

ATLAS Sites Jamboree

5 Mar 2019, 09:00 → 8 Mar 2019

<https://indico.cern.ch/event/770307/>

Changes expected in 2019 and beyond

- Evolution of the analysis model -> change the storage access patterns
 - Unique format for DAOD file
- Atlas will be using more Fast simulation (will change the CPU usage)
 - ~10x higher event throughput
- More events simulated, HITS to become “cheaper”, do we still need them on tape or can we just
- flush them away? Move towards a FastChain (eg EVNT --> AOD) model
- Monte Carlo overlay will change the way we run digi+reco jobs
 - Rather than use N files for pileup input, use only one pre-mixed file
 - Less disk I/)
- AthenaMT jobs expected on the grid soon

Containers

- Containers are being enabled
- Requirements
 - **At CentOS7 sites singularity is a requirement, should be installed on CentOS7 WNs by June 2019**
 - Current baseline: singularity 2.6.1
 - Default configuration works for ATLAS
 - 3.x series is not yet functional for ATLAS
 - Don't install it
 - No HC test yet, but we can run standalone container to test minimal functionality
 - If you have GPUs that you can put online let atlas know; they will be added to the current tests
 - Containers Deployment twiki
<https://twiki.cern.ch/twiki/bin/view/AtlasComputing/ADCContainersDeployment>
- Setuid/user namespaces
 - From CentOS7.6 user namespaces can be enabled without tweaking the kernel
 - Can be enable them with sysctl
echo "user.max_user_namespaces = 15000" > /etc/sysctl.d/90-max_user_namespaces.conf
sysctl -p /etc/sysctl.d/90-max_user_namespaces.conf
 - To use them in singularity you need to switch off setuid or remove the setuid binaries
 - If you remove setuid some functionality may not work anymore
- Overlay/underlay
 - Both mechanisms to bind a directory that doesn't exist in the image.
 - We cannot guarantee user images will have all the directories
 - Sites with caches will need them
 - ATLAS needs one of the two enabled or both
 - Overlay needs setuid enabled to work
 - Underlay works also in user namespaces
- AGIS
 - To enable the pilot containers functionality
 - **container_type: singularity:pilot**
 - In the future it will be possible to replace singularity with other rootless runtimes like podman or dockerd but for now atlas stick to it
 - **container_options: <options>**

Sites configuration - overview and suggestions

The responses from the sites were overwhelming, a lot of interesting information - a more detailed follow up is needed to process it all

- ATLAS Sites are much more diverse than we had expected
 - Especially in architecture, many sites are HPC like
- Most of the sites are following the ATLAS and WLCG recommendations
 - But it's hard to achieve 100%.
 - ATLAS might not be able to support everything in the future (eg using Centos 6 sites if production releases require it)
- New compute, storage and network technologies are coming fast. In addition to experimenting and testing, the recommendations and best practices need to follow soon.
- **Migration to CentOS 7**
 - Many sites already migrated in 2017 and 2018
 - Most of the remaining RH6 sites plan to migrate to CentOS 7 by June/July 2019
 - Many Big T1s and T2s are providing the resources through containers
 - transparently - OS can be chosen on demand
 - Few sites would like to migrate everything, but there are difficulties with sw availability, eg Storm on CC7
 - **Compute nodes are critical for ATLAS – containers**
- **Nodes, Batch**
 - Most of the sites have migrated to **HTCondor or SLURM**
 - some are planning to do so this or next year - in some cases it is coupled to migration to CentOS 7
 - **Other batch systems are considered deprecated by ATLAS**
 - Some sites are planning to migrate to Kubernetes/OpenStack
 - Both compute and services
 - Few sites use VAC
 - Under discussion how to use them with Harvester

Sites configuration

- **Computing Elements**
 - **The recommended CEs for ATLAS are CondorCE and ARC-CE**
 - There are many sites that still use Cream-CE, ~1/2
 - The migration to recommended CEs is non-trivial
 - EGI accounting
 - multi-VO support, customizations and custom gateways between CE and batch
 - Main reasons to migrate:
 - Support for modern features, eg cgroups, GPUs, containers...
 - Dynamic resource allocation: score vs mcore, and shares: analysis vs production
 - CE software development and support
- **Storage Elements**
 - Most of the sites use DPM, dCache and Storm
 - Some EOS and Ceph Objectstore
 - Objectstores have limited SE support for now
 - Custom developments to provide gridftp and xrootd
 - No clear recipe how to integrate them transparently in ATLAS DDM
 - Protocols:
 - Many sites support >5 access, transfer protocols
 - **ATLAS is using gridftp, https and xrootd**
- **Local storage**
 - Many sites have nodes with local disk without shared filesystem – WLCG standard architecture
 - Several large sites (~20) are shared with non-LHC activities and provide HPC-like architecture
 - GPFS, Lustre, CEPHFS
 - Typically with Storm if used for Storage Element
 - Some sites would like to use more modern storage, eg CEPH Objectstore or EOS
 - Not quite clear how to transparently integrate it in ATLAS
 - Funding concerns on RAW vs useful storage if erasure coding is not used

Sites configuration

- **IPv6 Deployment**
 - Most of the Storage Elements are accessible through IPv6
 - Nodes and grid services - less sites (~2/3)
 - This was not a WLCG requirement
 - Some sites have difficulties with deployment:
- **Network connectivity**
 - Most of the sites are connected with 10Gb/s or 20Gb/s WAN links or even faster,
 - Some sites reported frequent network saturation and expressed a wish to further optimize the transfers, data and job placement
 - However: many smaller sites did not report and they likely have 1Gb/s or less
 - Some sites will decommission LHCONE while upgrading to higher GPN throughput
 - Several sites have fast network (10 or 25Gb/s ethernet) or fast interconnects between the nodes and to the Storage
 - Typically on HPC-like shared facilities
 - Using MPI might be feasible on more sites than we expected
- **Containers**
 - Many sites have already deployed singularity, some docker:
 - Running pilots inside containers, eg modern OS with centos6 image/fs
 - Provisioning of batch nodes with docker
 - Some have it standby
 - ATLAS plans to:
 - Run pilot2 on host OS
 - Execute the data staging and payload execution inside middleware and custom payload Containers
 - Deploy singularity everywhere: (2.6.*)
 - Possibility to run it directly from cvmfs if OS supports unprivileged namespaces
- **Storage federations and diskless sites**
 - Most of the sites do not plan to federate the storage (yet?)
 - concerns on WAN saturation
 - concerns on high IOPs to affect storage servers if direct I/O is used
 - Some sites are diskless and connected to remote (close) Storage Element
 - Either direct access to SE from the nodes
 - Through ARC-cache and XCache - positive experience, but requires shared filesystem for
 - ARC-cache or dedicated storage servers for Xcache
 - Many sites participate in DOMA activities

DOMA TPC

- Phase 2 (deadline 30 June 2019): All sites providing > 3PB of storage to WLCG experiments are required to have one non-GridFTP endpoint in production.
- **xrootd**
 - Xrootd 4.9.0 is now in EPEL testing
 - Delegation protocol documentation for dcache has been released
 - DPM checksum
 - Native xrootd http bearer tokens
 - Instructions for different storages on how to enable it will be added to the twiki
 - For native xrootd, DPM, EOS will work out of the box
 - For ceph RAL is working on the plugin
- **http**
 - Works with bearer tokens
 - Functional tests quite green
 - Group of sites grown quite large >24 sites are participating
 - Hit kibana heat map display limits, we will have to split it
 - Still discussing how to split it
 - Paul Millar wrote a smoke test that performs a series of atomic tests for each DNS entry of a storage
 - Send a daily summary of the situation
 - Quite useful to chase sites
- **Policies**
 - For functional tests we will keep on adding sites that want to participate
 - However feeling that some sites are not proactive in fixing problems
 - Universal truth sites have to be chased but now we will give them a week to solve the problem if it can be solved otherwise they will be dropped
 - Testbeds welcome
 - For scale tests
 - We need production sites
 - For this phase large sites
 - They'll need to upgrade to the latest releases of their storages to support all the functionalities

DDM status

- Development of new tools to help reducing the number of transfer errors and detect errors proactively
 - Errors
 - **Lost files/Suspicious**
 - Lost/corrupted files has been a recurrent problem for many months. Produced from different sources :
 - Rucio (double submission to FTS)
 - FTS (2 FTS servers submitting the same transfers at the same time)
 - Panda/arc (double job submission)
 - Suspicious replicas (i.e. files that failed to be transferred many times) :
 - New WebUI to find the suspicious files/declare them bad
 - https://rucio-ui.cern.ch/suspicious_replicas
 - Automatic recovery enabled for suspicious files with more than 1 replica
 - Log files automatically declared bad if declared suspicious too frequently
 - **Files temporarily unavailable**
 - Still at an early stage, Tested on 2 sites so far
 - Feature requested by some sites last year. Available in the last Rucio feature release
 - Possibility to declare O(1M) files in less than a minute
 - One daemon (horizontally scalable) takes care of changing the state asynchronously
 - The temporary replicas can become automatically online after a defined timeout
 - Same permissions than the one to declare bad replicas
- https://twiki.cern.ch/twiki/bin/view/AtlasComputingDDMDarkDataAndLostFiles#Temporary_unavailable_files
- **Don't forget to provide monthly (or at least quarterly) storage dumps**
 - **Network**
 - Network seems to be one frequent issue (~10% of the failures are due to timeouts)
 - During the last months saturation observed leading to timeouts
 - Ideally, Atlas wants to put some minimum network requirements (as for the disk space) for the sites hosting data so that we have no blackholes
 - Sites not fulfilling these requirements could become lightweight sites
 - **Space reporting**
 - WLCG is now pushing for json space reporting.
 - ATLAS DDM infrastructure is able to use this new json as well as the old ATLAS one
 - If you want to deploy this json, please do it
 - for DPM sites it requires to run in DOME mode
 - **Space management**
 - During large deletion campaigns Atlas see problems because the deletion is not fast enough
 - Problem due to very small files O(10MB)
 - The deletion agent (reaper) in its current configuration cannot delete more than 16M files/days
 - Work ongoing to allow the reaper to be faster in case of high deletion acti
 - **Other problems** a detected by pROblem Detector (aka ROD)
 - IPv6 (Routing, firewall problems)
 - Deletion failures (in particular DPM)
 - Space reporting