Grids 4 Physics 101
An Overview of HEP related Grid Projects

Julian Bunn
Caltech
(with special thanks to Harvey Newman)
Heavy Books

This talk based on Chapter 39 "Data-Intensive Grids for High-Energy Physics", by Julian Bunn and Harvey Newman
HEP Computing Challenges: the need for Globally Distributed Data and Computing Systems
Computing Challenges: Petabytes, Petaflops, Global VOs

- Geographical dispersion: of people and resources
- Complexity: the detector and the LHC environment
- Scale: Tens of Petabytes per year of data

Major challenges associated with:
- Communication and collaboration at a distance
- Managing globally distributed computing & data resources
- Cooperative software development and physics analysis
- New Forms of Distributed Systems: Data Grids
LHC: Higgs Decay into 4 muons (Tracker only); 1000X LEP Data Rate

$10^9$ events/sec, selectivity: 1 in $10^{13}$ (1 person in 1000 world populations)
HEP Computing Challenges

• Rapid access to Exabyte data stores
• Secure, efficient, managed access to worldwide computing resources
• Tracking state and usage patterns
• Matching resources to demands
• Enable physicists in all world regions to collaborate on analysis
• Building high speed intercontinental networks to support the above
LHC Computing Planning ~1995

• In the mid-1990s, LHC Experiments began work on **Computing Models**

• "**The Grid**" was embryonic, and certainly not recognized by the experiments as being of significance

• Several projects were created to address the perceived challenges of building a **global computing system for LHC**

• The **Computing Technical Proposals** identified the need for R&D on the global systems
• Objects that cannot be recomputed (e.g. raw data) must be stored
• Physicists at any CMS institute should be able to easily work with the objects
• Resources should be used efficiently
• “... a geographically distributed federation of databases and heterogeneous computing platforms ... replication ... computing on demand ...”
• Thus a globally distributed computing and data serving system was already firmly on the table at this time
• Several R&D projects were begun to investigate: GIOD, MONARC, ALDAP etc.
The GIOD Project – 1997-1999

- Up to 240 clients reading simple objects from the database
- > 170 MB/sec writing LHC raw event data to the database
The MONARC Project

- Set up in 1998 to model/study the LHC Computing Models
- MONARC specified baseline Models representing cost-effective solutions to LHC Computing. Best candidate is Tiered System.
- LHC computing has a new scale and level of complexity.
  - A Regional Centre hierarchy of networked centres appears to be the most promising solution.
- A powerful simulation system and toolset was developed
- Synergy with other advanced R&D projects was identified and profited from.
Data Grid Hierarchy

CERN/Outside Resource Ratio ~1:2
Tier0/(Σ Tier1)/(Σ Tier2) ~1:1:1

Tier 0 +1
CERN 700k SI95
~1 PB Disk;
Tape Robot

Tier 1
~2.5 Gbps
IN2P3 Center
RAL Center
INFN Center

Tier 2
~2.5 Gbps
Tier2 Center

Tier 3
~2.5 Gbps
Institute
~0.25 TIPS

Tier 4
Physicists work on analysis “channels”
Each institute has ~10 physicists working on one or more channels
ALDAP – “Accessing Large Data archives in Astronomy and Particle physics”

- **Funded by NSF in 1999 for three years. Caltech + JHU + FNAL + Microsoft**
- **Purpose:** Investigate data organization and architecture issues for particle physics and astronomy
- **Findings:** To optimize for speed and flexibility, there needs to be a compromise between ordered, sequential storage and anarchic, random arrangements. Multidimensional clustering a useful paradigm to adopt.
- **ALDAP collaboration yielded several unforeseen benefits:** led to further large collaborations (GriPhyN, iVDGL, NVO), and fostered much better mutual understanding of data storage and analysis challenges between particle physicists and astronomers

---

**MBps vs Disk Config**

- mbps speed avg.
- mssql
- linear quantum
- 64bit/33MHz pci bus
- added 2nd ctr
- added 4th ctr
- 1 disk controller saturates
- 1 PCI bus saturates
- SQL saturates CPU

---

**SC2001 “Bandwidth Challenge” (1)**

**Bandwidth Greedy Grid-enabled Object Collection Analysis for Particle Physics**

Julian Bars, Ian Fisk, Keen Hoffmann, Harvey Newman, James Patton

The subject of this work is to design and implement a distributed grid-enabled object collection system for particle physics data storage. The grid infrastructure consists of several major components: the object collection system itself, the data management system, and the grid interface. The object collection system is designed to efficiently manage and distribute large datasets across multiple nodes in a distributed computing environment.

---

- **Using local Grid-enabled object collection system for particle physics data storage.**
- **Data management system handles large datasets across multiple nodes.**
- **Grid interface provides a scalable solution for accessing and analyzing particle physics data.**
- **Full event data, reduced to the bare essentials.**
The Grid Projects Era begins...
Scientific Grids

- From ~1999, we have seen the emergence of a plethora of Grid projects targeted at “big science”

- HEP is in the vanguard of these efforts, building on the solid foundations put in place by the LHC experiments’ CTPs, GIOD, MONARC, other earlier projects, and the pioneering work of running experiments like BaBar ...
PPDG – “Particle Physics Data Grid”

- **Funded by DoE in 1999.**
- Integrated Systems Used by Experiments – usually Experiment groups generate sub projects. PPDG resources provide liaisons to rest of experiment (>500 people)
- Grid-extend “mid-life” analysis systems – STAR, BaBar, D0 – adiabatically incorporate new functionality without destabilizing the systems. Constrains interface and component definition.
- For “young” experiments - CMS, Atlas - include grid services as soon as possible - production and analysis of simulated data a core activity which is needed throughout the experiment lifetime.
- Experiment/CS teams extend & develop grid software services as a result of project consensus.
- Robustness, fault tolerance, production use mean resources allocated for integration and operational deployment.

Showing the collaboration links between PPDG and the experiments and user communities
PPDG Goals

- **Integrated Systems Used by Experiments** – usually Experiment groups generate sub projects. PPDG resources provide liaisons to rest of experiment (>500 people).

- Grid-extend “mid-life” analysis systems – STAR, BaBar, D0 – adiabatically incorporate new functionality without destabilizing the systems. Constrains interface and component definition.

- For “young” experiments - CMS, Atlas - include grid services as soon as possible - production and analysis of simulated data a core activity which is needed throughout the experiment lifetime.

- Experiment/CS teams extend & develop grid software services as a result of project consensus.

- Robustness, fault tolerance, production use mean resources allocated for integration and operational deployment.
PPDG - Robust File Replication and Information Services

- Robust and performant file transfer (GridFTP, bbFTP)

- Replica catalogs (Globus LDAP, SRB)

- Replica management services (Globus, SRB)

- Intelligent replica management (historical and predictive)

- Production deployment support:
  - retrial, error recovery and reporting, failover

- Portal interface and data definitions (SRB, JLAB)
GriPhyN – “Grid Physics Network”

- Funded by NSF in 1999
- Collaboration of Computer Scientists and Physicists from:
  - ATLAS
  - CMS
  - LIGO
  - SDSS
- $11.9M (NSF) + $1.6M (matching) 17 universities, SDSC, 3 labs, ~80 people
- Integrated Outreach effort
- Focussing on creation of “Petascale Virtual Data Grids”
- Concept of “Virtual Data”: data whose properties are specified, and which may or may not exist in the Grid. Once the data are dereferenced, they are either fetched or created then fetched.

Dereferencing Virtual Data is a complex problem for the data Grid, because it involves
  - Local and global resource management
  - Policies
  - Security
  - Strategic decisions

A production Grid, as envisaged by GriPhyN, showing the strong integration of data generation, storage, computing and network facilities, together with tools for scheduling, management and security.
iVDGL – International Virtual Data Grid Laboratory

• “Part 2” of GriPhyN project
  – Much more application oriented than GriPhyN
  – Creation of computing Infrastructure
  – $15M, 5 years @ $3M per year
  – CMS + ATLAS + LIGO + SDSS/NVO + Computer Science

• Scope
  – Deploy Grid laboratory with international partners
  – Acquire Tier2 hardware, Tier2 support personnel
  – Integrate Grid software into applications
  – CS support teams to harden tools
  – Establish International Grid Operations Center (iGOC)
  – Deploy hardware at 3 minority institutions (Tier3)
iVDGL as a Laboratory

- **Grid Exercises**
  - “Easy”, intra-experiment tests first (10-30%, national, transatlantic)
  - “Harder” wide-scale tests later (30-100% of all resources)
  - CMS is already conducting transcontinental simulation productions

- **Operation as a facility**
  - Common software, central installation to ensure compatibility
  - CS teams to “harden” tools, support applications
  - iGOC to monitor performance, handle problems
iVDGL Participants

- U Florida
- Caltech
- UC San Diego
- Indiana U
- Boston U
- U Wisconsin, Milwaukee
- Penn State
- Johns Hopkins
- U Chicago
- U Southern California
- U Wisconsin, Madison
- Salish Kootenai
- Hampton U
- U Texas, Brownsville
- Fermilab
- Brookhaven
- Argonne Lab

CMS
CMS, LIGO
CMS, CS
ATLAS, iGOC
ATLAS
LIGO
LIGO
SDSS, NVO
CS
CS
Outreach, LIGO
Outreach, ATLAS
Outreach, LIGO
CMS, SDSS, NVO
ATLAS
ATLAS, CS

T2/Software
CS support
T3/Outreach
T1/Labs
iVDGL Summary

• **Principal components**
  – Tier1 sites (laboratories)
  – Tier2 sites (universities and other institutes)
  – Selected Tier3 sites (universities)
  – Fast networks: US, Europe, transatlantic
  – International Grid Operations Center (iGOC)
  – Computer Science support teams
  – Coordination, management

• **Complementary EU project:** DataTAG
  – Transatlantic network from CERN to StarLight (+ people)
  – Initially 2.5 Gb/s
Coordinating U.S. Projects: Trillium

- **Trillium**: GriPhyN + iVDGL + PPDG
  - Large overlap in project leadership & participants
  - Large overlap in experiments, particularly LHC
  - Joint projects (monitoring, etc.)
  - Common packaging, use of VDT & other GriPhyN software

- **Organization from the “bottom up”**
  - With encouragement from funding agencies NSF & DOE

- **DOE (OS) & NSF (MPS/CISE) working together**
  - DOE (labs), NSF (universities)
  - Collaboration of computer science/physics/astronomy encouraged
  - Collaboration strengthens outreach efforts
• **Funded by the European Union in 2001, for three years.**

• **Led by CERN, with**
  – European Space Agency
  – Centre for Scientific Research (CNRS, France)
  – National Institute of Physics (INFN, Italy)
  – Institute for HENP (NIKHEF, Holland)
  – Particle Physics and Astronomy Research Council (PPARC, United Kingdom)
  – 15 other less major parties

• **Purpose: set up a computational and data-intensive Grid of resources for analysis of scientific data**

• **Develop and test infrastructure for scientific “collaboratories”: sharing equipment, data and software**

Showing the structure of the EU DataGrid Project, and its component Work Packages
LCG – LHC Computing Grid

- Funded in 2002 for three years by the European Union.
- Managed by CERN.
- Goal: prepare the computing infrastructure for the simulation, processing and analysis of LHC data for all four experiments
- Application support
- Development and prototyping of computing services
- Computing Data Challenges (priority)
- Physics Data Challenges
- Second phase of Project (2006 to 2008) will oversee construction and operation of the initial LHC computing system

The organizational structure of the LHC Computing Grid, showing links to external projects and industry.
CrossGrid

- Funded in 2001 by the European Union. Eleven “Other” European countries
- Goal: Implement and exploit new Grid components for compute and data intensive science applications in the fields of:
  - Surgery
  - HEP
  - Flooding
  - Pollution
- Close collaboration with GGF and EU DataGrid (ensure interoperability)
Example Architectures and Applications
NSF TeraGrid
Caltech CMS TeraGrid Prototype

Grid Infrastructure for Caltech CMS Production on Alliance Resources

Caltech
- DAGMan
- CondorG

NCSA
- NCSA Linux Cluster
- UniTree

Wisconsin
- Condor Pool
- CMSIM
**GRAPPA (Atlas)**

- **GRAPPA** – “Grid Access Portal for Physics Applications”
- Access using Grid certificates
- **Active Notebooks describing tasks**
- **Integrated with the GriPhyN VDT**
# SAM (Do)

- **SAM** – “Sequential Access via Metadata”
- **Extract physics results from ~1PByte of measured and simulated data**
- **Being updated and “Gridified”**

<table>
<thead>
<tr>
<th>Client Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
</tr>
<tr>
<td>Command line</td>
</tr>
<tr>
<td>D0 Framework C++ codes</td>
</tr>
<tr>
<td>Python codes, Java codes</td>
</tr>
</tbody>
</table>

**Collective Services**

- Request Formulator and Planner
- Request Manager
- Cache Manager
- Job Manager
- Storage Manager

**“Dataset Editor”**

- SAM Resource Management

**“Project Master”**

- Batch Systems - LSF, PBS, Condor

**“Station Master”**

- Job Manager

**“Station Master”**

- Data Mover

**“File Storage Server”**

- “Stager”

**“Optimiser”**

- Significant Event Logger
- Naming Service
- Catalog Manager
- Database Manager

**Connectivity and Resource**

- CORBA
- UDP
- Catalog protocols
- File transfer protocols - ftp, bbftp, rcp
- GridFTP
- Mass Storage systems protocols e.g. encp, hpss

**Authentication and Security**

- SAM-specific user, group, node, station registration
- GSI
- Bbfip 'cookie'

**Fabric**

- Tape Storage Elements
- Disk Storage Elements
- Compute Elements
- LANs and WANs
- Code Repository
- Resource and Services Catalog
- Replica Catalog
- Meta-data Catalog

*Indicates component that will be replaced, enhanced or added using PPDG and Grid tools*

*Name in “quotes” is SAM-given software component name*
Clarens

- **The Clarens Remote Dataserver:** a WAN system for authenticated remote data analysis
  - Using the GriPhyN VDT
  - Several clients (ROOT, Objy)
- Clarens servers are deployed at Caltech, Florida, UCLA, FNAL
- SRB now installed as Clarens service on Caltech Tier2 (Oracle backend)
Web Services for Physics Analysis

- **Web Services**: easy, flexible, platform-independent access to data (Object Collections in Databases)
  - Well-adapted to use by individual physicists, teachers & students

- **SkyQuery Example** JHU/FNAL/Caltech with Web Service based access to astronomy surveys

- Can be individually or simultaneously queried via Web interface

- Simplicity of interface hides considerable server power (from stored procedures etc.)

- This is a “Traditional” Web Service, with no user authentication required
Simple HEP “Tag” Web Services

- ~180,000 Tag objects derived from di-jet ORCA events

- Each Tag:
  - Run & event no.
  - OID of ORCA event object
  - E,Phi,Theta,ID for 5 most energetic particles
  - E,Phi,Theta for 5 most energetic jets

- Tags are loaded into SQLServer (Caltech) and Oracle (CERN) databases

- Simple Web Service allows query, fetch etc. of Tags

- Web Service access can be from a Browser or other program.

- We added the necessary code to COJAC so that it could fetch the Tag events via the Web Services, rather than via Objectivity
COJAC: CMS ORCA Java Analysis Component: Java3D Objectivity JNI Web Services

- Object Collections
- Multiplatform, Light Client
- Interface to OODBMS

Demonstrated Caltech-Rio de Janeiro and Chile in 2002
Distributed System Services Architecture (DSSA): CIT/Romania/Pakistan

- Agents: Autonomous, Auto-discovering, self-organizing, collaborative
- “Station Servers” (static) host mobile “Dynamic Services”
- Servers interconnect dynamically; form a robust fabric in which mobile agents travel, with a payload of (analysis) tasks
- Adaptable to Web services: JINI → OGSA; and many platforms
- Adaptable to Ubiquitous, mobile working environments

Managing Global Systems of Increasing Scope and Complexity, In the Service of Science and Society, Requires A New Generation of Scalable, Autonomous, Artificially Intelligent Software Systems
MONARC SONN: 3 Regional Centres “Learning” to Export Jobs (Day 9)

- **CERN**: 30 CPUs
  - 1MB/s; 150 ms RTT
  - Simulations for Strategy Development; Self-Learning Algorithms for Optimization Are Key Elements of Globally Scalable Grids

- **CALTECH**: 25 CPUs
  - 1.2 MB/s; 150 ms RTT
  - \( \langle E \rangle = 0.83 \)

- **NUST**: 20 CPUs
  - 0.8 MB/s; 200 ms RTT
  - \( \langle E \rangle = 0.73 \)
  - \( \langle E \rangle = 0.66 \)

**Optimized:** Day = 9
**MonaLisa: A Globally Scalable Grid Monitoring System**

By I. Legrand (Caltech)

- Deployed on US CMS Grid
- Agent-based Dynamic information / resource discovery mechanism
- Talks w/Other Mon. Systems
- Implemented in
  - Java/Jini; SNMP
  - WDSL / SOAP with UDDI
- Part of a Global “Grid Control Room” Service
14308 Host Devices;
7664 Registered Users
in 64 Countries
45 Network Servers
Annual Growth 2 to 3X
Inter-Grid Coordination
HICB – HEP InterGrid Coordination Body

• Coordinates between
  – DataGrid
  – GriPhyN
  – iVDGL
  – PPDG
  – Other European Grid projects in UK, Italy, France...

• Deploy consistent, open-source standards-based global Grid infrastructure

• Main thrust: Ensure compatibility and interoperability
GLUE – Grid Laboratory Uniform Environment

- Sponsored by the HICB
- Focussing on interoperability between US and European Grid projects
- Define, assemble and test common core software components of the Grid middleware: the GLUE Suite
- Ensure that US and EU Grids can be configured to interoperate
GGF – Global Grid Forum

• An international consortium with a wide membership including most proponents of Grid technology

• Divided into Working Groups tasked with investigating:
  – Distributed Systems
  – Grid software implementation
  – Collaborative Environments
  – Remote Data Access
  – Etc. etc.
DataTag

- Funded in 2001 by the European Union for 3 years
- 9.8 M Euros: 90% for middleware and applications (HEP, EO and biology)
- Possible extensions (time and funds) on the basis of first successful results
  - DataTAG (2002-2003)
- Goals:
  - End to end Gigabit Ethernet performance using innovative high performance transport protocol stacks
  - Assess, develop & demonstrate inter-domain QoS and bandwidth reservation techniques.
  - Interoperability between major GRID projects middleware/testbeds in Europe and North America.
    - DataGrid, CrossGrid, possibly other EU funded Grid projects
    - PPDG, GriPhyN, IVDGL, Teragrid (USA)
    - LCG (LHC Computing Grid)

- CERN – International (Switzerland/France)
- CNRS - France
- ESA/ESRIN – International (Italy)
- INFN - Italy
- NIKHEF – The Netherlands
- PPARC - UK
DataTag Connectivity

Major 2.5/10 Gbps circuits between Europe & USA
DataTag R&D

• **Grid related network research (WP2, WP3):**
  – 2.5 Gbps transatlantic lambda between CERN (Geneva) and StarLight (Chicago) (WP1)
  – Dedicated to research (no production traffic)
  – Very unique multi-vendor testbed with layer 2 and layer 3 capabilities
    • in effect, a distributed transatlantic Internet Exchange Point.

• **Interoperability between European and US Grids (WP4)**
  – Middleware integration and coexistence - Applications
  – GLUE integration/standardization
    • GLUE testbed and demo
New and Recent Grid Initiatives
An Inter-Regional Center for High Energy Physics Research and Educational Outreach (CHEPREO) at Florida International University

- E/O Center in Miami area
- iVDGL Grid Activities
- CMS Research
- AMPATH network
- Int’l Activities (Brazil, etc.)

Status:
- Proposal submitted Dec. 2002
- Presented to NSF review panel Jan. 7-8, 2003
- Looks very positive
CAIGEE CMS Analysis – an Integrated Grid Enabled Environment

Plug-in Architecture based on Web Services

- Expose Grid Views of the “Global System” to physicists – at various levels of detail, with Feedback
- Supports Data Requests, Preparation, Production, Movement, Analysis of Physics Object Collections
- Initial Target: US-CMS physicists in California (CIT, UCSD, Riverside, Davis, UCLA)
- Expand to Include FIU, UF, FSU, UERJ
- Future: Whole US CMS then CMS
CAIGEE Architecture

Laptop

ROOT

Browser

PDA

Peer Group

Super Peer Group

Desktop

Clarens

Grid Services
Web Server

Catalog

Virtual Data
Catalog

Materialized
Data Catalog

Monitoring

Grid Processes

File Transfer

Planner 1

Planner 2

GDMP

Execution Priority
Manager

Grid Wide
Execution Service

CMS Apps

Others

From GriPhyn,
iVDGL, Globus etc

Caltech/CMS
Developments
GAE - Grid Analysis Environment

• The development of a physics analysis environment that integrates with the Grid systems, is where the real “Grid Challenge” lies
  – To be used by a large diverse community
  – 100s - 1000s of tasks with different technical demands
  – Needs priorities
  – Needs security

• How much automation is possible. How much is desirable?

• GAE is a key to “success” or “failure” for physics Grid applications:
  – Where the physics gets done
  – Where the Grid End-to-End Services and the Grid Application Software Layers get built
  – Where we learn how to collaborate remotely to do physics
High Speed Grids Already Here:
2 TBytes/day transfers

S.Ravot et al on the DataTAG
TCP/IP Getting Faster

FAST (Caltech): A Scalable, “Fair” Protocol for Next-Generation Networks: from 0.1 To 100 Gbps

Highlights of FAST TCP
- Standard Packet Size
- 940 Mbps single flow/GE card
- 9.4 petabit-m/sec 1.9 times LSR
- 9.4 Gbps with 10 flows
- 37.0 petabit-m/sec 6.9 times LSR
- 22 TB in 6 hours; in 10 flows

Implementation
- Sender-side modification
- Delay (RTT) based
- Stabilized Vegas

Internet: distributed feedback system

TCP

R_f(s)

AQM

R_p(s)

Next: 1 GB/sec disk to disk

URL: netlab.caltech.edu/FAST

C. Jin, D. Wei, S. Low
FAST Team & Partners
GAE Demonstrations at iGrid2002

Amsterdam Science & Technology
CERN (Geneva)
Amsterdam, The Netherlands
23-26 September 2002

iGrid 2002: the 3rd biannual International Grid applications-driven testbed event, challenges scientists and technologists to utilize multi-geography experimental optical networks, with special emphases on e-Science, LambdaGrid and Virtual Laboratory applications. The result is an impressive, coordinated effort by 28 teams representing 15 countries, showcasing how extreme networks, combined with application advancements and middleware innovations, can advance scientific research.

As computational sciences strive to better understand very complex systems — whether biological, environmental, atmospheric, geological or physics, from the atoms to the universe level, in both time and space, they will require fundamentally different computing architectures for tackling the computational requirements. Indeed, many important scientific disciplines require calculations far more complex than today’s largest parallel computers, which process ten trillion floating point operations per second (10 teraflops). An exascale is a billion gigabytes of storage, and today’s networks will eventually transmit data at one trillion bits per second. Some 20 million terabytes span the Internet connection.

Recent, major technological and cost breakthroughs in networking technology have...
“GECSR” - NSF Medium ITR 2003

“We propose to develop, prototype, test and deploy the first Grid-Enabled Collaboratory for Scientific Research (GECSR) on a global scale. A distinguishing feature of this proposal is the tight integration between the science of collaboratories, a globally scalable working environment built on the foundation of a powerful fully functional set of working collaborative tools, and an agent-based monitoring and decision-support system that will allow collaborating scientists to perform data intensive analysis tasks efficiently.”

Caltech, Maryland, Michigan and others

From the NSF Blue-Ribbon Advisor Panel on Cyberinfrastructure
Globally Enabled Analysis Communities: NSF “Large” ITR 2003

Develop and build Dynamic Workspaces

Build Private Grids to support scientific analysis communities
   Using Agent Based Peer-to-peer Web Services

Construct Autonomous Communities Operating Within Global Collaborations

Empower small groups of scientists (Teachers and Students) to profit from and contribute to international Big Science

Drive the democratization of science via the deployment of new technologies
Private Grids and P2P Sub-Communities in Global CMS

The Experiment Enterprise

- Individual Research
- Shared Resources
- Higgs Working Group
- Supersymmetry Working Group
- Trigger Studies Work Group
- New Phenomena Work Group

Pull In Outside Resources

- Physicist
- Computer
- Desktop
- Schema
- Storage
- Software
Grid projects have been a step forward for HEP and LHC: a path to meet the “LHC Computing” challenges.

The original Computational and Data Grid concepts are largely stateless, open systems: known to be scalable.

Analogous to the Web.

The classical Grid architecture has a number of implicit assumptions:
- The ability to locate and schedule suitable resources, within a tolerably short time (i.e. resource richness)
- Short transactions; Relatively simple failure modes

HEP Grids are data-intensive and resource constrained:
- Long transactions; some long queues
- Schedule conflicts; policy decisions; task redirection
- A Lot of global system state to be monitored + tracked
Conclusion: Future HEP Grid Architecture

Physics Reconstruction, Simulation and Analysis Code Layer

Experiments’ Software Framework Layer
- Modular and Grid-aware: Architecture able to interact effectively with the lower layers

Grid Applications Layer
- (Parameters and algorithms that govern system operations)
  - Policy and priority metrics
  - Workflow evaluation metrics
  - Task-site coupling proximity metrics

Global End-to-End System Services Layer
- (Mechanisms and services that govern long-term system operation)
  - Monitoring and Tracking component performance
  - Workflow monitoring and evaluation mechanisms
  - Error recovery and redirection mechanisms
  - System self-monitoring, evaluation and optimization mechanisms