

ATLAS SC3 Data Management

2005/06/14

ATLAS DDM

Miguel Branco - miguel.branco@cern.ch

Timetable

- Focusing mostly on Data Management
 - Main goal for SC3
- ATLAS will join in “Production mode” only in (mid) October
- 2 phases
- Phase I: Starting from now until end of August:
 - Test data transfer components and catalogs only, by extending existing ATLAS data management system (DQ): LFC (Fireman?), SRM, FTS
- Phase II: Starting from late September:
 - Integrate new ATLAS data management infrastructure (DQ2)

Phase I

- Plug DQ Reliable File Transfer into FTS
- Requires from LCG:
 - Deployment of FTS
 - Migrate EDG RLS to LFC
 - For Phase I only, a single central LFC catalog is sufficient!
- Exercise data transfers
 - Focus on SRM (v1.1): **Primary Goal for Phase I**
 - Great summary of baseline requirements from Nick Brook (presentation tomorrow)
 - Site availability, monitoring
- Data sizes: mostly files between 150 MB and 400 MB. Few files over 1 GB
 - Start with disk<->disk, later disk->tape
- If works better than DQ RFT, FTS will go into production ASAP

Phase I: Data Transfer

- Transfers (reuse few datasets from Rome production - not produced during SC3 “Phase I”!):
 1. Initially Tier0 to Tier1 disk
 2. From Tier1 to several Tier2
 3. Repeat Tier0 disk to Tier1 tape
- Main goals for Phase I
 - Throughput Tier 0 \Leftrightarrow Tier1
 - Disk to disk ~150 MB/s
 - Disk to tape ~60 MB/s
 - From Tier1 to Tier2 less clear; depends on Tier2? (transfer burst)
 - But NOT just throughput, also monitoring transfers; querying catalogs, and ...
 - support for site failures (FTS as a service not just a piece of s/w),
- Exercise parallel to SC3
 - Installing FTS on BNL and testing BNL to CERN transfers
 - This additional ATLAS testbed is used to develop on top of FTS

Phase II

- Phase II starts on (mid) October
- Automatic management of data transfers (from ATLAS perspective)
 - When new ATLAS ProdSys in place for SC3, will also use newly produced datasets
- Rates for Phase II (Production): see Gilbert's talk next
- Deploy DQ2
 - New dataset catalogs
 - VO-site services
- Will use data transfer components proven to work from Phase I
 - Expect to use FTS!
- This deployment will require from LCG:
 - Multiple LFC catalogs at sites (single central catalog no longer sufficient!)
 - Support for VO-site services...

Phase II: Grid Catalogs

- At each site (or tier):
 - LRC with POOL FC interface
 - (evaluating LFC and Fireman during Phase I)
 - Each LRC contains “site” files only
- ATLAS catalogs:
 - Global Dataset catalogues
 - Deployed centrally at CERN on ATLAS machines
 - An experiment-specific location catalog does mapping from GUIDs to Site LRC
 - Interfaces to RB and to ATLAS ProdSys

Phase II: VO-site services

Preliminary

- LCG: FTS, LRC with POOL FC interface
 - Expect LFC (Oracle at CERN, MySQL everywhere else?)
 - Messaging system: TBD - common with other experiments
- Experiment-specific S/W versions not defined yet: ATLAS will provide them later
- Set of experiment agents/services running at each site
- Require inbound and outbound connectivity
 - Restricted ports ok
- No data transfer in these services
- Designed to be “TCP traced”: XML messages. Minimal traffic

VO-site services

Preliminary

- Being designed to be deployed as jobs; (currently they run as cron jobs under a user account); BUT require a persistent area:
 - directory on machine for config files
 - MySQL server?
- Experiment should control deployment (e.g. “special job” that stays running indefinitely)
- Minimal disk space requirements: few megabytes
- WNs at site access these services (publish environment var)
- Services running centrally on ATLAS machines access these services as well
- Purpose of services: management of datasets, experiment-specific bookkeeping; Avoids outbound connectivity requirements for WNs; Greatly improves flexibility for the experiment to design Production infrastructure; Jobs running local have all required information available at the site

Phase II: Improvements

- Mostly improvements on ATLAS infrastructure, but...
- Data Transfer:
 - Currently movement is planned per individual file (although these can be grouped “randomly” for bulk operations)
 - Will switch to movement of datablocks
 - A datablock is the unit for all data access and data movement
 - Avoid small files
- Locating data and submitting jobs:
 - Based on datablocks
 - Will improve staging problems
- Predefined data flows:
 - DQ2 subscribes sites to datasets
- Sites will be able to monitor “expected data transfers”
 - Using subscriptions, they can know in advance how much data they will receive