

# Virtual Organizations and OSG Storage

**Abhishek Singh Rana**

OSG Storage Forum  
September 22, 2010

# Outline

- Categories
  - VOs successful with use of OSG storage.
  - VOs with specialized or unmet needs.
  - VOs expecting future use of storage.
- Feedback
  - Limitations due to different ‘need landscapes’.

Please note – LIGO discussion is covered by Kent’s presentation;  
ATLAS and CMS discussions are covered in the later sessions.

# Categories

# Category 1: VOs Successful with use of OSG Storage

- ATLAS, CMS, LIGO
- D0 Collaboration
- DOSAR
- HCC / Holland Computing Center
- Fermilab-VO and sub-VOs
- DES / Dark Energy Survey
- GridUNESP in Brazil

# Category 2: VOs with Specialized or Unmet Needs

- SBGrid / Structural Biology Grid
- IceCube Neutrino Telescope
- CDF Collaboration
- GEANT4 Collaboration
- NanoHUB
- NYSGrid

# Category 3: VOs expecting Future use of Storage

- ALICE LHC Experiment
- GLUE-X Collaboration
- DayaBay Reactor Neutrino Experiment
- JDEM / Joint Dark Energy Mission

# Feedback

# Feedback Sample

- Focus on small files: SBGrid, IceCube, NanoHUB, GEANT4.
- Need frequent caching, pre-fetching or pre-staging: SBGrid, IceCube.
- Need rapid-access low latency: Fermilab-VO's Intensity Frontier experiments/sub-VOs, SBGrid, NYSGrid.
- Need more security and robust privacy: SBGrid.
- Require ease of usability: SBGrid, GLUE-X.
- Resolve interference due to job pre-emption: All VO's relying on heavy opportunistic usage.



# Feedback Sample

- Help in evolution away from NFS: VOs using OSG\_DATA and hitting limitations.
- Need yet higher efficiency of data transfers: D0, CDF, NanoHUB, Dark Energy Survey.
- Guidance in workflow customization, database decomposition, data granularity: CDF, IceCube, CompBioGrid.
- Looking for tools to verify storage availability: VOs relying on multiple SEs across OSG.
- No current needs: DayaBay, JDEM.

# Feedback

- Current storage solutions in OSG have good capability and scalability to handle large-scale data movement and global storage.
- However, smaller VOs generally have a different landscape of storage needs: **Small files; Frequent caching or pre-fetching; Rapid-access low latency; More security and robust privacy; Ease of usability; Resolve pre-emption based interruptions; Evolution away from NFS; Higher efficiency of data transfers; Need for guidance with workflow customization and database decomposition; Tools to verify storage availability.**
- Successful usage of storage is case-by-case.