

SCALABLE GANGLIA SETUP FOR GRID ENVIRONMENTS

August 30, 2004

Nirmal Ranganathan
nrangana@cse.uta.edu

Department of Computer Science & Engineering
University of Texas at Arlington

Introduction:

Ganglia as a cluster monitoring tool can be further tuned to be scalable for a grid environment. This configuration requires the ganglia gmetad daemon and the web front end to run on every cluster thereby preventing the overhead of all gmond daemon data being collected at one central location.

Working:

The setup basically requires one central gmetad daemon and web front end which will link to other cluster sites. The other cluster sites will each run their own gmetad daemon and web front end. Explained below is the gmetad configuration for both the central server and the remote clusters.

Note: For installation of ganglia refer to HEP notes 25.

Central Server Configuration:

Modify the `/etc/gmetad.conf` file to reflect the following changes.

Blue	-	Comments
Maroon	-	Required tags
Green	-	Explanation

```
# This is an example of a Ganglia Meta Daemon configuration file
#
#       http://ganglia.sourceforge.net/
#
#-----
--
# Setting the debug_level to 1 will keep daemon in the foreground and
# show only error messages. Setting this value higher than 1 will make
# gmetad output debugging information and stay in the foreground.
# default: 0
# debug_level 10
#
#-----
--
# What to monitor. The most important section of this file.
#
# The data_source tag specifies either a cluster or a grid to monitor. If we
# detect the source is a cluster, we will maintain a complete set of RRD
# databases for it, which can be used to create historical graphs of the
# metrics. If the source is a grid (it comes from another gmetad),
# we will only maintain summary RRDs for it.
#
# Format:
# data_source "my cluster" [polling interval] address1:port addresses2:port
```

```

...
#
# The keyword 'data_source' must immediately be followed by a unique
# string which identifies the source, then an optional polling interval in
# seconds. The source will be polled at this interval on average.
# If the polling interval is omitted, 15sec is assumed.
#
# A list of machines which service the data source follows, in the
# format ip:port, or name:port. If a port is not specified then 8649
# (the default gmond port) is assumed.
# default: There is no default value
#
# data_source "my cluster" 10 localhost my.machine.edu:8649 1.2.3.5:8655
# data_source "my grid" 50 1.3.4.7:8655 grid.org:8651 grid-backup.org:8651

data_source "SWIFT-FARM" localhost:8649
data_source "LTU Physics" caps10.phys.LaTech.edu:8651
data_source "OCHEP" luhep02.lunet.edu:8651
data_source "UTA-DPCC" master.dpcc.uta.edu:8651
data_source "SPRACE" sprace.if.usp.br:8651
data_source "TATA-MCFarm" d0farm2.hecr.tifr.res.in:8651

# Enter all the data sources with the appropriate port number, default is
# 8651.
#
#-----
--
# Scalability mode. If on, we summarize over downstream grids, and respect
# authority tags. If off, we take on 2.5.0-era behavior: we do not wrap our
# output in <GRID></GRID> tags, we ignore all <GRID> tags we see, and always
# assume we are the "authority" on data source feeds. This approach does not
# scale to large groups of clusters, but is provided for backwards
# compatibility.
# default: on
# scalable on
#
#-----
--
# The name of this Grid. All the data sources above will be wrapped in a GRID
# tag with this name.
# default: Unspecified
# gridname "MyGrid"
#
gridname "D0 SAR (Southern Analysis Region)"

# Any suitable name for the grid
#-----
--
# The authority URL for this grid. Used by other gmetads to locate graphs
# for our data sources. Generally points to a ganglia/
# website on this machine.
# default: "http://hostname/ganglia/",
# where hostname is the name of this machine, as defined by gethostname().
# authority "http://mycluster.org/newprefix/"
#
authority "http://hepfm000.uta.edu/ganglia"

```

```

# This is the link for the ganglia web front end on the site.
#
#-----
--
# List of machines this gmetad will share XML with. Localhost
# is always trusted.
# default: There is no default value
# trusted_hosts 127.0.0.1 169.229.50.165 my.gmetad.org
#
#-----
--
# If you want any host which connects to the gmetad XML to receive
# data, then set this value to "on"
# default: off
all_trusted on
#
#-----
--
# If you don't want gmetad to setuid then set this to off
# default: on
# setuid off
#
#-----
--
# User gmetad will setuid to (defaults to "nobody")
# default: "nobody"
# setuid_username "nobody"
#
#-----
--
# The port gmetad will answer requests for XML
# default: 8651
# xml_port 8651
#
#-----
--
# The port gmetad will answer queries for XML. This facility allows
# simple subtree and summation views of the XML tree.
# default: 8652
# interactive_port 8652
#
#-----
--
# The number of threads answering XML requests
# default: 4
# server_threads 10
#
#-----
--
# Where gmetad stores its round-robin databases
# default: "/var/lib/ganglia/rrds"
# rrd_rootdir "/some/other/place"

```

Remote Cluster Configuration:

Configure the */etc/gmetad* file to reflect the changes shown below. If the machine is behind a firewall enable access to port **8651**. This is very essential for retrieving the remote gmetad information.

```
# This is an example of a Ganglia Meta Daemon configuration file
#
#       http://ganglia.sourceforge.net/
#
#-----
--
# Setting the debug_level to 1 will keep daemon in the foreground and
# show only error messages. Setting this value higher than 1 will make
# gmetad output debugging information and stay in the foreground.
# default: 0
# debug_level 10
#
#-----
--
# What to monitor. The most important section of this file.
#
# The data_source tag specifies either a cluster or a grid to monitor. If we
# detect the source is a cluster, we will maintain a complete set of RRD
# databases for it, which can be used to create historical graphs of the
# metrics. If the source is a grid (it comes from another gmetad),
# we will only maintain summary RRDs for it.
#
# Format:
# data_source "my cluster" [polling interval] address1:port addresses2:port
# ...
#
# The keyword 'data_source' must immediately be followed by a unique
# string which identifies the source, then an optional polling interval in
# seconds. The source will be polled at this interval on average.
# If the polling interval is omitted, 15sec is assumed.
#
# A list of machines which service the data source follows, in the
# format ip:port, or name:port. If a port is not specified then 8649
# (the default gmond port) is assumed.
# default: There is no default value
#
# data_source "my cluster" 10 localhost my.machine.edu:8649 1.2.3.5:8655
# data_source "my grid" 50 1.3.4.7:8655 grid.org:8651 grid-backup.org:8651

data_source "SWIFT-FARM" localhost:8649

# This is the data source for the various cluster nodes. And the default port
# is 8649.
#
#-----
--
# Scalability mode. If on, we summarize over downstream grids, and respect
# authority tags. If off, we take on 2.5.0-era behavior: we do not wrap our
# output in <GRID></GRID> tags, we ignore all <GRID> tags we see, and always
# assume we are the "authority" on data source feeds. This approach does not
```

```

# scale to large groups of clusters, but is provided for backwards
# compatibility.
# default: on
# scalable on
#
#-----
--
# The name of this Grid. All the data sources above will be wrapped in a GRID
# tag with this name.
# default: Unspecified
# gridname "MyGrid"
#
gridname "SWIFT-FARM"

# Enter a suitable grid name. This name is the one that will appear on the
# main
# ganglia server.
#-----
--
# The authority URL for this grid. Used by other gmetads to locate graphs
# for our data sources. Generally points to a ganglia/
# website on this machine.
# default: "http://hostname/ganglia/",
#   where hostname is the name of this machine, as defined by gethostname().
# authority "http://mycluster.org/newprefix/"
#
authority "http://hepfm000.uta.edu/ganglia"

# This is the ganglia web front end on the local cluster.
#-----
--
# List of machines this gmetad will share XML with. Localhost
# is always trusted.
# default: There is no default value
trusted_hosts 127.0.0.1 129.107.50.100

# In this case the trusted host is the Swift farm, which collects the summary
# of all the clusters. For detailed monitoring the link is redirected to the
# cluster thereby reducing the load on one single server.
#
#-----
--
# If you want any host which connects to the gmetad XML to receive
# data, then set this value to "on"
# default: off
# all_trusted on
#
#-----
--
# If you don't want gmetad to setuid then set this to off
# default: on
# setuid off
#
#-----
--
# User gmetad will setuid to (defaults to "nobody")

```

```

# default: "nobody"
# setuid_username "nobody"
#
#-----
--
# The port gmetad will answer requests for XML
# default: 8651
# xml_port 8651
#
#-----
--
# The port gmetad will answer queries for XML. This facility allows
# simple subtree and summation views of the XML tree.
# default: 8652
# interactive_port 8652
#
#-----
--
# The number of threads answering XML requests
# default: 4
# server_threads 10
#
#-----
--
# Where gmetad stores its round-robin databases
# default: "/var/lib/ganglia/rrds"
# rrd_rootdir "/some/other/place"

```

Conclusion:

The above setup ensures a scalable ganglia monitoring system. It reduces the load on a single central server and distributes it across all the clusters.

References:

1. Ganglia online documentation, <http://ganglia.sourceforge.net/docs/>
2. HEP Notes Ganglia installation guide, <http://hepfm000.uta.edu/Ganglia-installation.pdf>
3. THEGrid Ganglia setup, <http://thegrid.dpcc.uta.edu/ganglia-setup.html>