Lightweight sites
- introduction

GDB, 11th May 2016

Maarten Litmaath
CERN

v1.1

Session description

- http://indico.cern.ch/event/394782/
- Lightweight sites - ongoing activities, plans and ideas

One of the goals of WLCG Operations Coordination activities is to help simplify what the majority of the sites, i.e. the smaller ones, need to do to be able to contribute resources in a useful manner, i.e. with large benefits compared to efforts invested.

Classic grid sites may profit from simpler mechanisms to deploy and manage services. Moreover, we may be able to get rid of some service types in the end.

New sites may rather want to go into one of the cloud directions that we will collect and document.

There may be different options also depending on the experiment(s) that the site supports.

There is no one-size-fits-all solution. We will rather have a matrix of possible approaches, allowing any site to check which ones could work in its situation, and then pick the best.

In this session we present activities already ongoing or planned, and we will look for ideas worth pursuing in a task force.

Boundaries   (1/2)

- **This session is about small sites**
  - T0 and T1 are not directly targeted
  - But may profit from common simplifications

- **Build on earlier discussions and ideas**
  - Optimisation of operational costs, presented in Okinawa
  - Resource provisioning, ditto
  - Several sessions in the WLCG workshop in Lisbon

- **Storage and data access mostly handled elsewhere**
  - Various demonstrator projects e.g. presented in the April MB
  - Federations, caches, Ceph, …

Boundaries   (2/2)

- Here we are more concerned with services needed to enable computing at a site
  - CE
  - Batch system
  - Cloud setups
  - AuthZ system
  - Info system
  - Accounting
  - CVMFS Squid
  - Monitoring

- Storage service deployment may also profit from generic simplifications pursued here

T2 vs. T3 sites

- T3 sites typically dedicated to a single experiment → can take advantage of shortcuts
  - Can be pure AliEn / DIRAC / … sites
  - E.g. AliEn integration with OpenStack (Bergen Univ. Coll.)
  - …

- T2 sites have rules that apply
  - Accounting into EGI / OSG / WLCG repository
  - EGI: presence in the info system, at least for Ops VO
  - Security regulations
    - Mandatory OS and MW updates and upgrades
    - Isolation
    - Traceability
    - Security tests and challenges

T2 simplifications

- **Reduce** the catalog of required services, where possible

- **Replace** classic, complex services with new, simpler portfolio

- **Simplify** deployment, maintenance and operation of services

- Some sites could offer a partial portfolio, e.g. just "worker nodes"
  - Need to get their contributions properly recognized

- More in Mikalai's talk

CE and batch systems

- CE flavors
  - ARC
  - CREAM
    - Said to be the most complex
    - Maybe needed for other VOs supported by the site
  - HTCondor
    - On the rise

- Batch systems
  - More variety
  - May not be easy to change at a site
  - The majority of sites still use PBS/Torque
    - Still the default in EGI
  - HTCondor on the rise
    - Default in OSG

- Reductions in the phase space would help

Accounting

- Can it be simplified for classic grid sites?
  - A site's APEL host assembles records from CE and batch system
  - APEL would benefit from fewer CE and batch system flavors to support

- ARC CE publishes directly into central APEL service

- HTCondor CE?

- Transition toward cloud systems ought to help

Cloud systems

- Tap directly into local cloud deployments at sites
  - But OpenStack etc. may not be so easy for them either

- Paradigm shift: batch slots → VM instances
  - Need to ensure proper accounting

- Experiments can handle this today, but:
  - would like to see less variety, fewer interfaces to deal with
  - do not want to become sysadmins of the acquired resources

- Other supported VOs may be unable to use such resources
  - Sites may still need to run their classic setups instead
    - Or even in parallel

- More in Andrew's talk

AuthZ and Info systems

- Authorization systems
  - Argus now is supported under the aegis of Indigo-DataCloud
    - Version 1.7 is close to official release
      - Important improvements
      - Already beta tested at CERN and Budapest
  - GUMS is a cornerstone of OSG sites since many years

- Information systems
  - Information System TF explores simplifications for WLCG
  - BDII services may no longer be needed for WLCG
    - Still required for other VOs and EGI ops monitoring
    - Aiming for less work to support WLCG VOs

Configuration

- Slow but steady move towards Puppet
  - YAIM still being used for many services at many sites
  - Some sites prefer Ansible, CFEngine, Quattor, Salt, ...
    - But they presumably know what they are doing!

- Small sites also need help there

- Shared modules?
  - Not a lot of evidence yet?

- DPM comes with self-contained mini Puppet
  - An idea for other services?

Simplified deployment

- Can we provide VM images that require little local configuration?
  - The easiest test case may well be the Squid for CVMFS
  - Complex services would benefit more
  - But is the idea realistic for them?

- Would containers be better?
  - Could even work at sites without a (compatible) cloud infrastructure

- Ready-made (HW +) SW solutions?
  - Deployed in a DMZ like perfSONAR
  - Remotely operated by experts

- More in the US-ATLAS and US-CMS talks

Better documentation

- Online video tutorials?

- Or rather improve the textual documentation?
  - Screenshots in some places might be sufficient

- Cost-benefit analysis?

Monitoring

- ## Sites ought to have local fabric monitoring
  - Nagios, Ganglia, …

- ## SAM test failures ought to raise local alerts
  - And be understandable, well-documented
  - Supported for Nagios and used by some sites

- ## Experiment-specific monitoring could support alert subscriptions
  - RSS, Atom, e-mail
  - MonALISA does that

Experiment approaches and views

- Natural for dedicated sites

- Try to identify / encourage commonalities?

- More in Alessandra's talk