Report on Installed Resource Capacity

Flavia Donno
CERN/IT-GS

WLCG GDB,
CERN 10 December 2008

Outline

- Details about reporting installed capacity through the Information System for:
  - Computing
  - Storage
- The new version of the document is v1.8 publicly available
  - A few more comments received. They will be integrated in the next days
- Operational plan will be prepared starting in February 2009

Goals

- **Publishing information in order to provide the WLCG management with a view of the total installed capacity and resource usage by VOs at sites. Focus on four main use cases:**
  - Publishing information to provide the WLCG management with a view of the ***total installed capacity at a site***.
  - Publishing information to provide the WLCG and VO management with a view of the ***resource assignment per VO at a site***
  - Publishing information to allow VO operators to ***monitor the VO usage of the resources at a site***
  - Allowing clients (consumers of the information published) to ***select the correct resources and their characteristics***

Organization of the document

- **It includes an explanatory part where all involved Glue Schema v1.3 classes and attributes are explicitly described. Appendixes with specific instructions and examples.**
- **Logically organized in two parts.**
  - Computing (theory + appendix)
  - Storage (theory + appendix to come)
- **Explanation of commonly published attributes and recommendations**
- **Specific directions to publish specific Glue classes and attributes in order to satisfy the requirements stated in the goals.**

Computing Resources

- **Main Glue Classes. They are mandatory for WLCG sites**
  - CE <-> Cluster <- SubCluster (Host)
- **SubClusters:**
  - Are mostly hetherogeneous
  - For WMS matchmaking it is recommended to publish homogeneous queues using one Cluster and one SubCluster(Host)
- **Mandatory SubCluster attributes:**
  - PhysicalCPUs - Total number of real CPUs/physical chips in the Subcluster (includes offline/down nodes)
  - LogicalCPUs – Total number of cores/hyperthreaded CPUs in the SubCluster (includes offline/down nodes)
  - BenchmarkSI00 – Average SpecInt2000 rating per LogicalCPU

Computing Resources

- **Other recommended SubCluster attributes:**
  - ProcessorModel
  - ProcessorSpeed
  - SMPSize – number of Logical CPUs (cores) per WN
  - ProcessorOtherDescription: Cores=<typical number of cores per CPU>
  - RAMSize – Total physical memory of a WN in the SubCluster
  - VirtualSize – Total virtual memory of a WN in the SubCluster
- **CE attributes to be published**
  - AssignedJobSlots – Total number of assigned Job Slots in the queue available at a given moment
  - CECapability: CPUScalingReferenceSI00=<refCPU SI00> the CPU SI00 for which the published GlueCEMaxCPUTimes are valid
  - CECapability: Share=<VO>:<share> ; This value is used to express a specific VO share if VO shares are in operation (this attribute is not mandatory). Published only for WLCG VOs. Value between 0 and 100. The sum of the values over all WLCG VOs <= 100.

Computing Resources

- The ***installed computing capacity at a site*** is then expressed by the following formula:

$$F(WLCGSubCluster)_{Site} = \begin{cases} \sum_{WLCG\ VOs} GlueCECapability(Fairshare)/_{100} & if\ WLCG\ fairshare\ are\ published \\ 1 & if\ WLCG\ fairshare\ are\ not\ published \end{cases}$$

Then the total installed capacity at a site is given by the following formula:

$$Total\ Installed\ Computing\ Capacity_{Site}\ (KSI00) = \sum_{WLCG\ SubClusters} GlueHostBenchMarkSI00 * GlueSubClusterLogicalCPUs * \frac{F(WLCGSubCluster)_{Site}}{10^3}$$

Computing Resources: notes for admins

- **When a new head node is added at a site, YAIM configures automatically a new Cluster and SubCluster associated, causing double counting of resources.**
  - For the newly created Clusters and SubClusters insisting on the same resources, site administrators MUST set the GlueSubClusterPhysicalCPUs and GlueSubClusterLogicalCPUs attributes to 0.
- **Explicit instructions on how to set YAIM variables are given in section A.2 of Appendix A in the document.**
- **An explicit example of information published by a site with correction is given in section A.3 of Appendix A in the document.**

Storage Resources

- **Main Glue Classes. They are mandatory for WLCG sites**
  - SE <-> SA <- VOInfo
  - Service; SEControlProtocol; SEAccessProtocol
- **SE:**
  - UniqueID – The value SHOULD be the FQDN of the SRM host. One to one mapping between SE and SRM Service
  - Status – Production, Closed, Draining, Queueing
  - ImplementationName, ImplementationVersion
  - Architecture – disk, multidisk, tape, other
  - SizeTotal, SizeFree, TitalOnlineSize and UsedOnlineSize (GB)

Storage Resources

- **AccessProtocol**
  - LocalID
  - Type – Must be set to one of the allowed type. If secure, GSI MUST be published in *SupportedSecurity*
  - Version – Version of the protocol
- **ControlProtocol (MUST be published):**
  - Endpoint – The value MUST be the full SOAP endpoint for SRM: httpg://srm.example.com:8443/srm/managerv2
  - Type – SRM
  - Version – 1.1.0 or 2.2.0

Storage Resources

- **StorageArea (SA): describes a logical view of a portion of physical space that can include disks and tape resources. It can correspond to SRM static space reservations or not. SAs MAY overlap. One SA can be shared by several VOs.**
  - LocalID
  - AccessControlBaseRule – VO:<VO Name> or VOMS:<FQAN>
  - Reserved[Online|Nearline]Size (GB) – Space statically reserved to a VO. 0 if SA does not represent SRM reserved space
  - Total[Online|Nearline]Size (GB) – Less than or equal to ReservedSpace iff ReservedSpace > 0
  - Used[Online|Nearline]Size (GB) – Space occupied by available and accessible files not candidates for garbage collection. Used <= Total.
  - Free[Online|Nearline]Size (GB) – Total – Used.

Storage Resources

- **StorageArea: the Capability special attribute**
  - Installed[Online|Nearline]Capacity=<Size> (GB) – MUST be published. Introduced for accounting purposes only. It expresses the size of the physical space of a Storage Area. It can be 0 for overlapping SAs.
  - scratch – Storage Area is of type T0D0
  - stage – Storage Area is used for staging activities only

Storage Resources

- **VOInfo: describes VO specific atributes of a Storage Area. One SA can be associated to one or more VOInfo objects. *VOInfo represents the ability of a client/end-user to write into a particular Storage Area with a set of parameters specified by the VOInfo attributes*.**
  - LocalID – MUST be set
  - AccessControlBaseRule – VO:<VO Name> or VOMS:<FQAN>
  - VOInfoTag – Space Token Description
  - VOInfoPath– It describes the default path an application can use to write into the associated SA.

Storage Capacity

- The ***installed storage online capacity at a site*** is then expressed by the following formula:

Installed Online Capacity = $(\sum_{\text{WLCG GlueSA}}$ GlueSACapability(InstalledOnlineCapacitySize)

- The ***installed storage nearline capacity at a site*** is then expressed by the following formula:

Installed Nearline Capacity = $(\sum_{\text{WLCG GlueSA}}$ GlueSACapability(InstalledNearlineCapacitySize)

Validation procedures

- **Gstat:**
  - Computing resources sanity checks in preparation
  - Storage resources sanity checks already available but not yet in production
- **More accurate validation from executing jobs at sites**
- **SAM tests**
- **WNCM JobWrapper scripts**
- **Reports published**

Installed Capacity: status

- **The latest version of the document (v1.8-1) is published with all details, operational instructions for site administrators and explicative examples (OSG missing)**
  - https://twiki.cern.ch/twiki/pub/LCG/WLCGCommonComputingReadinessChallenges/WLCG_GlueSchemaUsage-1.8.pdf
- **Technical agreement on the essential content of the the document :**
  - OSG requested and was accorded one more month to digest the proposals.
- **Minor changes are still expected**
  - Operative instructions and examples from OSG needed

Installed Capacity: plan

- **Operational plan in preparation. We will start in February 2009 as soon as document formally approved.**
  - Precise instructions for site admins
  - Examples
  - Local test suites
  - Coordination for deployment
  - Many site admins ask for support. Support group needed.

- **Client tools: evaluating the effort for conformance to the document**
  - Gstat sanity checks.
  - GridMAP is being adapted.
  - Client tools gfal/lcg-utils will improve resource selection algorithm with next releases.
  - Reviewing SAM tests
  - Nothing yet done for APEL
  - OSG ?

Installed Capacity: plan

- **Coordination with WN working group**
  - Includes changes in YAIM and specific recommendations
- **Storage is mostly automatic**
  - GRIF and RAL have new info providers for DPM and CASTOR. dCache and StoRM expected in January.
  - dCache requires info about nearline storage to be set by hands.
  - *gstat sanity checks for storage ensure coherence of published information.*
- **Experience is needed in order to better plan**
  - A few sites are already publishing according to the document. This is really useful.