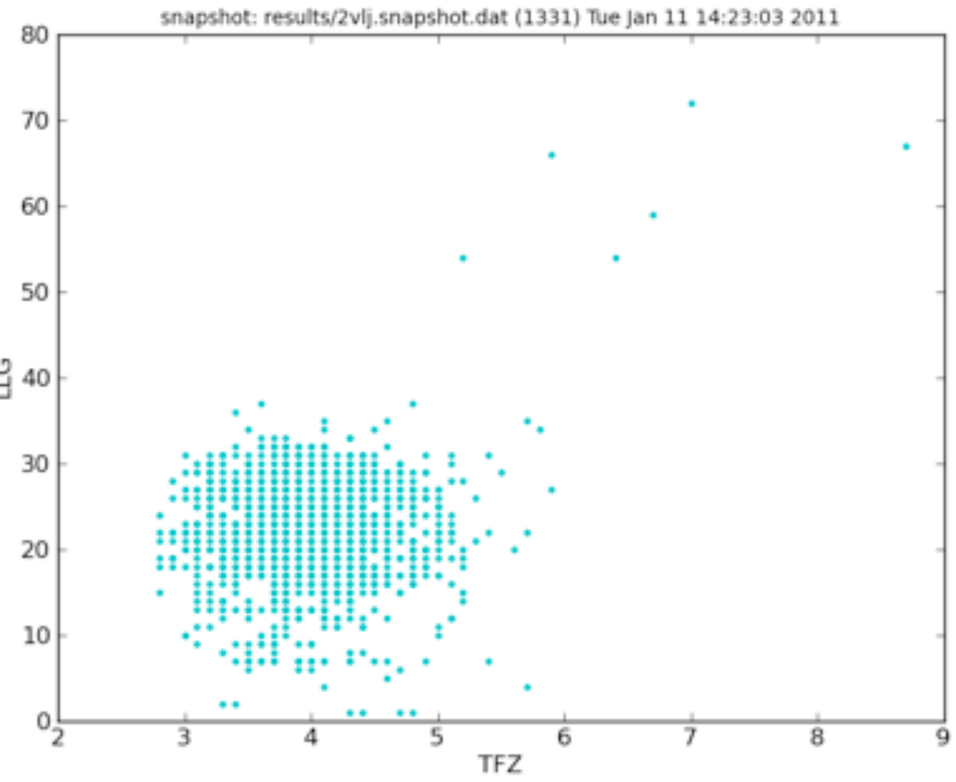
As such, we provide the following services and facilities:

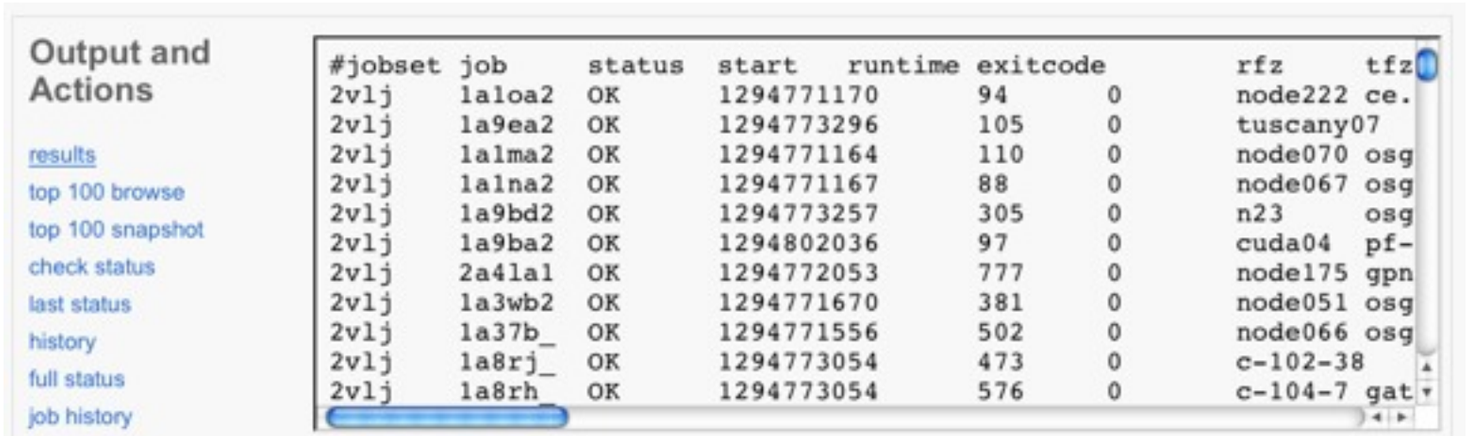
- Consulting expertise to enable advanced computational methods for researchers who are in need.

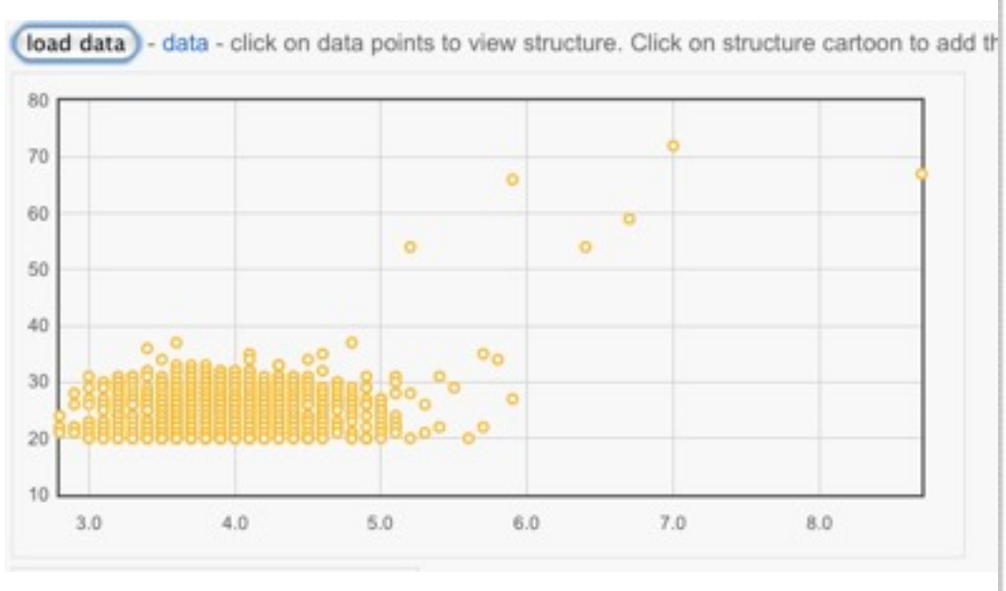
Consortium

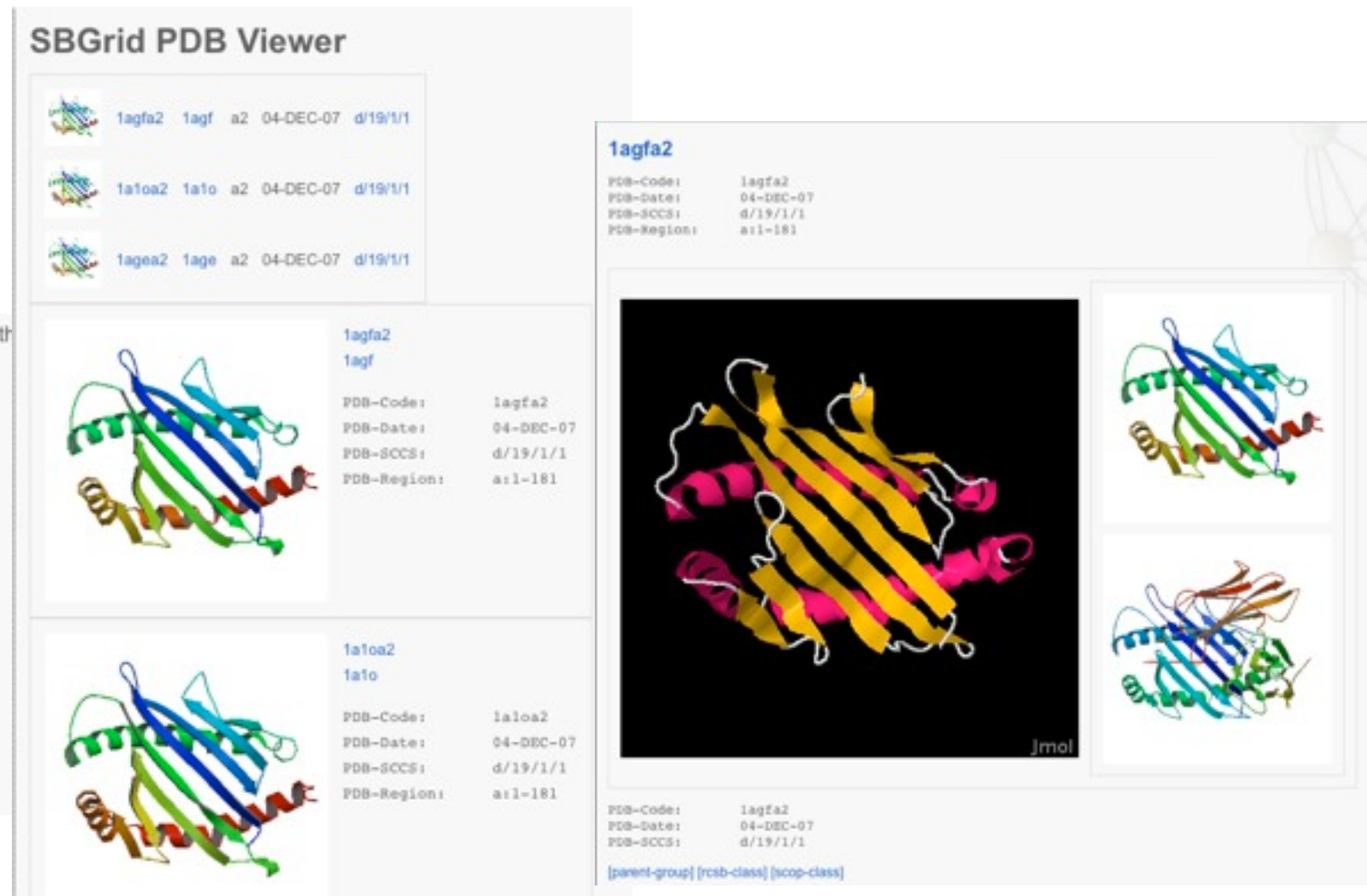
With funding from the National Science Foundation

Results Visualization and Analysis









REST

- Don’t try to read too much into the name
  - REpresentational State Transfer: coined by Roy Fielding, co-author of HTTP protocol and contributor to original Apache httpd server
- Idea
  - The web is the worlds largest asynchronous, distributed, parallel computational system
  - Resources are “hidden” but representations are accessible via URLs
  - Representations can be manipulated via HTTP operations GET PUT POST HEAD DELETE and associated state
  - State transitions are initiated by software or by humans
- Implication
  - Clean URLs (e.g. Flickr)

Cloud Computing: Industry solution to the Grid

- Virtualization has taken off in the past 5 years
  - VMWare, Xen, VirtualPC, VirtualBox, QEMU, etc.
  - Builds on ideas from VMS (i.e. old)
- (Good) System administrators are hard to come by
  - And operating a large data center is costly
- Internet boom means there are companies that have figured out how to do this really well
  - Google, Amazon, Yahoo, Microsoft, etc.
- Outsource IT infrastructure! Outsource software hosting!
  - Amazon EC2, Microsoft Azure, RightScale, Force.com, Google Apps
- Over simplified:
  - You can't install a cloud
  - You can't buy a grid

Is “Cloud” the new “Grid”?

- **Grid** is about mechanisms for federated, distributed, heterogeneous shared compute and storage resources
  - standards and software
- **Cloud** is about on-demand provisioning of compute and storage resources
  - services

No one buys a grid. No one installs a cloud.

The interesting thing about Cloud Computing is that we've redefined Cloud Computing to include everything that we already do. . . . I don't understand what we would do differently in the light of Cloud Computing other than change the wording of some of our ads.

Larry Ellison, Oracle CEO, quoted in the Wall Street Journal, September 26, 2008*

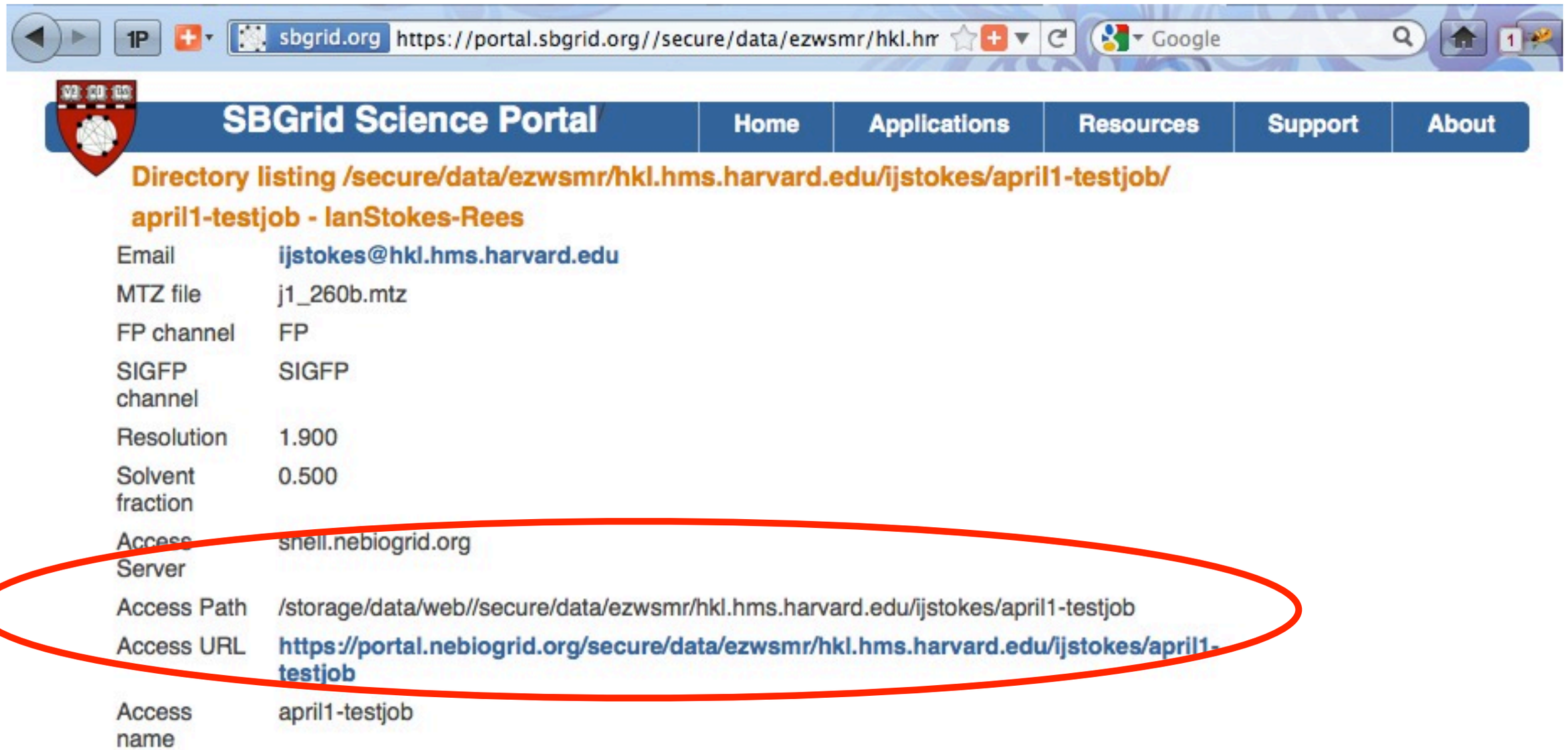
*http://blogs.wsj.com/biztech/2008/09/25/larry-ellisons-brilliant-anti-cloud-computing-rant/

When is cloud computing interesting?

- My definition of "cloud computing"
  - Dynamic compute and storage infrastructure provisioning in a scalable manner providing uniform interfaces to virtualized resources
- The underlying resources could be
  - "in-house" using licensed/purchased software/hardware
  - "external" hosted by a service/infrastructure provider
- Consider using cloud computing if
  - You have operational problems/constraints in your current data center
  - You need to dynamically scale (up or down) access to services and data
  - You want fast provisioning, lots of bandwidth, and low latency
  - Organizationally you can live with outsourcing responsibility for (some of) your data and applications
- Consider providing cloud computing services if
  - You have an ace team efficiently running your existing data center
  - You have lots of experience with virtualization
  - You have a specific application/domain that could benefit from being tied to a large compute farm or disk array with great Internet connectivity

Data Access

User access to results data



sbgrid.org https://portal.sbgrid.org//secure/data/ezwsmr/hkl.hms.harvard.edu/ijstokes/jw-fab/results

SBGrid Science Portal | Home | Applications | Resources

Directory listing /secure/data/ezwsmr/hkl.hms.harvard.edu/ijstokes/jw-fab/results/

[Parent directory]		
sifab82_22_data.final.augmented.dat	2552436309	16:14 Feb 11
sifab82_22_data.final.dat	1273469109	13:14 Feb 11
sifab82_22_data.final.raw.dat	1275014009	13:14 Feb 11
sifab82_22_data.inprogress.dat	1669562116	5014 Feb 11
sifab82_22_data.snapshot.dat	1658860714	26:14 Feb 11
sifab82_22_data.snapshot.dat.json	370383	14:26 Feb 11
sifab82_22_data.snapshot.dat.llg-vs-tfz.png	32841	14:26 Feb 11
sifab82_22_data.snapshot.top100.dat	16013	14:26 Feb 11
sifab82_22_data.snapshot.top100.html	834	14:26 Feb 11

You are /DC=org/DC=doegrids/OU=People/CN=Ian Stokes-Rees 411174
Manage directory . Switch to HTTP . Built with GridSite 1.5.18


```
[ijstokes@portal results]$ pwd
/storage/data/web/secure/data/ezwsmr/hkl.hms.harvard.edu/ijstokes/jw-fab/results
[ijstokes@portal results]$ ls --color=auto -Fla
total 82750
drwxr-xr-x 2 webportal webportal      363 Feb 14 09:13 ./
drwxr-xr-x 8 webportal webportal      376 Apr 30 08:29 ../
-rw-rw-r-- 1 webportal webportal 25524363 Feb 14 09:16 sifab82_22_data.final.augmented.dat
-rw-rw-r-- 1 webportal webportal 12734691 Feb 14 09:13 sifab82_22_data.final.dat
-rw-rw-r-- 1 webportal webportal 12750140 Feb 14 09:13 sifab82_22_data.final.raw.dat
-rw-r--r-- 1 webportal webportal 16695621 Feb 14 16:50 sifab82_22_data.inprogress.dat
-rw-rw-r-- 1 webportal webportal 16588607 Feb 14 14:26 sifab82_22_data.snapshot.dat
-rw-rw-r-- 1 webportal webportal   370383 Feb 14 14:26 sifab82_22_data.snapshot.dat.json
-rw-rw-r-- 1 webportal webportal    32841 Feb 14 14:26 sifab82_22_data.snapshot.dat.llg-vs-tfz.png
-rw-rw-r-- 1 webportal webportal    16013 Feb 14 14:26 sifab82_22_data.snapshot.top100.dat
-rw-rw-r-- 1 webportal webportal      834 Feb 14 14:26 sifab82_22_data.snapshot.top100.html
[ijstokes@portal results]$
```

Experimental Data Access

- Collaboration
- Access Control
- Identity Management
- Data Management
- High Performance Data Movement
- Multi-modal Access

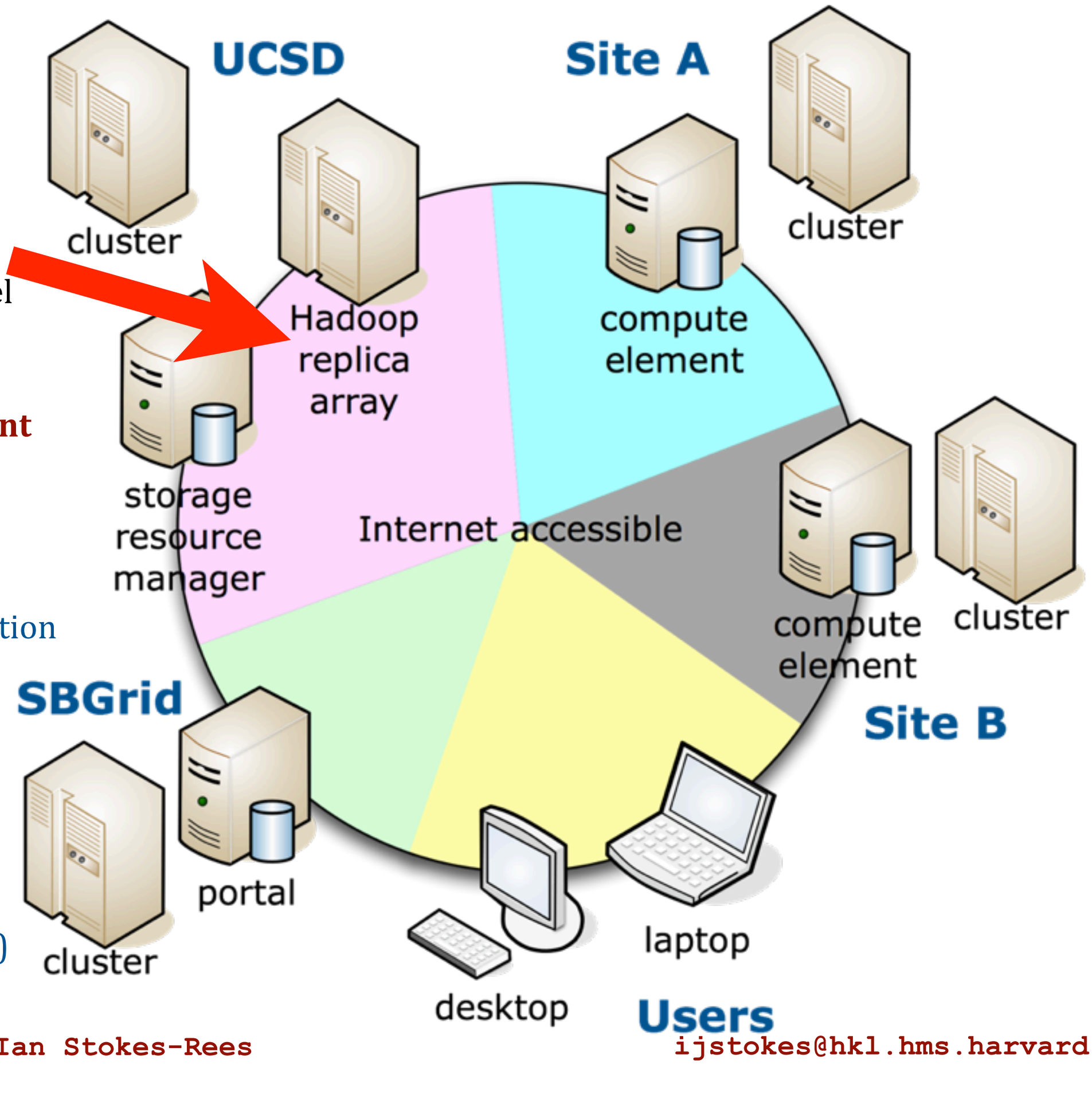
Data Model

- Data Tiers
  - **VO-wide**: all sites, admin managed, very stable
  - **User project**: all sites, user managed, 1-10 weeks, 1-3 GB
  - **User static**: all sites, user managed, indefinite, 10 MB
  - **Job set**: all sites, infrastructure managed, 1-10 days, 0.1-1 GB
  - **Job**: direct to worker node, infrastructure managed, 1 day, <10 MB
  - **Job indirect**: to worker node via UCSD, infrastructure managed, 1 day, <10 GB

About **2PB** with **100** front end servers for high bandwidth parallel file transfer

Data Management
quota
du scan
tmpwatch
conventions
workflow integration

Data Movement
scp (users)
rsync (VO-wide)
grid-ftp (UCSD)
curl (WNs)
cp (NFS)
htcp (secure web)



Globus Online: High Performance Reliable 3rd Party File Transfer

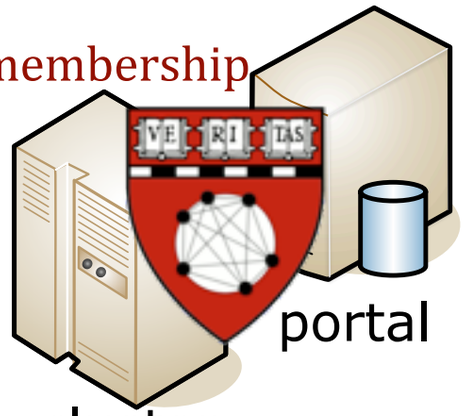
GUMS

MyProxy
Credential Management Service

DN to user mapping

VOMS

VO membership



portal

cluster

ESnet
Energy Sciences Network

DOEGrids

Certificate Authority

root of trust

Globus Online

file transfer service

lab file server

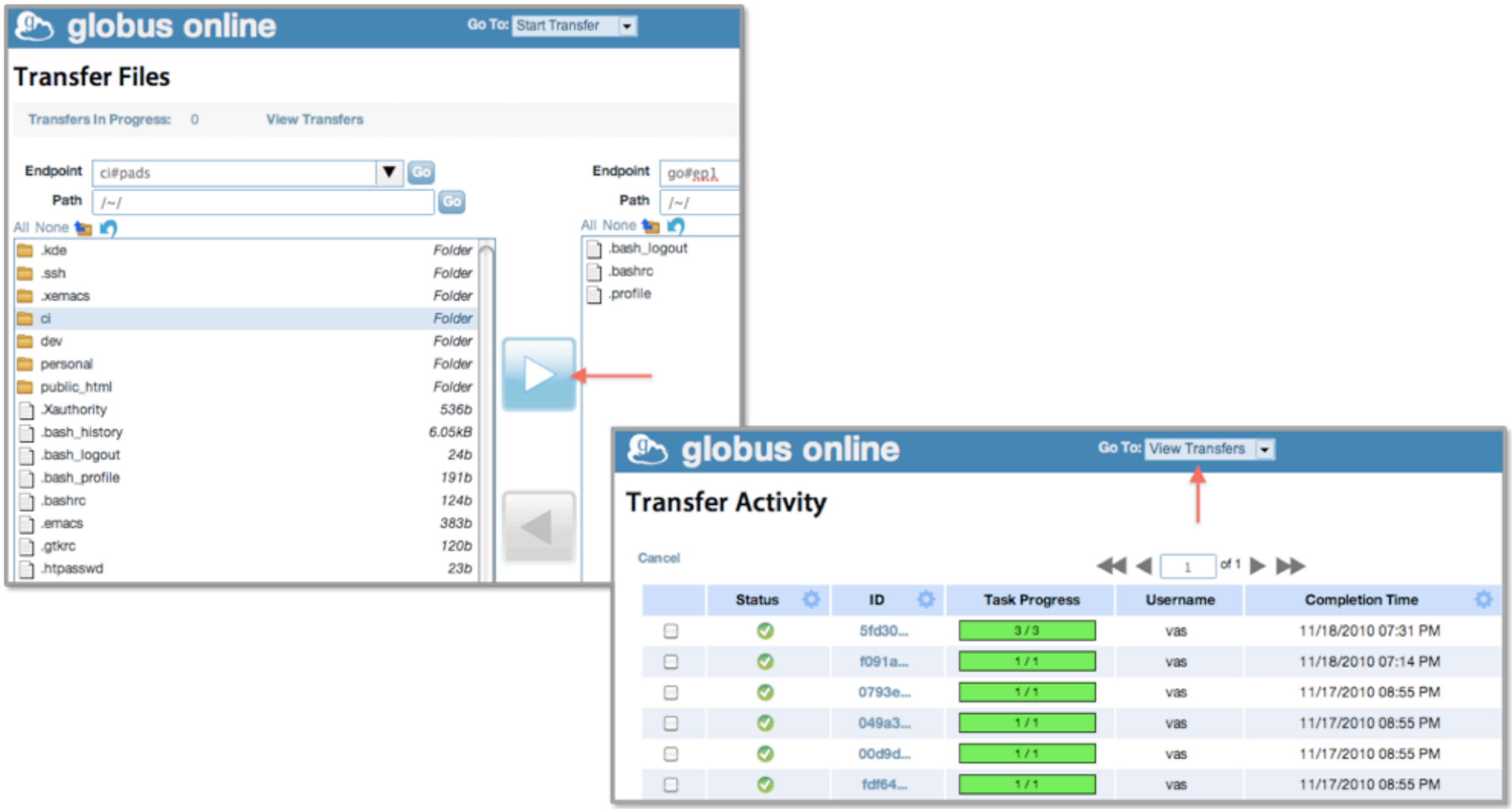
data collection facility

desktop

laptop

Architecture

- SBGrid
  - manages all user account creation and credential mgmt
  - hosts MyProxy, VOMS, GridFTP, and user interfaces
- Facility
  - knows about lab groups
    - e.g. “Harrison”, “Sliz”
  - delegates knowledge of group membership to SBGrid VOMS
    - facility can poll VOMS for list of current members
    - uses X.509 for user identification
  - deploys GridFTP server
- Lab group
  - designates group manager that adds/removes individuals
  - deploys GridFTP server or Globus Connect client
- Individual
  - username/password to access facility and lab storage
  - Globus Connect for personal GridFTP server to laptop
  - Globus Online web interface to “drive” transfers


globus online
Go To: Start Transfer
Transfer Files
Transfers In Progress: 0
View Transfers
Endpoint ci#pads
Go
Path /~/
Go
All None
.kde Folder
.ssh Folder
.xemacs Folder
ci Folder
dev Folder
personal Folder
public_html Folder
.Xauthority 536b
.bash_history 6.05kB
.bash_logout 24b
.bash_profile 191b
.bashrc 124b
.emacs 383b
.gtkrc 120b
.htpasswd 23b
Endpoint go#ep1
Path /~/
All None
.bash_logout
.bashrc
.profile
globus online
Go To: View Transfers
Transfer Activity
Cancel
1 of 1
Status
ID
Task Progress
Username
Completion Time
5fd30... 3 / 3 vas 11/18/2010 07:31 PM
f091a... 1 / 1 vas 11/18/2010 07:14 PM
0793e... 1 / 1 vas 11/17/2010 08:55 PM
049a3... 1 / 1 vas 11/17/2010 08:55 PM
00d9d... 1 / 1 vas 11/17/2010 08:55 PM
fdf64... 1 / 1 vas 11/17/2010 08:55 PM

Objective

- Easy to use high performance data mgmt environment
- Fast file transfer
  - facility-to-lab, facility-to-individual, lab-to-individual
- Reduced administrative overhead
- Better data curation

Challenges

- Access control
  - visibility
  - policies
- Provenance
  - data origin
  - history
- Meta-data
  - attributes
  - searching

User Credentials

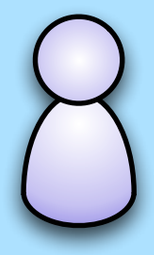
Unified Account Management

Hierarchical LDAP database

user basics

passwords

Standard schemas



Relational DB

user custom profiles

institutions

lab groups

Custom schemas

X.509 Digital Certificates

- Analogy to a passport:
  - Application form
  - Sponsor's attestation
  - Consular services
    - verification of application, sponsor, and accompanying identification and eligibility documents
  - Passport issuing office
- Portable, digital passport
  - fixed and secure user identifiers
    - name, email, home institution
  - signed by widely trusted issuer
  - time limited
  - ISO standard

Security

Access Control

- Need a strong Identity Management environment
  - individuals: identity tokens and identifiers
  - groups: membership lists
  - Active Directory/CIFS (Windows), Open Directory (Apple), FreeIPA (Unix) all LDAP-based
- Need to manage and communicate Access Control policies
  - institutionally driven
  - user driven
- Need Authorization System
  - Policy Enforcement Point (shell login, data access, web access, start application)
  - Policy Decision Point (store policies and understand relationship of identity token and policy)

Access Control

- What is a user?
  - .htaccess and .htpasswd
  - local system user (NIS or /etc/passwd)
  - portal framework user (proprietary DB schema)
  - grid user (X.509 DN)
- What are we securing access to?
  - Web pages?
  - URLs?
  - Data?
  - Specific operations?
  - Meta Data?
- What kind of policies do we enable?
  - Simplify to READ WRITE EXECUTE LIST ADMIN

Existing Security Infrastructure

- X.509 certificates
  - Department of Energy CA
  - Regional/Institutional RAs (SBGrid is an RA)
- X.509 proxy certificate system
  - Users self-sign a short-lived passwordless proxy certificate used for “portable” and “automated” grid processing identity token
  - Similarities to Kerberos tokens
- Virtual Organizations (VO) for definitions of roles, groups, attrs
- Attribute Certificates
  - Users can (attempt) to fetch ACs from the VO to be attached to proxy certs
- POSIX-like file access control (Grid ACL)

voms admin for VO: SBGrid Current user: Ian Stokes-Rees 411174

VO management | Subscriptions | Configuration | Other VOs on this server

Manage

- Users
- Groups
- Roles
- Attributes

Search roles

Create a new role

Roles:

RSV	delete
sbgrid	delete
VO-Admin	delete

1-3 of 3

GUMS 1.3.14

GRID User Management System

VOMS Servers

Home

Configuration
- Back Up/Restore
- Summary
- Persistence Factories
- VOMS Servers
- Account Mappers
  - Manage Pool Accounts
- User Groups
- Group To Account Mappings
- Host To Group Mappings
- Shortcut
- Merge Configuration

User Management
- Update VO Members
- Manual User Group Members
- Manual Account Mappings

View Mappings
- Map Grid Identity to Account

Configures VOMS servers.

edit / delete
VOMS Server: atlas
Description:
Base URL: https://voms.cern.ch:8443/voms/atlas/services/VOMSAdmin{remainder url}
Persistence Factory: mysql

edit / delete
VOMS Server: cdf
Description:
Base URL: https://voms-01.pd.infn.it:8443/voms/cdf/services/VOMSAdmin{remainder url}
Persistence Factory: mysql

edit / delete
VOMS Server: cdf1
Description:
Base URL: https://voms.cnaf.infn.it:8443/voms/cdf/services/VOMSAdmin{remainder url}
Persistence Factory: mysql

Gr ... lu

Directory listing /se/data/shared/biodb/scop-class/b/2/2/1/

1e5ba_ 1e5ca_ 1exga_ 1exha_ 1hehc_ 1hejc_ 1xbda_ 2xbda_

thumb	code	date	head	expdata
	1e5b	24-JUL-00	HYDROLASE	NMR
	1e5c	24-JUL-00	HYDROLASE	NMR, 5 STRUCTURES
	1exg	14-MAR-95	CELLULOSE DEGRADATION	NMR
	1exh	14-MAR-95	CELLULOSE DEGRADATION	NMR, 20 STRUCTURES
	1heh	22-NOV-00	HYDROLASE(XYLAN DEGRADATION)	NMR
	1hej	22-NOV-00	HYDROLASE(XYLAN DEGRADATION)	NMR, 5 STRUCTURES
	1xbd	16-OCT-98	HYDROLASE	NMR, 5 STRUCTURES
	2xbd	27-OCT-98	HYDROLASE	NMR

Grid Ove

Directory listing /se/data/shared/biodb/scop-class/g/9/1/2/

1h5oa_ 1z99a1

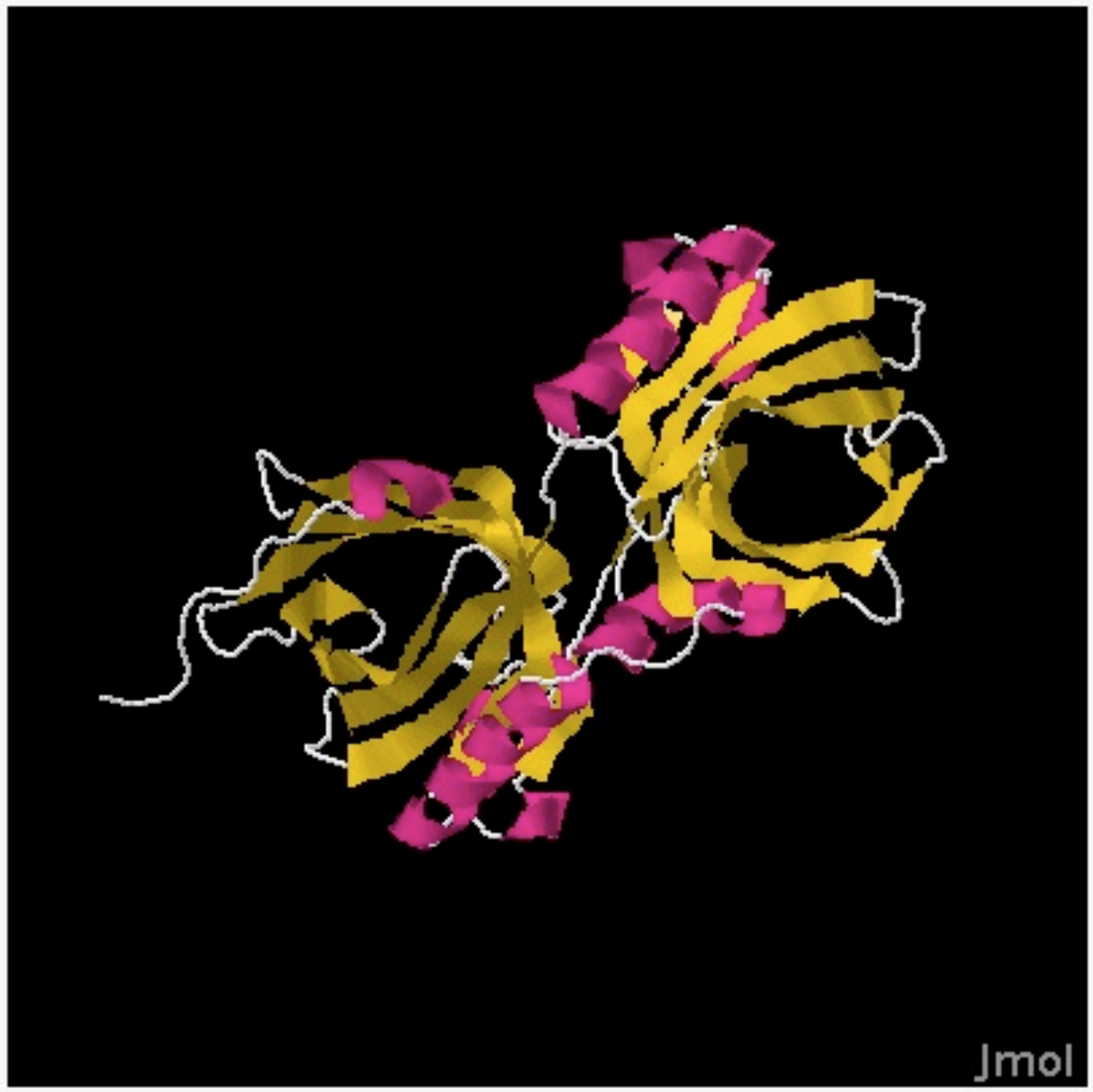
thumb	code	date	head	expdata
	1h5o	23-MAY-01	NEUROTOXIN	NMR, 26 STRUCTURES
	1z99	01-APR-05	TOXIN	NMR, 20 STRUCTURES

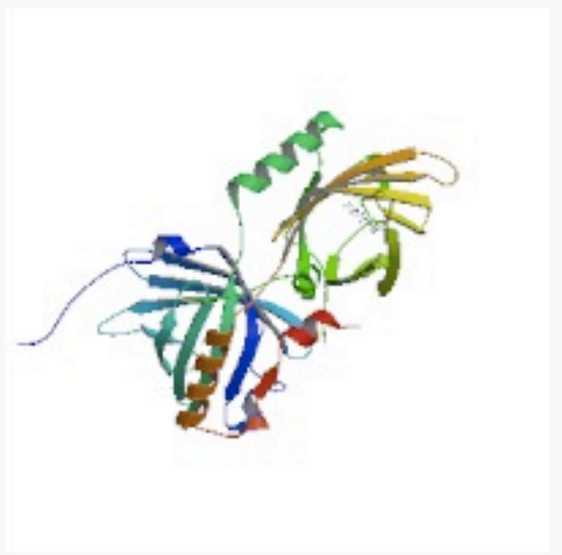
[Parent directory]

1h5oa_.pdb	1423274	14:57	4 Dec 07
1z99a1.pdb	1086839	16:55	4 Dec 07

You are /DC=org/DC=doegrids/OU=People/CN=Ian Stokes-Rees 411174

Manage directory . Switch to HTTP . Built with GridSite 1.6.0





```
PDB-Code:       1hn2
PDB-Header:     PROTEIN BINDING
PDB-Date:       05-DEC-00
PDB-Title:      CRYSTAL STRUCTURE OF BOVINE OBP COMPLEXED WITH  AMINOANTHRACENE
PDB-Keywords:   CRYSTAL STRUCTURE, OLFACTION, ODORANT BINDING PROTEIN,  AMINOANTHRACENE
PDB-ExpData:    X-RAY DIFFRACTION
PDB-Author:     F VINCENT S SPINELLI M TEGONI C CAMBILLAU
```

1hn2a_ 1hn2 a_ 04-DEC-07 b/60/1/1

1hn2b_ 1hn2 b_ 04-DEC-07 b/60/1/1

SBGrid Portal Browser Portal | File Browser | Queue Status | Poo

Manage directory /se/data/shared/biodb/s

[Parent directory]

Create .gacl

meta.html	1438	13:54	15 May 09	History Edit Delete Rename
1z99a1.pdb	1086839	16:55	4 Dec 07	Hi
1h5oa_.pdb	1423274	14:57	4 Dec 07	Hi

New name:

New name:

Upload file:

Local name:

You are /DC=org/DC=doegrids/OU=People/CN=Ian+Stokes-Rees+4

Back to directory . Built with GridSite 1.6.0

SBGrid Portal Browser Portal | File Browser | Queue Status | Pool Status

New Entry

Entry 1: Edit Entry Delete Entry

Credential No.	Attribute URI
1(Delete)	gacl:any-user <--
Add Credential	**Allowed:** read list **Denied:**

Entry 2: Edit Entry Delete Entry

Credential No.	Attribute URI
1(Delete)	dn:/DC=org/DC=doegrids/OU=People/CN=Ian+Stokes-Rees+411174 <--
Add Credential	**Allowed:** write admin **Denied:**

Entry 3: Edit Entry Delete Entry

Credential No.	Attribute URI
1(Delete)	dn:/DC=org/DC=doegrids/OU=People/CN=Peter+Doherty+869198
Add Credential	**Allowed:** write admin **Denied:**

```
<gacl>
  <entry>
    <any-user/>
    <allow>
      <read/><list/>
    </allow>
  </entry>
  <entry>
    <person>
      <dn>/DC=org/DC=doegrids/OU=People/CN=Ian Stokes-Rees 411174</dn>
    </person>
    <allow>
      <read/><list/><write/><admin/><exec/>
    </allow>
  </entry>
</gacl>
```

```
[ijstokes@abitibi jq]$ ls -Fla 1jqjd1.pdb
-r-xr-xr-x 1 datamgr hms 77123 Dec  4  2007 1jqjd1.pdb*
[ijstokes@abitibi jq]$ ls -Fla .meta/1jqjd1*
-rw-r--r-- 1 datamgr hms  1173 May 15 16:19 .meta/1jqjd1.html
-rw-r--r-- 1 datamgr hms 42532 May 14 13:28 .meta/1jqjd1-lg.png
-rw-r--r-- 1 datamgr hms    74 May 15 16:19 .meta/1jqjd1.pdb.meta
-rw-r--r-- 1 datamgr hms 16047 Jan 27  2009 .meta/1jqjd1.png
-rw-r--r-- 1 datamgr hms  1546 Jan 27  2009 .meta/1jqjd1-sm.png
-rw-r--r-- 1 datamgr hms   399 May 15 16:19 .meta/1jqjd1-summary.html
[ijstokes@abitibi jq]$ cat .meta/1jqjd1.pdb.meta
PDB-Code:       1jqjd1
PDB-Date:       04-DEC-07
PDB-SCCS:       a/80/1/1
```

Acknowledgements & Questions

- Piotr Sliz
  - Principle Investigator, head of SBGrid
- SBGrid Science Portal
  - Daniel O'Donovan, Meghan Porter-Mahoney
- SBGrid System Administrators
  - Ian Levesque, Peter Doherty, Steve Jahl
- Globus Online Team
  - Steve Tueke, Ian Foster, Rachana Ananthakrishnan, Raj Kettimuthu
- Ruth Pordes
  - Director of OSG, for championing SBGrid

Please contact me with any questions:

- Ian Stokes-Rees
- ijstokes@hkl.hms.harvard.edu
- ijstokes@spmetric.com

Look at our work

- portal.sbgrid.org
- www.sbgrid.org
- www.opensciencegrid.org