(Biased) report from the ATLAS sites Jamboree

Frédéric Derue, LPNHE Paris

Calcul ATLAS France (CAF) meeting
CC-IN2P3 Lyon, 1st April 2019

ATLAS sites jamboree and HPC strategy, 5-8th March, CERN [indico]
HSF and WLCG meeting, 18-22nd March, JLAB [*indico*]

HPC strategy and Sites Jamboree

- HPC strategy → first 1.5 days. (Too) large attendance for the room
  → can have a look at the summary on Friday by A. Filipcic [link]
  → dedicated session this afternoon on HPC France/Europe

- Sites Jamboree

Goals of this ATLAS Sites Jamboree

- Idea behind is to checkpoint on
  - where we are
  - what we are we doing
  - where we are going
- Share ideas - know which activities (production and R&Ds) are ongoing
  - Involve more people
  - Get critical mass
- Get people together
  - Knowing each others simplify the communication
  - (might) boost productivity
- Get feedback from Sites to ADC (and from ADC to Sites)
  - Understand and agree on directions
- ...and more...

For a nice resume of what has been done in 2018 see intro talk of J. Catmore and D. Constanzo [link]

French present or remote :

physicists : Fred, Laurent, Luc, Sabine, Stéphane

engineers : Edith, Emmanuel, Eric, Frédérique, Manoulis

ADC & sites configuration & recommendation (1/3)

https://twiki.cern.ch/twiki/bin/view/AtlasComputing/SitesSetupAndConfiguration

- see talk on wednesday by A. Filipcic [link]
  → questionnaire sent to sites (~55 sites answered – what about FR sites?)

- Migration to CentOS7
  → most of the remaining RH6 sites plan to migrate to CentOS 7 by June/July 2019

- 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

- Computing Elements
  → the recommended CEs for ATLAS are CondorCE and ARC-CE

- Storage Elements
  → most of the sites use DPM, dCache and Storm
  → ATLAS is using gridftp, https and xrootd

ADC & sites configuration & recommendation (2/2)

- IPv6 Deployment
  → most of the Storage Elements are accessible through IPv6
  → some sites have difficulties with deployment / cannot do it in the near future

- Network connectivity
  → most of the sites are connected with 10Gb/s or 20Gb/s WAN links or even faster, many considering upgrading to 100Gb/s in the next 2 years

- Non standard resources
  → surprisingly, a large number of sites have deployed or could offer access to GPUs (18 sites from few to few 100)
  → ATLAS encourages sites to provide GPUs

- Containers
  → Deploy singularity everywhere: (2.6.*)

- Storage federations and diskless sites
  → most of the sites do not plan to federate the storage (yet?)
  → some sites are diskless and connected to remote (close) Storage Element
  → many sites participate in DOMA activities

ADC & sites configuration & recommendation (3/3)

- Many sites said bulk of the problems are caused by hardware, power failures, network outages - not much can be done on ADC side to help with that

- Various comments:
  - Improve documentation
  - More automation to identify site and central service issues
  - Distinguish pilots running on sub-clusters (opportunistic, cloud bursting...)
  - Payload distribution optimizations, throttling high I/O or network
  - Providing containers centrally for site services, storage
  - Control the network and join the effort with other VOs, non-LHC
  - Support for AAI

- To be followed up by ADC

Storage evolution

Dedicated talk by S. Campana at WLCG 2019 *[link*]

We would like to understand how national efforts see storage evolving over the next 5 years

- Motivations and constraints:
  - Budget pressure increasing for HL-LHC (and before that …): higher lumi, flat funding at best, declining density laws
  - Recent technologies studies show opportunities for efficiency increase (caches, QoS)
  - Site dependencies and constraints (including non WLCG ones) to be considered evaluating cost effective solutions
- Storage Consolidation for T2 federations. What are the plans at national level? Options could be:
  1. Continue with the current scenario, growing up the size of storage of each center with flat budget
  2. Consolidate storage in fewer centers, supporting the whole federation. Some centers will run CPU only (+ caches?)
  3. Expose fewer storage instances but distributed across the federation. E.g. NDGF
  - Of course you can (and likely will) have mixed models

Storage evolution

Dedicated talk by S. Campana at WLCG 2019 [*link*]

Inquiry

- As part of this process, WLCG Operations Coordination would send out a set of questions to sites/federations
- Not concretely discussed, but questions might be along those lines:
  - How much effort is required to support storage at the site? What is storage capacity?
  - Which storage middleware implementation the site is using? Is there a plan to migrate to other implementation in the coming years?
  - Would the site be interested or is the site being pushed to follow one of the scenarios described at slide 3? If yes, which one?
  - Would the site be interested to take part in the prototyping/testing activity?

Lightweight sites (1/2)

- 4 presentations on this subject on thursday [*link*]:

- Recommendation (ICB-2018) to redirect funding from storage to CPUs for lightweight Grid site (2018 limit : 460 TB, 2019 (+15%): 520 TB)
  - Already implemented on volontary basis on 5 sites also had small amount of CPUs (<200 cores)
    - 3 T2s with pledge CPU resources (HEP-UIBK, RO-14,RO-16)
    - support is now close to zero (and no worry about future storage migrations)

  Question : Can this setup be extended to larger Grid sites (200-1000 cores)

- Focus activity to low IO production jobs
  - Lightweight Grid site represents ~15k (5%) of Grid CPU capacity
    → Transforming to diskless would not affect significantly high IO jobs processing
  - Low IO → Almost not affected by network occupancy → Lower operational burden
  - No more analysis queue (currently~2%)→No interest to keep DATADISK
  - **Migrating to ATLAS@Home/BOINC would be even simpler solution** (monitoring report issue solved recently)

- Up to the cloud or country coordination to drive reorganisation
  - Convince FA that it is optimal usage of funding

Lightweight sites (2/2)

- Actions from site admins (minimising changes requested to sites) :
  - Close local analysis queue and reallocate slots to Production
  - Close Grid storage
  - If site wants to keep storage for some time, reallocate all Grid storage to SCRATCHDISK and contribute to storage of second replicas of user outputs (lifetime of 2 weeks → storage decomissioning could be fast)

- Since more sites are expected to become diskless, necessary to setup ATLAS site configuration to enable reasonable level of high I/O jobs according to network connectivity
  - Stay diskless : All informations are available (network connectivity per experiment, data flow rate per job type) but requires implementation in the workflow
  - Use cache if necessary to smooth network usage (ARC-CE, xcache,...)
  - Usefull configuration for commercial cloud or HPC

- Other options
  - Federate storage :Reduction of manpower but responsibility of site admins to make it transparent to application (network, downtime)
  - Dynafed : One entry point for the experiment but sites keep their own management

DOMA R&D : Place to discuss/test/evaluate different options

ATLAS blacklisting

- Dedicated presentation by J. Caballero [*link*]

This talk is just a summary...

This talk only provides for an summary of policies, monitoring tools, etc. For more detailed information relevant for site admins, check the twikies or contact the experts:

HammerCloud:	https://twiki.cern.ch/twiki/bin/view/IT/HammerCloudTutorialATLASsiteAdmins
	atlas-adc-hammercloud-support@cern.ch
For DDM Support:	atlas-adc-ddm-support@cern.ch
Switcher:	https://twiki.cern.ch/twiki/bin/view/AtlasComputing/SwitcherBlacklisting
	atlas-adc-autoexclusion-support@cern.ch

Monitoring and analytics (1/2)

- 2 dedicated presentation by T. Beermann [*link*]

- DDM Transfer Monitoring
  - *DDM Transfers* is the main dashboard to be used in day-to-day to work
    - it includes the efficiency matrix, tables/ plots / details for transfer / staging / deletion
    - additionally, includes rucio and FTS submission queues
    - data has a 1 minute granularity with a 30 days retention
  - *DDM Transfers (Historical Data)*
    - includes the same plots as the main dashboard without the matrix and the queues
    - minimal granularity is 1h

- DDM Accounting
  - global accounting for all pledged DATADISK and TAPE (all DISK is coming) which are updated weekly and can be used for logical and physical accounting but there is no split by RSE available
    - *DDM Global Accounting (Snapshot)* shows the current accounting numbers for the past week
    - *DDM Global Accounting (Historical)* shows the evolution of volume and files. Currently data available since beginning of last year
  - Site accounting is meant as a replacement for the old DDM accounting
    - *DDM Site Accounting* provides daily physical accounting per RSE. A lot of additional groupings like datatype, account, campaign, topology available

Monitoring and analytics (2/2)

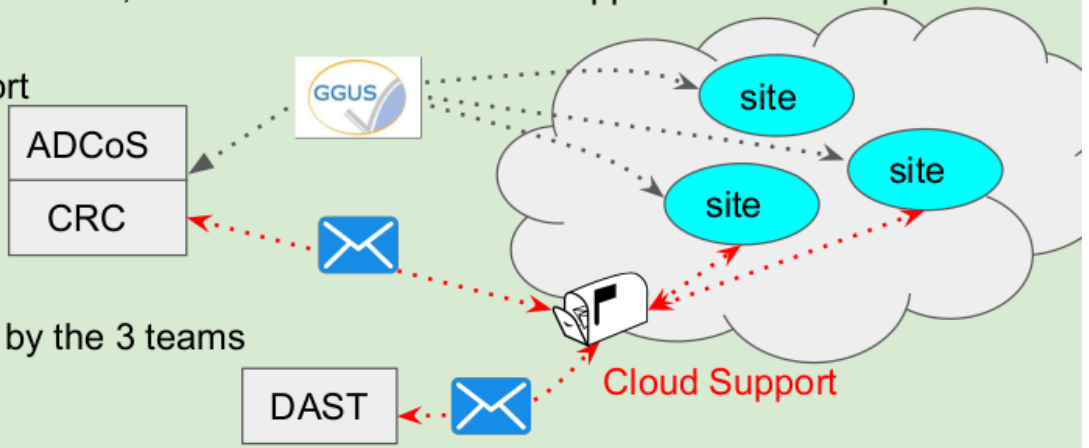
- Job Accounting
  - all plots available in one dashboard *Job Accounting (Historical Data)*
    - provides plots for submitted, pending, running, finalizing and completed jobs
    - pledges have been added
    - additional plots available for cpu consumption / efficiency, processed data, success/failure rates and resource utilisation

- *ADC live page* started to use new monitoring

ADC Shifts (1/2)

- Dedicated talk by M. Kataoka [*link*]

- ADC shift teams
  - CRC (Computing Run Coordinator)
    - coordinates daily ADC operations
  - ADCoS (ATLAS Distributed Computing operations Shifts)
    - checks/tickets distributed production (failing transfer, job/task errors, sites etc.)
  - DAST (Distributed Analysis Support Team)
    - provides the first line of support for ATLAS physicists doing distributed data analysis
- Communication between shifter and a site
  - CRC/ADCoS use GGUS for major issue , and CRC mails to Cloud Support and ADC experts if critical.
  - DAST uses email to Cloud Support
- In this talk,
  - Status of shift/support coverages by the 3 teams
  - Report from each team



CRC livepage [*link*]

ADC Shifts (2/2)

- CRC
  - To coordinate ADC operations, make excellent use of various communication tool (e.g. daily meeting,chat room, ggus, hypernews,elogs) not to miss any issues.
  - Build CRC livepage to ease the CRC shift.
- ADCoS
  - ready to transition to New DDM dashboard after many active discussion with developers.
  - Some discussion on shifter's tools with developers are on-going (SAM tests,GGUS,new SSB shifter view,new OTP Reports tool)
- DAST
  - Panda HowTo (pathena, prun)/FAQ twikis have been dramatically updated.
  - Suggestion for future improvement
    - To find/solve a problem by user-self, will provide some guides: flowchart, troubleshooting etc
    - A joint-tutorial by BigPanda and DAST for PandaClient and BigPanda Monitor in Physics Weeks.
    - To contact directly to a site admin, Can DAST escalate a site issue to ADCoS shifters who car ticket a GGUS ? (To be discussed with the ADCoS coordinators).

- contact cloud support only if their contribution is necessary and include site admins in the loop from the beginning ((emails in AGIS)
- or contacter directly the site via GGUS (or mail)

Database

- Dedicated talk on subject [*link*] and parallel session - thanks Andrea for the summary !

1) Activities for Run3

COOL will be in use for Run3 but we are preparing for the future, coordinating with system experts to simplify the present COOL storage methods.
Progress in the tools for the migration to Crest.
Studies and activities on going (based on Frontier monitoring results) to improve the caching of our infrastructure (Frontier/SQUID) by introducing more cacheable requests on the client side.

2) Frontier infrastructure updates since Dec 2018

New Squid-4 deployed at CERN
New infrastructure at RAL, with 3 new services which will replace the old ones (RAL will dismiss in the future the Frontier infrastructure because they abandon Oracle)
Proxy autoconfig in progress: it is to avoid “wrong” choice from clients (a recent problem was seen on CI jobs in Meyrin, choosing Wigner SQUIDs) . Backup proxy is in production.

3) Oracle

New version 18c will be deployed in Q3 2019.
Goldengate support under discussion: evaluate the impact in view of renewal of Oracle licensing for 2023.

4) Eventindex

Working generally smoothly with lots of data now. Developments on going to redesign part of the infrastructure in view of Run3.
A dedicated workshop at LAL is planned for beginning of June.

Network

- Dedicated talk on subject [*link*] → mostly on Run-3

Network Capacity Estimation for Run3

- Assume analysis on an HT-Core (job-slot) consumes 1.2 MBytes/sec
  - Implies job-slots need that level of network bandwidth to storage
  - WAN access to remote storage at 20% (ATLAS avg now)
    - 10 PBytes/day, 8 PB LAN, 2 PB WAN from Mario
  - **Minimal Tier-2**: 1000 job slots => 1.2 GBytes/sec, WAN 1.6 Gbits/sec
  - **Nominal Tier-2**: 5000 job slots => 6 GBytes/sec, WAN 9.6 Gbits/sec
  - **Leadership Tier-2**: 10000 job slots => 12 GBytes/sec, WAN 19.2 Gbits/sec
  - **NOTE**: Run-3 will have 3-4 times the data...have to either increase cores or improve average software throughput by that factor
- **Summary Network Capacity Recommendation:**
  - Average numbers above need a burst capability, assume x3
  - **Minimal Tier-2 WAN**: 1.6 Gbps x 3 = 4.8 Gbps => **10G link**
  - **Nominal Tier-2 WAN**: 9.6 Gbps x 3 = 28.8 Gbps => **40G link**
  - **Leadership Tier-2 WAN**: 19.2 Gbps x 3 = 57.6 Gbps => **80G link**

7