

Hawkeye Information Provider

Hawkeye Information Provider

Abstract

The Hawkeye Information Provider gathers Hawkeye data about Condor pool resources using the XML mapping of the GLUE schema and reports it to a GRAM4 service, which publishes it as resource properties. The official Condor site states the following:

Hawkeye utilizes the technologies already present in Condor and ClassAds to provide rich mechanisms for collecting, storing, and using information about computers. A Hawkeye system can be used to monitor various attributes of a collection of systems. The monitoring mechanism may also be used then to further the management of systems.

This information includes:

- basic host data (name, ID)
- processor information
- memory size
- OS name and version
- file system data
- processor load data
- other basic Condor host data

You can download a PDF version of Hawkeye Information Provider information [here](#)¹.

¹ [hawkeye.pdf](#)

Table of Contents

1. Release Notes	1
1. Component Overview	1
2. Feature Summary	1
3. Changes Summary	1
4. Bug Fixes	1
5. Known Problems	1
6. Technology Dependencies	2
7. Tested Platforms	2
8. Backward Compatibility Summary	2
9. Associated Standards	2
10. For More Information	2
2. Reference Guide	3
1. Overview	3
2. Prerequisites	3
3. Configuring	3
4. Resource Properties	4
5. Schema	4
6. Security Considerations	4
7. Testing	4
8. Troubleshooting	5
Glossary	6

Chapter 1. GT 4.2.1 Release Notes: Hawkeye Info Information Provider

1. Component Overview

The Hawkeye Information Provider gathers Hawkeye data about Condor pool resources using the XML mapping of the GLUE schema and reports it to a GRAM4 service, which publishes it as resource properties. The official Condor site states the following:

Hawkeye utilizes the technologies already present in Condor and ClassAds to provide rich mechanisms for collecting, storing, and using information about computers. A Hawkeye system can be used to monitor various attributes of a collection of systems. The monitoring mechanism may also be used then to further the management of systems.

This information includes:

- basic host data (name, ID)
- processor information
- memory size
- OS name and version
- file system data
- processor load data
- other basic Condor host data

2. Feature Summary

Features new in release 4.2.1:

- This provider publishes information such as the following: basic host data (name, ID), processor information, memory size, OS name and version, file system data, processor load data, other basic Condor host data.

3. Changes Summary

This is an existing information provider available from GT4 to GT 4.2.1.

4. Bug Fixes

There are ongoing fixed bugs for this information provider (see Bugzilla).

5. Known Problems

- There may currently be bugs for this information provider (see [Bugzilla](#)¹).

¹ <http://bugzilla.mcs.anl.gov/globus/>

6. Technology Dependencies

This information provider depends on the following GT components:

- [Java WS Core](#)
- [GRAM4](#)

This information provider depends on the following 3rd party software:

- A working Perl installation
- A working Condor installation

7. Tested Platforms

Tested Platforms for this information provider:

- N/A

Tested containers for this information provider

- Java WS Core container

8. Backward Compatibility Summary

This information provider works with all GT4 and WS MDS releases.

9. Associated Standards

Associated standards for this Information Provider:

- N/A

10. For More Information

See [Chapter 2, GT 4.2.1: Hawkeye Information Provider Reference](#) for more information about this information provider.

Chapter 2. GT 4.2.1: Hawkeye Information Provider Reference

1. Overview

The Hawkeye Information Provider gathers Hawkeye data about Condor pool resources using the XML mapping of the GLUE schema and reports it to a GRAM4 service, which publishes it as resource properties. The official Condor site states the following:

Hawkeye utilizes the technologies already present in Condor and ClassAds to provide rich mechanisms for collecting, storing, and using information about computers. A Hawkeye system can be used to monitor various attributes of a collection of systems. The monitoring mechanism may also be used then to further the management of systems.

This information includes:

- basic host data (name, ID)
- processor information
- memory size
- OS name and version
- file system data
- processor load data
- other basic Condor host data

This information provider is included in the GT4 toolkit and is used for reporting GLUE Computing Element (CE) information. The GLUE resource property (as used by GRAM) collects information from two sources: the scheduler and the cluster information system (for example Ganglia or Hawkeye). These will be merged to form a single output resource property in the GLUE schema.

2. Prerequisites

1. A working Perl installation
2. A working Condor installation

3. Configuring

The following configuration is required for this information provider:

1. This section describes configuring for use in GRAM. The scheduler does not need to be configured, as GRAM already knows the scheduler bundle to use through other means. GT4.0 comes with providers which collect data from two cluster monitoring systems: * Ganglia * Hawkeye These are configured in

```
$GLOBUS_LOCATION/etc/globus_wsrp_mds_usefulrp/gluerp.xml
```

To collect data from Hawkeye, set the defaultProvider option in this file to:

```
<defaultProvider>shell /scratch/gt4b/I/libexec/globus-mds-cluster-hawkeye</defaultProvider>
```

If no cluster monitoring in this file is available, set the defaultProvider as follows: (this is the way the toolkit comes as installed)

```
<defaultProvider>none</defaultProvider>
```

4. Resource Properties

- The data gathered is published as part of the GLUECE RP

4.1. Namespace URI

- The GLUE namespace is: `http://mds.globus.org/glue/ce/1.1`
- The CE namespace is: `http://mds.globus.org/glue/ce/1.1`
- The XML namespace is: `http://www.w3.org/2001/XMLSchema`

5. Schema

- `$GLOBUS_LOCATION/share/schema/mds/usefulrp/ce.xsd`

6. Security Considerations

General security considerations associated with the container and all MDS services apply. See: [Aggregator Framework](#).

6.1. WS MDS Aggregator Services (Index Service and Trigger Service) Security Considerations

By default, the *aggregator sources* do not use authentication credentials -- they retrieve information using anonymous SSL authentication or no authentication at all, and thus retrieve only publicly-available information. If a user or administrator changes that configuration so that a service's aggregator source uses credentials to acquire non-privileged data, then that user or administrator must configure the service's aggregator sink to limit access to authorized users.

7. Testing

N/A

8. Troubleshooting

Be sure that the `gluerp.xml` file is configured properly exactly as shown above. Simply having too much whitespace between the tags can cause errors.

Glossary

A

aggregator source

A Java class that implements an interface (defined as part of the Aggregator Framework) to collect XML-formatted data. WS MDS contains three aggregator sources: the query aggregator source, the subscription aggregator source, and the execution aggregator source.