

# Cloud pre-GDB Summary

Michel Jouvin  
LAL, Orsay  
[jouvin@lal.in2p3.fr](mailto:jouvin@lal.in2p3.fr)

March 2015

# Meeting Facts

- ⦿ Participation: 25 local, ~5 remote
  - > A few people could not come
  - > No experiment representatives
- ⦿ Agenda: <https://indico.cern.ch/event/319819/>
  - > Afternoon only
  - > Review of the work in progress since September
  - > Discussion with EGI FedCloud about possible collaboration
- ⦿ Summary will be produced
  - > <https://twiki.cern.ch/twiki/bin/view/LCG/CloudDiscussions>

# WG Goals

- ◉ WG is about exploring the possibility to use private/community clouds as a replacement for grid CEs
  - > No intention to cover all the aspects of cloud usage by experiments
  - > Focus on shared clouds rather than dedicated resources to one experiment
    - But many lessons can be learnt from private clouds...
  - > This perspective is part of the WLCG future directions presented/discussed at last WLCG workshops
    - Copenhagen and Barcelona
- ◉ Build on existing work in experiments: do not start a new huge R&D project
  - > No manpower available to do it
  - > Tackle all foreseen operational issues: scheduling, accounting, security...

# Target Shares...

- Ability to do a dynamic partitioning of a shared clouds
  - > Goal: achieve something similar to fairshare in batch systems
- 2 implementations presented at last pre-GDB (September 14)
  - > Vcycle (A. McNab): a service that is in charge of instantiating VMs according to VO target shares
    - Based on the Vac ideas, implemented as a cloud client
    - For pilot-based VOs: no attempt to contact the pilot framework, the VM will do it
    - Possibly cloud MW agnostic: will support EC2 and OCCl
  - > FairShareScheduler (Padova): a drop-in replacement for nova-scheduler, queuing of requests
    - OpenStack specific, based on SLURM scheduler
    - Trying to get it into mainstream OpenStack (BLAZAR)

# ... Target Shares

- Vcycle: impressive developments in the last 6 months
  - > Major redesign/rewrite for more commonalities with Vac
  - > Plugins for cloud backend: currently OpenStack and OCCl
  - > Images and “contextualization” data can be retrieved from a web server: no site intervention for image/params update
    - Include ssh key configuration to allow a site to log into the VM
  - > Target shares instantaneous: no history
    - To achieve target share on a long period, target must be adjusted
  - > Multiprocessor VM supported
  - > Accounting through APEL SSM: appear like a grid job
  - > Central syslog configured (machine features) for traceability
- FairShare Scheduler for OpenStack being redesigned as an External Manager
  - > Be less integrated into Nova code for better sustainability

# Security

- ◎ Cloud is a paradigm change compared to grid with respect to responsibilities sharing between sites and VOs
  - > VOs will have sysadmins responsibilities, e.g. VM image updates
  - > Change in liability: sites will not be liable for VM image problems
  - > WLCG: all experiments planning to use CernVM + CVMFS
    - Experiment specific components in CVMFS
- ◎ Cloud is not necessarily a major change in incident handling compared to today (pre-glexec) grid operations, **as long as traceability is ensured**
  - > Will need a cooperation between sites and VOs to get all the information needed for incident handling
    - Proved to work today in grid
  - > Security challenges
  - > Importance of discussion with VOs

# Traceability...

- ◎ TF about cloud traceability created last Fall
  - > Sites and experiments, propose “realistic” solutions
  - > First meeting (F2F) last month
- ◎ Focus on the WLCG use case
  - > Well controlled cloud usage: VMs connected to pilot factories
  - > Structured community
  - > Long standing trust relationship
- ◎ Learn from this experience to address more complex use cases later
  - > EGI FedCloud use case is much more complex

# ... Traceability

- Logging: the key component
  - > Should primarily depend on externally observable behaviours to allow correlation with other sources of information
    - Hypervisor, cloud MW, network
  - > Need ability to correlate and search logs: technology exists
  - > Sites with experience started to discuss usage of netflow for logging network traffic
  - > Quite some progress since 1 month
    - E.g. Vcycle work
- Quarantine of VM would help forinsics
  - > Not easy/possible with all cloud MW
  - > Storage impact may limit useability
  - > RAL looking at implementing quarantine in HTCondor



# Accounting...

- ◉ VM Benchmarking: required to normalize accounting
  - > Update on CERN approach based on a “lookup table”
  - > Issue: currently works only for batch WN
    - Ceilometer accounting doesn't provide any information on the VM machine type or hypervisor
- ◉ Status on APEL support for VM accounting
  - > Most of the work happened in the EGI FedCloud TF
  - > New Usage Record format adopted but not yet implemented
    - allows to store benchmarking metrics
    - Sites with several cloud instances
    - Report of image used
  - > Open issue: how to get periodic accounting information about long running VMs
    - Assign usage to the appropriate month

# ... Accounting

- ⦿ Accounting portal still requires some work to properly display cloud resources
  - > T1/T2 structure as for grid
  - > Combined grid/cloud views
    - Need to distinguish UR from grid and cloud
- ⦿ Accounting and pledges: matching them is a requirement for WLCG
  - > TF created in January, led by L. Field
  - > Use case and requirements defined
  - > A concrete work plan being finalized: should be ready soon
- ⦿ Potential issue: double counting between grid and cloud UR
  - > Vac particularly exposed: produces a grid-like UR
  - > Need to better define the use case and the problem

# DataBridge

- A new component developed for volunteer computing
  - > Give access from BOINC to a data storage common with grid
  - > In fact, an authentication bridge
- Based on http/Dav DynaFed + a S3 backend
  - > Apache manages the authentication: potentially any authentication method supported by Apache
    - Currently X509 (gridside), BOINC (mod\_auth\_sql, user/pwd)
    - Future: identity federation?
  - > S3 provides signed redirections: independent of the authentication method
    - Possible to use the same storage backend whatever authentication is used
  - > Trivial messaging system implemented as a web application
    - Allows to push requests to another component, e.g. FTS after job completion

# EGI FedCloud

- ◎ Part of EGI production infrastructure
  - > Based on standards for accessing/publishing the resources
  - > Integration with EGI security and operational tools
  - > Pervasive instead of focus on one specific use case/community
    - Services, HTC, workflows, high memory apps...
    - IaaS, PaaS, SaaS
- ◎ Area for possible collaborations with WLCG
  - > Accounting: already happening
    - APEL developments, cloud accounting TF
  - > Monitoring: current OCCl requirement is a potential issue
    - WLCG has currently no alternative monitoring infrastructure
  - > Federated identity: an important topic in EGI Engage and also a hot topic in WLCG
  - > VM image mgt: less interest for WLCG
    - CERNVM only: most things in CVMFS

# Conclusions

- Useful meeting: steady work in progress
- Dynamic sharing of resources: vcycle looks as a promising approach
  - > May be complemented by FairShare Scheduler for Open Stack sites
- Accounting: still a significant work to do
  - > Solutions identified/defined for most problems
  - > Double counting with grid accounting is a potential issue
- Traceability: work in progress, « affordable » solutions
- DataBridge offer a potential to bridge access to same storage from different infrastructures without perf penalties