

Draft Cloud Accounting Usage Record Schema v0.4

Draft 4 – 20/1/2015 – John Gordon

Some assumptions related to the version below:

1. Additions to v0.2 currently in use. There is a v0.3 in existence making proposals that have not yet been implemented.
2. CloudComputeService was chosen as a name to signify a level of granularity of cloud services within a site. cf ComputingService and StorageService in GLUE2.0. I am told it is the name currently in the working draft of an extension to GLUE 2 for clouds.
3. For normalization we took the name Benchmark from the OGF UR 2.0 <http://www.ogf.org/documents/GFD.204.pdf> For Grid CPU accounting, APEL currently uses ServiceLevel from CAR1.0. This means something else in UR 2.0.
4. We considered adding the profile name used when instantiating the VM but (in OpenStack) could not see how to retrieve this information. We believe that Memory, Disk, and cpucount captures the relevant information.
5. Imageld suggestions following a discussing within the Fed Cloud Task Force
6. This is a still a keyword-value pair version of the schema. Migration to a version based on OGF UR 2.0 is work for a future project.
7. No changes to time durations. Durations for running VMs are from starttime to the time the record was cut. This may change in future if we decide to change to a series of records covering a set of contiguous time periods in order to assign usage to (eg) months.

Assumptions in common with earlier versions

8. Many of the properties defined in the Usage Record may be set to NULL, this is based on the assumption that different Resource Providers may want to publish different sets of data, but we want only one Usage Record, usable by all.
9. The fields for GlobalUserName and FQAN relate to an existing method of authorizing grid users, if this authorization mechanism were to be used for cloud usage as well then these would provide a method of associating VM usage with a user and then querying usage in a portal could be based on user, VO etc.
10. I suggest that the VMUUID field contains a concatenation of CurrentTime, SiteName and MachineName. (Note: CurrentTime is the time the record was created on the client machine, Machine Name is the VMid)
11. We have job accounting and storage accounting systems and their related usage records already - we therefore account for jobs and storage using these existing records/methods.

The following table provides an revised draft schema v0.4 with semantic changes in bold and additional fields flagged NEW:

Cloud Usage Record Property	Type	Null	Definition
VMUUID	varchar(255)	No	Virtual Machine's Universally Unique Identifier concatenation of CurrentTime, SiteName and MachineName
SiteName	varchar(255)	No	GOCDDB SiteName - GOCDDB now has cloud service types and a cloud-only site is allowed.
CloudComputeService (NEW)	varchar(255)		Name identifying cloud resource within the site. Allows multiple cloud resources

			within a site. i.e. a level of granularity.
MachineName	varchar(255)	No	VM Id - the site name for the VM
LocalUserId	varchar(255)		Local user name
LocalGroupId	varchar(255)		Local group name
GlobalUserName	varchar(255)		Global identity of user (certificate DN)
FQAN	varchar(255)		Use if VOs part of authorization mechanism
Status	varchar(255)		Completion status - completed, started or suspended
StartTime	datetime		Must be set when Status = started
EndTime	datetime		Set to NULL until Status = completed
SuspendDuration	datetime		Set when Status = suspended (Timestamp)
WallDuration	int		WallClock time - actual time used
CpuDuration	int		CPU time consumed (Duration)
CpuCount	int		Number of CPUs allocated
NetworkType	varchar(255)		Needs clarifying
NetworkInbound	int		GB received
NetworkOutbound	int		GB sent
PublicIPCount (NEW)	int		Number of public IP addresses assigned to VM
Memory	int		Memory allocated to the VM
Disk	int		Size in GB allocated to the VM
BenchmarkType (NEW)	varchar(255)		Name of benchmark used for normalization of times (eg HEPSPEC06)
Benchmark (NEW)	Decimal		Value of benchmark of VM using ServiceLevelType benchmark'
StorageRecordId	varchar(255)		Link to other associated storage record Need to check feasibility
ImageId	varchar(255)		Every image has a unique ID associated with it. - For images from the EGI FedCloud AppDB this should be VMCATCHER_EVENT_AD_MPURI - for images from other repositories it should be a vmcatcher equivalent For local images - os_tpl mixin
CloudType	varchar(255)		Type of cloud infrastructure: OpenNebula; OpenStack; Synnefo; etc.