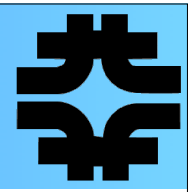


Grid2003 Activities

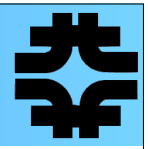
Ian Fisk

September 9, 2003

Grid Projects

Experiments

What is Grid2003?



Grid2003 is project to create a multi-VO Grid infrastructure in the US

- Collaborative effort of Trillium Grid Projects, US-ATLAS, US-CMS, LIGO and SDSS
- Natural extension of current testbed activities
- Pilot implementation of a persistent, national grid infrastructure.

Grid2003 is primarily an integration and deployment project

- The schedule and manpower do not permit a lot of development

Grid2003 Goals

Grid2003 is a project, done by a collaboration with well-defined stake holders

- contributions from the US ATLAS and US CMS S&C projects, iVDGL, GriphyN and PPDG

The purpose of the Grid3 is a project to build a grid environment to:

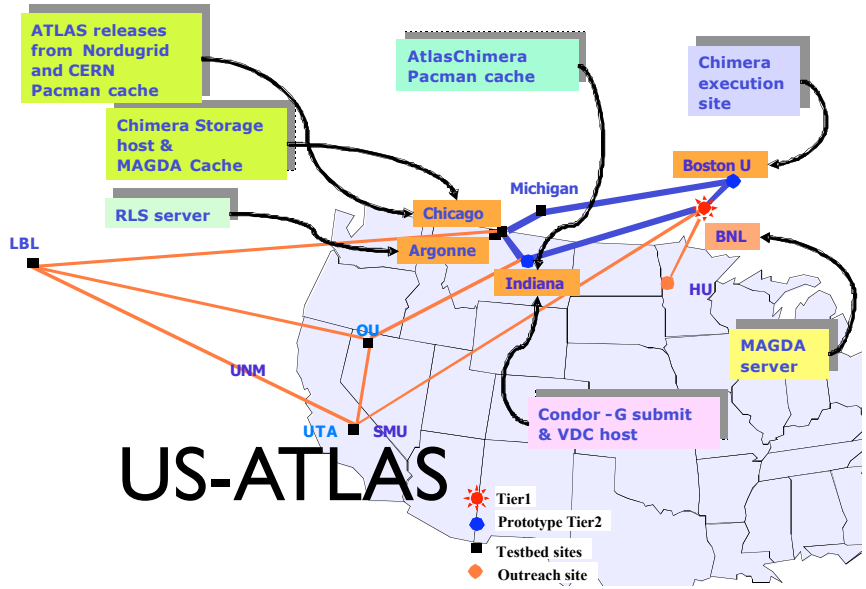
- Provide the next phase of the iVDGL Laboratory
- Provide the infrastructure and services needed for LHC production and analysis applications running at scale in a common grid environment
- Provide a platform for computer science technology demonstrators Provide a common grid environment for LIGO and SDSS applications, and possibly others

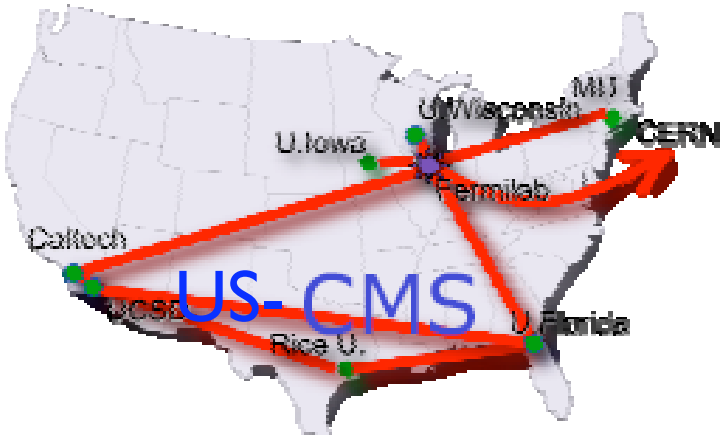
Building on Testbed Experience

US-ATLAS and US-CMS both have well developed testbed activities

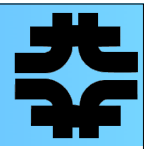
- Rolling prototypes for development
- Facilities at reasonable scales
- Successfully used for official simulated event production

Up to now these have been single VO installations





A "Federated" Grid2003



The Testbeds have been successful with a "federated" approach

- Local responsibility for facilities
  - Systems and support handled by local resources with well defined interfaces
- General Grid-wide services provided by some sites

Empower the sites

- Sites control the setup and management of systems, install and configure the middleware (generally thought VDT)
  - Publish attributes through information providers
- Sites control policy implementations

Empower the Grid

- Collaborative approach to enabling cross site services

Empower the Application Groups

- Application Groups responsible for working in the environment

Grid2003 Principles

Worked toward achieving a well defined set of metrics for success

- Definitions of functionality and scale

Define the minimum set of services needed on sites to support the VOs

- Simple services and common installations
  - Rely heavily on VDT
- Make as few requirements as possible on the sites themselves
  - Only put services on gateway systems
    - Separate privilege requiring installations from non-privileged
  - Should be no requirements except OS version on Worker Nodes

No components go into Grid2003 that haven't been evaluated and tested on individual VO testbeds

- Multi-VO environment and new scale is hard enough, without having to debug component functionality

Grid2003 Metrics

Metric (to be defined)	Target (to be defined)	Comments
Number of processors	up to 500	Possibly collected through MDS, archived and time-stamped
Data transferred per day	>2-3 TB	Data "flux" will need to be defined carefully.
Peak number of concurrent jobs	up to 450	Collect the total number of jobs running on Grid3, sorted by VO, archived so that time-dependent plots can be made.
Percentage of resources used	up to 90%	
Efficiency of job completion	up to 75%	Success to be defined. Other efficiency metrics could be identified.
Number of users	>10	Collect and sort by VO origin.
Number of different applications	>4	By an application "registry", such as proposed by the WorldGrid "project" mechanism, this could be simplified.
Number of sites running multiple applications	>10	Collect with time stamps, intervals to be determined.
Rate of Faults/Crashes	<1/hour	Measure, and perhaps categorize by degree of severity.
Operational Support Load of full demonstrator	<2 FTEs	How many people (and in which roles) were operating Grid3?

Grid2003 Packaging and Deployment

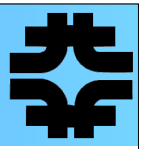
Grid2003 uses pacman for packaging and deployment.

- Currently the installation is broken into three area

1. Common Middleware
   - VDT Components, monitoring components not in VDT, signing policies not in VDT
2. Experiment specific "middleware"
   - Tools to populate configuration scheme, data server components, catalog components not in VDT
3. Experiment Applications
   - Since these generally do not require privileged access to install, experiments are being encouraged to install the applications dynamically at runtime.
     - US-CMS has switched to this method of running without incident
     -

Trying to keep as componentized as possible in section 1. Working to keep number 2 compatible to many VDT versions. Working to eliminate number 3

Packaging and Deployment (cont.)



Currently all these components are installed on a gateway node

- Privileged access is required for primary middleware (1) and parts of experiment middleware (2)
- Currently the Grid2003 site installation is 5 steps

1. Set up local scheduler
    - Default is Condor, but we expect installations of PBS, LSF, and FBSNG
2. Get Host certificate
3. Create Local Accounts
    - We are currently using generic VO accounts (uscms01,usatlas1,ligo1, sdss1)
4. Get pacman version
5. Install Grid2003 packages with pacman
    - Download, installation, and configuration is currently handled by pacman

- To install a new site should be very fast, but we are currently averaging a few hours of work on each site to get the configuration right

Grid2003 Requirements and Strategies

- Experiments must be able to effectively interoperate and run their applications on non-dedicated resources.
- Applications must be able to install themselves dynamically, thereby imposing minimum requirements on grid facility managers.
- We will work within a grid architecture consisting of
  - facilities (e.g., execution and storage sites),
  - services (e.g., a prototype operations center, an information system for resource publication and discovery, and so on),
  - applications (such as CMS and Atlas production and prototypical analysis).
- The middleware will be based on VDT 1.1.9 or greater, and will use components from other providers as appropriate.
- Set of used components of common middleware driven by each individual application.

Grid2003 R&S (cont.)

- Packaging, installation and configuration should be simple and easy for system managers — running a single Pacman script for example. Simple update capabilities should be available, without having to reinstall.
- An information service for resource publication and discovery based on MDS will be deployed. Grid3 plans to use Glue schema. Will further develop requirements for schema definitions and information providers to ensure correct operation of all intended services.
- A simple monitoring service will be deployed based on Ganglia, version 2.5.3 or greater (supporting hierarchy, grid-level collections), with collectors at one or more of the operations centers. Additionally, Ganglia information providers to MDS may be deployed if required, and native Ganglia installations may report to other collectors corresponding to other (logical) grids.
- Other monitoring systems, such as MonaLisa or a workflow monitoring system, may be deployed. Guidelines for information providers, sensor installation, and pacmanized versions of these with clear

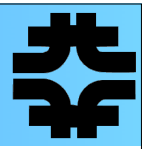
Grid2003 R&S (cont.)

- A consistent method or set of recommendations for packaging, installing, configuring, and creating run-time environments for applications will be provided, subject to review — for example:
  a) Pacman caches and instructions for pre-installation of application libraries
  b) Use of the Condor grid shell, job-wrapping mechanism
  c) Precise instructions for application environments.
  d) Use of the Chimera Gridlauncher.
  e) Use of the WorldGrid project mechanism.
- One or more VO management mechanisms for authentication and authorization will be used
  a) VOX project is integrating VOMS server based authentication, working with the EDG/DataTag VOMS people
  b) The WorldGrid project method as developed by Pacman.
  c) A fallback solution is to use LDAP VO servers, one for each VO containing the DN's of the expected application users.
  d) An acceptable use policy to be signed by all participants.

Grid2003 R&S (cont.)

- Support for replica location services as required by application components, such as Magda for ATLAS. An example would be an LRC, indexed to the experiment's RLS. To be discussed.
- A simple user-support model, and other service requirements such as maintaining a trouble reporting system, liaison functions, etc., will be developed and reviewed for the operations center.
- A common reporting and event logging tool or framework, possibly with a graphical user interface, to be used by the application teams, will be used to collect information about task workflow performance, statistics, etc. This will be used for archival purposes, grid-viewing data displays, feedback to the CS teams, and will form the basis for input to computing reviews, e.g., SC2003.
- Data replication and data movement services, for example a set of sites may deploy a storage element based on SRM, Gridftp, and dCache, as required by the applications.

Current Status



Currently

- Central Grid2003 Cache exists based on VDT 1.1.9
  - VO Testbeds have validated VDT 1.1.10 and are ready to upgrade
- Some application caches are published
- Grid2003 sites are publishing to VO GIIS, which in turn report to Grid2003 GIIS
- VOMS databases for the VOs are coming on-line
  - Currently a static grid-mapfile is shipped with Grid2003 core
- 11 Grid2003 sites have been installed and comissioned
  - 3CMS, 4 ATLAS, 1SDSS, 1 LIGO, and 2 from Grid Projects
  - Running "heart beat" scripts to verify the status of grid services
- Multi-VOs have run on a single site
- 1 VO has run on many sites
- No VO has run on all sites, all VOs have not yet run on all sites

Current Status

Last Updated on September 08, 2003 by Jerry Gieraltowski

NOTE 1: All tests were performed using a DOEgrids user certificate

Key

- No obvious problem detected in either the environment or grid services
- A problem was detected in either the environment or grid-services
  Check the *env and *gridsvcs links in the row entry for the node.

Grid3/2003 Sites						
Node Name		**VO**	**Pool ID**	**Domain Name**	**Environment Variables**	**Status of Grid Services**
Argonne Nat'l Lab-HEP Contact: Ed May	●	ATLAS	grid3-anlhep	atlas12.hep.anl.gov	grid3-anlhep-env	grid3-anlhep-gridsvcs
CalTech Contact: Suresh Singh	●	CMS	grid3-caltech	citgrid3.cacr.caltech.edu	grid3-caltech-env	grid3-caltech-gridsvcs
University of Indiana Contact: Leigh Grundhoefer	●	ATLAS	grid3-iu	iuatlas.physics.indiana.edu	grid3-iu-env	grid3-iu-gridsvcs
SDSS Contact: Vijay Sekhri	✸	SDSS	grid3-iu	tam01.fnal.gov	grid3-sdss-env	grid3-sdss-gridsvcs
ISI Contact: Nosa Olomu	✸	iVDGL	grid3-iu	skywalker.isi.edu	grid3-isi-env	grid3-isi-gridsvcs
Brookhaven Nat'l Lab Contact: Dantong Yu	●	ATLAS	grid3-bnl	spider.usatlas.bnl.gov	grid3-bnl-env	grid3-bnl-gridsvcs
University of Florida Contact: Jorge Rodriguez	●	CMS	grid3-ufl	ufgrid01.phys.ufl.edu	grid3-ufl-env	grid3-ufl-gridsvcs
University of Chicago Contact: Xin Zhao	●	ATLAS	grid3-uchicago	grid02.uchicago.edu	grid3-uchicago-env	grid3-uchicago-gridsvcs
San Diego Supercomputer Center	●	CMS	grid3-sdsc	uscmstb0.ucsd.edu	grid3-sdsc-env	grid3-sdsc-gridsvcs

Current Activities

Short term activities

- Working through September to run more VO applications on Grid2003 sites
- Upgrading to VDT 1.1.10 based Grid2003 core
- Demonstrate that the efficiency on Grid2003 is the same as efficiency on single VO testbeds
  - A conditions the experiments have placed before large resources are put into Grid2003
- Improving monitoring of both system components and configurations
- Move significant resources into Grid2003

On October 1st we hope Grid2003 to be at it's design scale

Grid2003 and the LCG

Grid3 should federate with the LCG

- We will have a working group within Grid3 to understand what is needed for basic interoperability, specifically submission of jobs and movement of files across the grids.

There will be other issues that may affect interoperability consistent replica management, virtual organization support and optimizing of resource usage across federated grids.

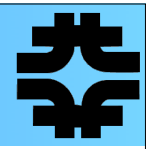
- We do not have the effort to address all these during the Grid3 project itself. We will identify, discuss and document such issues for collaborative work with the LCG.

Many of the working areas in Grid3 are already joint projects between the LCG and Trillium or the S&C projects.

- Additional collaboration in areas of monitoring and operations have been discussed over the past few months.

As we proceed to better understand the technical plans the expectation is that we will propose further areas of joint work.

LCG-1 Deployment in the US



Both US-ATLAS at BNL and US-CMS at FNAL are very close to a successful deployment of LCG-1

- US-CMS has used the supported LCFGNG installation
- US-ATLAS has used the unsupported "lite" installation

The FNAL site is publishing to the information providers for LCG-1

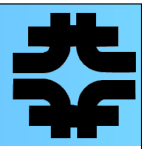
- Currently working through authentication and VO issues
  - Reasonable possibility of commissioned site today

The BNL site has been working through the "lite" installation

- This is currently unsupported but in the long term a much more acceptable installation scheme

Grid2003 has not reduced the manpower applied to LCG-1 installations

Grid2003 should become OSG-0



We are proposing to DOE and NSF a roadmap for the U.S. to build a national grid infrastructure for science, the **Open Science Grid** (OSG)
We propose a program of work to federate the U.S. LHC grid resources into a scalable, engineered and managed grid, the Open Science Grid.

- Much of the LHC grid infrastructure in Europe will be provided by a combination of CERN central resources and a consortium of European centers that propose to federate some of their resources in a grid for e-science in Europe EGEE. Centers in the U.S. and other parts of the world will federate with this European grid infrastructure in order to provide the global computing grid for LHC science.
- It is now time for the U.S. to also federate its LHC computing resources and in doing so to continue to lead the efforts towards a global grid infrastructure for LHC science and beyond. We propose to provide and operate these resources at the national laboratories and universities as the initial seed for the Open Science Grid

This new Peta-scale computational service will be built as an open national infrastructure, optimizing shared use of resources for diverse collaborative research. The Open Science Grid will serve as a backbone to merge grid computing efforts of allied experiments in particle and nuclear physics, and can be extended to other scientific communities

Federated Grids

This is a diagram from Ian Bird talk at the Lepton Photon Grid session

- LCG is beginning to approach the issue how large national and regional grid infrastructures interact and interoperate
- US-CMS sees GRID2003 as the pilot toward a regional grid infrastructure in the US