Grid2003 and Open Science Grid

Ruth Pordes
Fermilab  (contributes facilities and infrastructure for CDF, D0, SDSS, U.S CMS, BTeV..)
U.S. CMS
Trillium: PPDG Coordinator, iVDGL Management,
April 16th 2004

Reporting on the work and contributions of many people... including in this room... Ian
Fisk, Alain Roy, Peter Couvares, Alan DeSmet, Brian Moe, Scott Koranda, Tim Thomas,
Jason Smith, Gaurang Mehta, Dan Bradley, Anne Heavey + Miron
Outline

Theme: End to End, Shared, Integrated and Running, Grid based Systems


Roadmap: “Our” Experiments, Physics Grid Projects, Open Science Grid

mid2004->2005: Plans: Grid-Development, Grid-4, Open Science Grid-0,

Subtext: Partnerships in the U.S
Participation and collaboration with Europe
Partnership with Condor Project in many dimensions
Grid2003 - Project,
Grid3 - Environment,
Grid3+ - ad-hoc Continuation

www.ivdgl.org/grid2003
A Joint Project to Build a Significantly Scaled Shared Grid giving Science Benefit across Experiments, Computer Science Groups, Grid Projects, DOE, NSF

Grid2003 Project

- U.S. CMS S&C
- PPDG
- Trillium
- UofBuffalo
- iVDGL
- GriPhyN
- Grid Telemetry
- Virtual Data Toolkit (VDT)
- SDSS
- LIGO
- BTeV
- U.S. ATLAS Computing
- BTeV
Model and Method

End to End Usable and Supportable Grid Systems => Applications + Grid Infrastructure + Sites

- Particle and Nuclear Physics, Astrophysics, Gravitational Science, Computer Science
- The Applications Drive System Requirements and Implementations including Capabilities, Performance, Scale, Schedule and Operations
- ~10 applications all use Globus, Condor-G, some DAGMan/Virtual Data/Pegasus

Challenge in the # autonomous sites and teams, sustainability and use; Increase chance of success by:

- Matching Requirements<->Deliverables;
- Minimizing impact on existing Sites and Applications;
- Planning for Heterogeneity, Multiplicity, Opportunity, Dynamism;
- Incremental milestones involving development, integration and deployment, operation.

Risk Reduction with some expense in efficiency, overhead and overall goals

- Use existing single organization testbeds and pre-existing sites, support 4 batch systems.
- Incremental addition of sites, services, applications; Component & system tests & procedures.
- VDT releases & homogeneous grid middleware for core grid services + Pacman packaging / distribution + some Services responsibility of the Applications + Central Operations Center
- Minimal necessary metrics for success. New development restricted to necessary “glue” and “integration” components.
- All stakeholders participated in planning, execution and management.
Paradyne and Condor Week

Who

23 institutes

Argonne National Laboratory
Jerry Gieraltowski, Scott Gose, Natalia Maltsev, Ed May, Alex Rodriguez, Dinanath Sulakhe

Boston University
Jim Shank, Saúl Youssef

Brookhaven National Laboratory
David Adams, Rich Baker, Wensheng Deng, Jason Smith, Dantong Yu

Caltech
Iosif Legrand, Suresh Singh, Conrad Steenberg, Yang Xia

Fermi National Accelerator Laboratory

Hampton University
Keith Baker, Lawrence Sorrillo

Harvard University
John Huth

Indiana University
Matt Allen, Leigh Grundhoefer, John Hicks, Fred Luehring, Steve Peck, Rob Quick, Stephen Simms

Johns Hopkins University
George Fekete, Jan vandenBerg

Kyunpook National University / KISTI
Kihyeon Cho, Khwan Kwon, Dongchul Son, Hyeongwoo Park

Lawrence Berkeley National Laboratory
Shane Canon, Jason Lee, Doug Olson, Iowa Sakrejda, Brian Tierney

University at Buffalo
Mark Green, Russ Miller

University of California San Diego
James Leffts, Terrence Martin

University of Chicago
David Bury, Catalin Dumitrescu, Daniel Engh, Ian Foster, Robert Gardner*, Marco Mambelli, Yuri Smirnov, Jens Voeckler, Mike Wilde, Yong Zhao, Xin Zhao

University of Florida
Paul Avery, Richard Cavanaugh, Bockjoo Kim, Craig Prescott, Jorge L. Rodriguez, Andrew Zahn

University of Michigan
Shawn McKee

University of New Mexico
Christopher T. Jordan, James E. Prewett, Timothy L. Thomas

University of Oklahoma
Horst Severini

University of Southern California
Ben Clifford, Ewa Deelman, Larry Flon, Carl Kesselman, Gaurang Mehta, Nosa Olomu, Karan Vahi

University of Texas, Arlington
Kaushik De, Patrick McGuigan, Mark Sosebee

University of Wisconsin-Madison
Dan Bradley, Peter Couvares, Alan De Smet, Carey Kireyev, Erik Pauly, Alain Roy

University of Wisconsin-Milwaukee
Scott Koranda, Brian Moe

Vanderbilt University
Bobby Brown, Paul Sheldon

*Contact authors

~60 people working directly: 8 full time, 10 half time, 20 site admins _ time
Met Metrics in Nov 2003 since Nov ~ 6 months

Continuous use with mainly Application Effort

<table>
<thead>
<tr>
<th>Metric</th>
<th>Target</th>
<th>Grid2003 &quot;SC2003&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of CPUs</td>
<td>400</td>
<td>2762 (27 sites)</td>
</tr>
<tr>
<td>Number of users</td>
<td>&gt; 10</td>
<td>102 (16)</td>
</tr>
<tr>
<td>Number of Applications</td>
<td>&gt; 4</td>
<td>10</td>
</tr>
<tr>
<td>Number of site running concurrent applications</td>
<td>&gt; 10</td>
<td>17</td>
</tr>
<tr>
<td>Peak number of concurrent jobs</td>
<td>1000</td>
<td>1100</td>
</tr>
<tr>
<td>Data Transfer per day</td>
<td>&gt; 2-3 TB</td>
<td>4.4 TB (11.12.03)</td>
</tr>
</tbody>
</table>

Infrastructure is stable

US CMS 50% increase in throughput through opportunistic computing

GADU and SnB biology applications added without impact or significant overhead

Job mix changes without impact
Buts and Ands..

Need factor of ~5 in complexity:
Many Services missing or parochial
Few & Expert Users
Insufficient services to sustain system

Not Production Quality - in many aspects
Technology
Security
Management
Operation

Need increased scale for LHC alone:
~ 5 in 3-4 years for LHC startup
~10 in 8 years for LHC analysis

Homogeneity - not the Model
Sites serve many grid as well as local communities
Grid Infrastructure too prescriptive
Applications across grids - international, logical

LHC Enable Science for <2000 scientists in <400 institutions through <30 countries - who have to agree as a single Collaboration in publishing discoveries and results.

BaBar, Run II, STAR, rely today on remote processing to meet montecarlo and analysis with requirements ~20-50% and 5-7 years ahead of LHC.

Natural ongoing partnerships between producers and consumers of Services, between Computer Science and Application Science Communities, -> Methods that promote commonality and generality of solutions.
Open Science Grid White Paper

The initial phase of the Open Science Grid is to federate the LHC physics applications' grid services and computing resources in the U.S. into a global grid system, engineered and managed to serve the needs of the LHC scientific program.

National laboratories and universities participating in the U.S. LHC software and computing efforts will form the initial sites with special roles for the U.S. LHC Tier-1 centers at BNL and Fermilab. In the next phase, applications from other physics communities, specifically RunII experiments at Fermilab, RHIC experiments at BNL, and BaBar at SLAC, will move their resources and applications to the Open Science Grid. Other experiments and communities, like our PPDG and iVDGL partners, will join and further extend the Open Science Grid. In subsequent phases, other non-physics science applications will be included.

Each application will bring dedicated computing resources to be federated with the Open Science Grid. The construction of the Open Science Grid requires the following work elements:

1. Management and technical oversight.
2. Engineering and quality assurance: to supply the architectural foundation and engineering for the entire system.
3. International coordination: to assure interoperability with non-U.S. grids (e.g., LHC Computing Grid).
4. Education and Outreach: to provide opportunities for students to witness and participate in building this emerging national grid infrastructure.
5. Application Integration: an iterative program of work to integrate each specific application’s grid services and computing fabric.
High Data Throughput

- Pile-up events are stored across the pools.
- dCache server can trigger pool-to-pool replication and balance the access load.

LHC Events

- All charged tracks with $pt > 2$ GeV
- Reconstructed tracks with $pt > 25$ GeV

(+30 minimum bias events)
Physics “Services”

Experiment’s “Services” go end-to-end!

Experiment Layer
- Frameworks, Grid shells, Portals, ...
- Experiments, Physics Grid Projects, ...

Application Middleware
- Pool, Object Servers, Higher Level Services, LCG AA, Experiments, GAE, Grid Projects,

Grid Middleware
- Workload Systems, Replica Managers, ...
- VDT, LCG/EGEE, Grid Projects, GGF, ...

Facilities and Fabrics
- Worker Nodes, Storage Managers, ...
- LCG, Regional Centers, Grid Projects, ...

Physics "Data Services"
- CMS Physics Finder
- POOL Collection Manager, Object Server, ...
- Replica System, Data Access Optimizer, ...
- Storage Managers, Disk Caches, ...

Physics "Tasks Manager"
- Atlas Physics Portal
- MC Executer, Physics Prioritizer, ...
- Grid Executer, Workload Optimizer, ...
- Worker Node Managers, Queues, ...

Lothar A T Bauer Dick Fermilab ARDA Workshop
The Components

Application Communities

Grid Infrastructure

Sites
Grid-Development, Grid-4, Open Science Grid-0
Manage Updates

Add Services for Data Management, Maintainability/Usability, Understanding Problems

Development, Integration, Production Cycles

Information and Heterogeneity

Context Switching
Persistency and Sharing of the Workspace

Dynamic resource requirements

Mix of “official” experiment software and private user code

Validations

Input datasets not necessarily known a-prior and Possibly very sparse data access pattern

Large number of people submitting jobs concurrently and in an uncoordinated fashion resulting into a chaotic workload

Need for interactivity - importance of latency as well as throughput
Initial Phase of planning and demonstrations between U.S. LHC host labs (Tier-1s), EGEE/LCG extending to Tier-2s

Define Services

Issues of Production and Multi-VO use of Persistent, Durable, Volatile Storage

Try out what partnering means in practice.
Absolutely need a sustained ubiquitous production infrastructure - just like the network.

See significant benefits from the growing partnerships with the Computer Science community.

See significant benefits from contributing to and collaborating with common grids with other sciences.

Have successfully taken small simple steps for usable national shared application grids.

Need ~x5 in scale and complexity for initial LHC and ongoing Run II+ data analysis.

Plan annual and expanding deliverables to get to the needed scales and capabilities.

Specific reliance on Condor Project collaborations especially:

- Virtual Data Toolkit,
- Evolving and extending capabilities of Condor, Condor-G, DAGMan, ... to meet application needs,
- End-to-end problem solving.