UiO : University of Oslo

Summary of recommendations for sites

David Cameron

ATLAS Site Jamboree, 20 Jan 2017

Introduction

- Memory
- Batch systems
- Shares
- CEs
- WNs hardware
- Storage
- Agis
- Squid
- Traceability

- Harvester
- Hammer Cloud
- Virtualisation/containers
- CentOS7
- Monitoring
- DDM
  - Protocols
  - New movers
  - Tapes
  - Lightweight sites
- Event service
- Networking

A lot of high-level info can be found here:
https://twiki.cern.ch/twiki/bin/view/AtlasComputing/SitesSetupAndConfiguration

Worker nodes

- Memory - still 2GB/core
  - Don’t kill on vmem
- Scratch - 20GB/core or 100GB/8 cores
- Network – 0.25MB/s/core
- Software – HEPOS_libs and cvmfs is all you need

OS/Virtualisation/Containers

- CentOS7 is still not recommended ☹
  - ATLAS is ready but middleware (EMI WN) is not
- Many (most?) sites using some virtualisation like Openstack
  - We don’t see this so don’t worry about it
- Containers:
  - Singularity looks promising, useful to manage CentOS7 transition

Batch system/CE

- A modern batch system (SLURM, HTCondor) makes things easier for ATLAS
- Requested fair shares
  - Analysis: T2: 25%, T1: 5%
  - Production: T2: 75%, T1: 95%
    - SCORE: 20%
    - MCORE: 80%
  - Dynamic SCORE/MCORE partitions are recommended
- CEs
  - HTCondor-CE (US) and ARC CE (rest of world) are becoming standards

DDM/Storage

- Consolidation of small storage is encouraged
  - Use an alternative remote SE
  - Or become a cache site (xrootd/ARC cache)
- Non-SRM disk SE is now possible
  - ATLAS still asks for GridFTP, HTTP and Xrootd to be supported
  - A standard space reporting method is evolving
- Tapes
  - We will work on file sizes

Harvester/Event service

- ATLAS wants to better use the sites
- For this it needs more information
  - Info will be taken from the CE or pilot
  - This will be easier on the more modern CE/batch systems
- One consequence should be fewer (visible) panda queues per site
  - Ideally one
- Event service
  - Currently being commissioned on Grid sites
  - If you have preemptable queues participation is encouraged

Networking

- No significant changes in usage foreseen for the rest of Run-2
- Usage is heavily influenced by job brokering strategy, recently this has improved a lot
- MONARC has really gone, site “closeness” is based on actual measurements
- For Run-3, increase will be required, 100Gbps is probably ok for Tier-1s
- IPv6
  - Sites may provide IPv6-only worker nodes after 1st April 2017
  - All ATLAS services will be dual-stack by then
  - Full site IPv6-only: probably 4 years away at least

Monitoring

- Current dashboards will stay until new infrastructure is ready
  - But please try to use the new one before you are forced to
- Heavy development and commissioning of new dashboards ongoing
  - New framework is very flexible (maybe too flexible)
  - Custom dashboards are useful but 1 per person is probably too much
  - ADC will work on official/validated dashboards

Summary

- Thank you for your attendance and continued close cooperation with ADC!