

GT 4.2.0 Component Guide to Public Interfaces: Data Replication Service (DRS)

DRAFT

GT 4.2.0 Component Guide to Public Interfaces: Data Replication Service (DRS)

DRAFT

Table of Contents

1. APIs	1
1. Programming Model Overview	1
2. Component API	1
2. Services and WSDL	2
1. Protocol overview	2
2. Operations	2
3. Data Replication Service (DRS) Resource properties	3
4. Faults	3
5. WSDL and Schema Definition	4
I. DataRep Commands	?
globus-replication-create	6
globus-replication-start	8
globus-replication-stop	9
globus-replication-suspend	10
globus-replication-resume	11
globus-replication-finditems	12
3. Replication request file	13
1. Interface introduction	13
2. Syntax of the interface	13
4. Configuring	14
1. Configuration overview	14
2. Syntax of the interface	14
5. Environment variable interface	15
1. Environmental variables for DRS	15
A. Errors	16

List of Tables

1. Options	7
2. Options	8
3. Options	9
4. Options	10
5. Options	11
6. Options	12
A.1. Data Replication Service (DRS) Errors	16

DRAFT

Chapter 1. APIs

1. Programming Model Overview

The DRS is a WS-RF compliant service implemented using the Globus Java WS Core. It exposes a set of Resource Properties and operations to allow users to create replication resources, control replication resources' lifecycle, and inspect the state of replication resources' activities along with the success or failure of individual replicated data sets. In this release, the WSDL and the command-line clients are the primary public interfaces for developers. Two java interfaces exist on the service-side to allow developers and users to modify the source selection behavior of the DRS. These interfaces allow users to chose alternate schemes to select sources beyond the random selection provided by default.

2. Component API

Interfaces to influence source selection include:

- ReplicaCatalogFilter
- SourceSelector

Please see [service-side interfaces](#)¹ for documentation on these interfaces.

¹ <http://www.isi.edu/~schuler/drsdocs401/>

Chapter 2. Services and WSDL

1. Protocol overview

The DRS provides a set of Resource Properties and SOAP operations to create, manipulate and inspect replication activities. Users will begin by creating a replication resource (AKA "Replicator") by invoking the create operation and passing it a URL of the replication request file (described in the domain-specific interface section). Users may start, stop, suspend and resume the Replicator when necessary. Typically a user is expected to simply start the resource and allow it to run through completion. During and after the course of replication activities performed by the resource, users may invoke standard "get resource property" and DRS-specific "find" operations to inspect the state of the resource. When the resource finishes the replication activities and the user has satisfactorily inspected the resource state, the resource should be destroyed using the standard "destroy" operation.

2. Operations

Supported operations include:

- `createReplicator` creates the "Replicator" resource.
 - `[in] InitialTerminationTime` The requested initial termination time for the resource.
 - `[in] requestFileRequest` The request-file style request.
 - `credentialEPR` Endpoint Reference of the user's delegated credential.
 - `options` Replication options which include a set of options pertinent to the transfer stage of the request, such as concurrency, parallel streams, tcp buffer size, etc.
 - `autostart` A Boolean flag indicating whether the resource should be automatically started following resource creation.
 - `requestFileUri` The URI of the request file. Currently supported schemes include http, file, and ftp.
 - `format` The request file format (domain-specific). Currently, the service only supports a simple "Table" format.
 - `[out] EPR` The Endpoint Reference of the Replicator resource.
 - `[fault] fault` Indicates a general failure when attempting to create the Replicator resource.
- `start` starts the resource.
 - `[fault] invalidStateFault` Indicates the resource is in an invalid state to perform the operation.
- `stop` stops the resource.
 - `[fault] invalidStateFault` Indicates the resource is in an invalid state to perform the operation.
- `suspend` suspends the resource.
 - `[fault] invalidStateFault` Indicates the resource is in an invalid state to perform the operation.
- `resume` resumes the resource.

- [fault] `invalidStateFault` Indicates the resource is in an invalid state to perform the operation.
- `findItems` Finds state information for individual replication items.
 - [in] `byUri` Finds by replication URI (currently, this value must be the logical filename, LFN, rather than a properly formed URI). This param is mutually-exclusive with `byStatus`.
 - [in] `byStatus` Find by status, which includes `Pending`, `Finished`, `Failed`, and `Terminated`. This param is mutually-exclusive with `byUri`.
 - [in] `offset` An offset into the results set.
 - [in] `limit` A limit of results to be returned to the client.
 - [out] `items` An array of items to be returned to the client as a result of the find operation. Each item in the array contains the complete status of the replication item including its identifier, priority, status, error (if any), sources, and destinations.
- [fault] `internalErrorFault` Indicates that an internal error occurred.

3. Data Replication Service (DRS) Resource properties

Supported resource properties for `DataRep` include:

- `status`: The status of the resource, such as `Pending`, `Active`, `Suspended`, `Terminated`, `Destroyed`, etc.
- `stage`: The current stage or activity of the resource, such as `Discover`, `Transfer`, and `Register`.
- `result`: The final result (if any) of the resource, such as `Finished`, `Failed`, and `Exception`.
- `errorMessage`: A verbose description of an error (if any) encountered by the resource. The message may include error or exception information returned by one of the dependent components, such as `RLS` or `RFT`.
- `count`: An element containing counts of individual replication items pertaining to total, finished, failed, and terminated replication items.

4. Faults

Supported faults include:

- `CreateReplicatorFault` Indicates that the service failed to create the `Replicator` resource.
- `RequestBodyMissingFault` Indicates that the request body of the create message parameters was missing.
- `CredentialEprMissingFault` Indicates that the delegated credential EPR was missing from the create message.
- `InvalidStateFault` Indicates that the requested lifecycle operation (e.g., `start`, `stop`, `suspend`, `resume`) was performed on a resource that was not in the proper state for the operation to succeed (e.g., performing a `resume` operation on a non-suspended `Replicator` resource).
- `InternalErrorFaultType` Indicates that an internal error occurred (e.g., internal system failure, etc.).

5. WSDL and Schema Definition

For more information, please see the [Replicator Port Type](#)¹ or the complete list of [schemas](#)².

DRAFT

¹ http://viewcvs.globus.org/viewcvs.cgi/ws-replica/replicator/common/schema/replica/replicator/replicator_port_type.wsdl?rev=1.2.2.1&only_with_tag=globus_4_0_branch&content-type=text/vnd.viewcvs-markup

² http://viewcvs.globus.org/viewcvs.cgi/ws-replica/replicator/common/schema/replica/replicator/?only_with_tag=globus_4_0_branch

DataRep Commands

The DRS provides a set of command-line tools to control the creation and lifecycle of a given replication request. These command line tools are available on Unix and Windows platforms and will work in the same way (of course within the platform rules - the path syntax, variable definitions, etc.).

DRAFT

Name

`globus-replication-create` -- This tool is used to create a replication resource by submitting a replication request to the designated replication service.

```
globus-replication-create
```

Tool description

Use this tool to create replication resources (also referred to as "Replicator" resources). You must specify the URL of the ReplicationService where the resource will be created. You must submit the filename of a file containing an Endpoint Reference (EPR) to a delegated credential resource, which you must have previously created. Finally, you must submit the URL of a request file specifying the desired data replications. If the client is running local to the service container the URL may be a `file://` URL, whereas if the client is remote the URL may be a `http://` or `ftp://` URL. The request file adopts a table format structure where each line in the file represents a source-destination pair delimited by a single *tab* character. The source should be a logical filename (LFN) as found in a Replica Location Service (RLS) Replica Location Index (RLI) service. The destination should be a URL acceptable to the GridFTP server. Most likely, you will want to specify a filename in order to save the newly created Replicator resource's EPR. You may use the EPR for starting the resource and querying its resource properties.

Command syntax

```
globus-replication-create [options] request-file
```

Table 1. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
--binary <boolean>	Specifies binary data transfer
--blockSize <int>	Block size for data transfer
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-C,--delegatedCredential <file>	Loads Delegated Credential EPR from file
--concurrency <int>	Concurrency of data transfer
-d,--debug	Enables debug mode
--dataChannelAuth <boolean>	Data channel authentication for transfers
--destinationSubject <name>	Destination subject name for data transfer
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
--parallelStreams <int>	Parallel streams for data transfer
-s,--service <url>	Service URL
-S,--start	Starts the Replicator resource immediately
--sourceSubject <name>	Source subject name for data transfer
--subject <name>	Subject name for data transfer
--tcpBufferSize <int>	TCP buffer size for data transfer
--userName <name>	User name for data transfer
-V,--saveEpr <file>	Save EPR of newly created Replicator to file
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Name

globus-replication-start -- This tool starts the replication activities.

globus-replication-start

Tool description

Replication resources created with the `globus-replication-create` tool may be "started" by using this tool and passing the filename of the saved EPR as a parameter to the tool. The tool will indicate an error condition if the user attempts to start a resource that has been previously started.

Command syntax

```
globus-replication-start [options]
```

Table 2. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-d,--debug	Enables debug mode
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
-s,--service <url>	Service URL
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Name

globus-replication-stop -- This tool stops the replication activities.

globus-replication-stop

Tool description

Replication resources created with the `globus-replication-create` tool may be "stoped" by using this tool and passing the filename of the saved EPR as a parameter to the tool. The tool will indicate an error condition if the user attempts to stop a resource that has not been previously started, a resource that has been suspended, or a resource that has terminated or been destroyed.

Command syntax

```
globus-replication-stop [options]
```

Table 3. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-d,--debug	Enables debug mode
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
-s,--service <url>	Service URL
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Name

globus-replication-suspend -- This tool suspends the replication activities.

globus-replication-suspend

Tool description

Replication resources created with the `globus-replication-create` tool may be "suspended" by using this tool and passing the filename of the saved EPR as a parameter to the tool. The tool will indicate an error condition if the user attempts to suspend a resource that has not been previously started, a resource that has been suspended, or a resources that is done or has been destroyed.

Command syntax

```
globus-replication-suspend [options]
```

Table 4. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-d,--debug	Enables debug mode
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
-s,--service <url>	Service URL
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Name

globus-replication-resume -- This tool resumes the replication activities.

globus-replication-resume

Tool description

Replication resources created with the `globus-replication-create` tool may be "resumed" by using this tool and passing the filename of the saved EPR as a parameter to the tool. The tool will indicate an error condition if the user attempts to resume a resource that has not been previously suspended, or a resource that is done or has been destroyed.

Command syntax

```
globus-replication-resume [options]
```

Table 5. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-d,--debug	Enables debug mode
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
-s,--service <url>	Service URL
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Name

globus-replication-finditems -- This tool queries the replication resource to return the status of individual replication item activities.

globus-replication-finditems

Tool description

This tool provides the ability to query the status of individual replication items (e.g., replication of a specific file or files) managed by the given Replication resources. It is possible to query for the status of a specific named item or to query for the status of multiple items based on a particular status (e.g., Pending, Finished, Failed). In addition, to reduce potentially large overhead of returning a large results set to the client, the client may specify an offset and limit for the results set to be returned. The "name" or "status" option must be specified.

Command syntax

```
globus-replication-finditems [options] {-N name | -S status}
```

Table 6. Options

-a,--anonymous	Use anonymous authentication. (requires either -m 'conv' or transport (https) security)
-c,--serverCertificate <file>	A file with server's certificate used for encryption. Used in the case of GSI Secure Message encryption
-d,--debug	Enables debug mode
-e,--eprFile <file>	Loads EPR from file
-f,--descriptor <file>	Sets client security descriptor. Overrides all other security settings
-g,--delegation <mode>	Performs delegation. Can be 'limited' or 'full'. (requires -m 'conv')
-h,--help	Displays help
-k,--key <name value>	Resource Key
-l,--contextLifetime <value>	Lifetime of context created for GSI Secure Conversation (requires -m 'conv')
-L,--limit <num>	Limit on the size of the result set.
-m,--securityMech <type>	Sets authentication mechanism: 'msg' (for GSI Secure Message), or 'conv' (for GSI Secure Conversation)
-N,--byName <name>	Finds item by the Logical Filename (LFN) name.
-O,--offset <num>	Offset into the results set. Indexed by 0.
-p,--protection <type>	Sets protection level, can be 'sig' (for signature) can be 'enc' (for encryption)
-S,--byStatus <status>	Finds item(s) by status. Valid status values include "Pending", "Finished", "Failed", and "Terminated".
-s,--service <url>	Service URL
-z,--authorization <type>	Sets authorization, can be 'self', 'host' or 'none'

Chapter 3. Replication request file

1. Interface introduction

The DRS domain-specific interface defines the structure and expected contents of a request file used when creating a replication resource. When the client invokes the create operation of the DRS, it will be expected to submit a properly formatted request file. It is important to understand the structure of the request file and to ensure that the file is well-formed.

2. Syntax of the interface

For the present release, the DRS request file format is fairly trivial. The request file is structured as a "Table" style of rows and columns of text. Each row represent a requested replication item described in two columns. The first column contains the identifier of the data set which will be discovered and replicated. The identifier must be resolvable by the Replica Location Index (see the JNDI configuration for `defaultIndexUrl`). In most cases, it is expected that the identifier be a Logical Filename (LFN) per the Replica Location Service definition. The second column of the row contains the URL of the "destination" for the replication item. The two columns must be delimited by a TAB character and each row must be delimited by an EOL character.



Note

The service will not accept SPACE characters as a substitute for the TAB character.

The following example shows the output of a small request file.

```
% cat example.req
my-lfn-1      gsiftp://myhost:9001/sandbox/examples/files/my-pfn-1
my-lfn-2      gsiftp://myhost:9001/sandbox/examples/files/my-pfn-2
my-lfn-3      gsiftp://myhost:9001/sandbox/examples/files/my-pfn-3
my-lfn-4      gsiftp://myhost:9001/sandbox/examples/files/my-pfn-4
my-lfn-5      gsiftp://myhost:9001/sandbox/examples/files/my-pfn-5
```

Chapter 4. Configuring

This information is in addition to the basic configuration instructions in the [Installing GT 4.2.0](#). Aside from the basic configuration of GT 4.2.0, please review the following instructions:

1. Configuration overview

The DRS requires certain JNDI settings to be properly configured. The installed JNDI configuration file may be found at `$GLOBUS_LOCATION/etc/globus_wsrp_replicator/jndi-config.xml`. To view the default configuration file (shipped with the GT 4.2.0 release) from the Globus CVS repository [click here](#)¹.

2. Syntax of the interface

The settings are structured as name-value pairs. For example:

```
<parameter>
  <name>defaultIndexUrl</name>
  <value>rls://127.0.0.1:39281</value>
</parameter>
```

The following settings must be properly configured:

- `proxyfileDir`: the directory that you would like the DRS to temporarily store user proxies. No setting is necessary. This value may be empty.
- `requestfileDir`: the directory that you would like the DRS to temporarily store request files. No setting is necessary. This value may be empty.
- `defaultIndexUrl`: the connection URL for your installation of RLS running as a RLI service.
- `defaultRegistrationUrl`: the connection URL for your installation of RLS running as a LRC service.
- `defaultReliableTransferUrl`: the connection URL for your installation of the RFT ReliableFileTransferFactoryService.
- `proxyfileChangePermsCmd`: the platform-dependent command to change file permissions to user-only read-write permissions.
- The rest of the parameter/value pairs may retain the given default values.

¹ http://viewcvs.globus.org/viewcvs.cgi/ws-replica/replicator/service/java/source/deploy-jndi-config.xml?tab=globus_4_2_branch&content-type=text/vnd.viewcvs-markup

Chapter 5. Environment variable interface

1. Environmental variables for DRS

- `GLOBUS_LOCATION=/path/to/globus/install`

DRAFT

Appendix A. Errors

Table A.1. Data Replication Service (DRS) Errors

Error Code	Definition	Possible Solutions
Authorization failed. Expected <hostname1> target but received <hostname2>	Did not receive expected hostname	When authorization is enabled on the container, you may need to use the proper hostname when referencing the DRS service rather than using localhost.
org.globus.wsrp.ResourceException: Failed to create Replication: /scratch/testrun (No such file or directory)	Cannot find the request file	Ensure that the request file's filename is correct, that it is reachable by the DRS service, and that it has the appropriate permissions for the DRS service to access it.
org.globus.wsrp.ResourceException: Failed to create Replication: String index out of range: -1	The request file is malformed (for example by using spaces instead of a delimiting tab character) which is resulting in a runtime exception.	Make sure your request file is in the correct form as described here .