SRM+Cloud session

Paul Millar
Overview

• SRM+Cloud in five steps:
  • Description of the Topic, State of Play, Where Things Need to go, What Might go Wrong, What are the unasked questions?

• My job:
  • Stimulate discussion
Description of topic

- Protocols (and APIs) for controlling storage,
- Typical operations:
  - Adjust properties of a file: bring files online, prevent from garbage-collecting online copy.
  - Namespace: rename, move, delete files or directories.
  - Guaranteeing sufficient capacity ("space reservations")
- Ways of achieving this:
  - SRM
  - Cloud
  - WebDAV
State of Play: SRM

- Standard protocol
  - although standard is poorly written and there is various WLCG agreement codicils
- Supports everything needed:
  - Probably tries to do too much
- Implementation is [almost] uniformly available
  - Various states of “broken”
- Everyone in WLCG knows it; knows what to expect.
  - Completely unknown outside WLCG (+ some small communities)
State of Play: Cloud

- One major provider (Amazon S3) and lots of others
- No standard
  - CDMI shows promise, but Amazon not interested.
- Provides subset of features:
  - No true namespace
State of Play: WebDAV

- Standard protocol
- Fairly widely adopted for read/write access
  - e.g., adopted by Swestore for Sweden's national storage facility.
- Clients are various state of “broken”
  - No one client works perfectly
- Supports a subset of SRM functionality:
  - Missing facility to create and destroy space reservations.
Where things need to go

• Uncomfortable answer this question:
  • 1. do we need move “change for change's sake”
  • 2. I'm not an end-user: changes should be driven by end-user requirements

• Standardisation: move away from bespoke solutions towards industry standards:
  • Maintenance
  • Reusability of software (for other communities)

• Incremental changes on existing system
What might go wrong

- Everything stops working: big bang + smoke
- Change introduces something that doesn’t scale
  - Only find out when in production
- Effort is diverted from other activities (e.g. demonstrators)
  - Just ask what has already been testing → Industry
    - Curl, Apache
- Disenfranchise free resources (E.g. BestMan)
- Whatever you can think of, can go wrong.
What are the unasked questions?

• What will happen in the Cloud?
• What functionally do experiments really need?
  • This is HIGH LEVEL OVERVIEW, not protocol X
• Whether we can fix WebDAV clients