CrossGrid Installation Guide

SANTA-G

WP3.3 Grid Monitoring

Document Filename:	**installation.pdf**
Workpackage:	**WP3.3 Grid Monitoring**
Partner(s):	**TCD, CYFRONET, ICM**
Lead Partner:	**TCD**
Config ID:	**cg-santag-install**
Document classification:	**PUBLIC**

Abstract: This document explains what you need to know as an installer of the SANTA-G system.

Delivery Slip

	Name	Partner	Date	Signature
From	Stuart Kenny	TCD	August 2003	
Verified By				
Approved By				

Document Log

Version	Date	Summary of changes	Author
1-0	15 Aug 2003	First public version	Stuart Kenny
1-1	27 Aug 2003	Updated	Stuart Kenny
1-2	22 Sept 2003	Added SNORT sensor description	Stuart Kenny
1-3	20 Oct 2003	Updated	Stuart Kenny
1-4	13 Feb 2004	Added example deployment section	Stuart Kenny
1-5	19 Feb 2004	Updated	Stuart Kenny
1-6	09 Mar 2004	Added LCFG install section	Stuart Kenny
1-7	01 July 2004	Added extra troubleshooting item	Stuart Kenny
1-8	24 Aug 2004	Changed to new CrossGrid template	Stuart Kenny
1-9	08 Oct 2004	Added a description of the JIMS Plugin	Stuart Kenny
1-10	29 Nov 2004	Changed references to NFS to new FileServer	Stuart Kenny

Contents

Copyright Notice

This research is partly funded by the European Commission IST-2001-32243 Project CrossGrid.

1 About the Software

SANTA-G services are a specialized non-invasive complement to other more intrusive monitoring services. One application of these services would be in validation and calibration of both intrusive monitoring systems and systemic models, and also for performance analysis. The objectives are to:

1. allow information captured by external monitoring instruments to be introduced into the Grid information system,
2. support analysis of performance using this information.

The prototype illustrates these concepts with a NetTracer, a demonstrator that allows a user to analyse the network traffic on a site. In reality the underlying concepts have wider applicability; they allow information from a great variety of instruments to be accessed through the Grid information system.

1.1 Software Components

The SANTA-G NetTracer has three main components, the Query Engine, the Sensor and the Viewer. A Sensor can be run in one of three modes, static, dynamic or snort. A static sensor is used when there is a pre-existing set of log files, whereas a dynamic Sensor will start TCPdump at startup and dynamically generate log files. The Sensor informs the QueryEngine when a new file is started, or in the case of a static Sensor informs the QueryEngine about the set of files at startup. The sensor also provides access to the log files stored on its host machine by starting a fileserver which only the authorised QueryEngine has access to. The QueryEngine maintains tables of information about connected Sensors, and also the log files associated with each Sensor. It makes this information available to users, and the SANTA-G Viewer, through the R-GMA. It is the QueryEngine that receives queries from the R-GMA (by way of its CanonicalProducer interface), carries out the query on the appropriate log file, and returns ResultSets to the R-GMA. The Viewer allows a user to view, and query log files. The Viewer has two panels, a Packet View panel and a Query Panel. The Packet View panel displays packets from the log file graphically, whereas the Query panel allows a user to enter and execute an SQL query. The results of the query are displayed in a table. The Viewer also provides an SQL query builder to simplify the construction of complex queries. The Viewer sends the SQL queries to the SANTA-G QueryEngine through the R-GMA by using an R-GMA Consumer. Access to an R-GMA server that is running the R-GMA Servlets is required by both the QueryEngine component and the Viewer.

The following lists the SANTA-G RPMs that should be installed on the different node types within a CrossGrid site. The QueryEngine can be installed on any node that has access to the R-GMA host, and to which the Worker Nodes have access. The Viewer can be installed on any node that has access to the R-GMA host.

Worker Node: (hosts the SANTA-G Sensor component)

1. cg-santag-sensor[1]
   (a) cg-santag-common[1]
2. cg-lcfg-santag[2]

Storage Element: (hosts the SANTA-G QueryEngine)

1. cg-santag-queryengine[1]

   (a) cg-santag-common[1]

2. cg-lcfg-santag[2]

Workstation: (hosts the Viewer, can be any node with access to R-GMA Registry host machine via port 8080)

1. cg-santag-viewer[1]

   (a) cg-santag-common[1]

R-GMA Host: (node hosting the R-GMA Registry and Schema, possibly also the Producer Servlets, although these can be hosted on another separate node)

1. edg-rgma-servlets[3]

   (a) edg-rgma-common[3]
   (b) edg-rgma-api-java[3]
   (c) j2sdk[4]
   (d) MySQL[4]
   (e) tomcat4[4]

The following figure (Figure 1.1) summarises the suggested deployment of the SANTA-G system.

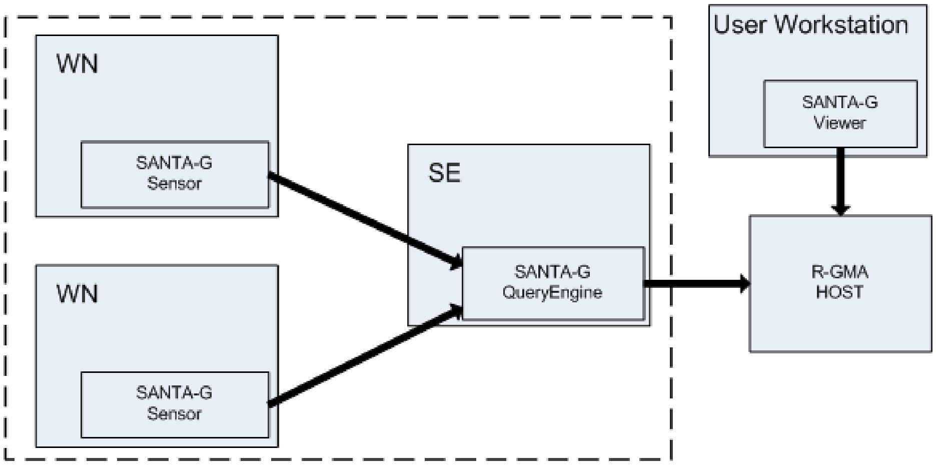

Figure 1.1: Suggested SANTA-G Deployment

1.2 Dependencies

The following section provides dependency lists for each of the SANTA-G components.

The EDG R-GMA RPM's, such as `edg-rgma-api-java`, have a number of further dependencies. RPM's for all of the R-GMA dependencies can be found on the R-GMA downloads page[4].

[1] https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp3/RPMS/
[2] https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp4/RPMS/
[3] http://hepunx.rl.ac.uk/edg/wp3/downloads/released/index.html
[4] http://hepunx.rl.ac.uk/edg/wp3/downloads/deps.html

1.2.1 Sensor Dependencies

The SANTA-G Sensor depends on Tcpdump. It uses this to gather network packets from the network and to store them in log files. The sensor provides access to the log files to the QueryEngine by using a FileServer that listens on a socket for file requests from the QueryEngine. All communcation on this socket is over SSL and the client must have a valid host certificate in order to be granted access to a file.

SANTA-G Sensor:

1. cg-santag-sensor[1]
    (a) cg-santag-common[1]
    (b) edg-java-security-client[4]
    (c) tcpdump[2] >= 3.7.2
2. snort[2]
3. j2sdk [4]

1.2.2 QueryEngine Dependencies

The QueryEngine depends on the R-GMA client packages. It only requires the Java API packages to be installed, and correctly configured. The QueryEngine host must have access to an R-GMA server, that is running the R-GMA Servlets.

SANTA-G QueryEngine:

1. cg-santag-queryengine[1]
    (a) cg-santag-common[1]
    (b) edg-rgma-api-java[3]
        i. edg-rgma-common[3]
    (c) edg-rgma-sqlutil[3]
2. j2sdk[4]

1.2.3 Viewer Dependencies

The Viewer also requires the R-GMA Java API packages to be installed and configured. It must have access to an R-GMA Servlets host on port 8080.

SANTA-G Viewer:

1. cg-santag-viewer[1]
    (a) cg-santag-common[1]
    (b) edg-rgma-api-java[3]
        i. edg-rgma-common[3]
    (c) edg-rgma-sqlutil[3]
2. j2sdk[4]

[1] https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp3/RPMS/
[2] https://savannah.fzk.de/distribution/crossgrid/crossgrid/wp3/wp3_3-moninfr/
[3] http://hepunx.rl.ac.uk/edg/wp3/downloads/releases/index.html
[4] http://hepunx.rl.ac.uk/edg/wp3/downloads/deps.html

2 Installation in the CrossGrid testbed

The CrossGrid testbeds are managed by the LCFG deployment support tool. This tool allows for the automatic installation of the software on all the required nodes.

2.1 RPM Lists for LCFG

This section provides RPM lists that should be included in the rpm list files of the components host machine.
santag-server-rpm:

1. cg-wp3.3.2-santag-common
2. cg-wp3.3.2-santag-queryengine
3. cg-wp3.3.2-santag-doc
4. cg-wp3.3.2-santag-examples
5. cg-lcfg-santag

This list should be included in the Storage Element rpm list.
santag-sensor-rpm:

1. cg-wp3.3.2-santag-common
2. cg-wp3.3.2-santag-sensor
3. cg-wp3.3.2-santag-doc
4. cg-wp3.3.2-santag-examples
5. cg-lcfg-santag

This list should be included in the Worker Node rpm list. The RPMS for the dependencies listed above should already be present in the testbed.

2.2 Profile Modifications for LCFG

In order to configure the SANTA-G components LCFG objects have been provided. These are available as two RPM packages:

1. cg-lcfg-santag
2. cg-lcfg-santag-defaults-s1

The package `cg-lcfg-santag` should be installed on any node hosting either a QueryEngine or Sensor component. The package `cg-lcfg-santag-defaults-s1` should be installed on the LCFG server.

In order to configure the SANTA-G system using LCFG a profile must be created. A default profile is provided as santag-1.def. A profile should be created for both the QueryEngine and Sensor components. The following two sections describe how to create these profiles.

2.2.1 QueryEngine LCFG Profile

Figure 2.1 provides an example LCFG profile for configuring the QueryEngine component.

```
/*
 SANTAG-server-cfg.h
 ======================================
 Configuration for SANTA-G QueryEngine
 */

/* OBJECT SANTA-G
-------------------------------------*/
EXTRA(profile.components)       santag
profile.version_santag          1
santag.ng_reconfig              configure
EXTRA(boot.services)            santag

/* Configuration. */
santag.cglocation                       /opt/cg
santag.Sensor                           no
santag.QueryEngine                      yes
santag.RgmaPropsLocation                /opt/edg/var/edg-rgma
santag.SiteId                           csTCDie
santag.QueryEngineHost                  cagnode47.cs.tcd.ie
santag.QueryEnginePort                  8998
santag.QueryEngineLogging               ON
santag.QueryEngineLogLocation           /opt/cg/var/log/santag
santag.QueryEngineMaxLogFileSize        5
santag.QueryEngineMaxNoLogFiles         5
```

Figure 2.1: SANTA-G QueryEngine Example LCFG Profile

The various parameters for configuring the QueryEngine are as follows:

cglocation the path to the CrossGrid installation. This will most likely be /opt/cg.

Sensor indicates that this configuration is for the Sensor component, should be set to *no*.

QueryEngine indicates that this configuration is for the QueryEngine component, should be set to *yes*.

RgmaPropsLocation the path to the directory that contains the rgma-defaults file. This will be `/opt/edg/var/edg-rgma` by default.

SiteId the ID of the site hosting the QueryEngine.

QueryEngineHost this should be set to the fully qualified domain name of the machine hosting the QueryEngine (i.e. the machine that this profile is being written for).

QueryEnginePort this is the port that the QueryEngine will use for communication with the Sensors and the R-GMA Server.

QueryEngineLogging determines whether QueryEngine logging should be turned on or off. Should be set to *ON* or *OFF*.

QueryEngineLogLocation the directory into which the QueryEngine log files will be written.

QueryEngineMaxLogFileSize the maximum size in MB of the QueryEngine log files.

QueryEngineMaxNoLogFiles the maximum number of log files that should be maintained.

2.2.2 Sensor LCFG Profile

Figure 2.2 provides an example LCFG profile for configuring the Sensor component.

```
/*
 SANTAG-sensor-cfg.h
 ===================================
 Configuration for SANTA-G Sensor
 */

/* OBJECT SANTA-G
-------------------------------------*/
EXTRA(profile.components)       santag
profile.version_santag          1
santag.ng_reconfig              configure
EXTRA(boot.services)            santag

/* Configuration. */
santag.cglocation                       /opt/cg
santag.SiteId                           csTCDie
santag.Sensor                           yes
santag.QueryEngine                      no
santag.QueryEngineHost                  cagnode47.cs.tcd.ie
santag.QueryEnginePort                  8998
santag.FsPort                           8999
santag.QueryEngineDn                    /C=IE/O=Grid-Ireland/OU=cs.tcd.ie
                                        /L=RA-TCD/CN=host/cagnode47.cs.tcd.ie
santag.CaCertificates                   /etc/grid-security/certificates/*.0
santag.SensorType                       DYNAMIC
santag.SensorTcpDumpArgs                -a tcp or udp or icmp
santag.SensorTracePath                  /opt/cg/var/log/santag-tcpdump
santag.SensorArchive                    DELETE
santag.SensorMaxFileSize                5
santag.SensorQueueSize                  5
santag.EdgTrustClientJar                /opt/edg/share/java/
                                        edg-java-security-trustmanager-client.jar
santag.Log4jJar                         /usr/share/java/log4j.jar
santag.BouncyCastleJar                  /usr/share/java/bcprov-jdk14.jar
```

Figure 2.2: SANTA-G Sensor Example LCFG Profile

The various parameters for configuring the Sensor are as follows:

cglocation the path to the CrossGrid installation. This will most likely be /opt/cg.

SiteId the ID of the site hosting the Sensor.

Sensor indicates that this configuration is for the Sensor component, should be set to *yes*.

QueryEngine indicates that this configuration is for the QueryEngine component, should be set to *no*.

SensorType this should specify the type of Sensor, should be set to either *STATIC*, *DYNAMIC*, or *SNORT*.

QueryEngineHost this should be set to the fully qualified domain name of the machine hosting the QueryEngine.

QueryEnginePort this is the port that the QueryEngine is listening on for incoming connections from the Sensor.

FsPort this is the port that the file server will listen on for file access requests from the QueryEngine.

QueryEngineDn this is the DN of the authorised QueryEngine. Access to files will only be allowed for the host with the DN specified here.

CaCertificates this is the path to the trusted CA certificates that will be used to valid the certificate presented to the FileServer.

SensorQueueSize this is the maximum number of Tcpdump log files that a dynamic sensor will maintain in the queue.

SensorMaxFileSize this is the maximum size in Mb of each Tcpdump logfile in the queue.

SensorArchive this tells the Sensor what to do when the maximum queue size is reached. It will either delete or archive the oldest file in the queue. Should be set to either *DELETE* or *ARCHIVE.*

SensorTcpDumpArgs these are the arguments to use when invoking Tcpdump.

SensorTracePath in the case of STATIC sensors this is the directory that contains the pre-obtained Tcpdump log files. For DYNAMIC sensors this is the directory into which the generated log files will be written. For SNORT sensors this is the directory in which the SNORT alerts file is stored along with the SNORT log file.

SensorStaticLogFileName if you are running a static sensor this is the base filename of the set of logfiles contained in the directory specified in the SensorTracePath, i.e. if the logfiles in the directory are: `hostname, hostname2, hostname3`, then this should be set to `hostname`.

SnortAlerts if configuring a SNORT Sensor this should be set to indicate the type of alert SNORT is generating. It should be set to either *FAST*, or *FULL.*

EdgTrustClientJar the path to the EDG Java Security trust client jar file.

Log4jJar path to the Log4J jar file.

BouncyCastleJar path to the Bouncy Castle jar file.

2.3 Manual Post Installation Steps

If the SANTA-G system is installed using LCFG no manual installation steps should be required.

It may be necessary, if this has not previously been carried out, to configure the R-GMA API on both the QueryEngine and Viewer host machines. To do this you should run the R-GMA configuration script:

```
$EDG_LOCATION/sbin/edg-rgma-config
```

$EDG_LOCATION is `/opt/edg` by default. Answer the questions and the script will configure the R-GMA. The defaults should be correct in most cases. You will need to enter the hostname of your R-GMA host. This the machine running the R-GMA Servlets. Information on the installation and use of the R-GMA can be found in [RGMAINSTALL] and [RGMAUSER].

3 Manual Installation

3.1 Download

SANTA-G RPMS:
https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp3/RPMS/

SANTA-G LCFG RPMS:
https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp4/RPMS/

SANTA-G Source:
https://savannah.fzk.de/cgi-bin/viewcvs.cgi/crossgrid/crossgrid/wp3/wp3_3-moninfr/wp3_3_2-santag/

SANTA-G JUnit Tests:
https://savannah.fzk.de/distribution/crossgrid/crossgrid/wp3/wp3_3-moninfr/

RGMA RPMS:
http://hepunx.rl.ac.uk/edg/wp3/downloads/releases/index.html

RGMA Dependencies:
http://hepunx.rl.ac.uk/edg/wp3/downloads/deps.html

Tcpdump:
https://savannah.fzk.de/distribution/crossgrid/crossgrid/wp3/wp3_3-moninfr/

SNORT:
http://www.snort.org

Ant:
http://hepunx.rl.ac.uk/edg/wp3/downloads/deps.html

3.2 Installation from RPM

To see if the system has been installed, type:

```
rpm -q cg-wp3.3.2-santag-common
```

and it will tell you the version of the rpm which was used. There are other rpms which may have been installed. The full set is:

cg-wp3.3.2-santag this is the full system. This RPM contains all of the SANTA-G components.

cg-wp3.3.2-santag-common this is the base rpm and is depended upon by all the others. It contains the configuration files.

cg-wp3.3.2-santag-queryengine contains the full set of files needed to run the SANTA-G QueryEngine.

cg-wp3.3.2-santag-sensor contains the full set of files needed to run the SANTA-G Sensor.

cg-wp3.3.2-santag-viewer contains the full set of files needed to run the SANTA-G Viewer.

cg-wp3.3.2-santag-doc contains the full set of SANTA-G documentation. Includes the Java APIDOC, aswell as the installation and user guides.

cg-wp3.3.2-santag-examples contains a set of example TCPdump logfiles which can be used to test your SANTA-G installation.

The following provides an example deployment. The QueryEngine can be installed on any node that has access to the R-GMA host, and to which the Worker Nodes have access. The Viewer can be installed on any node that has access to the R-GMA host.

Worker Node: (hosts the SANTA-G Sensor component)

1. cg-wp3.3.2-santag-sensor
   (a) cg-wp3.3.2-santag-common
2. cg-lcfg-santag
3. edg-java-security-client
4. Java
5. Tcpdump >= 3.7.2
6. Snort

Storage Element: (hosts the SANTA-G QueryEngine)

1. edg-rgma-api-java
   (a) edg-rgma-common
2. edg-rgma-sqlutil
3. cg-wp3.3.2-santag-queryengine
   (a) cg-wp3.3.2-santag-common
4. cg-santag-lcfg
5. Java

Workstation: (hosts the Viewer, can be any node with access to R-GMA Registry host machine via port 8080)

1. edg-rgma-api-java
   (a) edg-rgma-common
2. edg-rgma-sqlutil
3. cg-wp3.3.2-santag-viewer
   (a) cg-wp3.3.2-santag-common
4. Java

3.3 Installation from source

In order to build the SANTA-G NetTracer first the code must be obtained from the FZK CVS repository. Check the code out into a directory $SANTAG_LOCATION. Then follow the steps below. The only additional build time requirement is the java build tool Ant.

- To build the software, enter the command ant.

  ```
  ant
  ```

- To install the software type:

  ```
  ant -Dprefix=/path/to/install install
  ```

- The set of RPMs can be built by issuing the command:

  ```
  ant rpm
  ```

- To build the user documents:

  ```
  ant userdoc
  ```

- And to build the api documentation:

  ```
  ant apidoc
  ```

3.4 Configuration

If the R-GMA has not previously been configured then it is necessary to configure the system on both the QueryEngine and Viewer host machines. This is done by running the `edg-rgma-config` script found in $EDG_LOCATION/sbin (/opt/edg/sbin, by default) directory.

```
$EDG_LOCATION/sbin/edg-rgma-config
```

This script asks a number of questions. The default answers should be correct in most cases. The variable that must be set is the hostname of the machine hosting the Registry and Schema servlets. Installation [RGMAINSTALL] and user guides [RGMAUSER] for the R-GMA can be found on the R-GMA website:
http://hepunx.rl.ac.uk/edg/wp3/documentation/index.html.

A script is provided to ease the configuration of the SANTA-G system. It is contained in the `$CG_LOCATION/sbin` directory. To configure SANTA-G run the SANTA-G configuration script:

```
$CG_LOCATION/sbin/santag-config
```

You will be asked a number of questions. The first is the directory in which SANTAG is installed, (`/opt/cg` by default). When the script completes it will create two configuration files in the `$CG_LOCATION/etc/santag` directory, QueryEngine.conf and Sensor.conf. These files can be edited by hand if required. The answers given to the configuration script questions will be saved in a file named santag-defaults, also stored in the `$CG_LOCATION/etc/santag` directory. The script also creates the init.d scripts used to start the SANTA-G services. These can be found in the $CG_LOCATION/etc/init.d directory.

3.4.1 List of Configuration Files

QueryEngine: $CG_LOCATION/etc/santag/QueryEngine.conf

Sensor: $CG_LOCATION/etc/santag/Sensor.conf

3.4.2 Editing the Configuration Files

They files listed above are automatically generated by the configuration script, but may also be edited by hand if required. The following describes the entries in the files and their use.

QueryEngine.conf

This file contains configuration variables used by the SANTA-G QueryEngine.

SITE_ID this identifies the site on which the QueryEngine will run, e.g. csTCDie.

PORT the QueryEngine listens for communications from two sources, the CanonicalProducer Servlet, and connecting Sensors. It listens for these communcations on the port number specified here.

QUERYENGINE.LOGGING the QueryEngine can maintain logs of all the messages it receives from connecting Sensors and the RGMA, aswell as the time it received them. If this variable is equal to ON, then logs will be maintained. If OFF the QueryEngine will make no log entries.

QUERYENGINE.LOGLOCATION if logging is turned on, by setting the variable above, then this variable can be used to set the location in which to store the log files. You should ensure that the directory specified exists and that you have write permissions on it.

QUERYENGINE.LOGFILESIZE this specifies the maximum size of a log file in Mb. When this size is reached the file will be closed and a new log file started. Log files are created with the hostname and the date on which the QueryEngine was started. Subsequent log files are identified by a number being added to the base filename, 0, 1, 2, and so on.

QUERYENGINE.MAXNOFILES the logger writes to a rotating set of log files. This specifies the maximum number of files to use in the set.

Sensor.conf

This file contains configuration variables used by the SANTA-G Sensor.

SENSOR_TYPE a Sensor can be run in three modes, static, dynamic or snort. A static sensor (set the variable to STATIC) is used if there is a pre-exisiting set of log files. A dynamic sensor (set the variable to DYNAMIC) will start TCPdump at runtime which will dynamically generate log files. A SNORT sensor (set the variable to SNORT) is used to publish the SNORT alert file, and associated log file. The sensor types and their uses are described in more detail in the SANTA-G User Guide [SANTAGUSER].

QUERYENGINE_HOST this is the hostname of the machine running the QueryEngine.

QUERYENGINE_PORT the Sensor at startup registers with the QueryEngine to inform it that is has started and to obtain an ID from the QueryEngine. This variable should be set to the port number on which the QueryEngine is currently listening.

FILESERVER_PORT in order for the QueryEngine to access the log files stored on the sensors host machine the sensor starts a FileServer. This is the port that the FileServer will listen on for file requests from the QueryEngine. All communication with the FileServer is done over SSL connections. The sensor authenticates the QueryEngine based on its host certificate.

QUERYENGINE_DN this is the DN of the host certificate of the authorised QueryEngine host machine. Connections to the FileServer are only allowed for host with a valid certificate and files are only served to the machine with the DN specified here. The DN of the host machine can be obtained by running the following command:

```
openssl x509 -in /etc/grid-security/hostcert.pem -noout -subject
```

CA_CERTIFICATES in order to validate the certificate presented to it during a file access the sensor needs to know the path to the trusted CA certificates. This can be specified here. The default path is `/etc/grid-security/certificates/*.0`.

QUEUESIZE a dynamic sensor writes network traffic data to log files. When a file reaches a certain size it is closed and a new file opened. The sensor then continues writing to this new file. When the number of files has reached the level specified by this variable the sensor, when it closes the current file, opens a new file but then either deletes or archives the oldest file. In this way the number of files can be controlled.

MAXFILESIZE this specifies the maximum size of the file. When this size is reached the file is closed and a new file created.

SENSOR_ARCHIVE as described above when a certain number of files has been generated the sensor will either delete or archive the oldest file when it is starting a new log file. To delete the oldest file set this variable equal to DELETE, to archive set the variable to ARCHIVE. If set to archive the oldest file will be compressed and stored in an archive directory.

TCPDUMPARGS if you are running a dynamic sensor it will invoke TCPdump at startup. Here you can specify the arguments to use. Certain arguments are not allowed, such as -c. If used the Sensor will not start and an exception will be thrown. Refer to the TCPdump man pages, for more information on the available arguments. If you are editing the configuration file by hand then the arguments should be entered as they would be on the command line and quoted. For example: `TCPDUMPARGS="-a tcp or udp or icmp"`

SNORT_ALERTS if you are running a SNORT sensor then the Sensor needs to know the format of the alerts being logged by SNORT. SNORT has two types of alert, FULL and FAST. This variable should be set to one of these values, either FULL, or FAST.

TRACEPATH this is the full path to the directory that in the case of a static sensor stores the TCPdump log files. If you are running a dynamic sensor this is the directory into which the log files will be written. In the case of a SNORT sensor this should be set to the directory into which the SNORT alert file and binary log file are stored.

STATIC_FILENAMES this is the base filename of the set of static logfiles, contained in the directory specified in the TRACEPATH variable that you wish to publish. For example if the directory contains the following set of files: `hostname, hostname2, hostname3`, then this variable should be set to `hostname`. This is only used by a STATIC sensor.

3.4.3 Startup Scripts

The configuration script creates two init.d scripts. These can be found in the `$CG_LOCATION/etc/init.d` directory.

cg-santag-queryengine is used to start the queryengine, the usage is as follows:

```
$CG_LOCATION/etc/init.d/cg-santag-queryengine start | stop | status | restart
```

cg-santag-sensor is used to start the sensor, the usage is as follows:

```
$CG_LOCATION/etc/init.d/cg-santag-sensor start | stop | status | restart
```

Both are controlled by a single configuration file, **cg-santag.conf**, which can be found in the `$CG_LOCATION/etc` directory. When installed the configuration file is written to the etc directory as **cg-santag.conf.template**, so as not to overwrite any existing file. To use the file it should be renamed to **cg-santag.conf**.

3.4.4 Other requirements

If the Sensor is to run on a different machine to the QueryEngine, then the QueryEngine host machine must have a valid host certificate. This is used by the sensor when authorising requests for remote file access. Also if the Sensor is to be run in DYNAMIC mode it will need root privileges in order to start TCPdump.

Environment

No Environment variables need to be set.

Users

No special users need to be created.

Ports

The ports used for communication between the sensors and the QueryEngine are configurable. The QueryEngine will use the same port number for communication with the R-GMA server. The Viewer requires access to the R-GMA host on port 8080.

Certificates

The QueryEngine host machine must have a valid host certificate signed by a trusted CA.

Folders

No special folders are required.

4 Running and Testing

The following illustrates an example session in order to test the SANTA-G installation:

1. First start the NetTracer QueryEngine

   ```
   $CG_LOCATION/etc/init.d/cg-santag-queryengine start
   ```

2. Then start the Sensor

   ```
   $CG_LOCATION/etc/init.d/cg-santag-sensor start
   ```

3. Next run the Viewer script

   ```
   $CG_LOCATION/bin/startupViewer
   ```

   This will bring up the Viewer GUI.

4. The Viewer has two panels, the Packet View panel, and the Query Panel. The Packet View displays packets from the log file graphically. Pressing View will display the first packet in the file. The next and previous buttons allow navigation through the log file. The second panel, the Query Panel, allows the user to enter an SQL query and execute it. Results from the query are displayed in a table underneath the text area. SQL queries which can be entered are of the form:

   ```
   SELECT {* | [Table.]column_name [, [Table.]column_name...]} ]
   FROM Table
   [ WHERE [Table.]column_name { = | < | > } value
   [ AND [Table.]column_name { = | < | > } value, ...] ]
   ```

5. To simplify the construction of complex queries a Query Builder is provided. Pressing the Query Builder button opens this in another window. The table to query should first be chosen, then the fields of the table to be viewed. WHERE predicates can be added by pressing the add button in the Where panel. Pressing build will close the builder and display the constructed SQL query in the text area. Pressing execute will submit the query.

6. To stop the system, close the Viewer and then issue the following two commands:

   ```
   $CG_LOCATION/etc/init.d/cg-santag-sensor stop
   $CG_LOCATION/etc/init.d/cg-santag-queryengine stop
   ```

4.0.5 Testing

A number of JUnit tests, containing both system and unit tests have been provided in order to test the SANTA-G system. The tests require JUnit version 3.8.1, which is available from the R-GMA dependencies list.[1]. The tests, as they contain unit tests, require the SANTA-G software to be installed. They currently only need the cg-wp3.3.2-santag-queryengine package to be installed. Also required is the cg-wp3.3.2-santag-examples package. This package contains a set of sample log files, used to provide data for the tests.

To run the tests:

[1] http://hepunx.rl.ac.uk/edg/wp3/downloads/rpms-deps.html

1. Download the tests from https://savannah.fzk.de/distribution/crossgrid/crossgrid/wp3/wp3_3-moninfr/.
2. Un-compress and un-tar the tests archive.
3. Edit the santagtests-defaults file. This is located in the `test/etc/` directory.
4. Set the values for the RGMA_PROPS_LOCATION, this is location of the rgma-defaults file, the SANTAG_LOCATION, this is where the SANTA-G software is installed, usually /opt/cg, and the SANTAG_EXAMPLES_TRACEFILES_LOCATION, this is the full path to the directory that contains the SANTA-G example Tcpdump log files.
5. Build the tests by entering the test directory and typing `ant`. **NOTE: if the build fails at this point because the JUnit classes cannot be found this is most likely due to the JUnit path not being set correctly in the rgma-defaults file. To correct this edit the rgma-defaults file (in /opt/edg/var/edg-rgma by default) and correct the path in the JUNIT_JAR entry.**
6. When built start the QueryEngine, `$CG_LOCATION/bin/startupQueryEngine start`.
7. Then run the tests, from the test directory `./santag-tests swing`.

4.1 Log Files

The QueryEngine logs to the directory specified in the QueryEngine.conf configuration file. The filename is in the format hostname_log.date.n, where n is the file number. The log file contains all messages received from both the R-GMA server and the Sensors. Also any errors that occur are logged.

Two additional log files are maintained in the $CG_TMP directory, `santag_sensor.out` and `santag_queryengine.out`. These will contain any unexpected errors encountered, or errors that occurred during startup of the services.

5 EDG License Agreement

A Troubleshooting

The SANTA-G system is dependent on the R-GMA being installed and correctly configured. The R-GMA installation and user guides can be found on the DataGrid WP3 web site.[1]. These describe how to setup and configure the R-GMA. A number of common errors which effect SANTA-G are described below, most are related to an incorrectly configured R-GMA system:

1. **The R-GMA props directory has not been entered in the SANTA-G configuration script correctly.** SANTA-G needs to be able to find the R-GMA props directory (the directory which stores the R-GMA configuration files) in order to access the `rgma-defaults` file. This file contains the paths to the R-GMA jar files which SANTA-G needs. If the R-GMA props directory is incorrect then the following error will be reported when either the QueryEngine or the Viewer is started:

   ```
   Exception in thread "main" java.lang.NoClassDefFoundError:  org.edg.info.RGMAException
   ```

   To correct this first locate the R-GMA props directory. The default location is `/opt/edg/var/edg-rgma`. Re-run the `santag-config` script and make sure that this is the value entered for the R-GMA props directory location.

   If it is correct then the `rgma-defaults` file may not be correct. Enter the R-GMA props directory and edit the `rgma-defaults` file. Check that the file contains the correct path to the jar files. You should also re-run the `rgma-config` script to re-generate the file.

2. **The log4j system is incorrectly configured**. The R-GMA uses a java logging tool called log4j. If this is not correctly configured then error messages will be printed out at QueryEngine or Viewer startup. The error message is of the form:

   ```
   log4j:ERROR Could not find value for key log4j.appender.A1
   log4j:ERROR Could not instantiate appender named "A1".
   ```

   In order to correct this error you must edit the default `log4j.props` file which is installed with the R-GMA. This file can be found in the R-GMA props directory. Open the file and uncomment the following line:

   ```
   log4j.appender.A1=org.apache.log4j.ConsoleAppender
   ```

   For more information on log4j see the project website[2].

3. **The Viewer starts but displays the error message ”Could not collect file header information”**. This indicates that the Viewer was not able to obtain information from the log file. The most likely cause of this is that the QueryEngine was not able to access the remote file server running on the sensor host. You should ensure that the DN of the QueryEngine's host certificate has been correctly entered in the sensors configuration file, and that the QueryEngine has a valid host certificate. The other possibility is that a firewall is blocking the communication between the QueryEngine and the Sensor host machine.

[1] http://hepunx.rl.ac.uk/edg/wp3/

[2] http://jakarta.apache.org/log4j/docs//

B JIMS Plugin

A plugin is available for the SANTA-G QueryEngine that allows for the publication of a subset of the information captured by JIMS into the R-GMA. JIMS is a JMX-based Infrastructure Monitoring System developed within the CrossGrid Grid Monitoring task.

B.1 Plugin Installation

The plugin is distributed as an RPM that should be installed on the machine hosting the SANTA-G QueryEngine. The RPM can be obtained from the following URL:

https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc3.2.2/wp3/RPMS/

The RPM dependencies are as follows:

1. cg-wp3.3.2-santag-queryengine
2. cg-wp3.3.3-jims-client

B.1.1 LCFG Configuration

Once the plugin has been installed it must be configured. If you are using LCFG the following steps should be followed:

1. If the R-GMA has not yet been configured: `/etc/obj/rgma configure`
2. Re-configure the SANTA-G QueryEngine: `/etc/obj/santag configure`
3. Start the QueryEngine: `$CG_LOCATION/etc/init.d/cg-santag-queryengine start`

B.1.2 Manual Configuration

If you are not using LCFG then the following manual configuration steps must be carried out:

1. If the R-GMA has not yet been configured: `$EDG_LOCATION/sbin/edg-rgma-config`
2. Re-configure the SANTA-G QueryEngine: `$CG_LOCATION/sbin/santag-config`
3. Start the QueryEngine: `$CG_LOCATION/etc/init.d/cg-santag-queryengine start`

B.2 Plugin Configuration File

The plugin can be configured by editing the `JIMSPlugin.conf` file, found in the `$CG_LOCATION/etc/santag` directory. This file allows you to specify the update interval of the plugin. The update interval defines how often information should be gathered from the JIMS system and published to the R-GMA. The default value is every 15 minutes. The plugin streams the collected JIMS information to the central CrossGrid R-GMA server located in TCD. Here the information is aggregated and stored by using an R-GMA Archiver. Users can access the stored data by using an R-GMA Consumer. Information is stored for 24 hours. Any tuple of information older than this is deleted. If a particular site wishes to hold data for longer than this they should construct their own Archiver that holds data for a longer interval. Details of of how to create Archivers and Consumers can be found in the R-GMA User Guide [RGMAUSER].

In the configuration file it is also possible to define the location of the Soap Gateway host. This is the host you use to contact the JIMS system running in a particular site. The plugin will first attempt to determine this host by using the `$CG_LOCATION/etc/cg-site.cfg` file. The default location for the Soap Gateway is the computing element (CE) of the site. If this file is not present, or the Soap Gateway is located on a machine other than the CE, then the hostname has to be specified in the plugin configuration file.

Any errors that occur during startup, or operation of the plugin, will be logged to the QueryEngine logfile. This can be found in the directory specified in the QueryEngine configuration, `$CG_LOCATION/var/log/santag` by default.

Bibliography

[TEST] Jorge Gomes, LIP; **Middleware Test Procedure**; May 2002

[QAP] WP5, CYRFRONET; **Quality Assurance Plan**; Evolving document

[RGMAINSTALL] DataGrid WP3; **R-GMA Installation Guide**;
http://hepunx.rl.ac.uk/edg/wp3/documentation/doc/installation.pdf

[RGMAUSER] DataGrid WP3; **R-GMA User Guide**;
http://hepunx.rl.ac.uk/edg/wp3/documentation/doc/user.pdf

[SANTAGUSER] Stuart Kenny; **SANTA-G User Guide**;
http://gridportal.fzk.de/autobuild/i386-rh7.3-gcc3.2.2/wp3_3_2-santag/userdoc/user/user.pdf

[SANTAGDEVEL] Stuart Kenny; **SANTA-G Developers Guide**;
http://gridportal.fzk.de/autobuild/i386-rh7.3-gcc3.2.2/wp3_3_2-santag/userdoc/developer/developer.pdf