



U.S. CMS Software and Computing

Lothar A. T. Bauerdick
Project Manager
U.S. CMS Software and Computing

Tracker
HCAL
Solenoid
Magnet Yoke
 μ chambers



LHC is the most exciting accelerator ever

EW Symmetry Breaking, Origin of Mass,
First Scalars: Discovery of the Higgs Boson
Discovery of Super Symmetry (SUSY)
Big step in Energy and Luminosity \Rightarrow Much more

U.S. CMS will utilize and contribute to the full physics potential of CMS

Critical mass (500 U.S. physicists).
Good facilities (at Fermilab and Universities).
New technology for collaboration (Grid, Video Conference).

**It is the task of the U.S. LHC Research Program
— including Software and Computing —
to enable CMS and U.S. Universities
to take full advantage of this unique opportunity
by providing IT engineering and infrastructure support
for scientific discovery and research**



Scope and Deliverables



DOE/NSF sponsor LHC Research Program includes Software and Computing

- ➔ Computing Infrastructure in the U.S. and software engineering for CMS
 - ◆ Mission is to develop and build “User Facilities” for CMS physics in the U.S.
 - To provide the enabling IT infrastructure that will allow U.S. physicists to fully participate in the physics program of CMS
 - ◆ To provide the U.S. share of the framework and infrastructure software

Tier-1 center at Fermilab provides computing resources and support

- ➔ User Support for “CMS physics community”, e.g. software distribution, help desk
- ➔ Support for Tier-2 centers, and for physics analysis center at Fermilab

Five Tier-2 centers in the U.S.

- ➔ Together will provide same CPU/Disk resources as Tier-1
- ➔ Facilitate “involvement of collaboration” in S&C development
 - ◆ Prototyping and test-bed effort very successful
- ➔ Universities will “bid” to host Tier-2 center
 - ◆ taking advantage of existing resources and expertise
- ➔ Tier-2 centers to be funded through NSF program for “empowering Universities”
 - ◆ Proposal to the NSF submitted Nov 2001



Project Milestones and Schedules



Prototyping, test-beds, R&D started in 2000 “Developing the LHC Computing Grid” in the U.S.

- ➔ R&D systems, funded in FY2002 and FY2003
 - ◆ Used for “5% data challenge” (end 2003)
 - ⇒ release Software and Computing TDR (technical design report)
- ➔ Prototype T1/T2 systems, funded in FY2004
 - ◆ for “20% data challenge” (end 2004)
 - ⇒ end “Phase 1”, Regional Center TDR, start deployment

Deployment: 2005-2007, 30%, 30%, 40% costs

- ➔ Fully Functional Tier-1/2 funded in FY2005 through FY2007
 - ◆ ready for LHC physics run
 - ⇒ start of Physics Program

S&C Maintenance and Operations: 2007 on



Successful Base-lining Review



“The Committee endorses the proposed project scope, schedule, budgets and management plan”

Findings & Recommendations of the Project Management Subcommittee

- US CMS Project Management is in place and working well
- Project appears well defined
 - Scope can be achieved with proposed resources
 - Budget matches agency guidance profile
 - Schedule takes advantage of unofficial LHC slip - appears achievable (but with little or no margin)

Findings & Recommendations of the Project Management Subcommittee (2)

- US CMS has taken an excellent first step in defining what services they require from grid-develop SW packages
 - US CMS needs to implement the tracking procedure to assess grid-project progress and its impact on US CMS schedule
- Committee is concerned about the potential impact on US CMS of future design/specification decisions made by CERN, especially, e.g., in the area of data persistence models and GRID technology

Work with funding agencies, DOE and NSF, to successfully execute the project!



LHC Computing Grid Project



LCG Project at CERN for building the LHC computing environment at CERN

- ➔ Technology development and Data Challenges of the experiments from 2001 onwards leading to a Production Prototype at CERN and in MS and NMS by 2004
- ➔ Technical Design Report in 2004 describing
 - ◆ the necessary size and cost of the production facility at CERN and elsewhere
 - ◆ agreeing the relations between the distributed Grid nodes
 - ◆ their coordinated deployment and exploitation

Approved in the CERN Council Meeting of 20-Sept-01

- ➔ Project Leader, Les Robertson/CERN, Project Oversight by SC2, chair Matthias Kasemann/FNAL
- ➔ Project organisation frame setup and working
- ➔ Progress in identifying partners of LHC Computing Grid
- ➔ Contributions to Phase 1, hiring started, not yet complete (>50 FTE)
- ➔ Excellent start at LCG workshop March 11-15, to define work plan

The U.S. needs to be fully participating in the LCG program to ensure that we can fully participate in the LHC research program



U.S. CMS User Facilities Successful

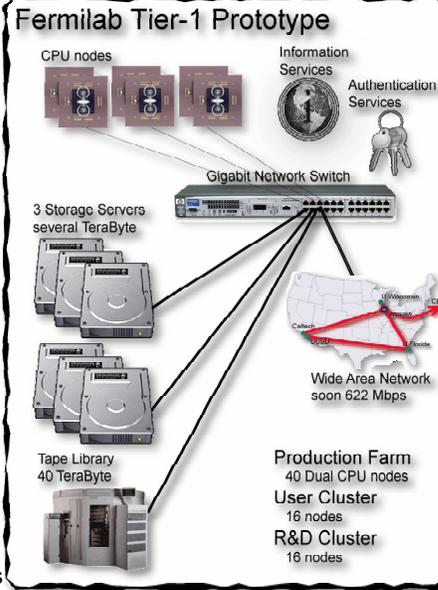


U.S. User Facilities established

- ➔ Facilities for event simulation, including reconstruction and sophisticated pile-up processing
- ➔ User cluster and hosting of large data samples for physics studies
- ➔ Facilities and Grid R&D

Excellent initial effort and DOE/NSF support for User Facilities

- ➔ Fermilab established as Tier-1 prototype and major Grid node for LHC computing
- ➔ Two Tier-2 prototypes established, strong interaction with Grid projects
- ➔ R&D to develop distributed environment for LHC physics research started
- ➔ Head-start for U.S. efforts pushed CERN commitment to support remote sites



U.S. CMS Prototypes Operational



University Wisconsin at Madison

Condor
High Throughput Computing

"Thousands of CPUs!"

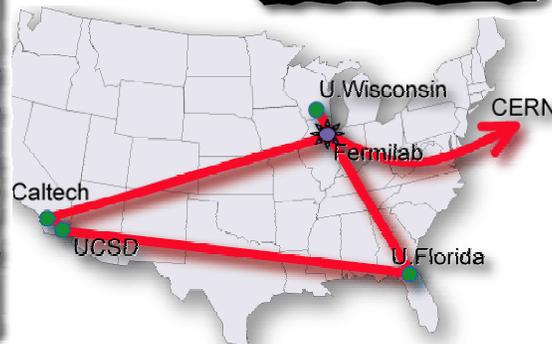
University of Florida

2nd U.S. CMS prototype Tier-2
72 CPU nodes
distributed disks + 1.5 TB RAID

California prototype Tier-2

Caltech UCSD

40 Duals
TB Storage Servers
Gbit Ethernet switches



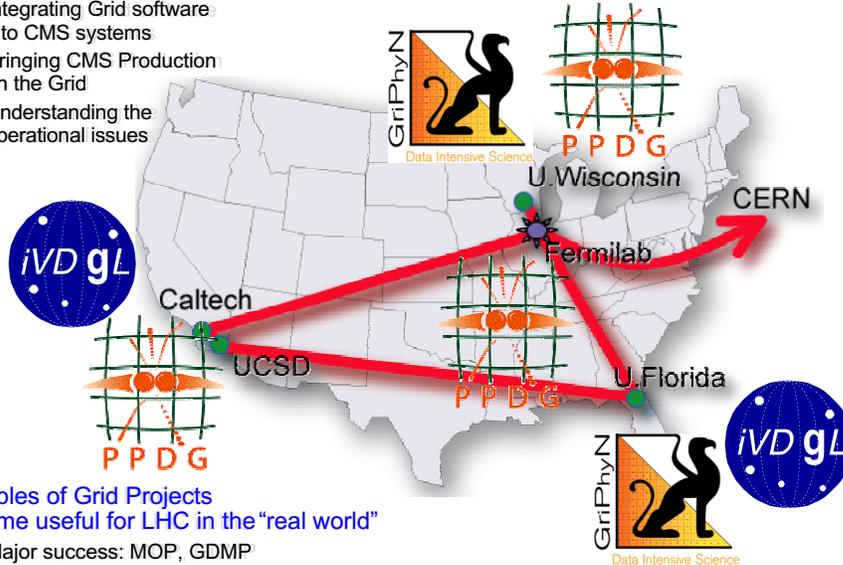


U.S. CMS Prototypes and Test-beds



All U.S. CMS S&C Institutions involved in DOE and NSF Grid Projects

- Integrating Grid software into CMS systems
- Bringing CMS Production on the Grid
- Understanding the operational issues



Deliverables of Grid Projects become useful for LHC in the "real world"

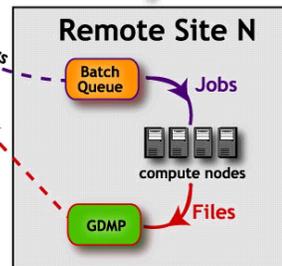
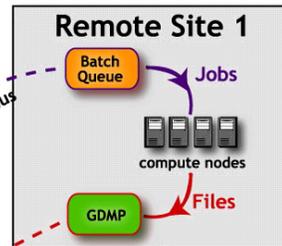
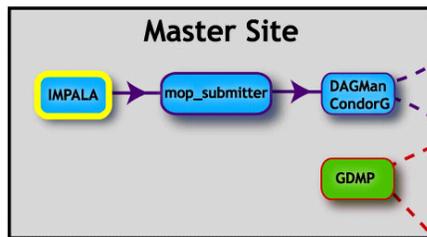
- Major success: MOP, GDMP



Very Successful Collaboration With US And EU Grid Projects!



MOP - Monte Carlo Distributed Production System for CMS



PPDG Grid-ware for CMS

- MOP (Fermilab, U.Wisconsin)
 - ◆ Remote job execution
- GDMP (Fermilab, European DataGrid WP2)
 - ◆ File replication and replica catalog

Successfully used on CMS test-bed

First use in CMS Production last week



U.S. Tier-1/2 Fully Integrated in CMS



CMS Grid Integration and Deployment on U.S. CMS Test Bed

Data Challenges and Production Runs on Tier-1/2 Prototype Systems

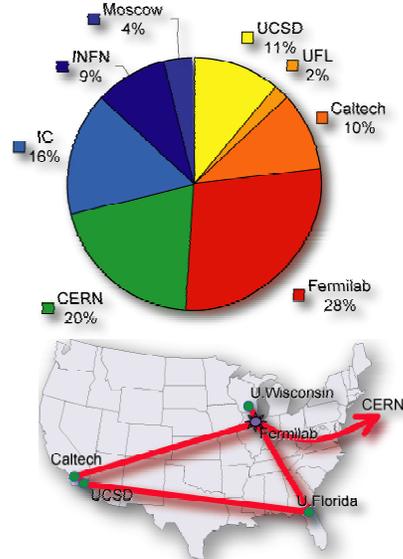
"Spring Production 2002"

Physics, Trigger, Detector studies

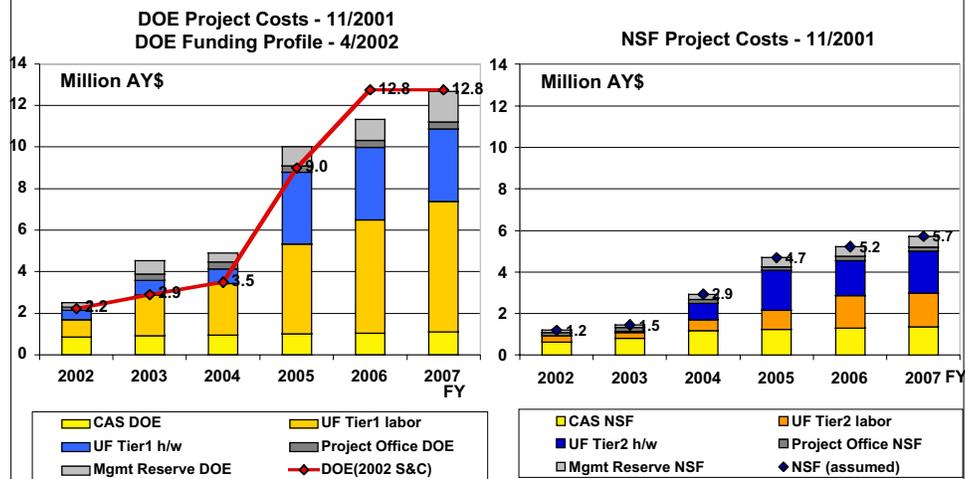
- Produce several $\times 10^6$ events fully simulated including pile-up fully reconstructed
- Status as of last week: 600GB produced
- Large assignment to U.S. CMS

Successful Production in 2001:

- 8.4M events fully simulated, including pile-up, 50% in the U.S.
- 29TB data processed 13TB in the U.S.



S&C Project Funding





FY2002 - FY2004 Are Critical



S&C base-line 2003/2004: Tier-1 effort needs to be at least \$1M/year above FY2002 to sustain the UF R&D and become full part of the LHC Physics Research Grid

- Need some some 7 additional FTEs, more equipment funding at the Tier-1-1
- Essential areas are insufficiently covered now, need to be addressed in 2003 the latest
 - ◆ Fabric management • Storage resource mgmt • Networking • System configuration management • Collaborative tools • Interfacing to Grid i/s • System management & operations support
- This has been strongly endorsed by the S&C baseline review Nov 2001
- All European RC (DE, FR, IT, UK, even RU!) have support at this level of effort

Country	#CMS collaborators	Tier-1 Equip	Tier-1 Personnel	Other
DE	95	9 kSI, 45 TB	5 Op, 8 Dev	5 Mio€ to CERN
FR	137	40 kSI, 55 TB	47 Op/Dev @ Lyon	
IT	262	20kSI, 80TB	13 Op, >> Dev	ItGrid
NL		9kSI, 20TB	10 Op	
RU	271	10kSI, 12TB	13 Op	
UK	98	20kSI, 50TB	13 Op, 20 Dev	GridPP: >25M€
USA	463	4kSI, 5TB	5.5 Op, 4 Dev	x2 for LHC total

We need to bring up our initial efforts to the required level— as planned otherwise U.S. seriously compromises ability to be part of the LHC Computing Grid
 But this is essential for successfully deliver the enabling technology and infrastructure to have U.S. Universities be fully participating in the LHC research program!



Conclusions



Prototypes and test-beds at Fermilab and U.S. Universities are operational now
R&D and test-bed efforts are essential to build the infrastructure for U.S. based LHC scientific research

LHC Computing Grid Project started at CERN, U.S. needs to be fully participating

U.S. CMS Prototype facilities are sizeable
already now play a major role for CMS user community in the U.S.
production efforts for physics studies and upcoming data challenges
user support, software distribution, hosting data samples

These Physics studies with full simulation including pile-up
are essential for trigger and detector understanding

We are looking for ways to bring the effort up to the required level
to ensure functionality of US managed sub-systems of the LHC Computing Grid

This essential for a successful project base-line
as has been endorsed by the November S&C review!