## GFD-R.022

Distributed Resource Management Application API (DRMAA) Working Group Hrabri Rajic, Intel Coorporation (maintainer)

Roger Brobst, Cadence Design Systems Waiman Chan, IBM Fritz Ferstl, Sun Microsystems Jeff Gardiner, John P. Robarts Res. Inst. Andreas Haas\*, Sun Microsystems Bill Nitzberg, Altair Grid Technologies Hrabri Rajic\*, Intel Coorporation Daniel Templeton, Sun Microsystems John Tollefsrud<sup>+</sup>, Sun Microsystems Peter Tröger\*, Hasso-Plattner-Inst. \*co-chairs <sup>+</sup>founding co-chair April, 2004 Updated Aug, 2007

# Distributed Resource Management Application API Specification 1.0

## Status of This Document

This document provides information to the Grid community regarding the specification of the Distributed Resource Management Application API. Distribution is unlimited.

## Updates to the Document

This document is an updated version of the document that completed public comment to become a Proposed Recommendation GFD-R-P.022 OGF in June 2004 [GFD.22]. Feedback from four independent DRMAA implementations as documented in the OGF DRMAA Experimental documents: [GFD.103], [GFD.104], [GFD.105], and [PBS/Torque] raised several issues. These issues and the resulting document changes have been recorded in the specified Gridforge Trackers. In summary, the changes fall into the following categories:

- Explicit specification of all error codes. (1171-1172,1175-1178, 1180-1181, 1783-1184, 1186-1191)
- Explicit and more self-sufficient text of the API semantics. (5884, 5882, 1786, 781, 1126, 1040, 1400)
  - Clarification of the multithreading issues. (1174, 1793, 5879)
  - More detailed explanation about job reaping. (1568, 1686)
  - Clarification of control routines behavior for sets of jobs. (1173, 1683, 1357)
  - Clarification of drmaa\_w\* routine semantics. (1125)
  - Listed drama\_w\* routines parameters with parameter comments. (1125)
  - JOB\_IDS\_SESSION\_ANY and JOB\_IDS\_SESSION\_ALL constants definition. (1289)
- Addition of missing text
  - Information that job submission calls are nonblocking. (5884)
  - In which attributes is the special index placeholder allowed. (5873)

## Copyright Notice

Copyright © Global Grid Forum (2002-2007). All Rights Reserved.

## Abstract

This document describes the Distributed Resource Management Application API (DRMAA), which provides a generalized API to distributed resource management systems (DRMSs) in order to facilitate integration of application programs.

The scope of DRMAA is limited to job submission, job monitoring and control, and retrieval of the finished job status. DRMAA provides application developers and distributed resource management builders with a programming model that enables the development of distributed applications tightly coupled to an underlying DRMS. For deployers of such distributed applications, DRMAA preserves flexibility and choice in system design.

# **Contents**

Abstract	
1. Introduction	
1.1 DRMAA Scope	
1.2 Language Issues	
1.3 Notational Conventions	
2. Interface Design and Implementation Consider	ations4
2.1 Late Binding and Portability	
2.2 Thread Safety	
2.3 Synchronization	
2.4 Distributed Application Environment	
	5
2.4.2 Native Specification	
2.5 Interface Routines General Description	
2.5.1 Init and Exit Routines	
2.5.5 Auxiliary Routines	
2.6 DRMAA Job State Transition Diagram	
3. API Specification	
3.1 Routines	
	11
3.1.2 DRMAA Sessions	11
	11
3.1.5 Site-Specific Requirements	
3.2 DRMAA API	
Mandatory attributes	
Optional Attributes	
3.2.4 Job Submission Routines	
3.3 List of DRMAA Errors	
4. Security Considerations	
5. Author Information	
6. Intellectual Property Statement	
7. Disclaimer	
8. Full Copyright Notice	
9. References	

# 1. Introduction

This document describes an API for the submission and control of jobs to one or more Distributed Resource Management Systems (DRMS). The specification encompasses the high-level functionality necessary for an application to consign a job to a DRMS, including common operations on jobs like termination or suspension. The objective is to facilitate the direct interfacing of applications to DRMS for application builders, portal builders, and independent software vendors. The specification abstracts the fundamental job interfaces of DRMS and provides an easy-to-use programming model, thereby encouraging adoption by both application builders.

# 1.1 DRMAA Scope

The scope of DRMAA 1.0 is limited to job submission, job monitoring and control, and retrieval of the finished job status.

# 1.2 Language Issues

The document authors maintain that the API should be described such that it can be implemented in multiple languages. Therefore, in this document DRMAA interfaces are described using an Interface Definition Language (IDL) - like language.

# 1.3 Notational Conventions

The key words "MUST," "MUST NOT," "REQUIRED," "SHALL," "SHALL NOT," "SHOULD," "SHOULD NOT," "RECOMMENDED," "MAY," and "OPTIONAL" are to be interpreted as described in RFC-2119 [BRADNER1].

The following abbreviations are used in this document:

API	Application Programming Interface
DRM	Distributed Resource Management
DRMS	Distributed Resource Management System
DRMAA	Distributed Resource Management Application API
ISV	Independent Software Vendor

# 2. Interface Design and Implementation Considerations

The DRMAA API has been developed to support what the authors believe will be desirable and common deployment scenarios of DRMAA implementations and applications. Specific attributes of library implementations of DRMAA should include, and the DRMAA specification anticipates:

# 2.1 Late Binding and Portability

DRMAA implementations SHOULD be provided as shared modules that could be interchangeably selected at the run time by the end user. DRMAA implementations MAY target one or more DRMSs. In the latter case a DRMAA-enabled application SHOULD be able to bind at run time to a specific DRMS via DRMAA library by setting an environment variable for the specific DRMS, or by providing the specific DRMS connection string in the initialization step.

# 2.2 Thread Safety

The authors expect that developers will link to a DRMAA library from serial and multithreaded codes; hence a DRMAA library SHOULD be thread-safe to allow a multithreaded application to use DRMAA interfaces without any explicit synchronization among the application threads. DRMAA implementers SHOULD label their implementations as thread safe if they meet the above criteria. Providers of non thread safe DRMAA implementations SHOULD document all the interfaces that are thread unsafe and provide a list of interfaces and their dependencies on external thread unsafe routines. Before a multithreaded application can use any of DRMAA interfaces, however, the DRMAA initialization routine SHOULD be called by only one thread, probably the main thread. The initialization routine is the only routine that MAY not be thread reentrant to still allow the implementers to call their implementation thread-safe. Similarly, the DRMAA library SHOULD be disengaged by only one thread. In case a standardized threads implementation such as POSIX threads exists it SHOULD be the preferred basis for thread-safe DRMAA implementation. Other threading implementations MAY be chosen, however, if documented accordingly.

# 2.3 Synchronization

DRMAA manages the asynchrony of job submission and job completion similarly to operating system process interfaces by blocking on the wait call for a specific job request.

# 2.4 Distributed Application Environment

DRMAA specifies mechanisms for submitting a job, monitoring and controlling it, and obtaining its final status. Ideally DRMAA implementations and distributed applications need not be concerned with a particular DRMS environment and DRMS site-specific policies. To facilitate deployments where this cannot be fully accomplished, Job Categories and Native Specification may be used to abstract or aggregate the site-specific policies into simple strings that are interpreted by DRMAA implementations.

# 2.4.1 Job Categories

DRMAA facilitates writing DRM-enabled applications even though the deployment properties, in particular the configuration of the DRMS, cannot be known in advance. This is a typical problem that has heretofore made writing DRM-enabled applications difficult for many Independent Software Vendors (ISVs), where the DRM system is selected by the end user.

Experience with integrations based on DRM command line interfaces show that even when the same ISV application is run as a job with the same DRMS, site-specific policies differ widely across users. These policies typically concern site-specific attributes such as what resources are to be used by the job, preferences where to run the job, and how the job should be scheduled relative to other jobs.

For supporting the variety of policies, job-specific requests expressed by DRMS submit options are common in the DRMS product space. Usually, however, these options do not affect the job from the perspective of the application or of the individual submitting the job request. This observation is the basis for "job categories," which insulate the application and individual requester from site-specific policies.

DRMAA 1.0 provides interfaces for "job categories" which encapsulate site-specific details, hiding these details from applications using the DRMAA interface. Site administrators may create a job category suitable for an application to be dispatched by the DRMS; the associated category name could be specified as a job submission attribute. The DRMAA implementation MAY then use the category name to manage site-specific resource and functional requirements of jobs in the category. Such requirements need to be configurable by the site operating a DRMS and deploying an application on top of it.

An example can help to illustrate this idea:

At site A, rendering application X is used in a heterogeneous clustered environment that is managed by a DRMS. Since application X is available only at a subset of these machines, the administrator sets up the DRMS so that the end users must put a -I X=true into their submit command line.

At site B, the same application is used in a homogeneous clustered environment with rendering application X supported at all machines managed by the DRMS. However, since X jobs do compete with applications Y sharing the same resources and X applications are to be treated with higher priority than Y jobs, end users need to put a -p 1023 into their submit command line for raising the dispatch priority.

An integration based on categories will allow submitting X jobs through the DRMAA interface in compliance with the policies of both sites A and B without the need to know about these policies. The ISV does this by specifying "X" as the category used for X rendering jobs submitted through the DRMAA interface and by mentioning this in the "DRMS integration" section of the X rendering software documentation.

The administrators at site A and site B read the documentation or installation instructions about the "X" DRMAA category. The documentation of their DRMS contains directions about the category support of their DRMAA interface implementation. From this documentation they learn how to configure their DRMS in a way that "-I X=true" is used for "X" jobs at site A while "-p 1023" is used at site B for those jobs.

DRMAA describes a mechanism for specifying the category. Associating the policy-related portion of the submit command line to the job is implementation specific.

# 2.4.2 Native Specification

The categories concept provides a means for completely hiding site-specific policy details to be considered with a DRMAA job submission for a whole class of jobs. One job category MUST be maintained for each policy to be used. In order to allow the DRMAA interface to also be used for the submission of jobs where job-specific policy specification is required "native specification" is supported. Native specification MAY be used without the requirement to maintain job categories, and submit options MAY be specified directly.

An example can help to illustrate this idea:

In order to implement the example from the previous section via native specifications, the native option string "-I X=true" has to be passed directly to the DRMAA interface while "-p 1023" has to be used at site B.

As far as the DRMAA interface specification is concerned, the native specification is an implementation-defined string and is interpreted by each DRMAA library. One MAY use job

categories and native specification with the same job submission for policy specification. In this case, the DRMAA library is assumed to be capable of joining the outcome of the two policy sources in a reasonable way. Care SHOULD be exercised to not change the job submission call semantics, pass options that conflict the already set attributes, or violate the DRMAA API in any way.

# 2.5 Interface Routines General Description

The interface routines are grouped in five categories: init and exit, job template handling, job submission, job monitoring and control, and auxiliary or informational routines that do not require initialization of a DRMAA session.

# 2.5.1 Init and Exit Routines

The calling sequence of the init routine allows all of the considered DRMS to be properly initialized, by interfacing either to the batch queue commands or to the DRMS API. Likewise, the exit routine requires parameters that will permit proper DRMS disengagement.

## 2.5.2 Job Template Routines

The remote jobs and their attributes SHALL be specified by the job template handle parameter. The job attributes SHALL be a string or a vector of string values.

The following job attributes are REQUIRED:

- Remote command to execute
- Remote command input parameters, a vector parameter
- Job state at submission
- Job environment, a vector parameter
- Job working directory
- Job category
- Native specification
- Standard input, output, and error streams
- Join output and error streams
- E-mail distribution list to report the job completion and status, a vector parameter
- E-mail suppression
- Job start time
- Job name to be used for the job submission

## 2.5.3 Job Submission Routines

Two job submission routines are described, one for submitting individual jobs and one for submitting bulk jobs.

## 2.5.4 Job Monitoring and Controlling Routines

The job monitoring and controlling API handles several functions:

- Job holding, releasing, suspending, resuming, and killing
- Checking the exit code of the finished remote job

- Checking the remote job status
- Waiting for the remote job till the end of its execution
- Waiting for all the jobs or a subset of the current session jobs to finish execution

The Unix like signals are replaced with the job control routines that have counterparts in DRMS. The only nontraditional feature is the passing of DRMAA\_JOB\_IDS\_SESSION\_ALL string as job\_id parameter to indicate operations on all job Id's in the current session.

The remote job SHALL be in one of the following states:

- System hold
- User hold
- System and user hold simultaneously
- Queued active
- System suspended
- User suspended
- Running
- Finished (un)successfully

A rejected job is not assigned a job Id and consequently SHALL NOT have a state.

In a distributed system it may not be possible for the DRMAA implementation to determine the status of the remote job at all times.

## 2.5.5 Auxiliary Routines

The auxiliary routines are needed to obtain a textual representation of errors and other DRMAA implementation-specific information.

# 2.6 DRMAA Job State Transition Diagram

Figure 1 shows the DRMAA job state transition diagram in Harel notation:

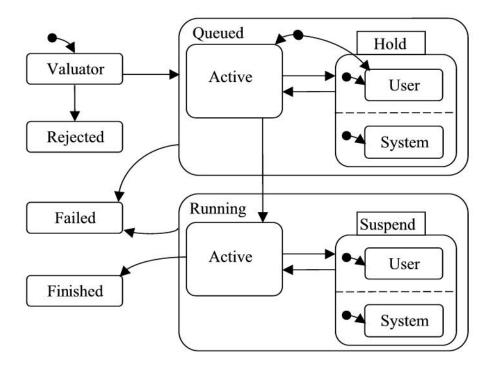


Figure 1: DRMAA job state transition diagram

# 3. API Specification

The API uses an IDL-like language to generalize discussion of allocation and deallocation, which have language specific implementations. The interface parameters could be IN, OUT, or INOUT parameters. Readers who are familiar with the C programming language should think of the parameters as being passed by value or by reference. Furthermore, the parameters could be scalar or vector values. The vector values are clearly documented.

# 3.1 Routines

In order to prevent interface name collisions all the routines have a prefix "drmaa."

# 3.1.1 Error Codes

All of the interfaces return an error code on exit. Successful return SHALL be indicated by return value DRMAA\_ERRNO\_SUCCESS. All internal errors SHALL be indicated with DRMAA\_ERRNO\_INTERNAL\_ERROR error. Out of memory errors SHALL be marked as DRMAA\_ERRNO\_NO\_MEMORY. An invalid argument SHALL be flagged as DRMAA\_ERRNO\_INVALID\_ARGUMENT. The return codes are specified and listed in Section 3.3.

The error code MAY be provided to the drmaa\_strerror routine to retrieve a textual representation of the error. Routines MUST output a context-specific error string that MAY be used in addition to the textual representation obtained from the error code. Zero length string indicates that such information is not available. This string is undefined on normal returns. The parameter used to convey the context specific error SHALL be ignored by the routine whenever success is returned. The length of any output context-specific error string, including the ending null character SHALL NOT exceed DRMAA\_ERROR\_STRING\_BUFFER.

# 3.1.2 DRMAA Sessions

An application process SHALL open only one DRMAA session at a time. Another session MAY be opened only after the current one is closed. Nesting of sessions SHOULD NOT be possible. It is RECOMMENDED that the DRMAA library SHALL free all the session resources, although this is not guaranteed, so it is RECOMMENDED that old session resources not be used later. Job Id's SHALL remain valid from one session to another. Job control routines SHOULD work correctly if a job Id came from a previous DRMAA session, provided the current DRMAA session knows how to resolve this job Id. The burden is on the user to match previous job Id's with appropriate DRMAA sessions (i.e., DRMAA implementations). It is RECOMMENDED that restartable applications make job Id's persistent in order to access the already submitted jobs. Succesful drmaa\_wait() and drmaa\_synchronize(), with dispose = true parameter, calls will make job id's invalid by reaping the job run usage data.

# 3.1.3 Run Usage Data

A DRMAA implementation SHALL collect remote run usage data (rusage variable) after the remote job run and job finish information (stat variable). The user MAY reap this data only once. The implementation is free to "garbage collect" the reaped data at a convenient time. Only the data from the current session's job Id MUST be available. Reaping data from other session job Id's MAY be supported in a DRMAA implementation.

# 3.1.4 Precedence Rules

The attributes set by using API routines SHALL be set at the compile time. The attributes set by job categories SHALL be set at installation time. The attributes set by the native specification SHALL be set at the run time. In principle these should determine the precedence rules, but these ideal precedence rules are not always achievable in practice because of complex interaction of attributes. Moreover, certain attributes in job categories may not be allowed to be overridden. The precedence rules are therefore implementation specific.

# 3.1.5 Site-Specific Requirements

Job categories and native specifications are two means for describing site-specific requirements. Setting of job categories is implementation specific. On the other hand, setting the native specification, while straightforward in the user code, could be a challenge if the user needs to provide a complex set of options. Quotation marks are especially problematic if only one variable is used for a set of native specification options.

The following are RECOMMENDED to developers to use this feature effectively:

- For each class of remote jobs, give end users a chance to specify site-specific environments, such as a queue where to send remote jobs or architecture(s) where the remote applications are available.
- Let users specify native specifications in a file if the distributed application has several classes of jobs to submit or several DRMAA sessions.
- Applications with a graphical user interface could have a dedicated dialog for this purpose.

## 3.1.6 Job Valuator

Before a submitted job enters a queue, it SHALL be passed through a valuator that determines whether the job attributes as specified are valid. If yes, a job Id SHALL be returned, and the job is successfully queued. If not, the job is rejected, job Id SHALL NOT be returned, and no job state is possible.

## 3.2 DRMAA API

For convenience, the API is divided in its five logical sections: init/exit, job template handling, job submission, job monitoring and control, and auxiliary routines.

/\* ------ Major Assumptions/Restrictions ------\*/

No explicit file staging.

Job Id Uniqueness -- "As unique as the underlying DRM makes them"

/\*Global constants \*/

DRMAA_ERROR_STRING_BUFFER	= 1024
DRMAA_JOBNAME_BUFFER	= 1024
DRMAA_SIGNAL_BUFFER	= 32
DRMAA_TIMEOUT_WAIT_FOREVER	= -1
DRMAA_TIMEOUT_NO_WAIT	= 0
JOB_IDS_SESSION_ANY	= "DRMAA_JOB_IDS_SESSION_ANY"

## 3.2.1 Initialization and Exit Routines

## drmaa\_init(contact, drmaa\_context\_error\_buf)

IN contact /\* contact information for DRM system (string) \*/ OUT drmaa\_context\_error\_buf /\*Contains a context-sensitive error upon failed return\*/

Initialize DRMAA API library and create a new DRMAA session. 'Contact' is an implementation-defined string that MAY be used to specify which DRM system to use. This routine MUST be called before any other DRMAA calls, except for drmaa\_version(), drmaa\_get\_DRM\_system(), drmaa\_get\_DRMAA\_implementation(), drmaa\_strerror(), or drmaa\_get\_contact(). If 'contact' is NULL, the default DRM system SHALL be used provided there is only one DRMAA implementation in the provided binary module. When there is more than one DRMAA implementation in the binary module, drmaa\_init() SHALL return the DRMAA\_ERRNO\_NO\_DEFAULT\_CONTACT\_STRING\_SELECTED error. drmaa\_init() SHOULD be called by only one of the threads. The main thread is RECOMMENDED. A call by another thread SHALL return DRMAA\_ERRNO\_ALREADY\_ACTIVE\_SESSION.

drmaa\_init routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_CONTACT\_STRING, DRMAA\_ERRNO\_ALREADY\_ACTIVE\_SESSION, DRMAA\_ERRNO\_NO\_DEFAULT\_CONTACT\_STRING\_SELECTED, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_DRMS\_INIT\_FAILED, DRMAA\_ERRNO\_DRMS\_COMMUNICATION\_FAILURE, or DRMAA\_ERRNO\_DEFAULT\_CONTACT\_STRING\_ERROR.

## drmaa\_exit(drmaa\_context\_error\_buf)

OUT drmaa\_context\_error\_buf /\*Contains a context-sensitive error upon failed return\*/

Disengage from DRMAA library and allow the DRMAA library to perform any necessary internal cleanup.

This routine SHALL end the current DRMAA session but SHALL NOT affect any jobs (e.g., queued and running jobs SHALL remain queued and running). drmaa\_exit() SHOULD be called by only one of the threads. The first call to call drmaa\_exit() by a thread will operate normally. All other calls from the same and other threads SHALL fail, returning a DRMAA ERRNO NO ACTIVE SESSION error code.

drmaa\_exit routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRMS\_EXIT\_ERROR, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, or DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION.

## 3.2.2 Job Template Routines

#### drmaa\_allocate\_job\_template( jt, drmaa\_context\_error\_buf )

OUT jt /\* job template (implementation-defined handle) \*/ OUT drmaa context error buf /\*Contains a context-sensitive error upon failed return\*/

Allocate a new job template.

drmaa\_allocate\_job\_template() SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE.

#### drmaa\_delete\_job\_template(jt, drmaa\_context\_error\_buf)

INOUT jt /\* job template (implementation-defined handle) \*/ OUT drmaa\_context\_error\_buf /\*Contains a context-sensitive error upon failed return\*/

Deallocate a job template. This routine has no effect on jobs. drmaa\_delete\_job\_template() SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_AUTH\_FAILURE, or DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE.

#### drmaa\_set\_attribute(jt, name, value, drmaa\_context\_error\_buf )

INOŪT jt	/* job template (implementation-defined handle) */
IN name	/* attribute name (string) */
IN value	/* attribute value (string) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Adds ('name', 'value') pair to list of attributes in job template 'jt'. Only non-vector attributes SHALL be passed.

drmaa\_set\_attribute routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_FORMAT, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_VALUE, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, or DRMAA\_ERRNO\_CONFLICTING\_ATTRIBUTE\_VALUES.

## drmaa\_get\_attribute(jt, name, value, drmaa\_context\_error\_buf )

IN jt IN name OUT value /\* job template (implementation-defined handle) \*/ /\* attribute name (string) \*/ /\* attribute value (string) \*/ OUT drmaa\_context\_error\_buf /\*Contains a context sensitive error upon failed return\*/

If 'name' is an existing non-vector attribute name in the job template 'jt', then the value of 'name' SHALL be returned; otherwise, NULL MUST be returned.

drmaa\_get\_attribute routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

### drmaa\_set\_vector\_attribute(jt, name, values, drmaa\_context\_error\_buf)

	/* job template (implementation-defined handle) */
IN name	/* attribute name (string) */
IN values	<pre>/* vector of attribute value (string vector) */</pre>
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Adds ('name', 'values') pair to list of vector attributes in job template 'jt'. Only vector attributes SHALL be passed.

drmaa\_set\_vector\_attribute routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_FORMAT, DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_VALUE, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, or DRMAA\_ERRNO\_CONFLICTING\_ATTRIBUTE\_VALUES.

#### drmaa\_get\_vector\_attribute(jt, name, values, drmaa\_context\_error\_buf )

IN jt	/* job template (implementation-defined handle) */
IN name	/* attribute name (string) */
OUT values	<pre>/* vector of attribute value (string vector) */</pre>
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

If 'name' is an existing vector attribute name in the job template 'jt', then the values of 'name' SHALL be returned; otherwise, NULL MUST be returned.

drmaa\_get\_vector\_attribute routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

#### drmaa\_get\_attribute\_names( names, drmaa\_context\_error\_buf )

OUT names	<pre>/* vector of attribute name (string vector) */</pre>
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

SHALL return the set of supported attribute names whose associated value type is String. This set SHALL include supported DRMAA reserved attribute names and native attribute names.

drmaa\_get\_attribute\_names routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

### drmaa\_get\_vector\_attribute\_names( names, drmaa\_context\_error\_buf )

OUT names/\* vector of attribute name (string vector) \*/OUT drmaa\_context\_error\_buf/\* Contains a context sensitive error upon failed return\*/

SHALL return the set of supported attribute names whose associated value type is String Vector. This set SHALL include supported DRMAA reserved attribute names and native attribute names.

drmaa\_get\_vector\_attribute\_names routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

## 3.2.3 Job attributes

#### Mandatory attributes

The following reserved attributes SHALL be available in all implementations of DRMAA. Vector attributes are marked with a 'V':

#### remote command to execute (string)

It is relative to the execution host. It is evaluated on the execution host. No binary file management is done. The attribute name is drmaa\_remote\_command.

## V input parameters (vector of strings)

These parameters SHALL be passed as arguments to the job. The attribute name is drmaa\_v\_argv.

## job state at submission (string value)

This might be useful for a rather rudimentary, but very general job-dependent execution. The states SHALL be drmaa\_hold and drmaa\_active: drmaa\_active means job has been queued, and is eligible to run drmaa\_hold means job has been queued, but it is NOT eligible to run The attribute name is drmaa\_js\_state.

#### V job environment (vector of strings)

The environment values that define the remote environment. Each string SHALL comply with the format <name>=<value>. The values override the remote environment values if there is a collision. If above is not possible, it is implementation dependent. The attribute name is drmaa v env.

### job working directory (string)

This attribute specifies the directory where the job is executed. If not set, it is implementation dependent. Evaluated relative to the execution host.

A \$drmaa\_hd\_ph\$ placeholder at the begin denotes the remaining portion of the directory\_name as a relative directory name resolved relative to the job users home directory at the execution host.

The \$drmaa\_incr\_ph\$ placeholder MAY be used at any position within the directory\_name of parametric job templates and SHALL be substituted by the underlying DRM system with the parametric jobs' index.

The directory\_name MUST be specified in a syntax that is common at the host where the job is executed.

If set and no placeholder is used, an absolute directory specification is expected. If set and the job was successfully submitted and the directory does not exist, the job enters the state DRMAA\_PS\_FAILED. The attribute name is drmaa\_wd.

## job category (string)

An implementation-defined string specifying how to resolve site-specific resources and/or policies. Detailed description is provided in section 2.4.1. The attribute name is drmaa\_job\_category.

#### native specification (string)

An implementation-defined string that is passed by the end user to DRMAA to specify site-specific resources and/or policies. Detailed description is provided in section 2.4.2. The attribute name is drmaa\_native\_specification.

## V e-mail address (vector of strings)

It is used to report the job completion and status. The new values replace old values, The attribute name is drmaa\_v\_email.

#### e-mail suppression (string)

It is used to block sending e-mail by default, regardless of the DRMS setting. 1 block 0 do not block. The attribute name is drmaa\_block\_email

job start time ( string )

This attribute specifies the earliest time when the job may be eligible to be run. The attribute name is drmaa\_start\_time

The value of the attribute SHALL be of the form [[[[CC]YY/]MM/]DD] hh:mm[:ss] [{-|+}UU:uu] where

CC is the first two digits of the year (century-1) YY is the last two digits of the year MM is the two digits of the month [01,12] DD is the two-digit day of the month [01,31] hh is the two-digit hour of the day [00,23] mm is the two-digit minute of the day [00,59] ss is the two-digit second of the minute [00,61] UU is the two-digit hours since (before) UTC uu is the two-digit minutes since (before) UTC If the optional UTC-offset is not specified, the offset associated with the local timezone SHALL be used. If the day (DD) is not specified, the current day SHALL be used unless the specified hour:mm:ss has already elapsed, in which case the next day SHALL be used.

Similarly for month (MM), year (YY), and century-1 (CC). If seconds (ss) value is omitted it is set to 0.

## Example:

The time: Sep 3 4:47:27 PM PDT 2002, could be represented as: 2002/09/03 16:47:27 -07:00

## job name

A job name SHALL comprise alphanumeric and \_ characters. The drmaa-implementation SHALL NOT provide the client with a job name longer than DRMAA\_JOBNAME\_BUFFER -1 (1023) characters. The drmaa-implementation MAY truncate any client-provided job name to an implementation-defined length that is at least 31 characters. The attribute name is drmaa\_job\_name

## input stream ( string )

Specifies the jobs' standard input.

Unless set elsewhere, if not explicitly set in the job template, the job is started with an empty input stream.

If set, specifies the network path of the jobs input stream file of the form [hostname]:file\_path

When the drmaa\_transfer\_files job template attribute is supported and contains the character 'i', the input file SHALL be fetched by the underlying DRM system from the specified host or from the submit host if no hostname is specified.

When the drmaa\_transfer\_files job template attribute is not supported or does not contain the character 'i', the input file is always expected at the host where the job is executed, irrespective of a possibly hostname specified. The \$drmaa\_incr\_ph\$ placeholder can be used at any position within the file\_path of parametric job templates and SHALL be substituted by the underlying DRM system with the parametric jobs' index.

A \$drmaa\_hd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved

relative to the job users home directory at the host where the file is located. A \$drmaa\_wd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved relative to the jobs working directory at the host where the file is located. The file\_path MUST be specified in a syntax that is common at the host where the file is located.

If set and the job was successfully submitted and the file can't be read, the job enters the state DRMAA\_PS\_FAILED.

The attribute name is drmaa\_input\_path.

## output stream (string)

Specifies how to direct the jobs' standard output.

If not explicitly set in the job template, the whereabouts of the jobs output stream is not defined.

If set, specifies the network path of the jobs output stream file of the form [hostname]:file\_path

When the drmaa\_transfer\_files job template attribute is supported and contains the character 'o', the output file SHALL be transferred by the underlying DRM system to the specified host or to the submit host if no hostname is specified.

When the drmaa\_transfer\_files job template attribute is not supported or does not contain the character 'o', the output file is always kept at the host where the job is executed irrespectively of a possibly hostname specified. The \$drmaa\_incr\_ph\$ placeholder can be used at any position within the file\_path of parametric job templates and SHALL be substituted by the underlying DRM system with the parametric jobs' index.

A \$drmaa\_hd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved relative to the job users home directory at the host where the file is located. A \$drmaa\_wd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved relative to the jobs working directory at the host where the file is located. The file\_path MUST be specified in a syntax that is common at the host where the file is located.

If set and the job was successfully submitted and the file can't be written before execution the job enters the state DRMAA\_PS\_FAILED. The attribute name is drmaa output path.

## error stream (string)

Specifies how to direct the jobs' standard error.

If not explicitly set in the job template, the whereabouts of the jobs error stream is not defined.

If set, specifies the network path of the jobs error stream file of the form [hostname]:file\_path

When the drmaa\_transfer\_files job template attribute is supported and contains the character 'e', the output file SHALL be transferred by the underlying DRM system to the specified host or to the submit host if no hostname is specified.

When the drmaa\_transfer\_files job template attribute is not supported or does not contain the character 'e', the error file is always kept at the host where the job is executed irrespectively of a possibly hostname specified. The \$drmaa\_incr\_ph\$ placeholder can be used at any position within the file\_path of parametric job templates and SHALL be substituted by the underlying DRM system with the parametric jobs' index.

A \$drmaa\_hd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved relative to the job users home directory at the host where the file is located. A \$drmaa\_wd\_ph\$ placeholder at the begin of the file\_path denotes the remaining portion of the file\_path as a relative file specification resolved relative to the jobs working directory at the host where the file is located. The file\_path MUST be specified in a syntax that is common at the host where the file is located.

If set and the job was successfully submitted and the file can't be written before execution the job enters the state DRMAA\_PS\_FAILED. The attribute name is drmaa\_error\_path.

## join files ( string )

Specifies if the error stream should be intermixed with the output stream. If not explicitly set in the job template the attribute defaults to 'n'. Either 'y' or 'n' can be specified.

If 'y' is specified the underlying DRM system SHALL ignore the value of the drmaa\_error\_path attribute and intermix the standard error stream with the standard output stream as specified with drmaa\_output\_path. The attribute name is drmaa\_join\_files.

## **Optional Attributes**

The following reserved attribute names are OPTIONAL in a conforming DRMAA implementation. For attributes that are implemented, the meanings are REQUIRED to be as follows:

Note that the list of attributes that are implemented may be programmatically obtained by using the drmaa\_get\_attribute\_names and drmaa\_get\_vector\_attribute\_names routines.

## transfer files (string)

Specifies how to transfer files between hosts. If not explicitly set in the job template the attribute defaults to ". Any combination of 'e', 'i' and 'o' MAY be specified. Whether the character 'e' is specified impacts the behavior of the drmaa\_error\_path attribute. Whether the character 'i' is specified impacts the behavior of the drmaa\_input\_path attribute. Whether the character 'o' is specified impacts the behavior of the drmaa\_output\_path attribute. The attribute name is drmaa\_transfer\_files.

## absolute job termination time ( string )

Specifies a deadline after which the DRMS will terminate a job. This is a reserved attribute named drmaa\_deadline\_time The value of the attribute SHALL be of the form [[[[CC]YY/]MM/]DD] hh:mm[:ss] [{-|+}UU:uu] where

CC is the first two digits of the year (century-1) YY is the last two digits of the year MM is the two digits of the month [01,12] DD is the two digit day of the month [01,31] hh is the two digit hour of the day [00,23] mm is the two digit minute of the day [00,59] ss is the two digit second of the minute [00,61] UU is the two digit hours since (before) UTC uu is the two digit minutes since (before) UTC

If an optional portion of the time specification is omitted, then the termination time SHALL be determined based upon the job's earliest start time.

If the day (DD) is not specified, the earliest start day for the job SHALL be used unless the specified hour:mm:ss precedes the corresponding portion of the job start time, in which case the next day SHALL be used.

Similarly for month (MM), year (YY), and century-1 (CC). If seconds (ss) value is omitted it is set to 0.

## Example:

The time: Sep 3 4:47:27 PM PDT 2002, could be represented as: 2002/09/03 16:47:27 -07:00

#### wall clock time limit (string)

This attribute specifies when the job's wall clock time limit has been exceeded. The DRMS SHALL terminate a job that has exceeded its wall clock time limit. Suspended time SHALL also be accumulated here. This is a reserved attribute named drmaa\_wct\_hlimit The value of the attribute SHALL be of the form

[[h:]m:]s where

h is one or more digits representing hours m is one or more digits representing minutes s is one or more digits representing seconds

#### Example:

To terminate a job after 2 hours and 30 minutes, any of the following MAY be passed: 2:30:0, 1:90:0, 150:0

## soft wall clock time limit (string)

This attribute specifies an estimate as to how long the job will need wall clock time to complete. Note that the suspended time is also accumulated here. This attribute is intended to assist the scheduler. If the time specified in insufficient, the drmaa-implementation MAY impose a scheduling penalty. This is a reserved attribute named drmaa\_wct\_slimit The value of the attribute SHALL be of the form [[h:]m:]s where h is one or more digits representing hours m is one or more digits representing minutes s is one or more digits representing seconds

## job run duration hlimit ( string )

This attribute specifies how long the job MAY be in a running state before its limit has been exceeded, and therefore is terminated by the DRMS. This is a reserved attribute named drmaa\_run\_duration\_hlimit The value of the attribute SHALL be of the form [[h:]m:]s where

h is one or more digits representing hours m is one or more digits representing minutes s is one or more digits representing seconds

## job run duration slimit ( string )

This attribute specifies an estimate as to how long the job will need to remain in a running state to complete. This attribute is intended to assist the scheduler. If the time specified in insufficient, the drmaa-implementation MAY impose a scheduling penalty.

This is a reserved attribute named drmaa run duration slimit

The value of the attribute SHALL be of the form

[[h:]m:]s where

h is one or more digits representing hours m is one or more digits representing minutes

s is one or more digits representing seconds

# 3.2.4 Job Submission Routines

# drmaa\_run\_job(job\_id, jt, drmaa\_context\_error\_buf )

OUT job_id	/* job identifier (string) */
IN jt	/* job template (implementation-defined handle) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Submit a job with attributes defined in the job template 'jt'. The job identifier 'job\_id' is a printable, NULL terminated string, identical to that returned by the underlying DRM system. The call SHOULD return after the DRMAA implementation submits the job request to the underlying DRM system.

drmaa\_run\_job routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_TRY\_LATER, DRMAA\_ERRNO\_DENIED\_BY\_DRM, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

## drmaa\_run\_bulk\_jobs(job\_ids, jt, start, end, incr, drmaa\_context\_error\_buf )

OUT job_ids	/* job identifiers (array of strings) */
IN jt	/* job template (implementation-defined handle) */

IN start/\* beginning index (unsigned integer)\*/IN end/\* ending index (unsigned integer) \*/IN incr/\* loop increment (integer)\*/OUT drmaa\_context\_error\_buf/\* Contains a context sensitive error upon failed return\*/

Submit a set of parametric jobs, dependent on the implied loop index, each with attributes defined in the job template 'jt'.

The job identifiers 'job\_ids' SHALL be all printable,

NULL terminated strings, identical to those returned by the underlying

DRM system. Nonnegative loop bounds SHALL NOT use file names

that start with minus sign like command line options. The call SHOULD return after

the DRMAA implementation submits the job request to the underlying DRM system.

The special index placeholder is a DRMAA defined string

drmaa\_incr\_ph /\* == \$incr\_pl\$ \*/

that is used to construct parametric job templates. Due to DRM system differing implementations job working directory,

input stream, output stream, and error stream are the only attributes that are allowed to have the index placeholder.

For example:

drmaa\_set\_attribute(pjt, "stderr", drmaa\_incr\_ph + ".err" ); /\*C++/java string syntax used \*/

drmaa\_run\_bulk\_jobs routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_TRY\_LATER, DRMAA\_ERRNO\_DENIED\_BY\_DRM, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_AUTH\_FAILURE.

# 3.2.5 Job Control Routines

## drmaa\_control(job\_id, action, drmaa\_context\_error\_buf)

IN job_id	/* job identifier (string) */
IN action	/* control action (const) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Start, stop, restart, or kill the job identified by 'job\_id'. If 'job\_id' is DRMAA\_JOB\_IDS\_SESSION\_ALL, then this routine SHALL act on all jobs <u>submitted</u> during this DRMAA session, at the moment it is called. If 'job\_id' is DRMAA\_JOB\_IDS\_SESSION\_ALL, then a call on an empty session SHALL result to DRMAA\_ERRNO\_SUCCESS for all control operations. To avoid thread races in multithreaded application the user of DRMAA implementation should explicitly synchronize this call with any other job submission call or control call that changes the number of remote jobs. The legal values for 'action' and their meanings SHALL be:

DRMAA_CONTROL_SUSPEND:	stop the job,
DRMAA_CONTROL_RESUME:	(re)start the job,
DRMAA_CONTROL_HOLD:	put the job on-hold,

DRMAA\_CONTROL\_RELEASE: DRMAA\_CONTROL\_TERMINATE: release the hold on the job, and kill the job.

This routine SHALL return once the action has been acknowledged by the DRM system, but does not necessarily wait until the action has been completed. If there are multiple kinds of errors the routine SHALL return INTERNAL\_ERROR error. The error description string SHOULD contain a good explanation of the problems.

drmaa\_control routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_RESUME\_INCONSISTENT\_STATE, DRMAA\_ERRNO\_SUSPEND\_INCONSISTENT\_STATE, DRMAA\_ERRNO\_HOLD\_INCONSISTENT\_STATE, DRMAA\_ERRNO\_RELEASE\_INCONSISTENT\_STATE, DRMAA\_ERRNO\_RELEASE\_INCONSISTENT\_STATE, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_INVALID\_JOB.

#### drmaa\_synchronize(job\_ids, timeout, dispose, drmaa\_context\_error\_buf)

IN job_ids	/* job identifiers (array of strings) */
IN timeout	/* how long we block in this call (signed long) */
IN dispose	/* dispose reaping information (boolean)*/
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Wait until all jobs specified by 'job\_ids' have finished

execution. If an invalid or already waited job id is specified, the routine SHALL fail with a DRMAA ERRNO INVALID JOB error. If 'job ids' is DRMAA JOB IDS SESSION ALL, then this routine SHALL wait for all jobs submitted during this DRMAA session, at the moment it is called. If the session contains no jobs, the routine MUST immediately return with DRMAA\_ERRNO\_SUCCESS, regardless of the dispose parameter value. Because a DRMAA implementation is not required to retain information about jobs which have been reaped, the routine is not required to, but MAY distinguish between non-existent and reaped jobs. The routine MUST return DRMAA ERRNO INVALID JOB if a provided 'job ids' contains a job that is unrecognized. If the routine successfully validates a 'job id' for an already reaped job. it MAY return DRMAA ERRNO SUCCESS. To avoid thread races in multithreaded application the user of DRMAA implementation should explicitly synchronize this call with any other job submission call or control call that changes the number of remote jobs. To prevent blocking indefinitely in this call, the caller MAY use timeout specifying after how many seconds to time out in this call. The value DRMAA TIMEOUT WAIT FOREVER (-1) MAY be specified to wait indefinitely for a result. The value DRMAA TIMEOUT NO WAIT (0) MAY be specified to return immediately if no result is available. If the call exits before timeout, all the jobs have been waited on or there was an interrupt.

If the invocation exits on timeout, the return code is DRMAA\_ERRNO\_EXIT\_TIMEOUT. The caller SHOULD check system time before and after this call in order to check how much time has passed. The dispose parameter specifies how to treat reaping of the remote job's system resources consumption and other statistics. If dispose is set to false, the job's information remains available and can be retrieved through drmaa\_wait(). If dispose is set to true, the job's information is not retained.

drmaa\_synchronize routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_EXIT\_TIMEOUT, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_INVALID\_JOB.

drmaa_wait(job_i	id, job_id_out, stat, timeout, rusage, drmaa_context_error_buf )
IN ioh id	/* iob identifier (string) or

This routine SHALL wait for a job with job\_id to fail or finish execution. If the special string DRMAA\_JOB\_IDS\_SESSION\_ANY is provided as the job\_id, this routine SHALL wait for any job from the session. In a multithreaded environment, only the active thread gets the status of the finished or failed job in that case, while the rest of the threads continue waiting. drama\_wait would wait on the submitted and not yet done/failed jobs at the moment it has been issued, so possible indeterminism is possible in multithreaded environments.

If there are no more running or completed jobs the routine SHOULD return DRMAA\_ERRNO\_INVALID\_JOB error. If an invalid or already waited job id is specified, the routine SHOULD fail with a DRMAA\_ERRNO\_INVALID\_JOB error.

The timeout value is used to specify the desired behavior when a result is not immediately available.

The value DRMAA\_TIMEOUT\_WAIT\_FOREVER (-1) MAY be specified to wait indefinitely for a result. The value DRMAA\_TIMEOUT\_NO\_WAIT (0) MAY be specified to return immediately if no result is available. Alternatively, a number of seconds MAY be specified to indicate how long to wait for a result to become available.

If the call exits before timeout, either the job has been waited on successfully or there was an interrupt.

If the invocation exits on timeout, the return code is DRMAA\_ERRNO\_EXIT\_TIMEOUT. The caller SHOULD check system time before and after this call in order to check how much time has passed.

The routine reaps jobs on a successful call, so any subsequent calls to drmaa\_wait SHOULD fail returning an error DRMAA\_ERRNO\_INVALID\_JOB meaning that the job has been already reaped. This error is the same as if the job was unknown. Failing due to an elapsed timeout has an effect that it is possible to issue drmaa\_wait multiple times for the same job\_id. When successful, the rusage information SHALL be provided as an array of strings, where each string complies with the format <name>=<value>.

The string portion <value> contains the amount of resources consumed by the job and is implementation-defined.

drmaa\_ wait routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_NO\_RUSAGE, DRMAA\_ERRNO\_EXIT\_TIMEOUT, DRMAA\_ERRNO\_NO\_RUSAGE, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_INVALID\_JOB.

The 'stat' drmaa\_wait parameter is used in a series of functions, defined below, for providing more detailed information about job termination if it is available. If those functions: drmaa\_wcoredump(), drmaa\_wexitstatus(), drmaa\_wifaborted(), drmaa\_wifexited(), drmaa\_wifsignaled(), and drmaa\_wtermsig() are called when they should not have been, DRMAA\_ERRNO\_INVALID\_ARGUMENT SHALL be returned and a descriptive message SHOULD be provided in drmaa\_context\_error\_buf.

## drmaa\_wifexited(OUT exited, IN stat, OUT drmaa\_context\_error\_buf)

OUT exited	/* more diagnosis indication value (integer) */
IN stat	/* status code of job (integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Evaluates into 'exited' a non-zero value if stat was returned for a job that either failed after running or finished after running (see section 2.6). More detailed diagnosis can be provided by means of drmaa\_wexitstatus().

A zero result for the 'exited' parameter either indicates that

1) although it is known that the job was running, more information is not available

2) it is not known whether the job was running

In both cases drmaa\_wexitstatus() SHALL NOT provide exit status information.

The function SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

#### drmaa\_wexitstatus(OUT exit\_status, IN stat, OUT drmaa\_context\_error\_buf)

OUT exit_status	/* exit code of the job (integer) */
IN stat	/* status code of job (integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

If the OUT parameter 'exited' of drmaa\_wifexited() is non-zero, this function evaluates into 'exit\_code' the exit code that the job passed to \_exit() (see exit(2)) or exit(3C), or the value that the child process returned from main.

The function SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

## drmaa\_wifsignaled(OUT signaled, IN stat, OUT drmaa\_context\_error\_buf)

OUT signaled	/* receipt of signal indication (integer) */
IN stat	/* status code of job (integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

Evaluates into 'signaled' a non-zero value if stat was returned for a job that terminated due to the receipt of a signal. A zero value can also indicate that although the job has terminated due to the receipt of a signal the signal is not available or that it is not known whether the job terminated due to the receipt of a signal. In both cases drmaa\_wtermsig() SHALL NOT provide signal information. A non-zero value returned in 'signaled' parameter indicates signal information can be retrieved by means of drmaa\_wtermsig.

The function SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

## drmaa\_wtermsig(OUT signal, IN stat, OUT drmaa\_context\_error\_buf)

OUT signal	/* returned signal (string) */
IN stat	/* status code of job (integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

If the parameter 'signaled' of drmaa\_wifsignaled is

non-zero, this function evaluates into signal which is a string representation of the signal that caused the termination of the job. For signals declared by POSIX, the symbolic names SHALL be returned (e.g., SIGABRT, SIGALRM). For signals not declared by POSIX, any other string MAY be returned.

The function SHALL return  $\mathsf{DRMAA\_ERRNO\_SUCCESS}$  on success, otherwise

DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

#### drmaa\_wcoredump(OUT core\_dumped, IN stat, OUT drmaa\_context\_error\_buf)

OUT core_dumped	/* core image indication value (integer) */
IN stat	/* status code of job (integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

If the parameter 'signaled' of drmaa\_wifsignaled is non-zero, this function evaluates into 'core\_dumped', a non-zero value if a core image of the terminated job was created. The function SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

## drmaa\_wifaborted( OUT aborted, IN stat, OUT drmaa\_context\_error\_buf )

OUT aborted/\* job aborted indication value (integer) \*/IN stat/\* status code of job (integer) \*/OUT drmaa\_context\_error\_buf/\* Contains a context sensitive error upon failed return\*/

Evaluates into 'aborted', a non-zero value if 'stat' was returned for a job that ended before entering the running state.

The function SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_MEMORY, or DRMAA\_ERRNO\_INTERNAL\_ERROR.

# drmaa\_job\_ps( IN job\_id, OUT remote\_ps, OUT drmaa\_context\_error\_buf );

IN job_id	/* job identifier (string) */
OUT remote_ps	/* program status (constant) */
OUT drmaa context error buf	/*Contains a context sensitive error upon failed return*/

Get the program status of the job identified by 'job\_id'. The possible values returned in 'remote\_ps' and their meanings SHALL be:

DRMAA_PS_UNDETERMINED	= 00H : process status cannot be determined
DRMAA_PS_QUEUED_ACTIVE	= 10H : job is queued and active
DRMAA_PS_SYSTEM_ON_HOLD	= 11H : job is queued and in system hold
DRMAA_PS_USER_ON_HOLD	= 12H : job is queued and in user hold
DRMAA_PS_USER_SYSTEM_ON_HOLD	= 13H : job is queued and in user and
	system hold
DRMAA_PS_RUNNING	= 20H : job is running
DRMAA_PS_SYSTEM_SUSPENDED	= 21H : job is system suspended
DRMAA_PS_USER_SUSPENDED	= 22H : job is user suspended
DRMAA_PS_DONE	= 30H : job finished normally
DRMAA_PS_FAILED	= 40H : job finished, but failed.

DRMAA SHOULD always get the status of job\_id from DRM system, unless the previous status has been DRMAA\_PS\_FAILED or DRMAA\_PS\_DONE and the status has been successfully cached. Terminated jobs get DRMAA\_PS\_FAILED status.

drmaa\_synchronize routine SHALL return DRMAA\_ERRNO\_SUCCESS on success, otherwise

DRMAA\_ERRNO\_NO\_MEMORY, DRMAA\_ERRNO\_INTERNAL\_ERROR, DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE, DRMAA\_ERRNO\_AUTH\_FAILURE, DRMAA\_ERRNO\_INVALID\_ARGUMENT, DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION, or DRMAA\_ERRNO\_INVALID\_JOB.

## 3.2.6 Auxiliary Routines

## error\_string drmaa\_strerror (IN errno, OUT error string);

IN errno	/* Errno number (integer) */
OUT error_string	/* Readable text version of errno (constant string) */

SHALL return the error message text associated with the errno number. The routine SHALL return null string if called with invalid ERRNO number.

### drmaa\_get\_contact(OUT contacts, OUT drmaa\_context\_error\_buf);

OUT contacts /\* Default contacts information for DRM systems (string) \*/ OUT drmaa context error buf /\* Contains a context sensitive error upon failed return \*/

If called before drmaa\_init(), it SHALL return a comma delimited default DRMAA implementation contacts string, one per each DRM system provided implementation. If called after drmaa\_init(), it SHALL return the selected contact string. The output (string) is Implementation dependent.

drmaa\_get\_contact routine SHALL return DRMAA\_ERRNO\_SUCCESS on success.

## drmaa\_version(OUT major, Out minor, OUT drmaa\_context\_error\_buf)

OUT major	/* major version number (non-negative integer) */
OUT minor	/* minor version number (non-negative integer) */
OUT drmaa_context_error_buf	/*Contains a context sensitive error upon failed return*/

SHALL return the major and minor version numbers of the DRMAA library; for DRMAA 1.0, 'major' is 1 and 'minor' is 0.

drmaa\_version routine SHALL return DRMAA\_ERRNO\_SUCCESS on success.

# drmaa\_get\_DRM\_system(OUT drm\_systems, OUT drmaa\_context\_error\_buf)

OUT drm\_systems /\* DRM systems information (string) \* OUT drmaa\_context\_error\_buf /\* Contains a context sensitive error upon failed return\*/

If called before drmaa\_init(), it SHALL return a comma delimited DRM systems string, one per each DRM system provided implementation. If called after drmaa\_init(), it SHALL return the selected DRM system. The output (string) is implementation dependent.

drmaa\_get\_DRM\_system routine SHALL return DRMAA\_ERRNO\_SUCCESS on success.

# drmaa\_get\_DRMAA\_implementation(OUT drmaa\_implementations,

# OUT drmaa\_context\_error\_buf)

OUT drmaa\_implementations /\* DRMAA implementations information (string) \* OUT drmaa context error buf /\* Contains a context sensitive error upon failed return\*/

If called before drmaa\_init(), it SHALL return a comma delimited DRMAA implementations string, one per each DRM system provided implementation. If called after drmaa\_init(), it SHALL return the selected DRMAA implementation. The output (string) is implementation dependent and COULD contain the DRM system as its part.

drmaa\_get\_DRM\_implementation routine SHALL return DRMAA\_ERRNO\_SUCCESS on success.

## 3.3 List of DRMAA Errors

------ these are relevant to all sections ------

DRMAA\_ERRNO\_SUCCESS Routine returned normally with success.

DRMAA\_ERRNO\_INTERNAL\_ERROR Unexpected or internal DRMAA error like system call failure, etc.

DRMAA\_ERRNO\_DRM\_COMMUNICATION\_FAILURE Could not contact DRM system for this request.

DRMAA\_ERRNO\_AUTH\_FAILURE The specified request is not processed successfully due to authorization failure.

DRMAA\_ERRNO\_INVALID\_ARGUMENT The input value for an argument is invalid.

DRMAA\_ERRNO\_NO\_MEMORY The system is unable to allocate resources.

----- init and exit specific -----

DRMAA\_ERRNO\_INVALID\_CONTACT\_STRING Initialization failed due to invalid contact string.

DRMAA\_ERRNO\_DEFAULT\_CONTACT\_STRING\_ERROR DRMAA could not use the default contact string to connect to DRM system.

DRMAA\_ERRNO\_NO\_DEFAULT\_CONTACT\_STRING\_SELECTED No defaults contact string was provided or selected. DRMAA requires that the default contact string is selected when there is more than one default contact string due to multiple available DRMAA implementations.

DRMAA\_ERRNO\_DRMS\_INIT\_FAILED

drmaa-wg@gridforum.org

Initialization failed due to failure to init DRM system.

DRMAA\_ERRNO\_ALREADY\_ACTIVE\_SESSION Initialization failed due to existing DRMAA session.

DRMAA\_ERRNO\_NO\_ACTIVE\_SESSION Exit routine failed because there is no active session.

DRMAA\_ERRNO\_DRMS\_EXIT\_ERROR DRM system disengagement failed.

------ job attributes specific ------DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_FORMAT The format for the job attribute value is invalid.

DRMAA\_ERRNO\_INVALID\_ATTRIBUTE\_VALUE The value for the job attribute is invalid.

DRMAA\_ERRNO\_CONFLICTING\_ATTRIBUTE\_VALUES The value of this attribute is conflicting with a previously set attributes.

------ job submission specific ------

DRMAA\_ERRNO\_TRY\_LATER Could not pass job now to DRM system. A retry MAY succeed however (saturation).

DRMAA\_ERRNO\_DENIED\_BY\_DRM The DRM system rejected the job. The job will never be accepted due to DRM configuration or job template settings.

------ job control specific -----

DRMAA\_ERRNO\_INVALID\_JOB The job specified by the 'jobid' does not exist.

DRMAA\_ERRNO\_RESUME\_INCONSISTENT\_STATE The job has not been suspended. The RESUME request SHALL NOT be processed.

DRMAA\_ERRNO\_SUSPEND\_INCONSISTENT\_STATE

drmaa-wg@gridforum.org

The job has not been running, and it cannot be suspended.

DRMAA\_ERRNO\_HOLD\_INCONSISTENT\_STATE The job cannot be moved to a HOLD state.

DRMAA\_ERRNO\_RELEASE\_INCONSISTENT\_STATE The job is not in a HOLD state.

## DRMAA\_ERRNO\_EXIT\_TIMEOUT

We have encountered a time-out condition for drmaa\_synchronize or drmaa\_wait.

## DRMAA\_ERRNO\_NO\_RUSAGE This error code is returned by drmaa\_wait() when a job has finished but no rusage and stat data could be provided.

## 4. Security Considerations

The DRMAA API does not specifically assume the existence of GRID Security infrastructure. The scheduling scenario described herein presumes that security is handled at the point of job authorization/execution on a particular resource. It is assumed that credentials owned by the process using the API are used by the DRMAA implementation to prevent abuse of the interface. In order to not unnecessarily restrict the spectrum of usable credentials, no explicit interface is defined for passing credentials.

It is conceivable an authorized but malicious user could use a DRMAA implementation or a DRMAA enabled application to saturate a DRM system with a flood of requests. Unfortunately for the DRM system this case is not distinguishable from the case of an authorized good-natured user that has many jobs to be processed. For this case DRMAA defines the DRMAA\_ERRNO\_TRY\_LATER return code to allow a DRM system to reject requests and properly indicate DRM saturation.

DRMAA implementers should guard against buffer overflows that could be exploited through DRMAA enabled interactive applications or web portals. Implementations of the DRMAA API will most likely require a network to coordinate subordinate DRMS, however the API makes no assumptions about the security posture provided the networking environment. Therefore, application developers should further consider the security implications of "on-the-wire" communications.

For environments that allow remote or protocol based DRMAA clients DRMAA should consider implementing support for secure transport layers to prevent man in the middle attacks. DRMAA does not impose any security requirements on its clients.

## 5. Author Information

Roger Brobst rbrobst@cadence.com Cadence Design Systems, Inc 555 River Oaks Parkway San Jose, CA 95134

Waiman Chan waimanc@us.ibm.com International Business Machines Corporation 2455 South Road Poughkeepsie, NY 12601

Fritz Ferstl fritz.ferstl@sun.com Sun Microsystems GmbH Dr.-Leo-Ritter-Str. 7 D-93049 Regensburg Germany

Jeffrey T. Gardiner gardiner@imaging.robarts.ca Robarts Research Institute PO Box 5015, 100 Perth Drive London ON, N6A 5K8 Canada

Andreas Haas andreas.haas@sun.com Sun Microsystems GmbH Dr.-Leo-Ritter-Str. 7 D-93049 Regensburg Germany

Bill Nitzberg nitzberg@pbspro.com Altair Grid Technologies 2685 Marine Way, Suite 1209 Mountain View, CA 94043

Hrabri L. Rajic hrabri.rajic@intel.com Intel Americas Inc. 1906 Fox Drive Champaign, IL 61820

Daniel Templeton dan.templeton@sun.com Sun Microsystems 18 Network Circle, UMPK18-117 Menlo Park, CA 94025

John Tollefsrud j.t@sun.com Sun Microsystems 18 Network Circle, UMPK18-211 Menlo Park, CA 94025 Peter Tröger peter.troeger@hpi.uni-potsdam.de Hasso-Plattner-Institute at University of Potsdam Prof.-Dr.-Helmert-Str. 2-3 D-14482 Potsdam Germany

# Contributors

We gratefully acknowledge the contributions made to this specification by Nicholas Geib, Tim Harsch, Krzysztof Kurowski, Ignacio M. Llorente, and Ruben S. Montero.

# Acknowledgements

We are grateful to numerous colleagues for discussions on the topics covered in this document, in particular (in alphabetical order, with apologies to anybody we've missed) Guillaume Alleon, Ali Anjomshoaa, Matthieu Cargnelli, Karl Czajkowski, Paul Foley, Becky Gietzel, Ancor Gonzalez Sosa, Greg Hewgill, Rayson Ho, Eduardo Huedo, Peter G. Lane, Miron Livny, Andre Merzky, Steven Newhouse, Michaell Primeaux, Greg Quinn, Alexander Saar, Martin Sarachu, Jennifer Schopf, Enrico Sirola, Chris Smith, and Douglas Thain.

# 6. Intellectual Property Statement

The GGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the GGF Secretariat.

The GGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the GGF Executive Director.

# 7. Disclaimer

This document and the information contained herein is provided on an "As Is" basis and the GGF disclaims all warranties, express or implied, including but not limited to any warranty that the use of the information herein will not infringe any rights or any implied warranties of merchantability or fitness for a particular purpose.

# 8. Full Copyright Notice

Copyright (C) Global Grid Forum (2002-2007). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the GGF or other organizations, except as needed for the purpose of

developing Grid Recommendations in which case the procedures for copyrights defined in the GGF Document process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the GGF or its successors or assignees.

## 9. References

[BRADNER1] Bradner, S. Key Words for Use in RFCs to Indicate Requirement Levels, RFC 2119. March 1997.

- [GFD.22] Roger Brobst, Waiman Chan, Fritz Ferstl, Jeff Gardiner, Andreas Haas, Bill Nitzberg, Hrabri Rajic, and John Tollefsrud. Distributed Resource Management Application API Specification 1.0. http://forge.ggf.org/projects/drmaa-wg/, 2004
- [GFD.103] P. Tröger, B. Gietzel. Condor DRMAA 1.0 Implementation Experience Report, http://www.ogf.org/gf/docs//documents/GFD.103.pdf, 2007
- [GFD.104] J. Herrera, E. Huedo, R. Montero, I. Llorente. GridWay DRMAA 1.0 Implementation -Experience Report, http://www.ogf.org/gf/docs//documents/GFD.104.pdf, 2007 [GFD.105] D. Templeton, A. Haas. N1<sup>TM</sup> Grid Engine DRMAA 1.0 Implementation - Experience
- Report, http://www.ogf.org/gf/docs//documents/GFD.105.pdf, 2007
- [PBS/Torque] L. Ciesnik, P. Domagalski, K. Kurowski, P. Lichocki. PBS/Torque DRMAA 1.0 Implementation - Experience Report, http://www.ogf.org/Public Comment Docs/Documents/Aug-2007/torgue opendsp DRMAA experience report.pdf, 2007