The ARC Job Description Internal Representation

Contents

1 Introduction

The next generation ARC middleware comes with support for the JSDL, XRSL and JDL job description languages. In order to provide support for the three languages listed above, an internal job description data strucuture is needed which is the union of the supported languages. This document describes the internal job description data structure of the next generation ARC middleware. Additional documentation can be found in the respective language documentation[1][2][3], but also in the job description mapping document[4] which documents how the three languages are mapped in and out from the ARC internal job description data structure.

2 Data Structure

The ARC internal job description is made up of a JobDescription C++ class which is made up from a set of other classes: JobIdentification, Application, Resources, DataStaging and JobMetaData. The JobIdentification class is used to identify the job using various types, tags and names. The Application class is used to describe explicitly the executable which should be executed at the CE, environment variables, logging, standard in and out, credential service, etc. The Resources class contains information about the type of execution node prefered by the user. The DataStaging class contains information about input files and output files/directories created by the job. Finally the job description may contain meta data, describing the context in which the description was generated, and this information is stored in the JobMetaData class.

In the following sections detailed definitions of the classes are given.

2.1 Types

In addition to the atomic types like string, integer, datetime, URL etc. we introduce the following resuable complex types.

2.1.1 ExecutableType

In order to generalize the description of any executable this new complex base type has been introduced.

2.1.1.1 Path

This mandatory string element specify the path of the executabe relative to the session directory. Multiplicity is one.

2.1.1.2 Argument

This optional element is a string specifying an argument for the executable. The multiplicity is zero or more with strict ordering. There is no default value of this element.

Comment: introduced complex structure and semantic changes in comparison to the Argument element of the GFD 136.

2.1.2 SoftwareRequirement

The SoftwareRequirment element provides the general envelope to express logical relation of

1 JDSL: http://www.ogf.org/documents/GFD.56.pdf
2 XRSL: http://www.nordugrid.org/documents/xrsl.pdf
3 JDL: EGEE-JRA1-TEC-592336 & EGEE-JRA1-TEC-590869-JDL-Attributes-v0-8
4 http://svn.nordugrid.org/trac/nordugrid/browser/arc1/trunk/doc/tech_doc/client/job_description_mapping.pdf

Software requests. The SoftwareRequirement element specifies the software requirements of a job. It may contain multiple other Software elements. This element may in turn contain a boolean that specifies that all or only one of its child elements has to be satisfied. The multiplicity of this element is zero or more. There is no default value of this element.

Comment: new. The proposed new structure replaces ApplicationName, ApplicationVersion, OperatingSystem, OperatingSystemType, OperatingSystemVersion, OperatingSystemName of GFD136.

2.1.2.1 Software

The Software is a triplet of strings. The multiplicity is zero or more. There is no default value of this element. The structure is composed of Family, Name and Version string elements.

Comment: The proposed new structure replaces ApplicationName, ApplicationVersion, OperatingSystem, OperatingSystemType, OperatingSystemVersion, OperatingSystemName of GFD136.

2.1.3 ScalableTime

This new optional complex structure is used to define scalable time request depending on benchmark results. Multiplicity zero or one. There is no default value of this element.

2.1.3.1 Value

Mandatory integer range. Multiplicity is one. There is no default value of this element.

2.1.3.2 BenchmarkType

Optional string element which specifies the type of the benchmark used to obtain the BenchmarkValue. Multiplicity is zero or one. There is no default value of this element. It is an open enumeration with values of the GLUE2 Benchmark_t type:

•bogomips

•cfp2006

•cint2006

•linpack

•specfp2000

•specint2000

2.1.3.1 BenchmarkValue

Optional integer element which specifies the result of the benchmark defined by BenchmarkType. Multiplicity zero or one. There is no default value of this element.

2.1.4 Range

Simplified structure to express required minimum and maximum values. The range has Min and Max sub elements.

2.2 JobIdentification

Stores the name of the job and defines its type. The multiplicity of this element is zero or one.

2.2.1 JobName

This optional string element may be specified by a user to name the job. It may not be unique to a particular job description, which means that a user may specify the same JobName for multiple job descriptions. Some project defines their own format for the JobName in order to categorize and explicitly define the particular version of job they run. However the recommended way to attach user specified categories to the job is to use UserTag element. This element has no default value. The multiplicity of this element is zero or one.

Comment: Same as JobName of the GFD 136.

2.2.2 Description

This optional string element may contain a longer textual description of job. This element has no default value. The multiplicity of this element is zero or one.

2.2.3 JobType

This optional element provides a classification of the job. The default value of this element is *single*. The multiplicity of this element is zero or one. The type of this element is an enumeration with the following elements:

- *collectionelement*: a job submitted as part of a collection of individual jobs which do not communicate among them,
- *parallelelement*: a job submitted as part of a collection of individual jobs which communicate among them,
- *single*: an individual stand-alone job,
- *workflownode*: a job submitted as part of a workflow.

2.2.1 UserTag

This optional element is for human readable comments, tags for free grouping or identifying different jobs. This element has no default value. The multiplicity of this element is zero or more. The type of this element is string.

2.2.2 JobVOName

This element should be used to indicate the Virtual Organisation of the job. If this optional element is not present then it is not defined. The multiplicty of this element is zero or one. The type of this element is string. It is recommended not to use this information to access contriol because authenticity is not gratanteed.

Comment: Similar information is also found in the JobProject of GFD 136.

2.3 Application

Is used to describe the executed application and its sofware environment. The section is **mandatory** and its multiplicity is one.

2.3.1 Executable

This mandatory complex element with ExecutableType is specifying the main executable of the job. Multiplicity is one.

Comment: Introduced complex structure and semantic changes in compare to the Executable element of the GFD 136.

2.3.2 Input

This optional element is a string specifying the input (Standard Input) for the Executable. The Input element is a filename which should be relative to the session directory of the job. There is no default value of this element. The multiplicity is zero or one.

Comment: Same as Input in the GFD 136.

2.3.3 Output

This optional element is a string specifying the output (Standard Output) for the Executable. The Output element is a filename which should be relative to the session directory of the job. There is no default value of this element.. The multiplicity is zero or one.

Comment: Same as Output in the GFD 136.

2.3.4 Error

This optional element is a string specifying the error output (Standard Error) for the Excutable. The Error element is a filename which should be relative to the session directory of the job. There is no default value of this element. The multiplicity is zero or one.

Comment: Same as Error in the GFD 136.

2.3.5 Join

This is an optional boolean element. If true it specifies that standard output and error should be joined into the file specified by Output element. If it is not defined then the default value is false, and joining is not done. Multiplicity is zero or one.

2.3.6 Environment

This optional element specifies environment variables which should be defined at the execution service in the execution environment of the job. It consists of a name/value pair of strings. The multiplicity of this elemenet is zero or more with strict ordering. There is no default value of this element.

Comment: Introduced complex structure in comparison to the Environment element of the GFD 136.

2.3.6.1 Name

This mandatory string element defines the name of the environment variable. Multiplicity is one. There is no default value of this element.

2.3.6.2 Value

This mandatory string element defined the value of the environment variable. Multiplicity is one. There is no default value of this element. It is not recommended to use system specific notion, macro etc as a value of this element.

2.3.7 Prologue

This optional ExecutionType element specifies the path of the executable and its arguments which will be run before the actual executable is run. If the prologue fails the job execution terminates. There is no default value of this element.

2.3.8 Epilogue

This optional ExecutionType element specifies the path of the executable and its argumnets which will be run after the actual executable is run. If the epilogue fails the job execution terminates. There is no default value of this element.

2.3.9 LogDir

This optional string element defines the name of the directory relative to the session directory containing grid-specific diagnostics per job. This directory is kept in the session directory of the job to be available for retrieval. Multiplicity is zero or one. There is no default value of this element.

2.3.10 RemoteLogging

The optional elements specifies an URL for a logging service to send reports about job. There is no default value of this element. It is up to a user to make sure the requested logging service accepts reports from the set of computing service he or she intends to use. Multiplicity is zero or more.

2.3.11 Rerun

An optional integer specifying the number of possible reruns the user can initiate in case of a system failure. If it is not defined, the default value is 0. The actual reruns can take place in different parts of the executation chain. Multiplicity is zero or one.

2.3.12 ExpiryTime

The ExpiryTime element is optional and specifies the date and time after which the processing of a passive job can be terminated. Multiplicity is zero or one. There is no default value of this element.

2.3.13 StartOnHold

Ths optional boolean element defines wheter the created job should be put on hold and processing suspended until explicit state change request arrives. The default value of this element is false. Multiplicity is zero or one.

2.3.14 ProcessingStartTime

This optional datetime element defines the requested start date and time of the job as passed through to the underlying LRMS. There is no default value of this element. Multiplicity is zero or one.

2.3.15 Notification

This optional string element defines the request in custom format for e-mail notifications on job status change. Multiplicity is zero or more. There is no default value of this element.

2.3.16 CredentialService

This optional URL elements specifies an endpoint which may be used to contact a server to renew/extend delegated proxy credential used by a submitted job. There is no default value of this element but unexpected behaviour may occur because of expired proxies. Multiplicity is zero or more.

2.3.17 AccessControl

Free format place holder for policy description.

2.4 Resources

The optional complex resource element describe the resource requirements of the job. It specifies the operating system requirement to be provided for the grid job. Its type is SoftwareRequirement. Multiplicity is zero or more. There is no default value of this element.

2.4.1 OperatingSystem

In case of OperatingSystem the Family element of the Software structure embeded in the SoftwareRequirement is open enumeration with values of the GLUE2 OSFamily_t type:

- *linux*: Family of operating systems based on Linux kernel
- *macosx*: Family of operating systems based on MacOS X
- *solaris*: Family of operating systems based on Solaris
- *windows:* Family of operating systems based on Windows
- *aix*: AIX
- *centos*: CentOS
- *debian*: Debian
- *fedoracore*: RedHat Fedora
- *gentoo*: Gentoo Linux
- *leopard*: Mac OS X 10.5 (Leopard)
- *linux-rocks*:
- *mandrake*: Mandrake
- *redhatenterpriseas*: RedHat Enterprise Server
- *scientificlinux*: Scientific Linux
- *scientificlinuxcern*: Scientific Linux CERN
- *suse*: SUSE
- *ubuntu*: Ubuntu
- *windowsvista*: Microsoft Windows Vista
- *windowsxp*: Microsoft Windows XP

In case of OperatingSystem the Name element of the Software structure embeded in the SoftwareRequirement is open enumeration with values of the GLUE2 OSName_t type:
•*aix*: AIX
•*centos*: CentOS
•*debian*: Debian
•*fedoracore*: RedHat Fedora
•*gentoo*: Gentoo Linux
•*leopard*: Mac OS X 10.5 (Leopard)
•*linux-rocks*:
•*mandrake*: Mandrake
•*redhatenterpriseas*: RedHat Enterprise Server
•*scientificlinux*: Scientific Linux
•*scientificlinuxcern*: Scientific Linux CERN
•*suse*: SUSE
•*ubuntu*: Ubuntu
•*windowsvista*: Microsoft Windows Vista
•*windowsxp*: Microsoft Windows XP

Comment: new structure obsolates OperatingSystem, OperatingSystemType, OperatingSystemVerion, OperatingSystemName of the GFD136.

2.4.1 Platform

Optional element which speficies the platform architecture requirement to be provided for the grid

job. Multiplicity is one or one. There is no default value of this element. Its is an open enumeration with a values of the GLUE2 Platform_t type:

- *amd64*: AMD 64bit architecture
- *i386*: Intel 386 architecture
- *itanium*: Intel 64-bit architecture
- *powerpc*: PowerPC architecture
- *sparc*: SPARC architecture

Comment: This element obsolates the CPUArchitecture and CPUArchitectureName of the GFD136.

2.4.1 NetworkInfo

This optional element specifies the type of the interconnect, the internal network connection available inside the computing element. Multiplicity is zero or one. There is no default value of this element. Multiplicity is zero or one. Its is an open enumeration with the values of GLUE2 NetworkInfo_t type:

- *100megabitethernet*: Network based on 100 MBit/s Ethernet technology
- *gigabitethernet*: Network based on 1 GBit/s Ethernet technology
- *infiniband*: Network based on Infiniband technology
- *myrinet*: Network based Myrinet technology

2.4.1 NodeAccess

The optional element defines the required connectivity of the executation node. Multiplicity is zero or one. If it is not defined than network connection is not required for grid job. It is closed enumeration with the following values:

- *inbound*: inbound network is required to grid job
- *outbound*: outbond network is required to grid job
- *inoutbound*: both direction is required to grid job

2.4.1 IndividualPhysicalMemory

This optional element is a range value specifying the amount of physical memory required to be available on every node of the computing element used by (a multi slot) grid job. The amount is given in bytes. Multiplicity is zero or one.

Comment: Sematically modified version of IndividualPhysicalMemory of the GFD136. This element together with the Slots is sufficient to request physical memory therefore TotalPhysicalMemory of the GFD136 is not used.

2.4.2 IndividualVirtualMemory

This optional element is an integer range value specifying the amount of virtual memory required to be available on every node of the computing element used by (a multi slot) grid job. The amount is given in bytes. Multiplicity is zero or one.

Comment: sematically modified version of IndividualVirtualMemory of the GFD136. This element together with the Slots is sufficient to request physical memory therefore TotalVirtualMemory of the GFD136 is not used.

2.4.3 DiskSpaceRequirement

This optional complex element desrcibes the space requirements of the grid job in terms of total size and its distribution over cache and session area. There is no defiault value of this element. Multiplicity is zero or one.

Comment: New structure which obsolates IndividualDiskSpace, TotalDiskSpace, DiskSpace and Filesystem of the GFD136.

2.4.3.1 DiskSpace

This mandatory range value type speficies the total required disk space of the grid job in bytes. The optional CacheDiskSpace and SessionDiskSpace elements can refine the distribution of this disk space over cache and session areas. There is no default value of this element. Multiplicity is one.

2.4.3.2 CacheDiskSpace

This optional integer element contains the requested amount of disk space in bytes on a longer lifetime cache territory. There is no default value of this element. Multiplicity is zero or one.

2.4.3.3 SessionDiskSpace

This optional integer element contains the requested amount of disk space in bytes on a session directory of the job. There is no default value of this element. Multiplicity is zero or one.

2.4.4 SessionLifetime

An optional duration element specifying the maximal time to keep the session directory of the job on the computing service upon job completion. There is no default value of this element. Multiplicity is zero or one.

2.4.5 SessionAccessMode

An optional element specifying the mode how the user may want to access the session directory on the Computing Service. Multiplicity is zero or one. If it is not defined that the user not interested to access session directory in any way. It is closed enumeration with the following values:

- *readwrite:*both read and write access is required
- *read*: only read access is required.

2.4.1 IndividualCPUTime

This is an optional scalable time element where the Value sub-element specifies the number of CPU seconds required individually on every node of the computing element used by (a multi slot) grid job. There is no default value of this element. Multiplicity is zero or one.

Comment: New structure which obsoletes the IndividualCPUTime of the GFD136.

2.4.2 TotalCPUTime

This is an optional scalable time element where the Value sub-element specifies the total number of CPU seconds required, across all CPUs of the computing element used by (a multi slot) grid job. There is no default value of this element. Multiplicity is zero or one. For singe slot jobs it is recommended to use IndividualCPUTime instead of this element.

Comment: New structure which obsoletes the TotalCPUTime of the GFD136.

2.4.3 IndividualWallTime

This is an optional scalable time element where the Value sub-element specifies the wall clock time of single slot job or every element of a multi slot job. There is no default value of this element. Multiplicity is zero or one.

Comment: Obsoletes the WallTimeLimit of the Posix of the GFD 136.

2.4.4 TotalWallTime

This is an optional scalable time element where the Value sub-element specifies the wall clock time required for multi slot job from the start of the first process until the completion of the last process. There is no default value of this element. Multiplicity is zero or one. For singe slot jobs it is recommended to use IndividualWallTime instead of this element.

2.4.5 Homogeneous

This optional boolean element defines wether the job is required to be run on computing element with a single type of Execution Environment as defined in the GLUE2 Homogenous attribute. Multiplicity zero or one. The default value is false.

2.4.6 Benchmark

This complex optional element enables the request of clusters with specific benchmark results advertised according to the GLUE2 Benchmark element. Multiplicity is on or more. If multiple elements are given they are connected by OR logical operation.

2.4.6.1 Type

Mandatory element specify the type of the benchmark used to obtain the Value. Multiplicity is one. It is an open enumeration with values of the GLUE2 Benchmark_t type:

- bogomips
- cfp2006
- cint2006
- linpack
- specfp2000
- specint2000

2.4.6.1 Value

Mandatory integer range element specify the requested result range of the benchmark defined by Type. Multiplicity is one.

2.4.7 CEType

This optional SoftwareRequirement element defines the requested computing service implementation type. There is no default value of this element the Family sub element of the Software sub-element of the SoftwareRequirement is not used in the CEType. Multiplicity is one or more.

2.4.8 SlotRequirement

This optional complex element specify the requested count of slots and its optional distribution for multi-slot grid jobs. The required NumberOfSlots subelement MAY used for machmaking but the optional ProcessPerHosts, ThreadsPerProcesses are only meant to be passed to the underlying LRMS for scheduling. The ProcessPerHost, ThreadsPerProcesses elements can refine distribution of NumberOfSlots and the optional SPMDVariation characterizes the parallel application.

Comment: Obsolates TotalCPUCount of the GFD 136 and reuses elements from the SPMD profile GFD 115.

2.4.8.1 NumberOfSlots

This mandatory integer range element specify the total number of slots where the separate instances of the executable of the parallel application (multi slot job) requested to run.

Slot is a portion of exectubale time offerd by a logical CPU. Logical CPU correspondes to a CPU as visible by the Operating System running on either real or virtual machine. The slots can be thread, process etc. Multiplicity is one.

2.4.8.2 ProcessPerHost

This optional integer element specifies the number of instances of the executable that the consuming system MUST start per host. There is no default value of this element. Multiplicity is zero or one.

2.4.8.3 ThreadsPerProcesses

This optional integer element specifies the number of threads per process (i.e., per instance of the executable). There is no default value of this element. Multiplicity is zero or one.

2.4.8.4 SPMDVariation

This optional element defines the type of multi-slot application. There is no default value of this element. It is a open enumeration with values of from the GFD 115.

- *MPI*: Any MPI environment
- *GridMPI* :The GridMPI environment
- *IntelMPI*: The Intel MPI environment
- *LAM-MPI*: The LAM/MPI environment
- *MPICH1*: The MPICH version 1 environment
- *MPICH2*: The MPICH version 2 environment
- *MPICH-GM*: The MPICH-GM environment
- *MPICH-MX:* The MPICH-MX environment
- *MVAPICH*: The MVAPICH (MPI-1) environment
- *MVAPICH2*: The MVAPICH2 (MPI-2) environment
- *OpenMPI*: The Open MPI environment
- *POE:* The POE (IBM MPI) environment
- *PVM:* A Parallel Virtual Machine environment

2.4.1 CandidateTarget

This optional complex element represents the computing service prefered to be used. Multiplicity is zero or more. There is no default value of this element.

Comment: New structure which obsolates the CandidateHosts, HostName of the GFD 136.

2.4.1.1 EndPointURL

This mandatory element contains the computing element endpoint's URL. Multiplicity is one.

2.4.1.2 QueueName

This optional string element defines the name of the preferred queue. Multiplicity is zero or one. There is no default value of this element.

2.4.2 RunTimeEnvironment

This optional SoftwareRequirement element defines the runtime environment required by the job. Multiplicity is zero or one. There is no default value of this element.

Comment: Complex element which obsolates the Application, ApplicationName, ApplicationVersion of the GFD 136.

2.5 DataStaging

Data staging is a optional complex element which describes all the files and directories should be transferred to the computing element (stage in) and the files and directories that should be transfered from the computing element (stage out). The data movement can be performed by both the client and execution service. The Data staging element is composed of File and Directory sub-elements which order is not significant. Multiplicity is zero or one. There is no default value of this element.

Comment: Largely redefines the DataStaging of the GFD 136.

2.5.1 File

This is an optional complex element which defines stage in or stage out request of a file. Multiplicity is zero or more. The ordering of the Target and Source sub-elements not significant. The file data staging element MUST contain at least one Source or Target.

2.5.1.1 Name

This mandatory string element defines the name of the staging object on the execution service. Multiplicity is one. The name is given as a relative path to the session directory or cache.

2.5.1.2 Source

This optional complex element specifies the source location of the stage in data transfer of a file. Multiplicity zero or more. In case of multiple sources it up to the computing service implementation how utilize them. There is no default value of this element.

2.5.1.2.1 URI

This mandatory URI element defines the source location of the file. Multiplicity is one. It is up to a user to make sure the computing service or the client is able communicate to the given data source.

2.5.1.2.2 DataIndexingService

This optional URL element contains the location of the data indexing service. Multiplicity is zero or more. There is no default value of this element. It is up to a user to make sure the computing service

is able communicate to the given data indexing service.

2.5.1.2.3 Threads

This optional integer element defines the number of threads which can be optionally used by execution service during the stag-in process. The default value is 1. Multiplicity is zero or one.

2.5.1.3 Target

This optional complex element specifies the target location of the stage out data transfer of a file. Multiplicity zero or more. In case of multiple targets the execution service MUST upload the file to all the mandatory targets or at least one of the targets in case there was no mandatory element defined. There is no default value of this element.

2.5.1.3.1 URI

This mandatory URI element defines the target location of the file. Multiplicity is one. It is up to a user to make sure the computing service or the client is able communicate to the given data target.

2.5.1.3.2 Threads

This optional integer element defines the number of threads which can be optionally used by execution service during the stag out process. The default value is 1. Multiplicity is zero or one.

2.5.1.3.3 Mandatory

This optional boolean element defines if the given Target must be used during the stage out data transfer or not. Multiplicity is zero or one. There is no default value of this element.

2.5.1.3.4 DataIndexingService

This optional URL element contains the location of the data indexing service. Multiplicity is zero or more. There is no default value of this element. It is up to a user to make sure the computing service is able communicate to the given data indexing service.

2.5.1.3.5 NeededReplica

This optional integer element can be used to define the required count of the replicas on a given target storage system. The default value is 1. Multiplicity is zero or one.

2.5.1.4 KeepData

This optional boolean element specify wether the file required to be keep in the session directory after successful stage out. Multiplicity is zero or one. The default value is false.

2.5.1.5 IsExecutable

This optional boolean element specify wether the executable flag has to be put on the file or not. Multiplicity is zero or one. The default value is false.

2.5.1.6 DownloadToCache

This optional boolean value specifies if the file is to be staged in to the cache instead of the session directory. Multiplicity is zero or one. The default value is false.

2.5.2 Directory

This is an optional complex element which defines stage in or stage out request of an entire

directory. Multiplicity is zero or more. The ordering of the Target and Source sub-elements not significant. The directory data staging element MUST contain at least one Source or Target.

Separate element for directory required because the handling at the computing element side of the directory may differs from the files. The involved data transfer operations are always assumed to be recursive.

Currently the Directory element has the same set of sub-elements as the File element.

2.6 JobMetaData

This is an optional complex element which describe the context in which the job description was generated and provide some further instruction how the specified requirements can be interpretered.

Comment: new.

2.6.1 Author

This optional string element may contain the identifier of the creator of the job description. Multiplicity zero or one. There is no default value of this element.

2.6.2 DocumentExpiration

This optional date element contains a date and time when the job request became irrelevant. Multiplicity is zero or one.There is no default value of this element.

2.6.3 Rank

This optional string element states how to rank computing elements that have already met the other requirements. Multiplicity is zero or one. There is no default value of this element.

2.6.4 FuzzyRank

This optional boolean element enables fuzzyness in the ranking computation. Multiplicity is zero of one. The default value is false.

3 XML Rendering

Below the Root element and high level structure of the XML representation of the data model.

```
<JobSpecification>
  <JobIdentification... />?
  <Application .../>?
  <Resources .../>?
  <DataStaging .../>?
       <File... />*
       <Directory... />*
  </DataStaging>
  <Meta/>?
</JobSepcification>
```