Small Fish in a Big Data Pond

Data Storage and Distribution Strategies for new VOs.

Tuesday, 9 April 2013
Itinerary

• The Problem

• What you can do to help yourselves.

• What Sites / Devs can do to help you.
The Problem

• Grid Middleware is hard to use.

• Data Management is hard to do.

  • Data locality, integrity, access controls...

• Data Management with Grid Middleware is therefore (hard$^2$).
The Big VO Problem

- Big VOs make the situation worse.
- Lots of manpower
  - “Fix” Grid Middleware & Data Management
  - ...for themselves.
- “Basic” infrastructure sees little of the development effort.
Coming to Terms with Reality

• Small VO policies, therefore:
  • Must cope with “base” middleware
  • Adopt strategies to minimise risk
  • Make use of experience where present.
  • Prefer techniques where experience is wide.
The Big VO Advantage

Experience

• Big VOs have a lot of experience in managing a lot of data.

• We can learn general lessons that can be applied to any VO with data needs on the Grid.

• Don’t make mistakes that ATLAS, CMS, et al have already discovered...
What You can do to help yourselves.

Resilience

• Book-keeping, book-keeping, book-keeping
  • eg. T2K, biomed issues

• Help Us to Help You
  • Talk to Sites, or site-representing bodies
    • e.g. UK has a GridPP Storage Group Experience

Tuesday, 9 April 2013
“Generic” Data model

- Data Types
  - Simulation
    - programmatically generated, reproducible with same time input
  - Experimental
    - from physical source, not reproducible
- Precious
Access Controls

- Decide *early on*. who needs to access what.
- Build this into your VO structure.
- Tradeoff:
  - Ease of Use
  - Ease of Consistency
  - Err on side of Consistency.

Resilience
Filenames and Metadata

• Lessons from ATLAS, CMS
• Regularise LFNs, SURLs to common pattern (so SURL can be derived from LFN)
• Consider encoding some metadata in filename to allow recovery (ie LFN path)
• Use LFC for central file catalog (don’t reinvent the wheel).
Protocols

- Try to use popular standards.
  - Historically an issue with Grid storage
  - WebDAV, pNFS now/soon supported on all major Grid storage platforms.
  - Allows use of existing tools, experience from outside the merely EGI community.
Managing Data

- Tracking
  - Dark Data is a problem even for Big VOs
  - Enforce LFC registration of files (VO discipline)

- Deleting
  - Central Policies
  - “Cache” spaces can be cleaned up by sites/different reliability/capacity tradeoff.
Simulation Data

• Two models:
  • Write to site job runs at (eg ATLAS, NA62) & register in common catalogue
  • Write to common silo (eg LHCb)
  • Former is complex to track & gather
  • Latter scales poorly with VO growth
• Need multiple copies / alg. redundancy
  • Currently, best approach is 3+ copies at Tier-2 sites
  • Importance of Tape is overstated.
  • Jobs should go to data → distribution also balances job load
“Tier-3s”

Standards

- “Final stage” analysis often off Grid ("Tier-3,4")
- Copying to local storage wasteful if near Tier-2 resource.
- Open standards help - mount WebDAV/pNFS from Tier-2 to local resource.
- Maximise use of existing infrastructure!
Moving Data

- FTS is the only reliable file transfer system in the default infrastructure.
- Care needed: does not register replicas in LFC.
- For all it’s faults, it is still the best thing we have (still used by Big VOs).
What We can do to help you.

• Sites can’t rewrite middleware.

• We **can** and **should** provide assistance where needed.

• We **can** and **should** make sure our config is reliable

• We **can** and **should** consider validating our integrity
What We are doing for you.

- GridPP effort (via QMUL) for LFC/SE consistency checking tool.
  - Bookkeeping, resilience
- Develop Tools for Storage
  - DPM-centric tools from GridPP
What We should do for you

• Method of setting deletion ACL separately from creation ACL.
  • Enforce central deletion

• Provide metric for “total storage used/avail” per VO
  • Easiest with Spacetoken per VO
Summary

- Reuse **Experience**
- Use **Standards** whenever possible
- Design for **Resilience**