Grid, cloud, and science: Accelerating discovery

A View and Practice from University of Chicago

Ian Foster

Presented by Ioan Raicu

Computation Institute

Argonne National Lab & University of Chicago

April 28th, 2010

The Forum on Cloud in Academia at Illinois Institute of Technology

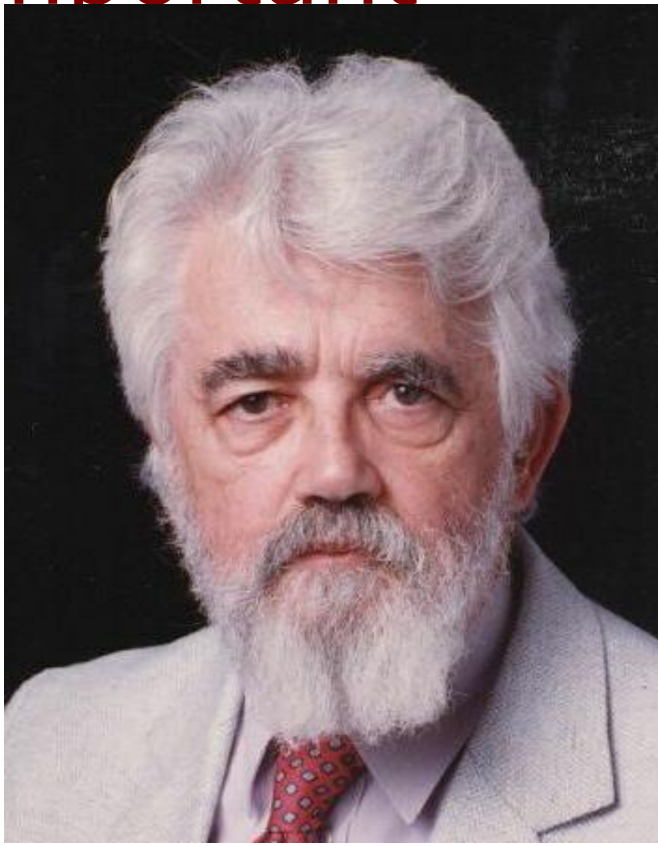
On-demand access to information and computation spurs creativity

- "By combining Google maps and Twitter restaurant reviews, I can help people find the best new restaurants"
- "By combining insurance records with flood predictions, I can calculate my company's exposure to a Katrina-class storm"
- "By integrating all new genome sequences, as they are produced, I can accelerate research on genetic diseases"

Reducing cycle time accelerates discovery

"Computation may someday be organized as a public utility … The computing utility could become the basis for a new and important industry."

John McCarthy (1961)

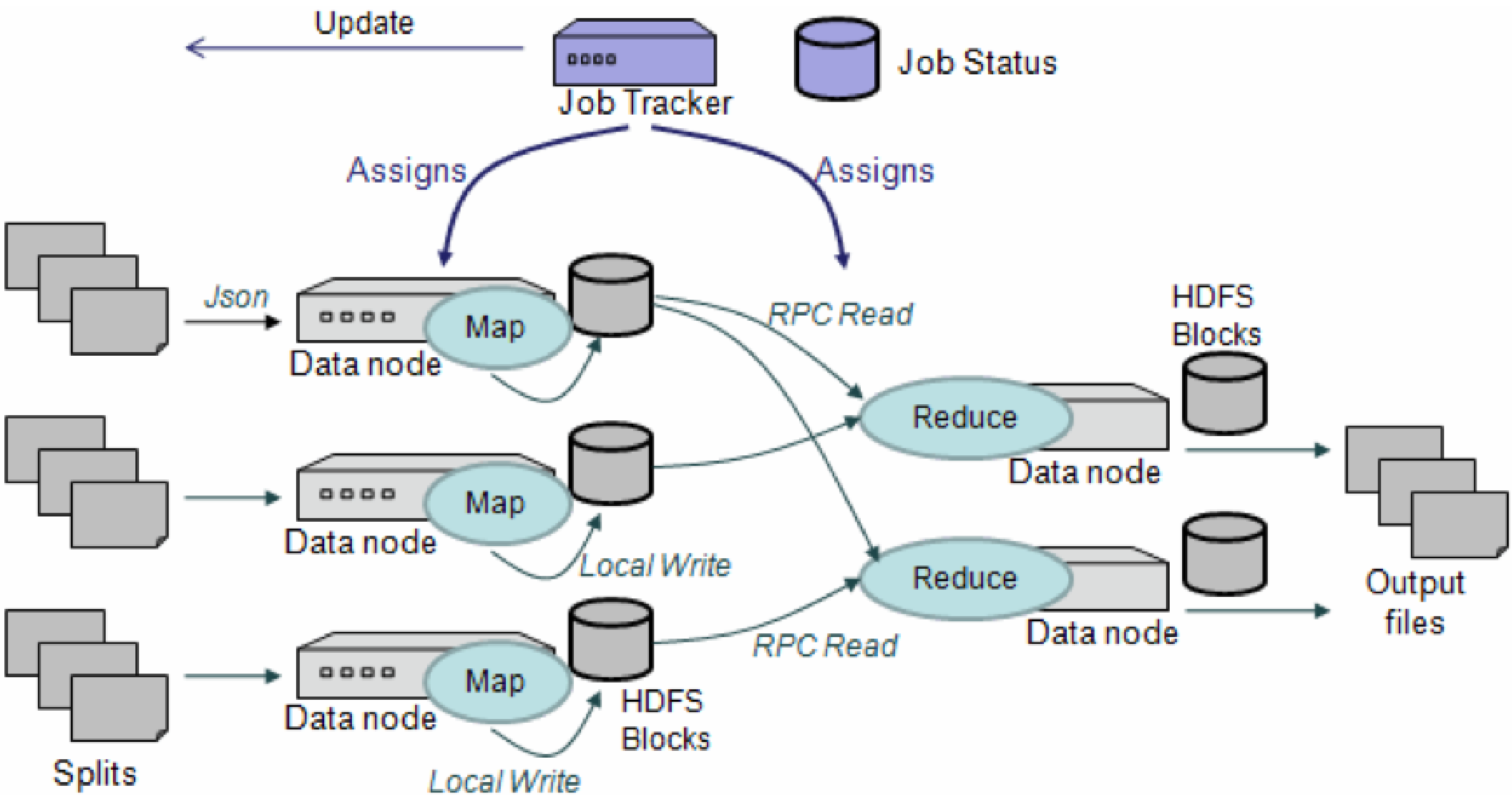


"I've been doing cloud computing since before it was called grid."

Grid = federation

Cloud = hosting

How data analysis happens at data-intensive computing workshops



How data analysis really happens in scientific laboratories

% foo file1 > file2

% bar file2 > file3

% foo file1 | bar > file3

% foreach f (f1 f2 f3 f4 f5 f6 f7 … f100)

foreach? foo $f.in | bar > $f.out

foreach? end

%

% Now where on earth is f98.out, and how did I generate it again?

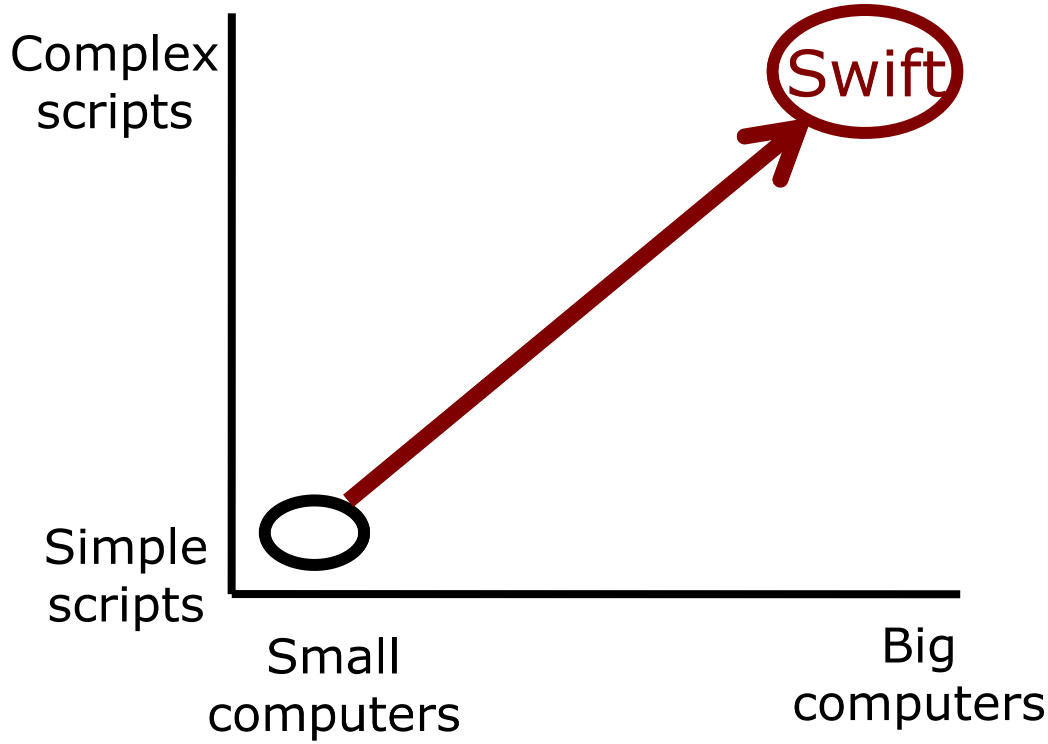
Now: command not found.

%

Parallel scripting

Ioan Raicu | Zhao Zhang | Mike Wilde

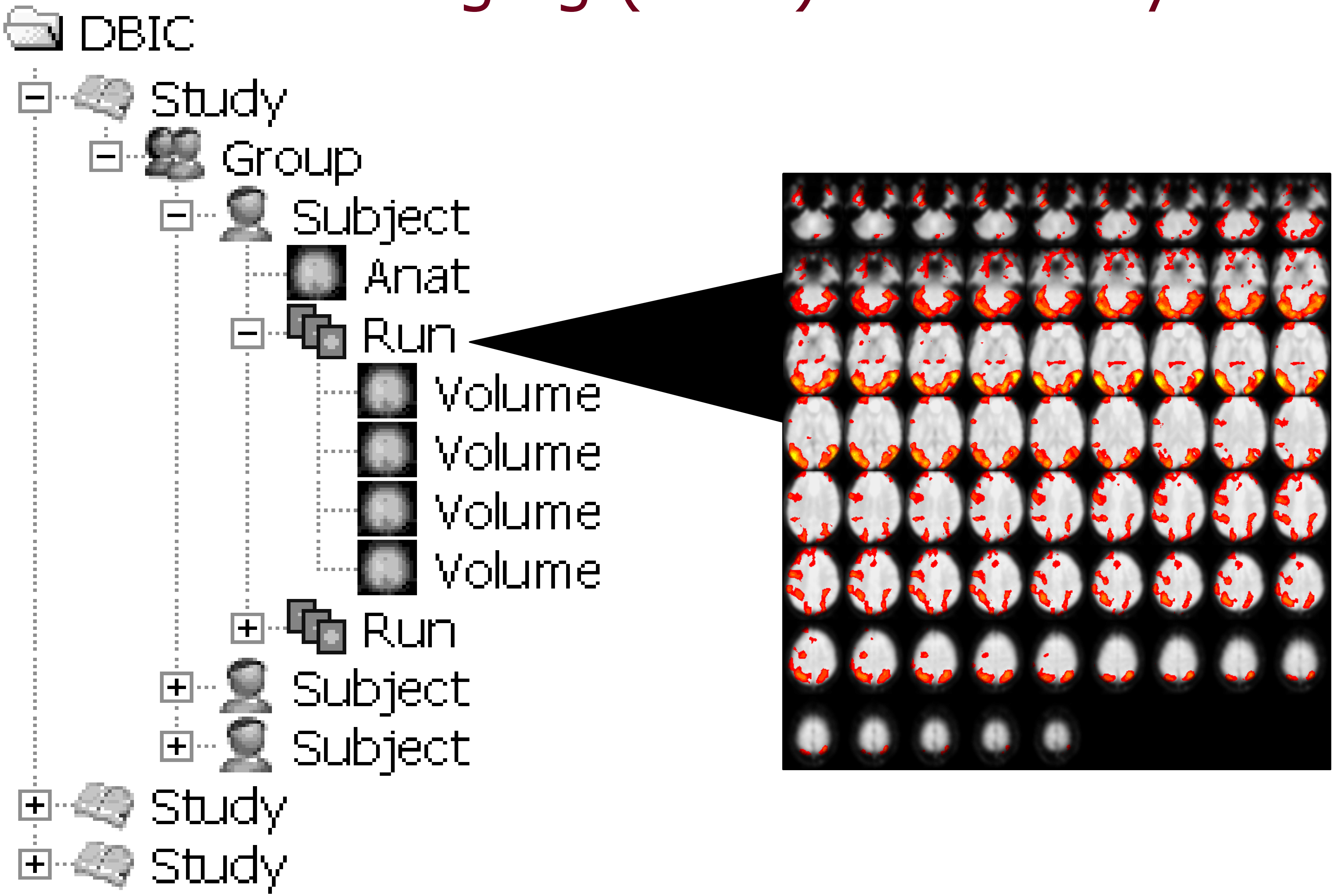
Many activities
Numerous files
Complex data
Data dependencies
Many programs



Preserving
file system semantics,
ability to call
arbitrary executables

Many processors
Storage hierarchy
Failure
Heterogeneity

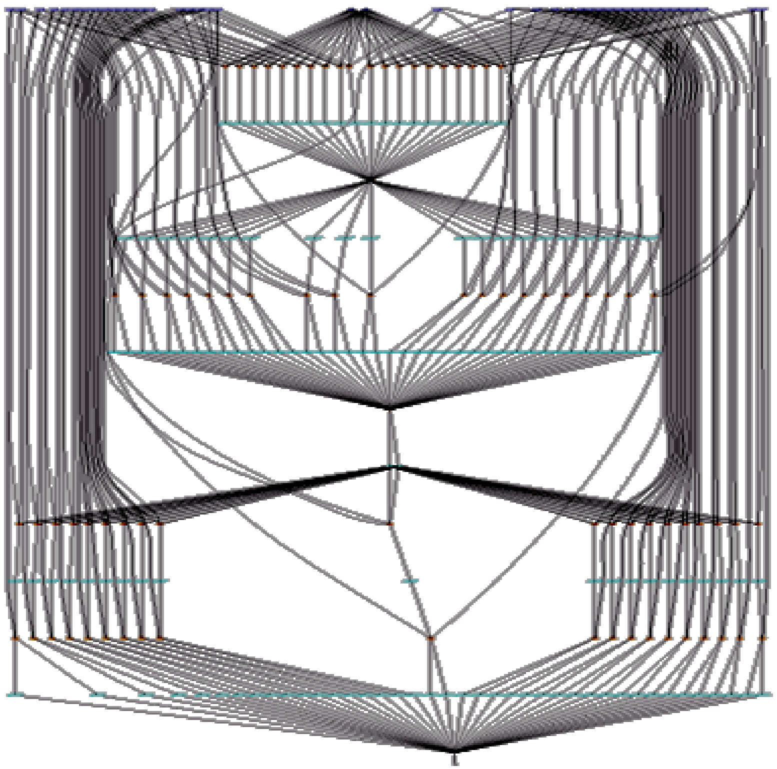
Functional magnetic resonance imaging (fMRI) data analysis

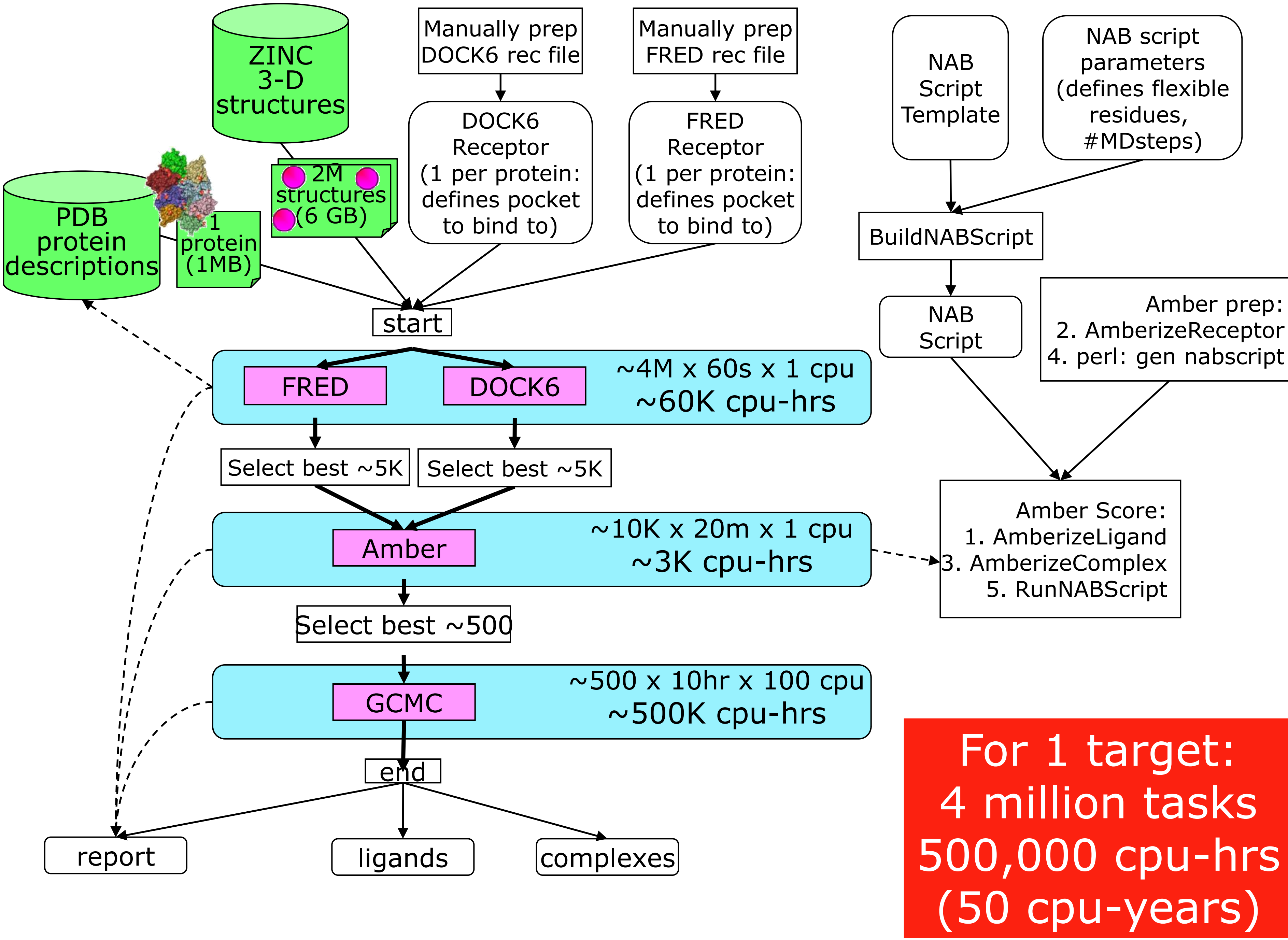


AIRSN program definition

```
(Run or) reorientRun (Run ir,
                      string direction) {
    foreach Volume iv, i in ir.v {
        or.v[i] = reorient(iv, direction);
    }
}
```

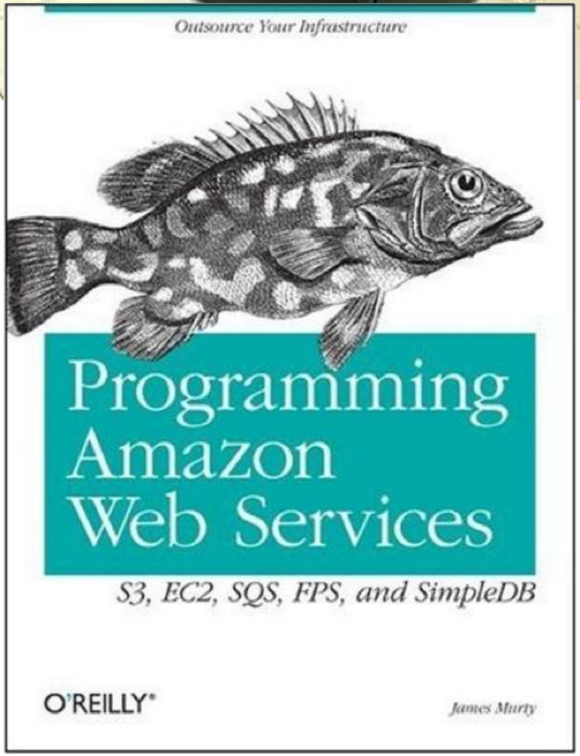
```
(Run snr) functional ( Run r, NormAnat a,
                       Air shrink ) {
    Run yroRun = reorientRun( r , "y" );
    Run roRun = reorientRun( yroRun , "x" );
    Volume std = roRun[0];
    Run rndr = random_select( roRun, 0.1 );
    AirVector rndAirVec = align_linearRun( rndr, s
    );
    Run reslicedRndr = resliceRun( rndr, rndAirVe
    Volume meanRand = softmean( reslicedRndr,
    Air mnQAAir = alignlinear( a.nHires, meanRa
    Warp boldNormWarp = combinewarp( shrink,
    Run nr = reslice_warp_run( boldNormWarp, r
    Volume meanAll = strictmean( nr, "y", "null"
    Volume boldMask = binarize( meanAll, "y" );
    snr = gsmoothRun( nr, boldMask, "6 6 6" );
}
```



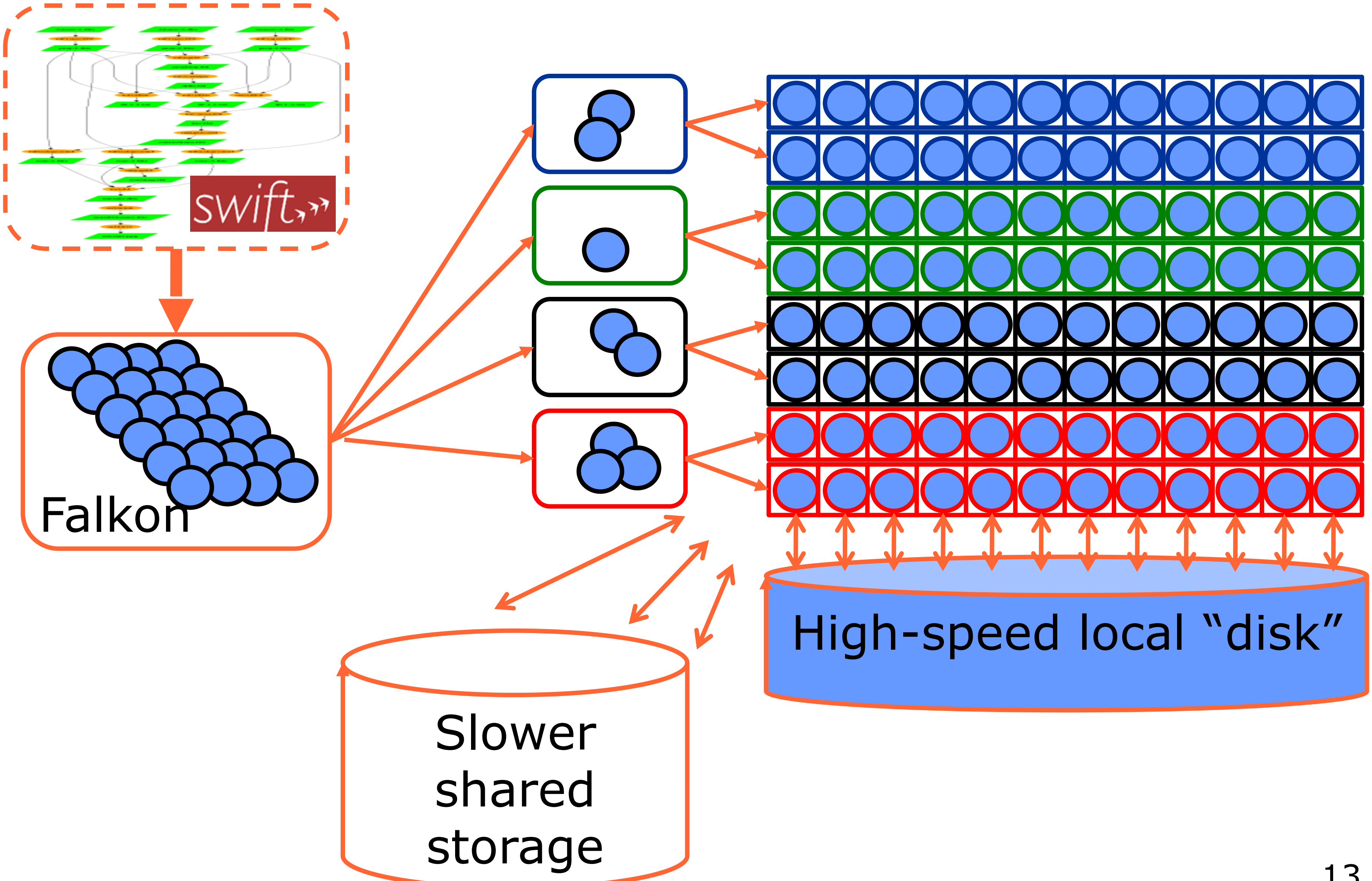

ZINC 3-D structures
Manually prep DOCK6 rec file
Manually prep FRED rec file
NAB Script Template
NAB script parameters (defines flexible residues, #MDsteps)
PDB protein descriptions
2M structures (6 GB)
1 protein (1MB)
DOCK6 Receptor (1 per protein: defines pocket to bind to)
FRED Receptor (1 per protein: defines pocket to bind to)
BuildNABScript
NAB Script
Amber prep:
2. AmberizeReceptor
4. perl: gen nabscript
start
FRED
DOCK6
~4M x 60s x 1 cpu
~60K cpu-hrs
Select best ~5K
Select best ~5K
Amber
~10K x 20m x 1 cpu
~3K cpu-hrs
Amber Score:
1. AmberizeLigand
3. AmberizeComplex
5. RunNABScript
Select best ~500
GCMC
~500 x 10hr x 100 cpu
~500K cpu-hrs
end
report
ligands
complexes
For 1 target:
4 million tasks
500,000 cpu-hrs
(50 cpu-years)


IBM
Blue Gene/P

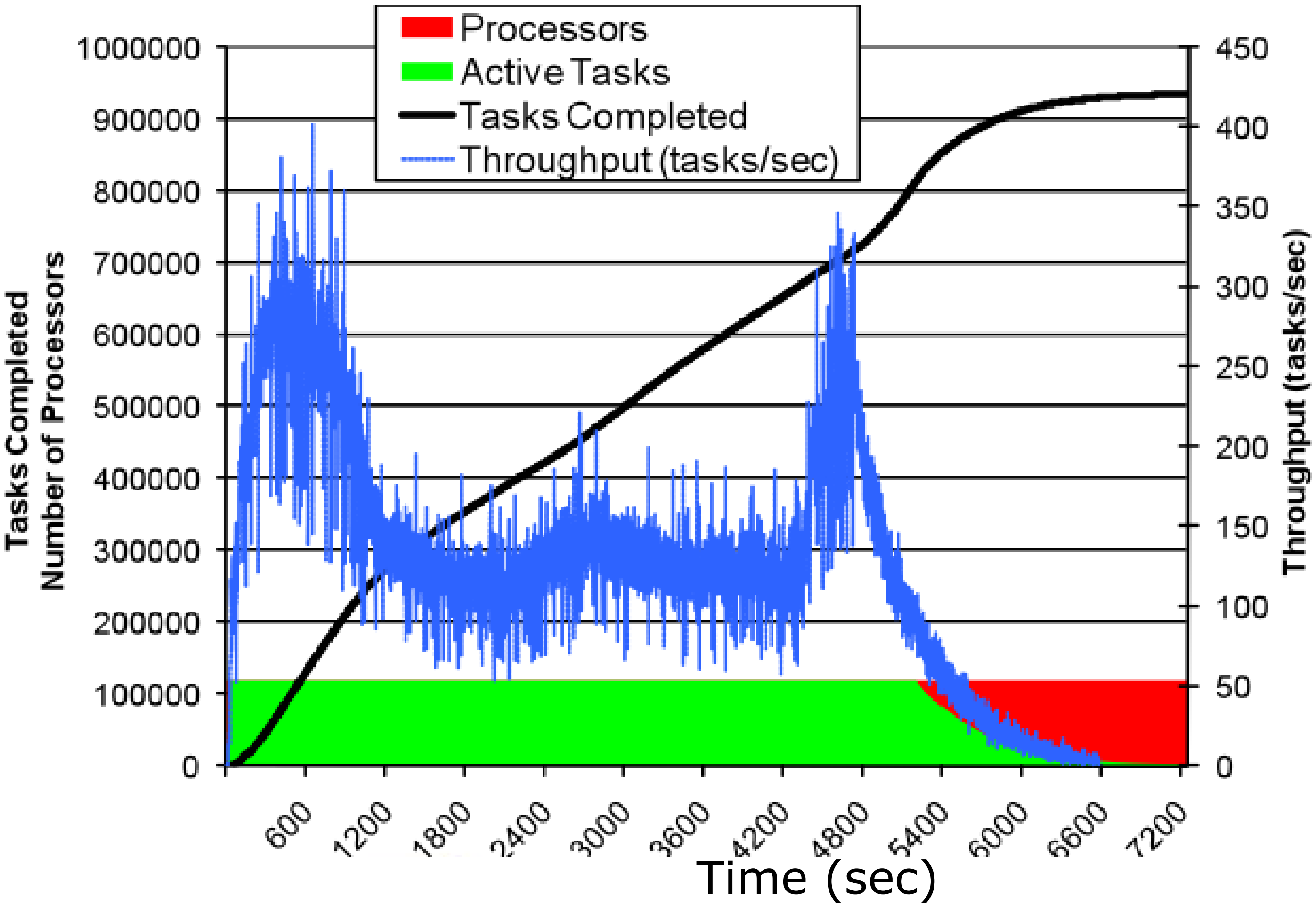

TERAGRID


Outsource Your Infrastructure
Programming Amazon Web Services
S3, EC2, SQS, FPS, and SimpleDB
O'REILLY
James Murty

Managing 160,000 cores



DOCK on BG/P: ~1M tasks on 119,000 CPUs

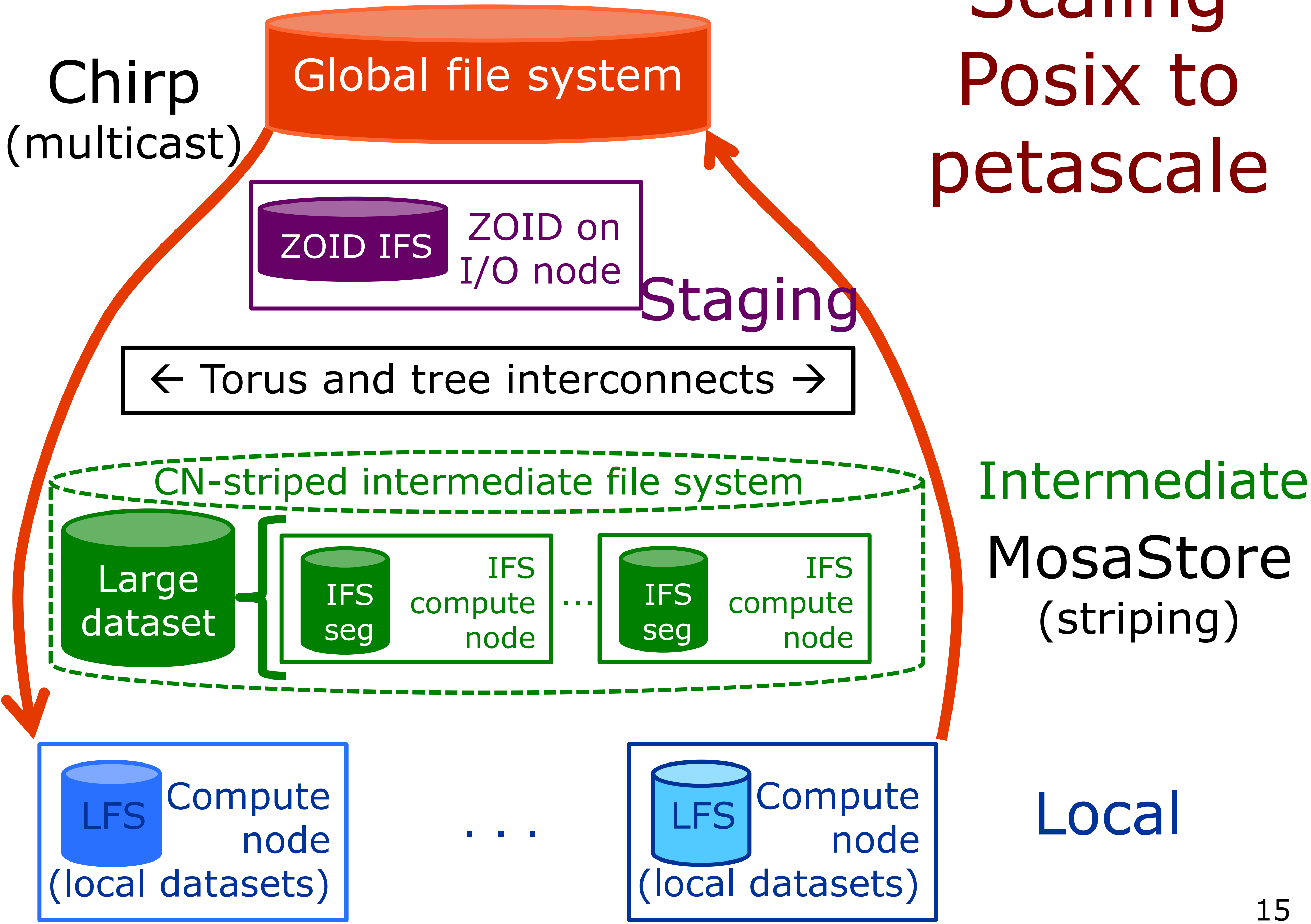


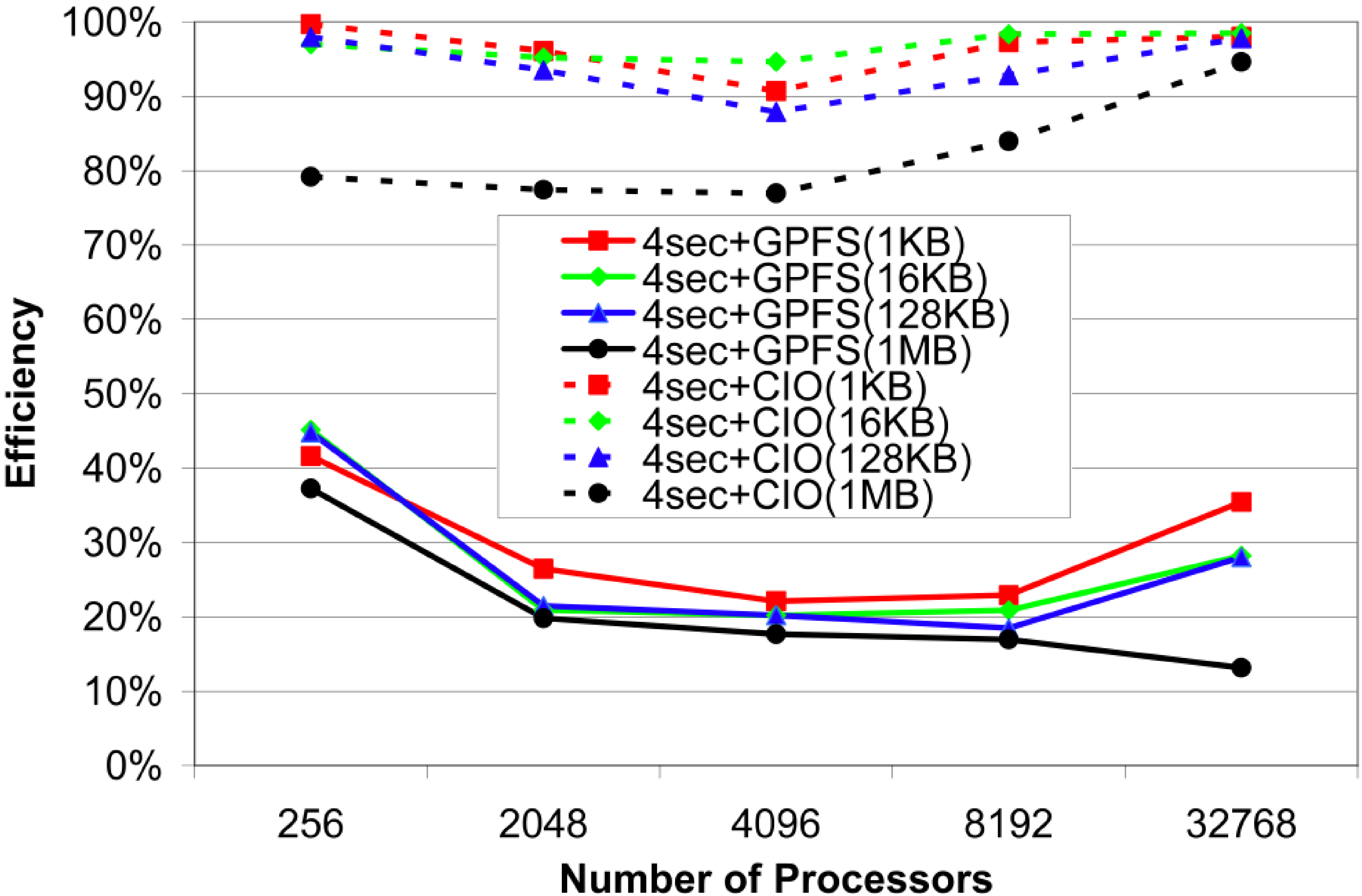
- 118784 cores
- 934803 tasks
- Elapsed time: 7257 sec
- Compute time: 21.43 CPU years
- Average task: 667 sec

Relative efficiency 99.7% (from 16 to 32 racks)
Utilization: 99.6% sustained, 78.3% overall

Ioan Raicu et al.

Scaling Posix to petascale

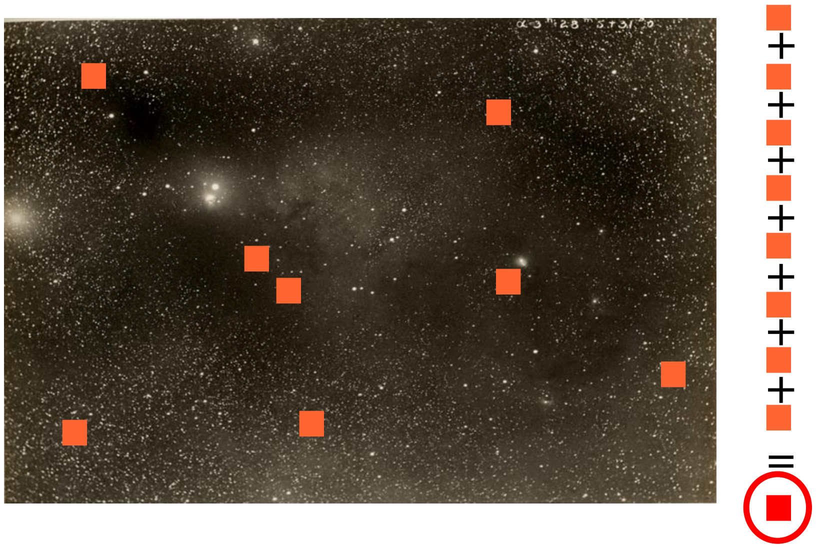


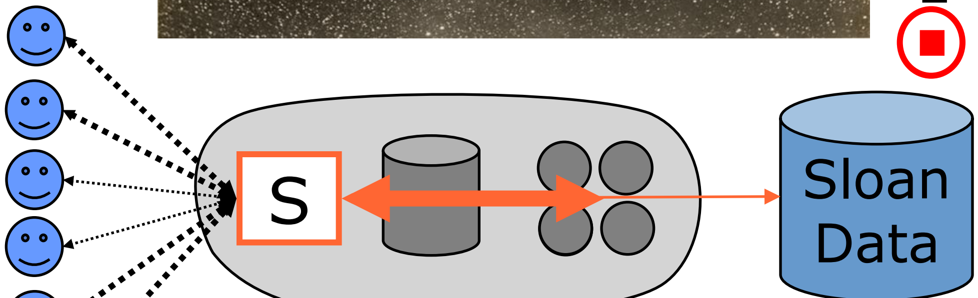


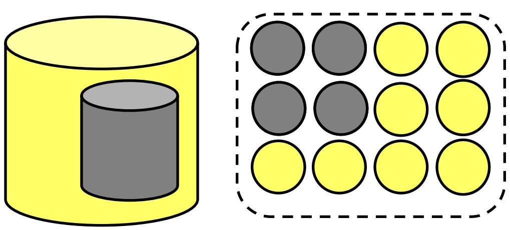
Efficiency for 4 second tasks and varying data size (1KB to 1MB) for CIO and GPFS up to 32K processors

Provisioning for data-intensive workloads

- Example: on-demand "stacking" of arbitrary locations within ~10TB sky survey
- Challenges
  - Random data access
  - Much computing
  - Time-varying load
- Solution
  - Dynamic acquisition of compute & storage





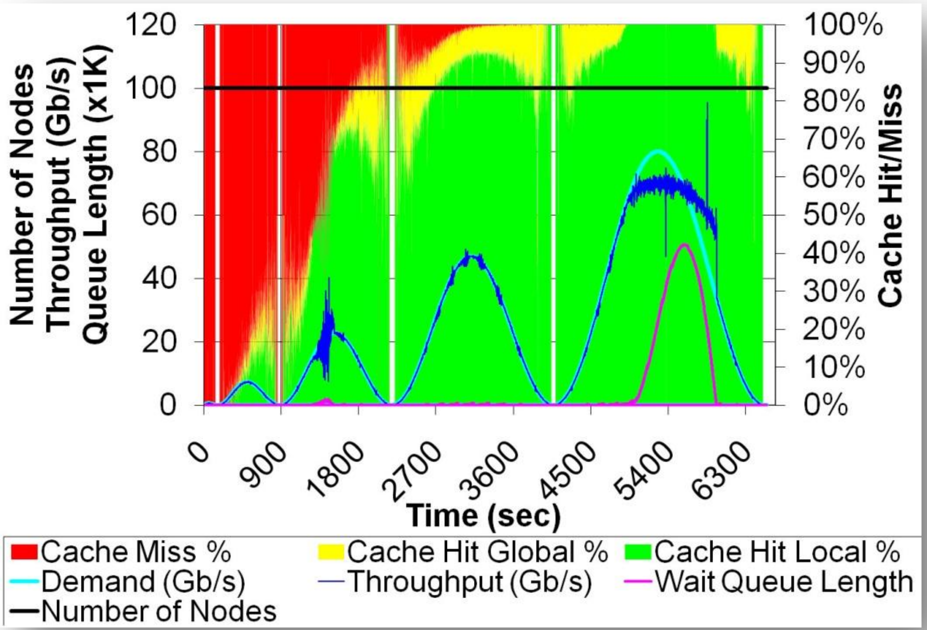


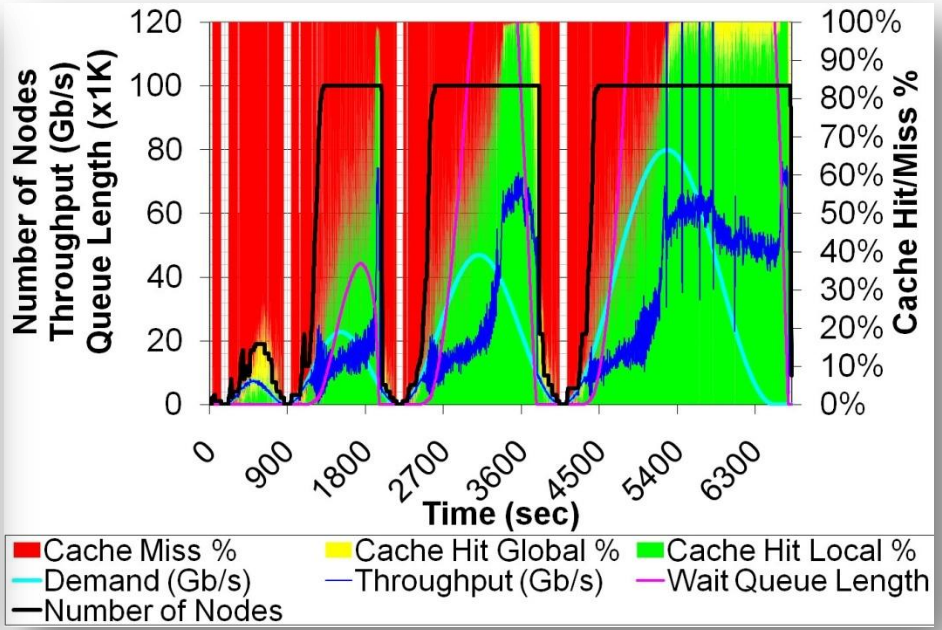
Ioan Raicu

Data diffusion

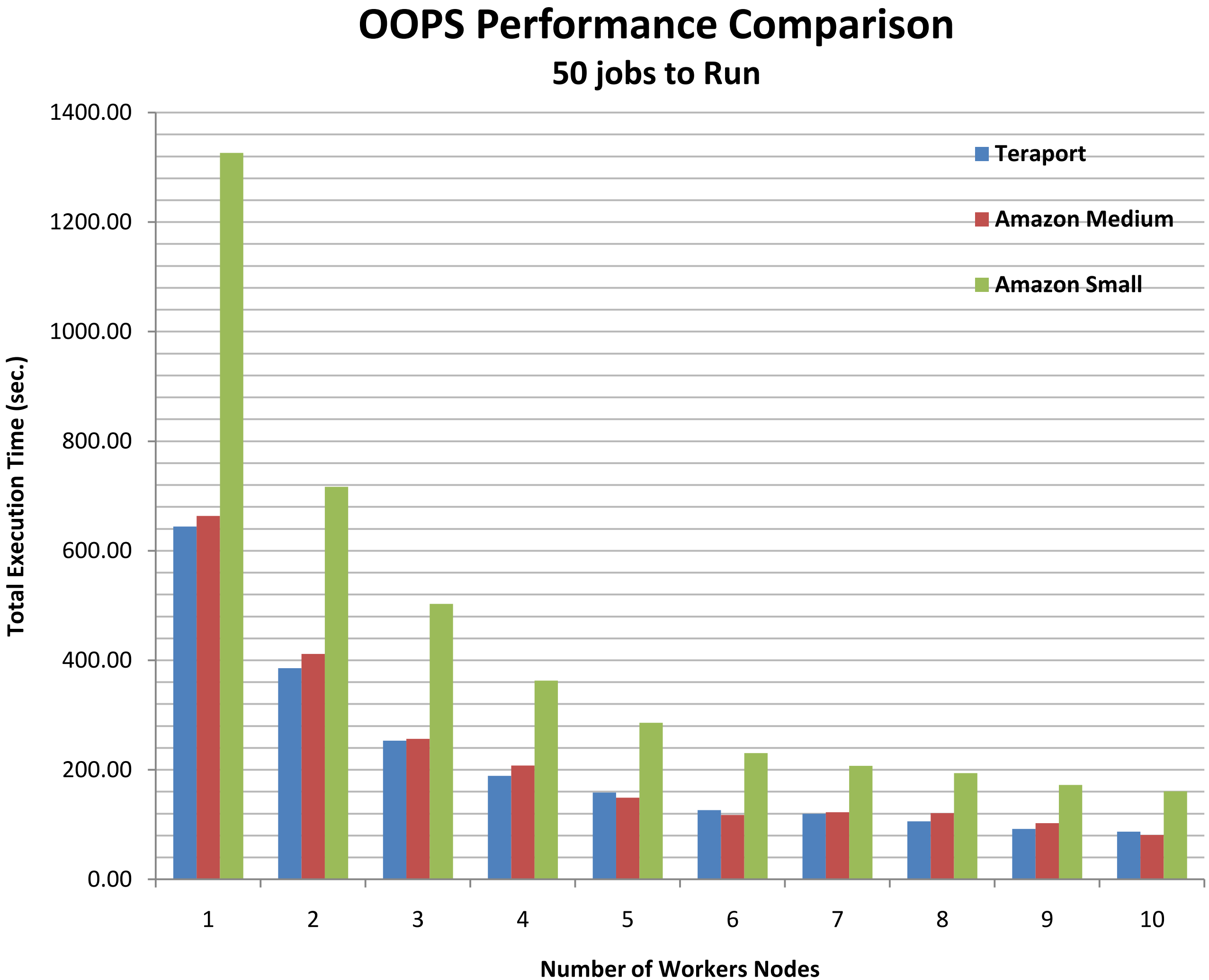
Data diffusion sine-wave workload: Summary

- GPFS → 5.70 hrs, ~8Gb/s, 1138 CPU hrs
- DD+SRP → 1.80 hrs, ~25Gb/s, 361 CPU hrs
- DD+DRP → 1.86 hrs, ~24Gb/s, 253 CPU hrs





Yi Zhu



Teraport:

2 X AMD Opteron 248 2.2GB

4GB Memory

Amazon Standard Small:

1 EC2 Compute Unit*

1.7GB Memory

$0.10 per hour

Amazon High-CPU Medium:

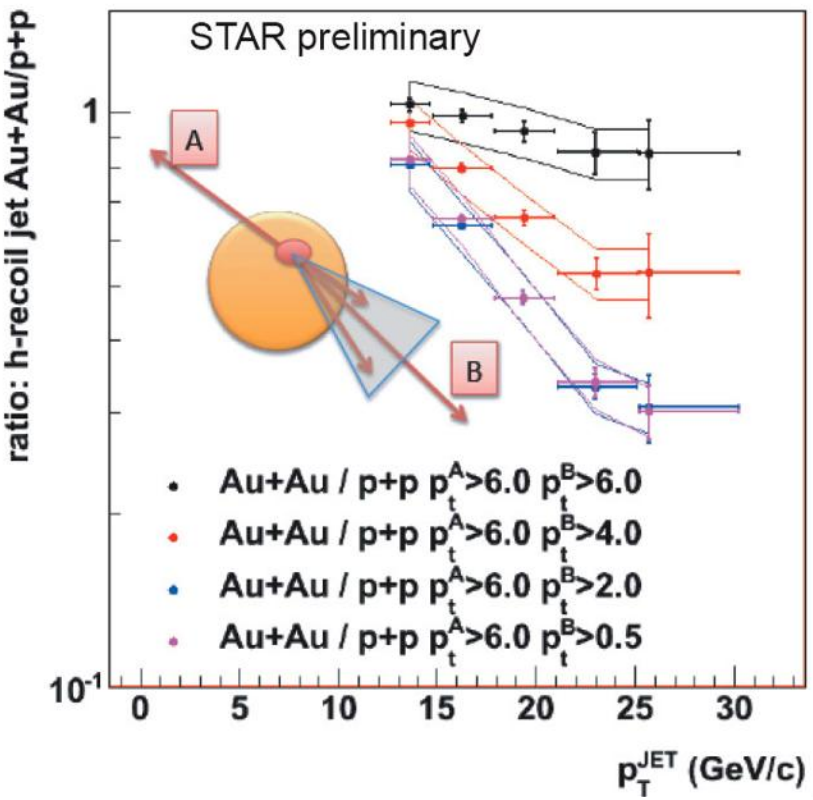
2 X 2.5 EC2 CU

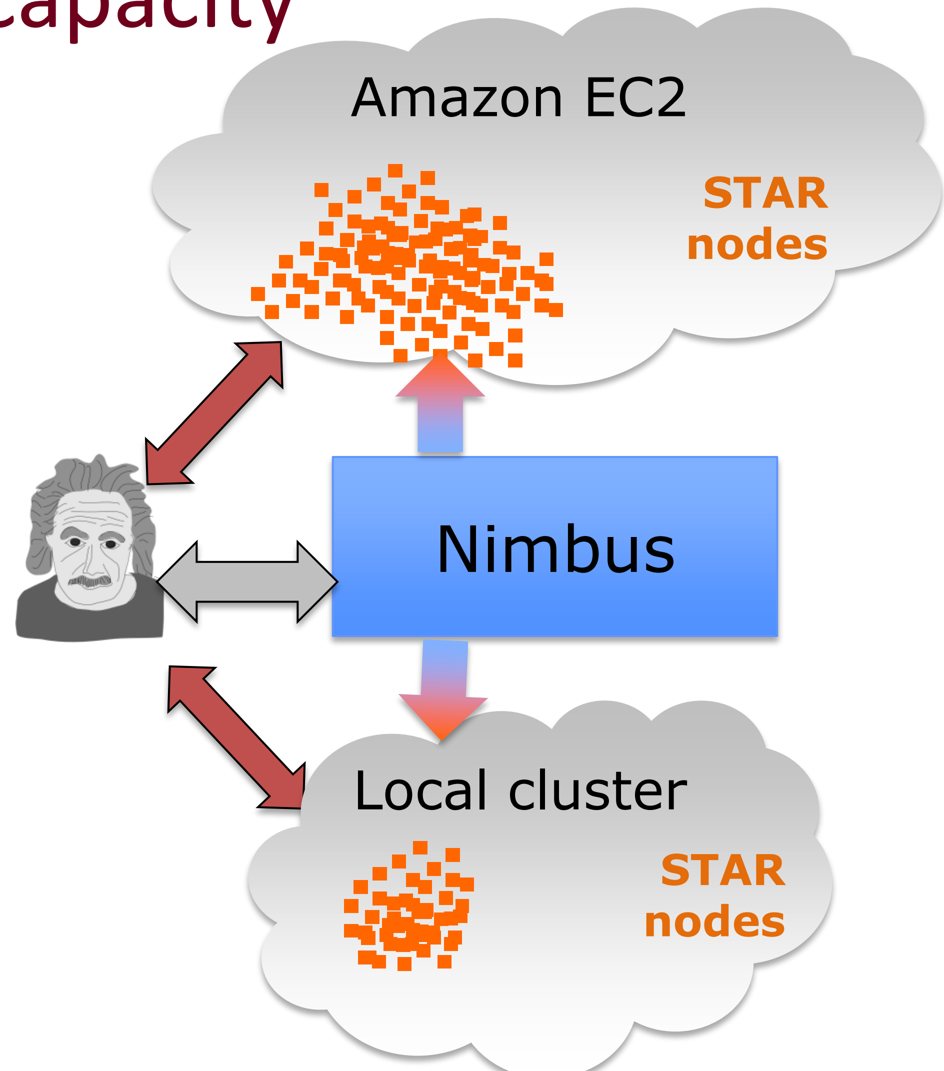
1.7GB Memory

$0.20 per hour

One EC2 Compute Unit equals 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor

Elastic capacity

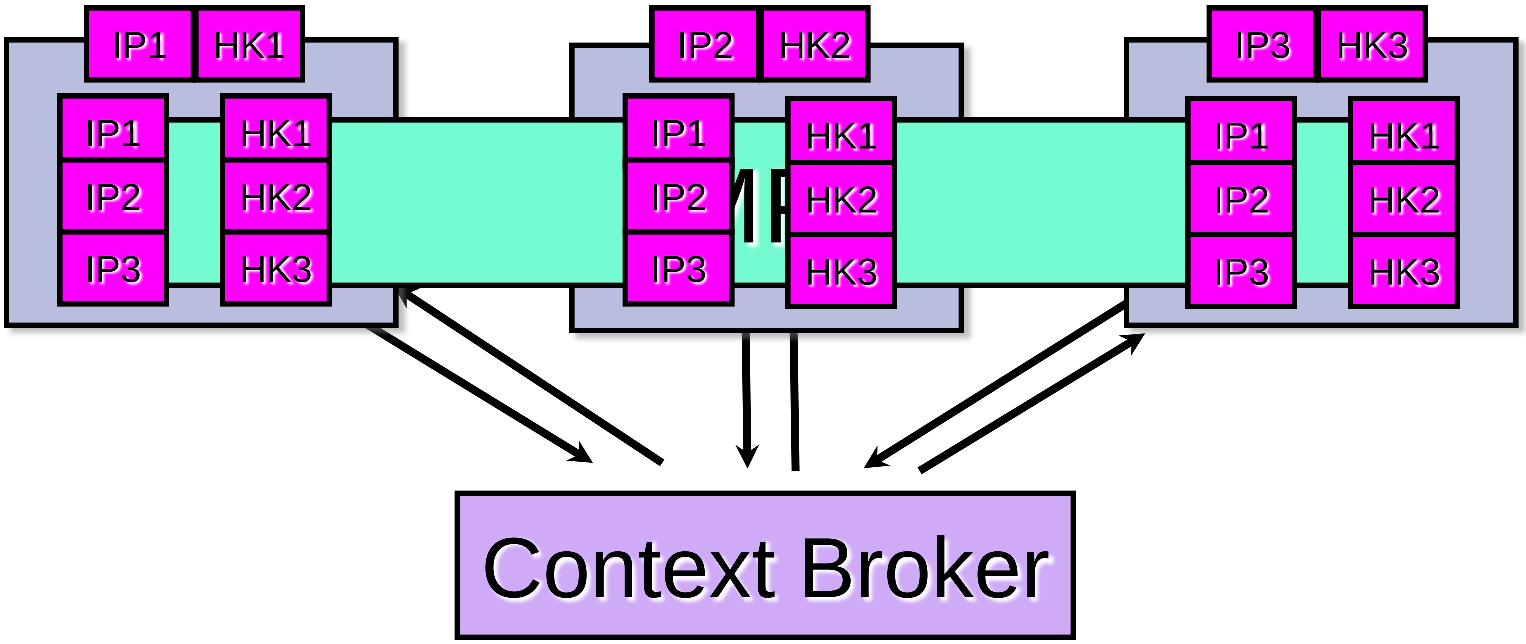




Jérôme Lauret, Kate Keahey et al.

Turnkey Virtual Clusters

- Turnkey, tightly-coupled cluster
  - ◆ Shared trust/security context
  - ◆ Shared configuration/context information

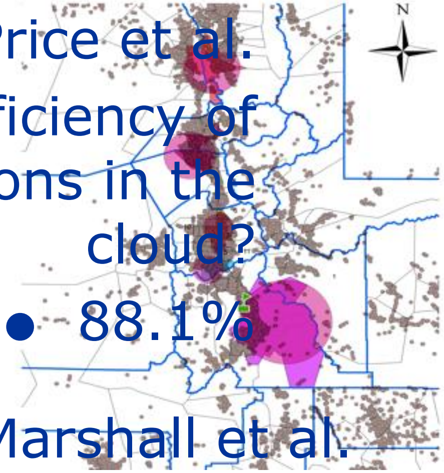


Examples of Nimbus Applications

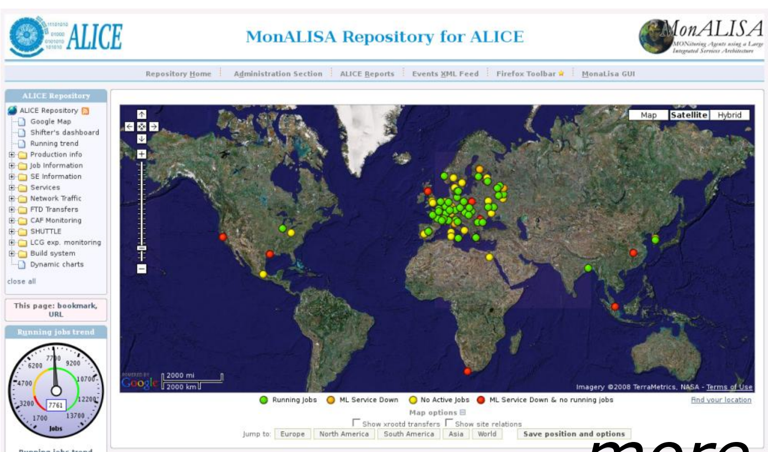
- Jerome Lauret et al.
- Are Science Clouds for me?

STAR

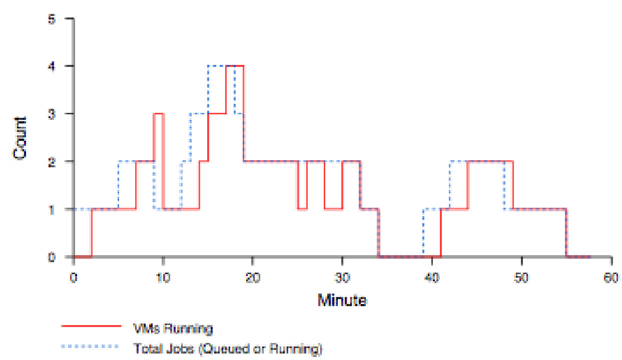
- Ron Price et al.
- What is the efficiency of our simulations in the cloud?
  - 88.1%



- Harutyunyan et al.
- Can we elastically extend a production testbed?



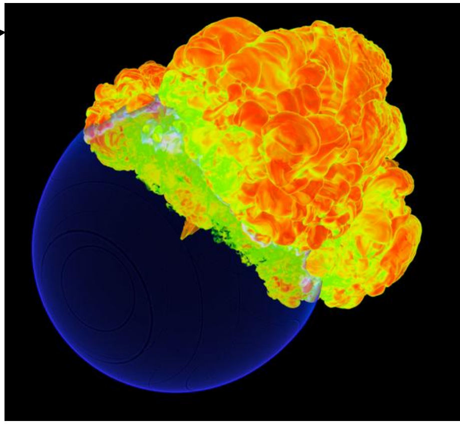
- Marshall et al.
- How reactive are such extensions going to be?

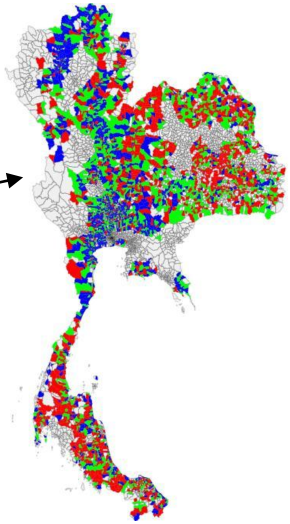


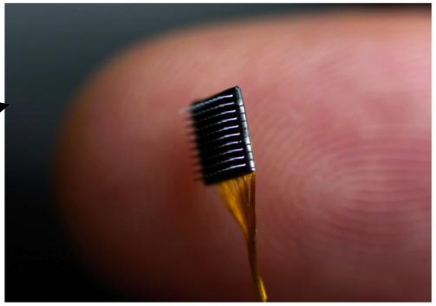
more at www.nimbusproject.org

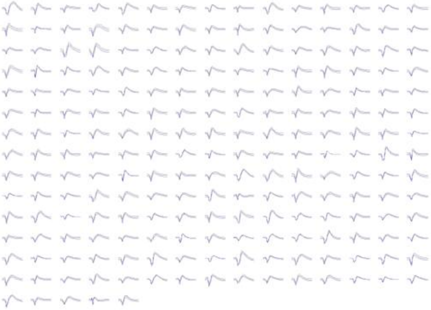
Data-intensive computing @ Computation Institute: Example applications

- Astrophysics
- Cognitive science
- East Asian studies
- Economics
- Environmental science
- Epidemiology
- Genomic medicine
- Neuroscience
- Political science
- Sociology
- Solid state physics







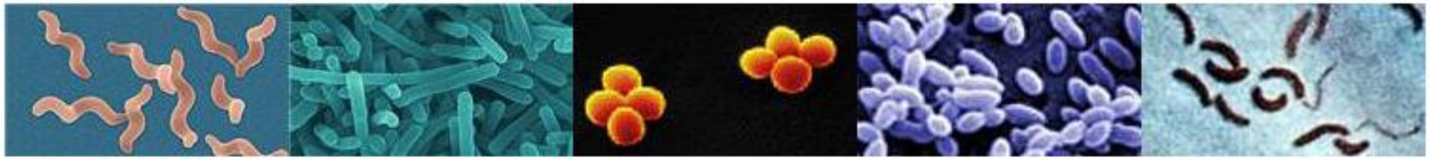


NMPDR Home

http://www.nmpdr.org/

MacZealots ...ss on Tiger | Nullriver So...e ~ Products | Scopus | Zimbra | UC Proxy | The SEED: Entry Page | MCS Mac Software | Google | Radio-Locator

NMPDR — National Microbial Pathogen Data Resource Center



- Search
- About Us
- How to Use NMPDR
  - FAQs
  - Searching NMPDR
  - Navigating NMPDR
  - Teaching with NMPDR New
  - Citing NMPDR
  - Database Design
- Organism Data Summaries
  - Campylobacter
  - Listeria
  - Staphylococcus
  - Streptococcus
  - Vibrio
- Subsytem Summaries
- Subsystem-Based Genome Annotation Service New
- Signature Genes Tool
- Essential Genes
- Drug, Antitoxin, Vaccine Targets
- Publications
- Events
- News and Journals
- Links
  - Cells, Reagents, Databases, Images
  - Bioinformatic Resource Centers
  - Regional Centers of Excellence
  - The SEED
- Downloads
- Contact Us

National Microbial Pathogen Data Resource

Welcome to NMPDR — an environment for the comparative functional analysis of genomes and biological subsystems, with an emphasis on pathogenic species of *Campylobacter*, *Listeria*, *Staphylococcus*, *Streptococcus*, and *Vibrio*.

Announcing **Summer workshop for undergraduate educators.** NMPDR is hosting a collaborative workshop for the development of teaching materials that make use of NMPDR resources for teaching undergraduate microbiology, July 9 - 12. More info >>

Version of October 8, 2008

Includes 45 archaeal, 660 bacterial, and 29 eukaryal genomes.

612 active subsystems, 809925 new annotations, 384028 new features.

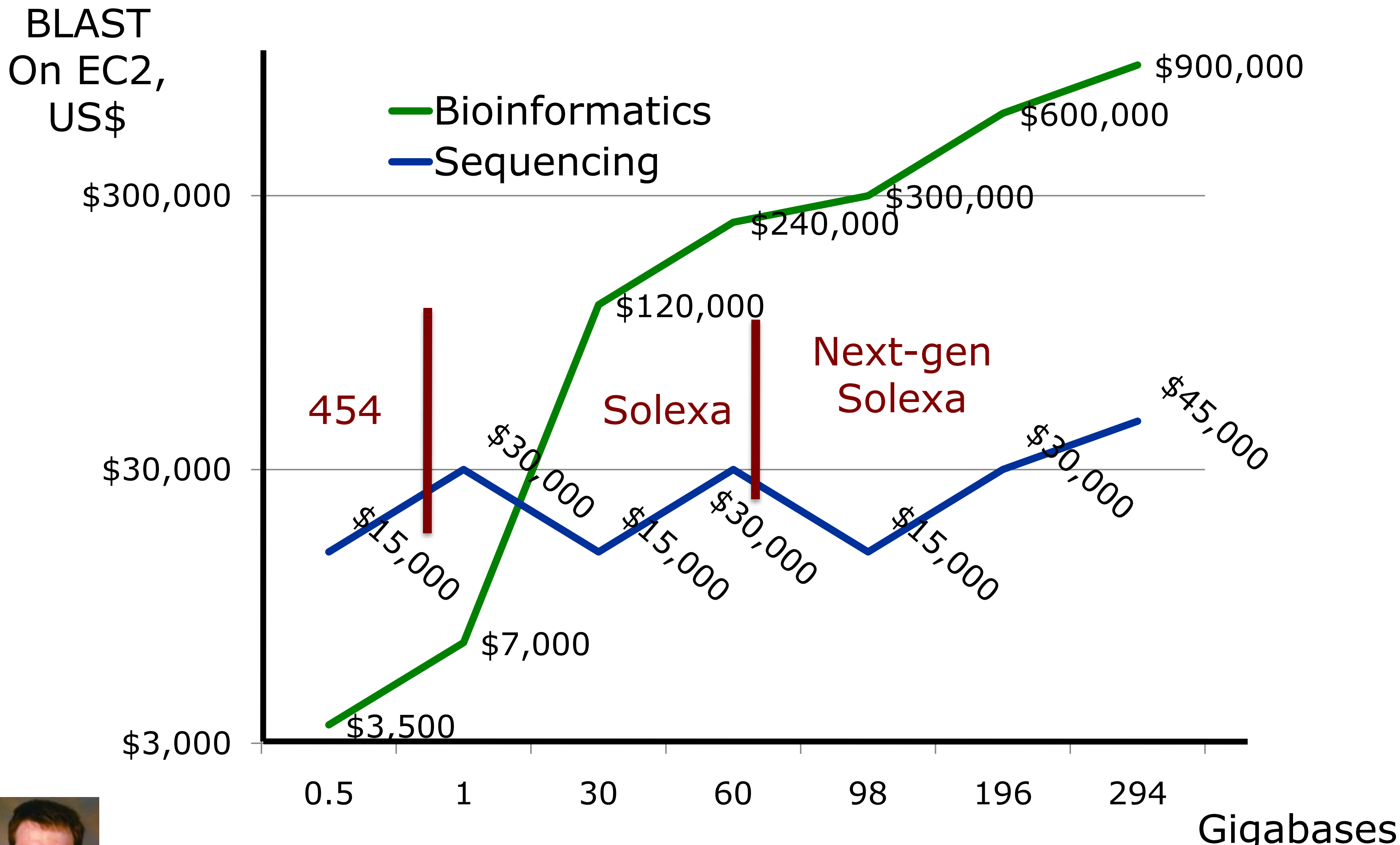
Using a strategy of subsystems annotation, NMPDR provides researchers with corrected functional annotations in a structured biological context. Hundreds of distinct subsystems have been developed to describe central and secondary metabolism, complex structures, and virulence or other phenotypes.

Find gene or protein keyword(s) or numerical id **Go**

BLAST | Genes | Subsystems | Organisms

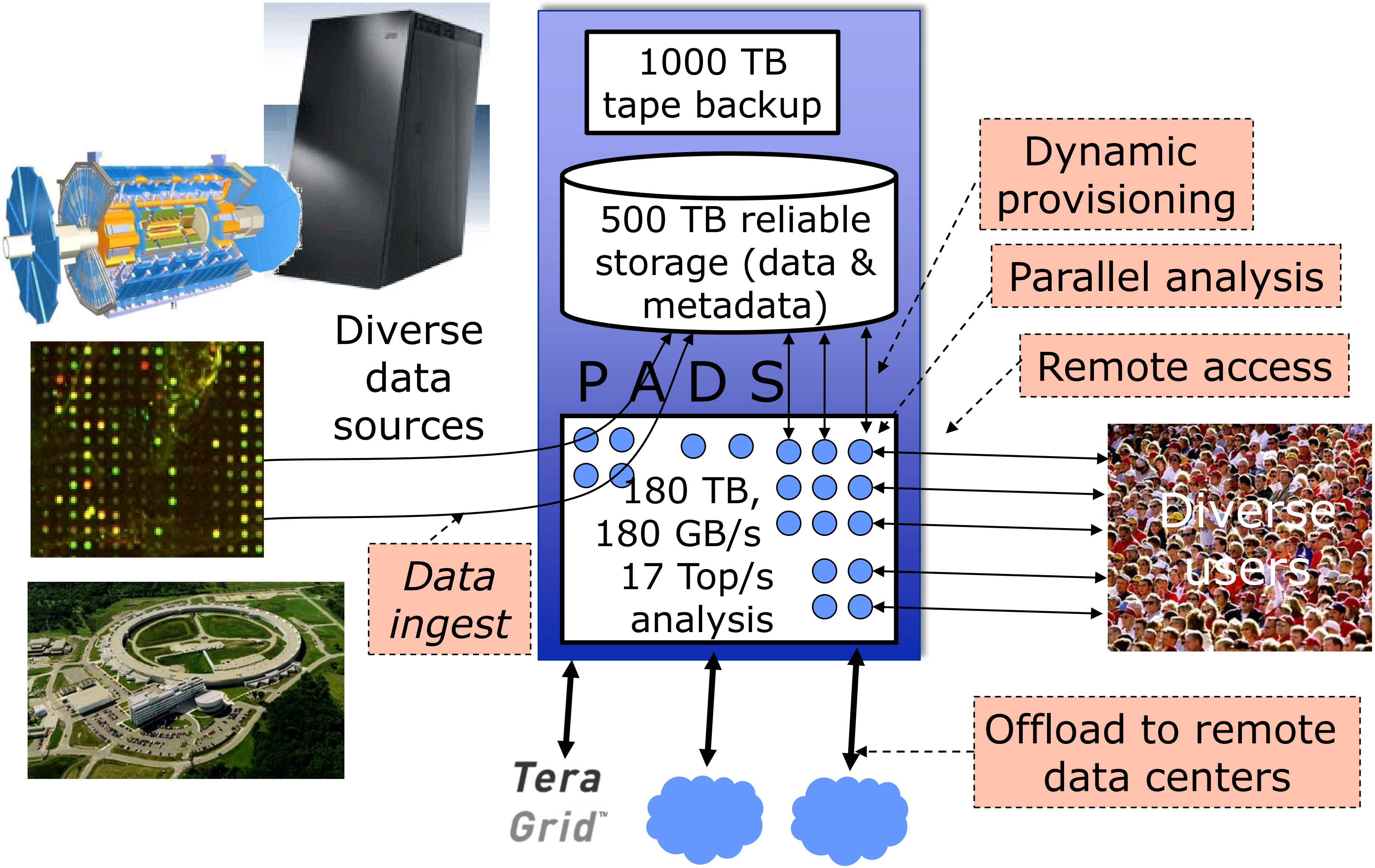
Quick-start guide to using NMPDR

Sequencing outpaces Moore's law



Folker Meyer, Computation Institute

Data-intensive computing @ Computation Institute: Hardware

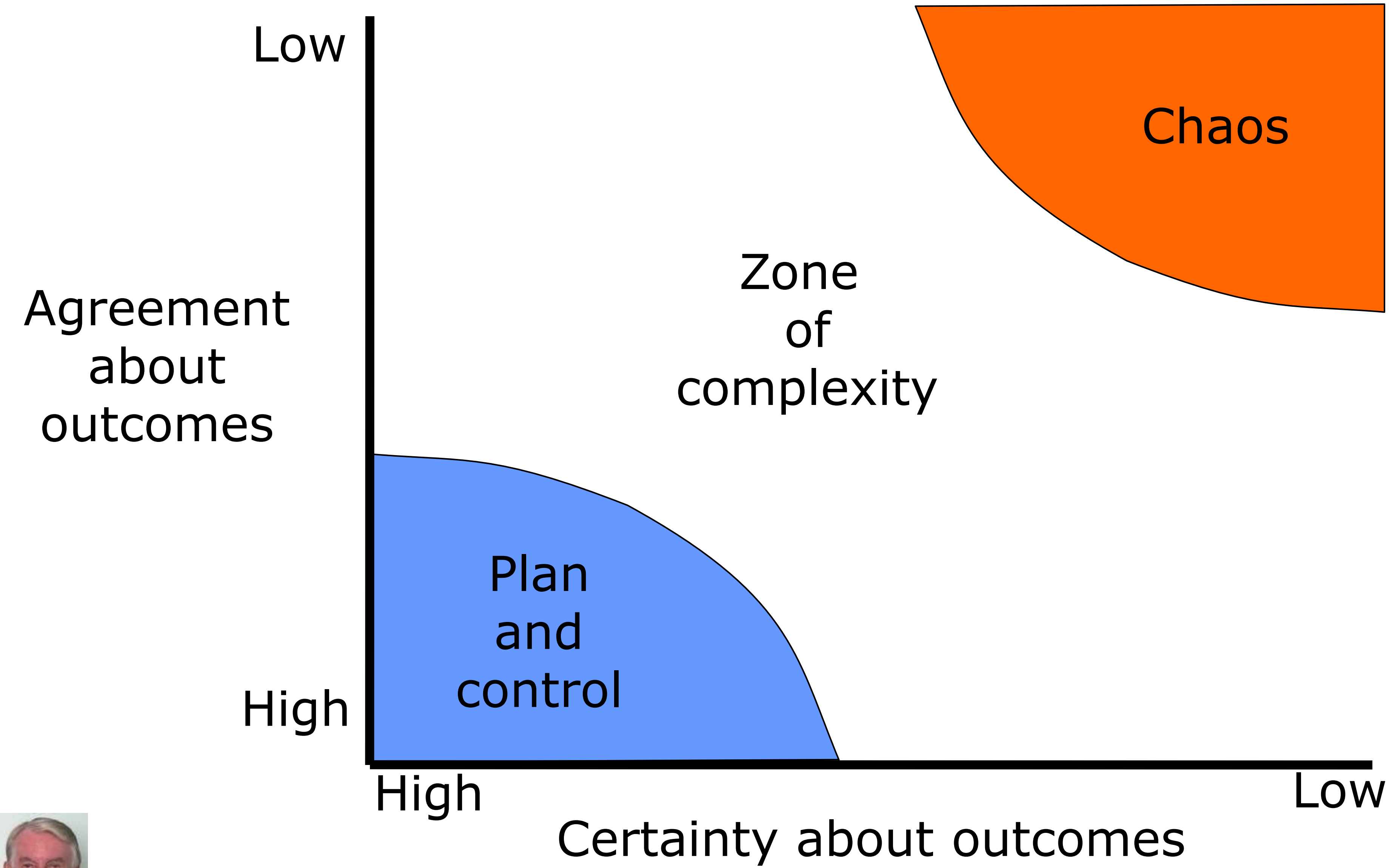


PADS: Petascale Active Data Store (NSF MRI)

Data-intensive computing @ Computation Institute: Software

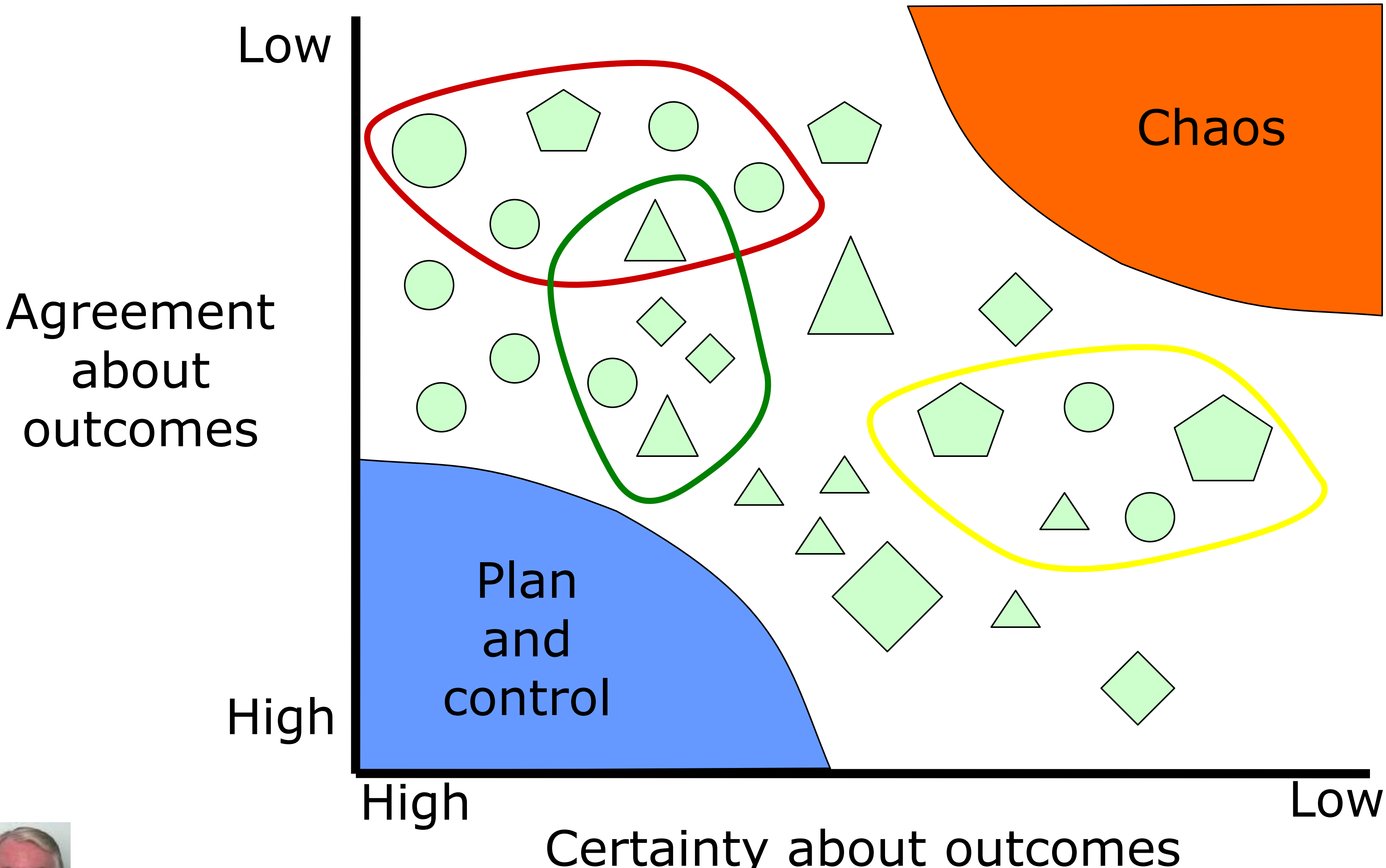
- HPC systems software (MPICH, PVFS, ZeptOS)
- Collaborative data tagging (GLOSS)
- Data integration (XDTM)
- HPC data analytics and visualization
- Loosely coupled parallelism (Swift, Hadoop)
- Dynamic provisioning (Falkon)
- Service authoring (Introduce, caGrid, gRAVI)
- Provenance recording and query (Swift)
- Service composition and workflow (Taverna)
- Virtualization management (Workspace Service)
- Distributed data management (GridFTP, etc.)

Data-intensive computing is an end-to-end problem



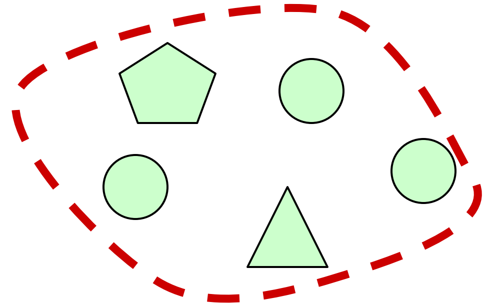
Ralph Stacey, *Complexity and Creativity in Organizations*, 1996

We need to function in the **zone of complexity**

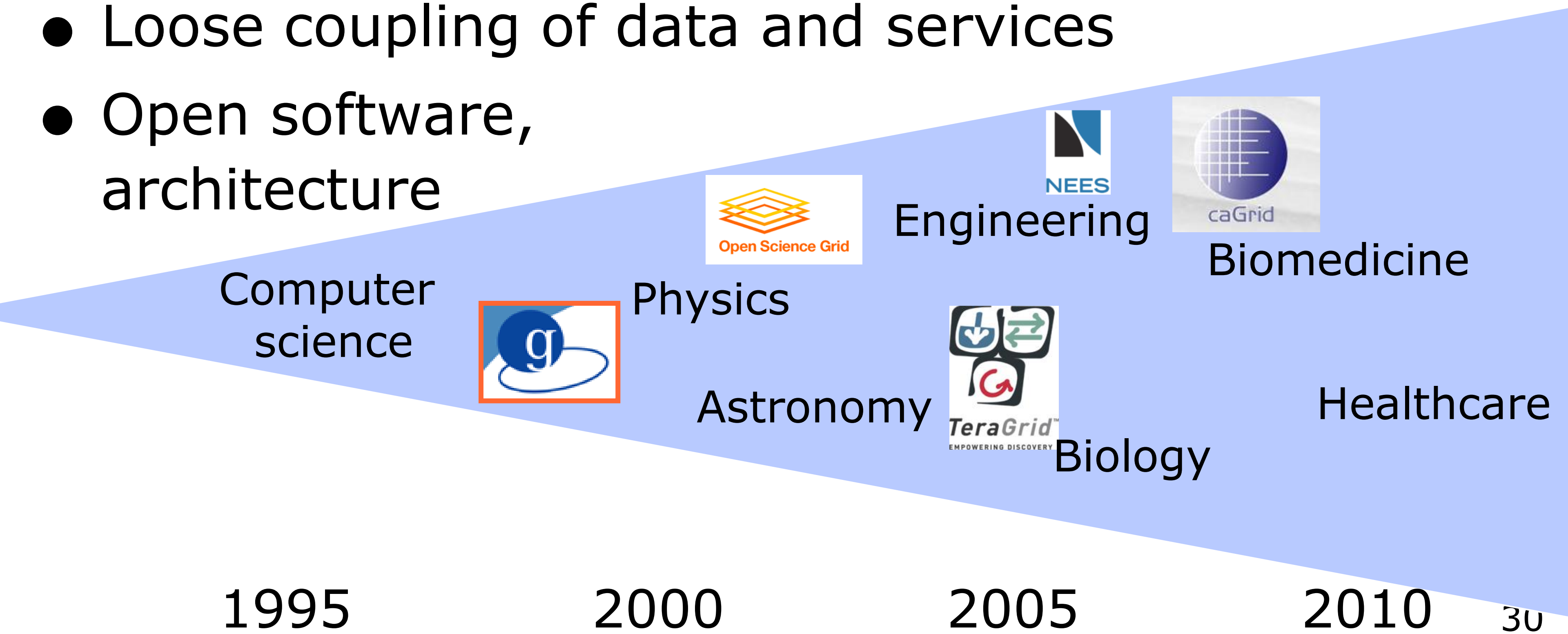


Ralph Stacey, *Complexity and Creativity in Organizations*, 1996

The Grid paradigm



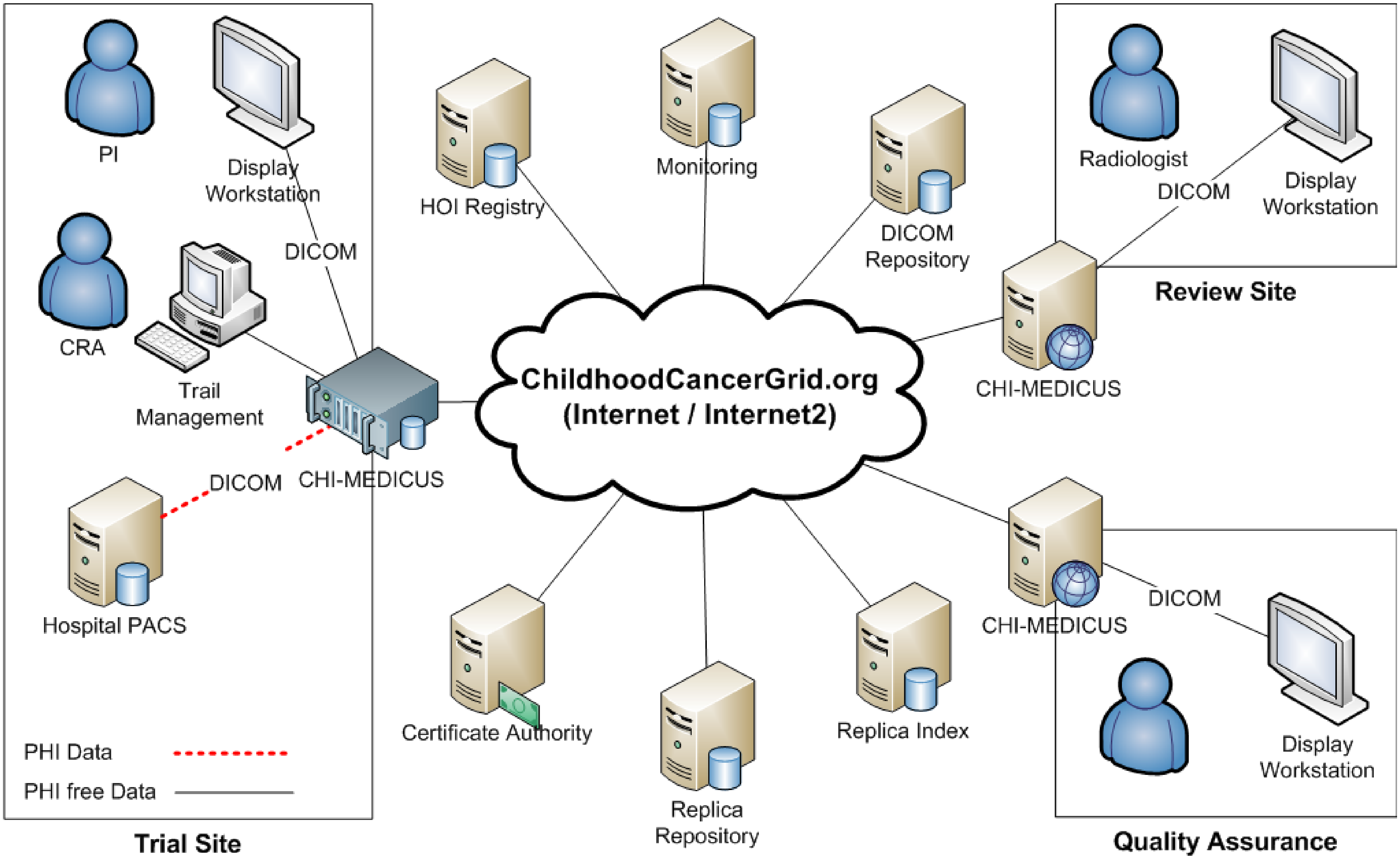
- Principles and mechanisms for dynamic virtual organizations
- Leverage service oriented architecture
- Loose coupling of data and services
- Open software, architecture



As of Oct 19, 2008: → 122 participants

105 services → { 70 data, 35 analytical }

Multi-center clinical cancer trials image capture and review

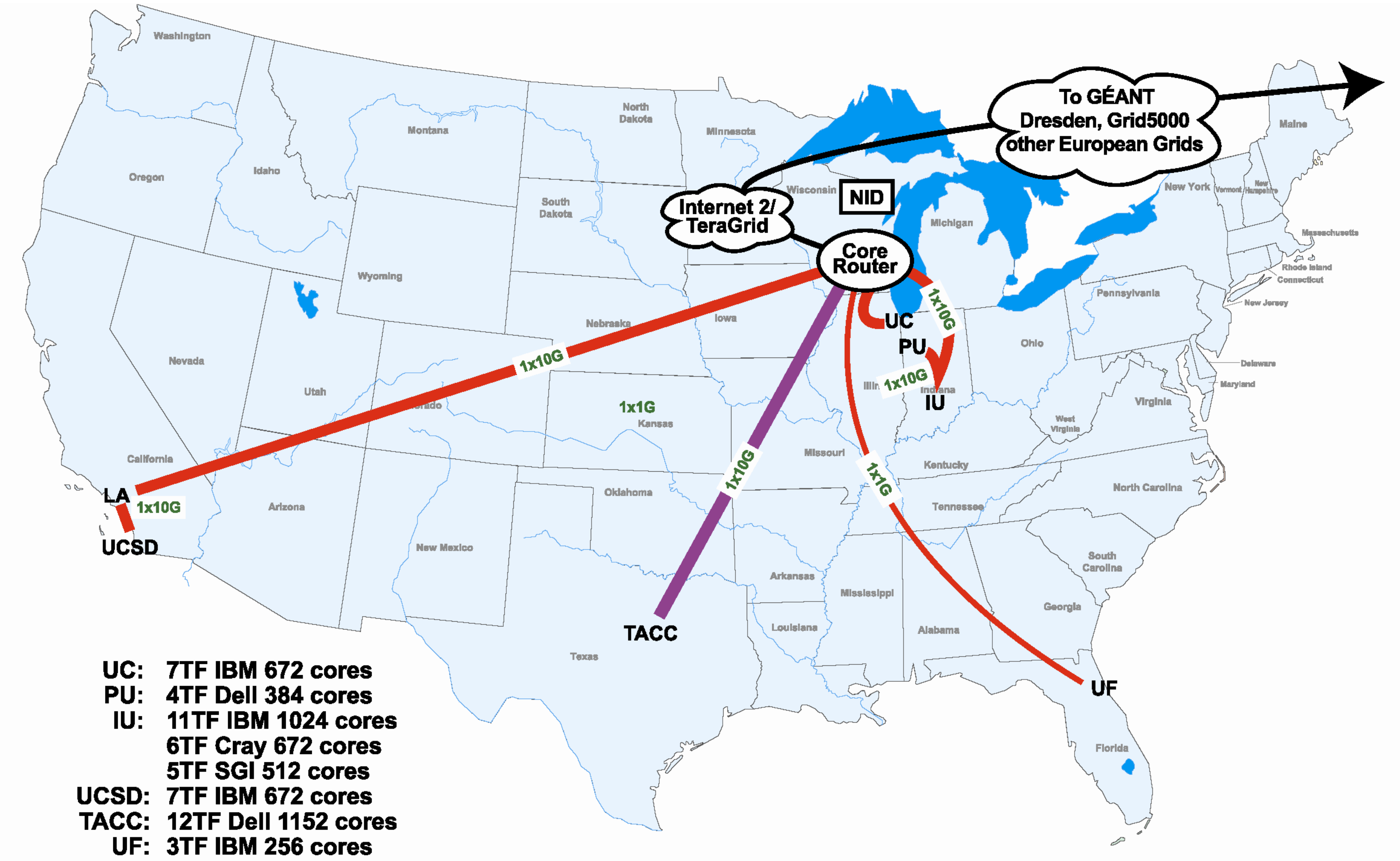


(Center for Health Informatics)

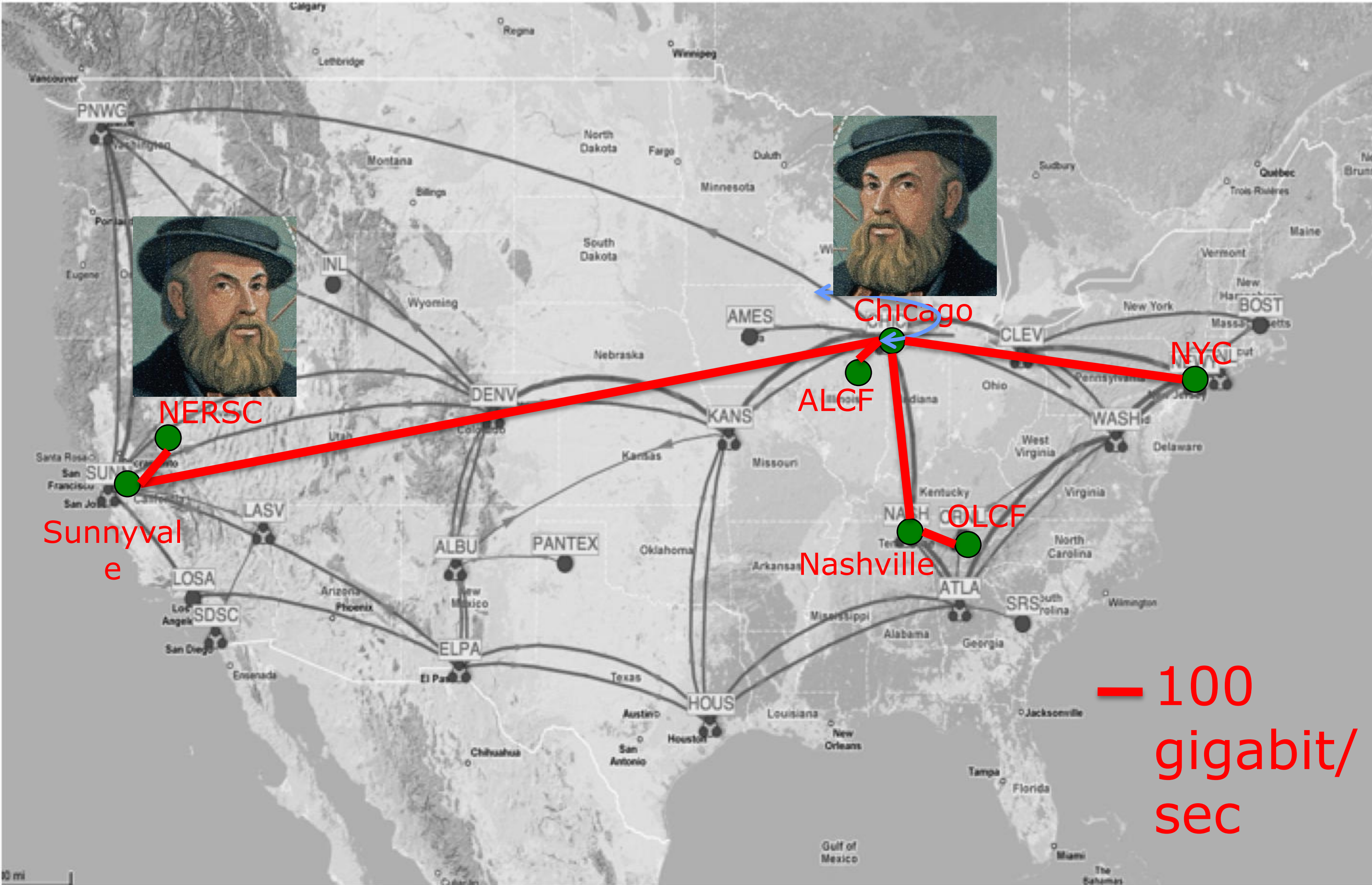
What is FutureGrid?

- *An HPC experimental infrastructure to support the research on the future of distributed, grid, and cloud computing.*
- NSF-funded project (2009)
  - 4 years @ ~$15M ($10M+$5M)
- Proposal team:
  - PIs: Geoffrey Fox (IU), Jose Fortes (UFL), Andrew Grimshaw (UV), Kate Keahey (UC/ANL), Warren Smith (TACC)
  - Funded partners: SDSC, University of Southern California, University of Tennessee, University of Virginia
  - Unfunded partners: Purdue, Technische Universtität Dresden

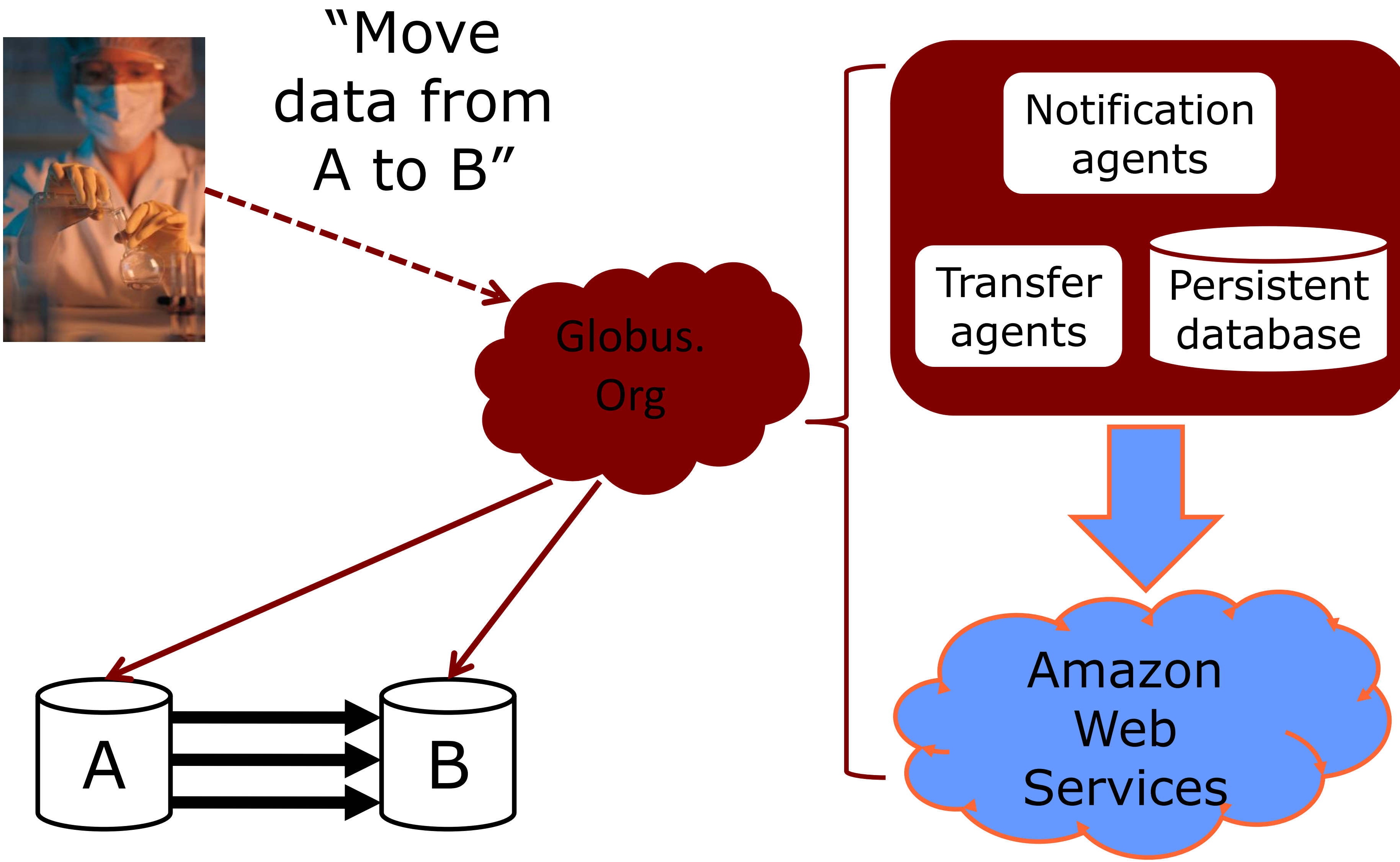
FutureGrid Hardware



Magellan + DOE's Advanced Network Initiative



Globus.Org services: Data replication



Summary

- Grid = federation, cloud = hosting
- Globus.org: hosting grid services on cloud
- Parallel scripting as a tool for rapid creation of computationally demanding applications
- Intrepid + PADS + Magellan + cloud …

Thank you!

Computation Institute

www.ci.uchicago.edu