GLUE Schema: LDIF to old classad mapping

Gabriele Garzoglio

Computing Division, Fermilab

May 31, 2006

Overview

- Status
- Limitations
- Basic rules
- Extended rules
- Conclusions

Status

- On Mar 14, 2006 we have decided to work collaboratively on a document to map the GLUE schema to old classad format

- Since then, the document has undergone 3 reviews

- The document is divided in 2 set of rules for the mapping
  - Basic: Consider CE, Cluster, and Subcluster entities ONLY
  - Extended: Extend the basic mapping with rules for VOView and Location entities.

Limitations

- We are working on a mapping to facilitate resource selection

- We are NOT focusing on monitoring information (Job entity, etc.)

- We are currently NOT focusing on Storage Resources

Basic Rules (1)

- Attribute names are built using the same rules as for the LDIF mapping

- Each combination of inter-related ComputingElement, Cluster, and Subcluster entities forms a classad. Resources at a site are represented by a set of classads.

- Attributes of the CE, Cluster, and Subcluster entities that **CANNOT have more than one value** are translated as AttributeName = Value

Basic Rules (2)

- Attributes that **can have more than one value** appear in the classad only once. The multiple values are represented as a single comma-separated string of the original values.

- Every classad has the following three attributes added after the mapping is complete: MyType, TargetType, and Name (useful for Condor Match Making).
  [Should we keep this rule?]

Basic Rules (3)

- In some GIP implementations, the ContactStringInfo attribute of CE is not a GRAM URL and may have the form <GRAM-URL>-<VOName>. The <VOName> is stripped off.

- Values are quoted if they are of type "string" (see detailed rules).

Example: GLUE in LDIF format (extract)

dn: GlueCEUniqueID=rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf,mds-vo-name=local,o=grid
objectClass: GlueCETop
objectClass: GlueCE
objectClass: GlueSchemaVersion
objectClass: GlueCEAccessControlBase
objectClass: GlueCEInfo
objectClass: GlueCEPolicy
objectClass: GlueCEState
objectClass: GlueInformationService
objectClass: GlueKey
GlueCEHostingCluster: rsgrid3.its.uiowa.edu
GlueCEName: cdf
GlueCEUniqueID: rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf
GlueCEInfoGatekeeperPort: 2119
GlueCEInfoHostName: rsgrid3.its.uiowa.edu
GlueCEInfoLRMSType: condor
GlueCEInfoLRMSVersion: 6.7.13
GlueCEInfoTotalCPUs: 1
GlueCEInfoJobManager: condor
GlueCEInfoContactString: rsgrid3.its.uiowa.edu:2119/jobmanager-condor
GlueCEInfoApplicationDir: /export/grid/app
GlueCEInfoDataDir: /export/grid/data
GlueCEInfoDefaultSE: rsgrid3.its.uiowa.edu
GlueCEStateEstimatedResponseTime: 0
GlueCEStateFreeCPUs: 0
GlueCEStateRunningJobs: 0
GlueCEStateStatus: Production
GlueCEStateTotalJobs: 0
GlueCEStateWaitingJobs: 0
GlueCEStateWorstResponseTime: 0
GlueCEStateFreeJobSlots: 1
GlueCEPolicyMaxCPUTime: 0
GlueCEPolicyMaxRunningJobs: 0
GlueCEPolicyMaxTotalJobs: 0
GlueCEPolicyMaxWallClockTime: 0
GlueCEPolicyPriority: 0
GlueCEPolicyAssignedJobSlots: 1
GlueCEAccessControlBaseRule: VO:cdf
GlueForeignKey: GlueClusterUniqueID=rsgrid3.its.uiowa.edu
GlueInformationServiceURL: ldap://rsgrid3.its.uiowa.edu:2135/mds-vo-name=local,o=grid
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

dn: GlueClusterUniqueID=rsgrid3.its.uiowa.edu,mds-vo-name=local,o=grid
objectClass: GlueClusterTop
objectClass: …
GlueClusterName: rsgrid3.its.uiowa.edu
GlueClusterService: rsgrid3.its.uiowa.edu
GlueClusterUniqueID: rsgrid3.its.uiowa.edu
GlueForeignKey: GlueCEUniqueID=rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf
GlueForeignKey: GlueCEUniqueID=rsgrid3.its.uiowa.edu:2119/jobmanager-condor-fermilab
GlueForeignKey: GlueSiteUniqueID=rsgrid3.its.uiowa.edu
GlueInformationServiceURL: ldap://rsgrid3.its.uiowa.edu:2135/mds-vo-name=local,o=grid
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

dn: GlueSubClusterUniqueID=rsgrid3.its.uiowa.edu, GlueClusterUniqueID=rsgrid3.its.uiowa.edu, mds-vo-name=local,o=grid
objectClass: GlueClusterTop
objectClass: …
GlueChunkKey: GlueClusterUniqueID=rsgrid3.its.uiowa.edu
GlueHostApplicationSoftwareRunTimeEnvironment: OSG-0.3.6
GlueHostApplicationSoftwareRunTimeEnvironment: ATLAS_LOC_903
GlueHostApplicationSoftwareRunTimeEnvironment: LCG-2_6_0
GlueHostArchitectureSMPSize: 2
GlueHostBenchmarkSF00: 380
GlueHostBenchmarkSI00: 400
GlueHostMainMemoryRAMSize: 512
GlueHostMainMemoryVirtualSize: 1024
GlueHostNetworkAdapterInboundIP: FALSE
GlueHostNetworkAdapterOutboundIP: TRUE
GlueHostOperatingSystemName: linux-rocks-3.3
GlueHostOperatingSystemRelease: Rocks Linux
GlueHostOperatingSystemVersion: 3.3
GlueHostProcessorClockSpeed: 1000
GlueHostProcessorModel: Pentium III (Coppermine)
GlueHostProcessorVendor: GenuineIntel
GlueSubClusterName: rsgrid3.its.uiowa.edu
GlueSubClusterUniqueID: rsgrid3.its.uiowa.edu
GlueSubClusterPhysicalCPUs: 1
GlueSubClusterLogicalCPUs: 2
GlueSubClusterTmpDir: /export/grid/data
GlueSubClusterWNTmpDir: /tmp
GlueInformationServiceURL: ldap://rsgrid3.its.uiowa.edu:2135/mds-vo-name=local,o=grid
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

Example: 1 of 2 Old Classads (basic rules)

Computing Element Entity (1) [Classad for 2nd CE not shown]
GlueCEHostingCluster = "rsgrid3.its.uiowa.edu"
GlueCEName = "cdf"
GlueCEUniqueID = "rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf"
GlueCEInfoGatekeeperPort = 2119
GlueCEInfoHostName = "rsgrid3.its.uiowa.edu"
GlueCEInfoLRMSType = "condor"
GlueCEInfoLRMSVersion = "6.7.13"
GlueCEInfoTotalCPUs = 1
GlueCEInfoJobManager = "condor"
GlueCEInfoContactString = "rsgrid3.its.uiowa.edu:2119/jobmanager-condor"
GlueCEInfoApplicationDir = "/export/grid/app"
GlueCEInfoDataDir = "/export/grid/data"
GlueCEInfoDefaultSE = "rsgrid3.its.uiowa.edu"
GlueCEInfoGRAMVersion = "1.7"
GlueCEStateEstimatedResponseTime = 0
GlueCEStateFreeCPUs = 0
GlueCEStateRunningJobs = 0
GlueCEStateStatus = "Production"
GlueCEStateTotalJobs = 0
GlueCEStateWaitingJobs = 0
GlueCEStateWorstResponseTime = 0
GlueCEStateFreeJobSlots = 1
GlueCEPolicyMaxCPUTime = 0
GlueCEPolicyMaxRunningJobs = 0
GlueCEPolicyMaxTotalJobs = 0
GlueCEPolicyMaxWallClockTime = 0
GlueCEPolicyPriority = 0
GlueCEPolicyAssignedJobSlots = 1
GlueCEAccessControlBaseRule = "VO:cdf"
GlueInformationServiceURL = "ldap://rsgrid3.its.uiowa.edu:2135/mds-vo-name=local,o=grid"
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

Cluster Entity
GlueClusterName = "rsgrid3.its.uiowa.edu"
GlueClusterService = "rsgrid3.its.uiowa.edu"
GlueClusterUniqueID = "rsgrid3.its.uiowa.edu"
GlueClusterTmpDir = "/tmp"
GlueClusterWNTmpDir = "/scratch"

SubCluster Entity
GlueHostApplicationSoftwareRunTimeEnvironment =
    "OSG-0.3.6,ATLAS_LOC_903,LCG-2_6_0"
GlueHostArchitectureSMPSize = 2
GlueHostArchitecturePlatformType = "undefined"
GlueHostBenchmarkSF00 = 380
GlueHostBenchmarkSI00 = 400
GlueHostMainMemoryRAMSize = 512
GlueHostMainMemoryVirtualSize = 1024
GlueHostNetworkAdapterInboundIP = FALSE
GlueHostNetworkAdapterOutboundIP = TRUE
GlueHostOperatingSystemName = "linux-rocks-3.3"
GlueHostOperatingSystemRelease = "Rocks Linux"
GlueHostOperatingSystemVersion = 3.3
GlueHostProcessorClockSpeed = 1000
GlueHostProcessorModel = "Pentium III (Coppermine)"
GlueHostProcessorVendor = "GenuineIntel"
GlueHostProcessorVersion = "undefined"
GlueHostProcessorInstructionSet = "i686"
GlueHostProcessorOtherDescription = "i686"
GlueSubClusterName = "rsgrid3.its.uiowa.edu"
GlueSubClusterUniqueID = "rsgrid3.its.uiowa.edu"
GlueSubClusterPhysicalCPUs = 1
GlueSubClusterLogicalCPUs = 2
GlueSubClusterTmpDir = "/export/grid/data"
GlueSubClusterWNTmpDir = "/tmp"

Extended Rules

- Each combination of inter-related ComputingElement, Cluster, Subcluster and VOView entities forms a classad.

- VOView attributes that have the same names as the ones for CE overwrite the CE values (e.g. GlueCEStateFreeJobSlots).

- For every subcluster, every attribute in the Location entity appears once with the list of values expressed as a (ordered) list.

Example: GLUE in LDIF format (additions from prev.)

dn: GlueVOViewLocalID=cdf,GlueCEUniqueID=rsgrid3.its.uiowa.edu:2119/
jobmanager-condor-cdf,mds-vo-name=local,o=grid
objectClass: GlueCETop
objectClass: GlueVOView
objectClass: GlueCEInfo
objectClass: GlueCEState
objectClass: GlueCEAccessControlBase
objectClass: GlueCEPolicy
objectClass: GlueKey
objectClass: GlueSchemaVersion
GlueVOViewLocalID: cdf
GlueCEAccessControlBaseRule: VO:cdf
GlueCEStateRunningJobs: 0
GlueCEStateWaitingJobs: 0
GlueCEStateTotalJobs: 0
GlueCEStateFreeJobSlots: 1
GlueCEStateEstimatedResponseTime: 0
GlueCEStateWorstResponseTime: 0
GlueCEInfoDefaultSE: rsgrid3.its.uiowa.edu
GlueCEInfoApplicationDir: /export/grid/app
GlueCEInfoDataDir: /export/grid/data
GlueChunkKey: GlueCEUniqueID=rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

dn: GlueSiteUniqueID=rsgrid3.its.uiowa.edu,mds-vo-name=local,o=grid
objectClass: GlueTop
objectClass: GlueSite
objectClass: GlueKey
objectClass: GlueSchemaVersion
GlueSiteUniqueID: rsgrid3.its.uiowa.edu
GlueSiteName: UIOWA-ITB
GlueSiteDescription: OSG Site
GlueSiteUserSupportContact: mailto: shaowen-wang@uiowa.edu
GlueSiteSysAdminContact: mailto: shaowen-wang@uiowa.edu
GlueSiteSecurityContact: mailto: shaowen-wang@uiowa.edu
GlueSiteLocation: Iowa City , US
GlueSiteLatitude: 41.67
GlueSiteLongitude: -91.55
GlueSiteWeb: http://rsgrid3.its.uiowa.edu
GlueForeignKey: GlueSiteUniqueID=rsgrid3.its.uiowa.edu
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

OSG_SITE_READ, rsgrid3.its.uiowa.edu, rsgrid3.its.uiowa.edu, local, grid
dn: GlueLocationLocalID=OSG_SITE_READ,
GlueSubClusterUniqueID=rsgrid3.its.uiowa.edu,
GlueClusterUniqueID=rsgrid3.its.uiowa.edu,mds-vo-name=local,o=grid
objectClass: GlueClusterTop
objectClass: GlueLocation
objectClass: GlueKey
objectClass: GlueSchemaVersion
GlueLocationLocalID: OSG_SITE_READ
GlueLocationName: OSG_SITE_READ
GlueLocationPath: /export/grid/data
GlueChunkKey: GlueClusterUniqueID=rsgrid3.its.uiowa.edu
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

OSG_SITE_WRITE, rsgrid3.its.uiowa.edu, rsgrid3.its.uiowa.edu, local, grid
dn: GlueLocationLocalID=OSG_SITE_WRITE,
GlueSubClusterUniqueID=rsgrid3.its.uiowa.edu,
GlueClusterUniqueID=rsgrid3.its.uiowa.edu,mds-vo-name=local,o=grid
objectClass: GlueClusterTop
objectClass: GlueLocation
objectClass: GlueKey
objectClass: GlueSchemaVersion
GlueLocationLocalID: OSG_SITE_WRITE
GlueLocationName: OSG_SITE_WRITE
GlueLocationPath: /export/grid/data
GlueChunkKey: GlueClusterUniqueID=rsgrid3.its.uiowa.edu
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2

Example: 1 of 2 Old Classads (extended rules)

Computing Element Entity (1) [Classad for 2nd CE not shown]
GlueCEHostingCluster = "rsgrid3.its.uiowa.edu"
GlueCEName = "cdf"
GlueCEUniqueID = "rsgrid3.its.uiowa.edu:2119/jobmanager-condor-cdf"
GlueCEInfoGatekeeperPort = 2119
GlueCEInfoHostName = "rsgrid3.its.uiowa.edu"
GlueCEInfoLRMSType = "condor"
GlueCEInfoLRMSVersion = "6.7.13"
GlueCEInfoTotalCPUs = 1
GlueCEInfoJobManager = "condor"
GlueCEInfoContactString = "rsgrid3.its.uiowa.edu:2119/jobmanager-condor"
GlueCEInfoGRAMVersion = "1.7"
GlueCEStateFreeCPUs = 0
GlueCEStateStatus = "Production"
GlueCEPolicyMaxCPUTime = 0
GlueCEPolicyMaxRunningJobs = 0
GlueCEPolicyMaxTotalJobs = 0
GlueCEPolicyMaxWallClockTime = 0
GlueCEPolicyPriority = 0
GlueCEPolicyAssignedJobSlots = 1
GlueInformationServiceURL =
    "ldap://rsgrid3.its.uiowa.edu:2135/mds-vo-name=local,o=grid"
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2
Attributes overwritten by VOView Entity
GlueVOViewLocalID: cdf
GlueCEAccessControlBaseRule = "VO:cdf"
GlueCEStateRunningJobs = 0
GlueCEStateWaitingJobs = 0
GlueCEStateTotalJobs = 0
GlueCEStateFreeJobSlots = 1
GlueCEStateEstimatedResponseTime = 0
GlueCEStateWorstResponseTime = 0
GlueCEInfoDefaultSE = "rsgrid3.its.uiowa.edu"
GlueCEInfoApplicationDir = "/export/grid/app"
GlueCEInfoDataDir = "/export/grid/data"

Cluster Entity
GlueClusterName = "rsgrid3.its.uiowa.edu"
GlueClusterService = "rsgrid3.its.uiowa.edu"
GlueClusterUniqueID = "rsgrid3.its.uiowa.edu"
GlueClusterTmpDir = "/tmp"
GlueClusterWNTmpDir = "/scratch"
SubCluster Entity
GlueHostApplicationSoftwareRunTimeEnvironment =
    "OSG-0.3.6,ATLAS_LOC_903,LCG-2_6_0"
GlueHostArchitectureSMPSize = 2
GlueHostArchitecturePlatformType = "undefined"
GlueHostBenchmarkSF00 = 380
GlueHostBenchmarkSI00 = 400
GlueHostMainMemoryRAMSize = 512
GlueHostMainMemoryVirtualSize = 1024
GlueHostNetworkAdapterInboundIP = FALSE
GlueHostNetworkAdapterOutboundIP = TRUE
GlueHostOperatingSystemName = "linux-rocks-3.3"
GlueHostOperatingSystemRelease = "Rocks Linux"
GlueHostOperatingSystemVersion = 3.3
GlueHostProcessorClockSpeed = 1000
GlueHostProcessorModel = "Pentium III (Coppermine)"
GlueHostProcessorVendor = "GenuineIntel"
GlueHostProcessorVersion = "undefined"
GlueHostProcessorInstructionSet = "i686"
GlueHostProcessorOtherDescription = "i686"
GlueSubClusterName = "rsgrid3.its.uiowa.edu"
GlueSubClusterUniqueID = "rsgrid3.its.uiowa.edu"
GlueSubClusterPhysicalCPUs = 1
GlueSubClusterLogicalCPUs = 2
GlueSubClusterTmpDir = "/export/grid/data"
GlueSubClusterWNTmpDir = "/tmp"
Location Entity
GlueLocationLocalID = "OSG_SITE_READ,OSG_SITE_WRITE"
GlueLocationName = "OSG_SITE_READ,OSG_SITE_WRITE"

Open questions

- What should we do with the Site Entity?
- Shall we get rid of Condor-specific rules in the mapping?
- Should we include the Path attribute of Location in the mapping? (It is not in the examples)
- There may be more than one Location entity. Location attributes appear as ordered lists of values from different Location entities. How should the match making function that match location values look like?

    matchOrderedPair(GlueLocationLocalID, GlueLocationName,
                      "LocalID-val", "Name-val")

Conclusions

- The document is almost finalized. With your feedback, I hope we can agree on a final version in a couple of weeks.

- An implementation of the *non-finalized* extended rules is deployed with CEMon v1.7.2 (VDT 1.3.11).

- What is the time scale to provide an implementation of the finalized rules?