

# CMS Drivers

Ian Fisk  
GDM Meeting  
April 12, 2005



# 2005/2006 Drivers



Starting the Tier-2 program, including DISUN Tier-2C program

- ➔ goal of full set of Tier-2s for start of physics

Ramping facilities to be ready for physics in 2007

- ➔ scaling the facilities, CPU, storage, networks to physics needs

Baseline software and baseline services -> “Magnet Test”, DC06

- ➔ core software, event model, data management & bookkeeping, data bases, physics analysis environment, work flows, etc.
- ➔ integration activity with meaningful milestones and challenges

Analysis data processing in the distributed environment

- ➔ placement of datasets, configuring and running jobs, serving data, managing & monitoring the system, implementing collaboration policies & priorities, secure, open & friendly environment

Supporting U.S. CMS community, particularly the LPC as driver

Moving Open Science Grid forward and interoperating with LCG



# Starting the Tier-2 Program



## Workshop during the first week of May

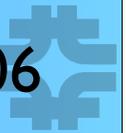
- ➔ We have a basic understanding of how to offer grid-enabled processing resources
  - Reasonable success from Grid3 and CMS distributed event production
  - Working on reasonable interface for users through MCPS (related project to the CMS RunJob effort)
- ➔ Storage Resources are more challenging
  - Improvements in operations for physical infrastructure
  - Improvements in operational expertise for dCache/SRM
  - Improvements in the SRM functionality
    - Resource reservation
    - Priority implementation
    - Authorization Improvements



# Ramping Facilities to be ready for Physics

## Scaling the Facilities and Services

- ➔ Need to physically increase the processing capacity of the farms
  - Total physical systems are similar to CDF or DZero farms
  - Improvements in auditing accounting and monitoring are needed
  - Better understanding and development of site policy and enforcement technology is required.
    - Big International Distributed Collaboration
- ➔ Need to continue increasing the scale of data serving and archiving
  - CMS expected IO requirements are between 800 - 1600 MB/s of active storage IO
  - Currently routinely achieving 200-400MB/s
  - Need to improve operational depth of dCache
- ➔ Need to achieve production networking to accommodate CMS TI requirements.
  - We have demonstrated efficient use of research link



## Re-engineering of Framework and Event Data Model

## Development of Data Management Services

- ➔ Continued Deployment of data transfer tools
- ➔ Dataset bookkeeping prototype expected by summer

## Improvements in the physics analysis environment

- ➔ Integration with publishing and bookkeeping services
- ➔ Software environment improvements

## Workflow development

- ➔ User analysis workflow and submission
- ➔ Simulation Production workflow on OSG and beyond



# Analysis Data Processing



## Placement of datasets

- ➔ Experiment policy and integration with placement and transfer services

## Configuring and submission of jobs

- ➔ Workflow and submission improvements
- ➔ **Monitoring and Debugging improvements**

## Serving data

- ➔ Integration with dataset bookkeeping

## Managing & monitoring the system

- ➔ **Improvements in auditing, accounting, and tracking**
- ➔ **Policy definition and enforcement**
  - implementing collaboration policies & priorities



# Moving OSG Forward and Interoperability

## OSG Deployment

- Support of OSG VOs
  - US-CMS
- Support for OSG Services
  - Privilege
  - VOMSRS and VOX
  - SRM/dCache (Selected sites)
- Operations

## Interoperability

- OSG LCG interoperability challenges
  - Technical issues
    - Information providing, VO managements, etc
  - Policy issues
    - How to account and identify sites