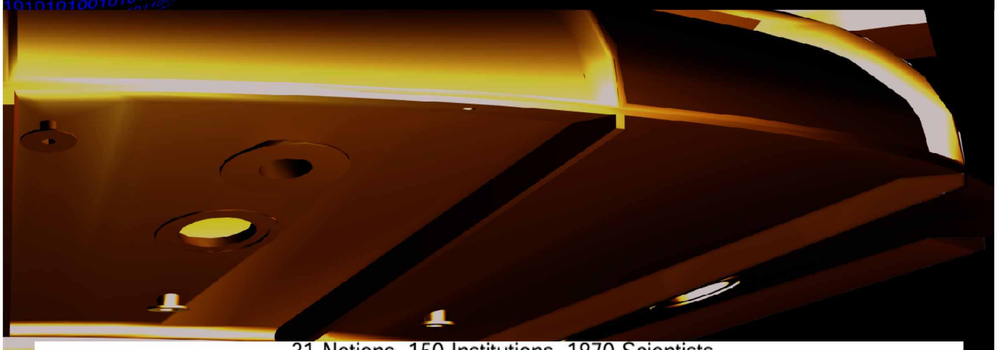
Data Intensive Science



31 Nations, 150 Institutions, 1870 Scientists

TRIGGER & DATA ACQUISITION
Austria, CERN, Finland, France, Greece, Hungary, Italy, Korea, Poland, Portugal, Switzerland, UK, USA

TRACKER
Austria, Belgium, CERN, Finland, France, Germany, Italy, Japan*, Switzerland, UK, USA

CRYSTAL ECAL
Belarus, CERN, China, Croatia, Cyprus, France, Italy, Japan*, Portugal, Russia, Switzerland, UK, USA

PRESHOWER
Armenia, Belarus, CERN, Greece, India, Russia, Taiwan (PC), Uzbekistan

RETURN YOKE
Barrel: Czech Rep., Estonia, Germany, Greece, Russia
Endcap: Japan*, USA

SUPERCONDUCTING MAGNET
All countries in CMS contribute to Magnet financing in particular:
Finland, France, Italy, Japan*, Korea, Switzerland, USA

FEET
Pakistan
China

FORWARD CALORIMETER
Hungary, Iran, Russia,Turkey, USA

HCAL
Barrel: Bulgaria, India, Spain*, USA
Endcap: Belarus, Bulgaria, Russia, Ukraine
HO: India

MUON CHAMBERS
Barrel: Austria, Bulgaria, CERN, China, Germany, Hungary, Italy, Spain,
Endcap: Belarus, Bulgaria, China, Korea, Pakistan, Russia, USA

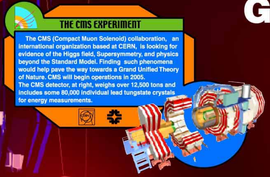
Total weight	:	12500 T
Overall diameter	:	15.0 m
Overall length	:	21.5 m
Magnetic field	:	4 Tesla

* Only through industrial contracts

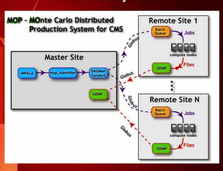
GRID COMPUTING FOR THE FUTURE

THE CMS EXPERIMENT

The CMS (Compact Muon Solenoid) collaboration, an international organization based at CERN, is looking for evidence of the Higgs field, Supersymmetry, and physics beyond the Standard Model. Finding such phenomena would help pave the way towards a Grand Unified Theory of Nature. CMS will begin operations in 2005.
The CMS detector, at right, weighs over 12,500 tons and includes some 80,000 individual lead tungstate crystals for energy measurements.



Distributed Systems

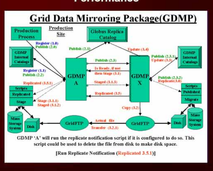


Above are shown block diagrams and flowcharts detailing the software solutions needed to serve the data management needs of a large international collaboration such as CMS.

Grid Tools: *Tying It All Together*

Jobs for Monte Carlo Production are distributed among N sites using MOP; a product that uses Globus and Condor-G. Large data transfers among sites are achieved using GDMP.

Performance

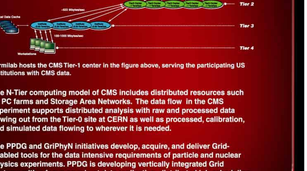


GDMP "A" will run the replicate notification script if it is configured to do so. This script could be used to delete the file from disk to make disk space.
[See Replicate Notification (Replicated 3.5.1)]

Above are shown graphs and diagrams that show the use and performance of the early grid tools in distributed CMS Monte Carlo production and analysis.

Grid Initiatives and High Energy Physics

The CMS experiment is expected to yield several Petabytes of stored raw and processed data in its first few years of operation. A comparable amount of Monte Carlo simulated data must also be generated, stored, and analyzed.



Fermilab hosts the CMS Tier-1 center in the figure above, serving the participating US institutions with CMS data.

The N-Tier computing model of CMS includes distributed resources such as PC farms and Storage Area Networks. The data flow in the CMS experiment supports distributed analysis with raw and processed data flowing out from the Tier-0 site at CERN as well as processed, calibration, and simulated data flowing to wherever it is needed.

The PPDG and GriPhyN initiatives develop, acquire, and deliver Grid-enabled tools for the data intensive requirements of particle and nuclear physics experiments. PPDG is developing vertically integrated Grid systems with a focus on robust data replication, distributed job scheduling, and monitoring services. GriPhyN is developing and instantiating Data Grid architecture via tools that support location and materialization transparency, request planning, and execution.