

Group Management Service Configuration in Glassfish V3.1

1. Introduction

1.1. Project/Component Working Name:

Group Management Service (GMS-Shoal project)

1.2. Name(s) and e-mail address of Document Author(s)/ Supplier:

Sheetal Vartak : sheetal.vartak@oracle.com

Joe Fialli : joe.fialli@oracle.com

1.3. Date of This Document:

05/18/2010 - Joe, Update with feedback from April 29th admin meeting.

04/06/2010 - Sheetal, Initial version

2. Project Summary

2.1. Project Description:

GMS configuration

2.2. Risks and Assumptions:

No risks. But there is a plan to change certain property names to make them easier to understand. So there needs to be backward compatibility in place.

3. Problem Summary

3.1. Problem Area:

Shoal GMS requires configuration information to join a dynamic cluster. This information is configured in DAS domain.xml under cluster and group-management-service elements. The information is read and used to join the group in cluster/gms-adapter.

3.2. Justification:

The main theme for Glassfish 3.1 is Clustering and HA. GMS being an important piece of this theme needs to be configured correctly. Hence the need for support of the GMS configuration through domain.xml.

4. Technical Description:

4.1. Details:

The work mainly comprises of bringing 3.1 to 2.1.1 parity. All the GMS related attributes/properties that were supported in GF 2.1.X domain.xml will be brought over to 3.1. There are some

attributes whose names need to be changed just so that the names reveal the real reason behind the attribute and are not confusing to understand. For this, backward compatibility will need to be preserved. Some properties that were introduced in 2.1.1 need to be promoted to attributes. A few new attributes will need to be introduced due to support for Grizzly as the underlying transport layer.

4.2. Bug/RFE Number(s):
None so far.

4.3. In Scope:
Everything explained in 4.1.

4.4. Out of Scope:

4.5. Interfaces:
The Config API will need to be modified in order to incorporate the new GMS config changes.

- GMS config values in GF 2.1.1 structure in domain.xml

```
<cluster config-ref="dev-cluster-config"
  name="dev-cluster"
  heartbeat-enabled="true",
  heartbeat-address="any value in the range of 225.0.0.0 to
231.255..255.255",
  heartbeat-port="integer larger 2000"
  <property name="gms-bind-interface-address"
    value=${GMS_DEV-CLUSTER_BIND_INTERFACE_ADDR}"/>
  <property name="gms-initial-host-list"
    value="tcp://ipAddr1:port,tcp://ipAddr2:port,..."/>
  <property name="gms-member-type"
    value=${GMS_DEV-CLUSTER_MEMBER_TYPE}/>
</cluster>
```

cluster element notes:

In Glassfish v2.1.1 asadmin create-cluster created the above cluster element in domain.xml, The creation command ensured that cluster.heartbeat-address attribute value was unique for each cluster listed in the domain.xml. heartbeat-address has to be a valid multicast address.

```
<config dynamic-reconfiguration-enabled="true"
  name="dev-cluster-config">
```

```

...
<group-management-service fd-protocol-max-tries="3"
  fd-protocol-timeout-in-millis="2000"
  merge-protocol-max-interval-in-millis="10000"
  merge-protocol-min-interval-in-millis="5000"
  ping-protocol-timeout-in-millis="5000"
  vs-protocol-timeout-in-millis="1500">
  <!-- property below configures gms so when it attempts
    -- to connect to a suspected failed server instance,
    -- the tcp socket creation timeout should be
    -- set to 3 seconds. This value is probably too small
    -- but was necessary to achieve goal of detecting hw
    -- failure within 15 seconds.Default value of 10 seconds
    -- detects hw failure in 28 seconds.
  -->
  <property name="failure-detection-tcp-retransmit-timeout"
    value="3000"/>
</group-management-service>
</config>

```

See <http://appserver.sfbay.sun.com/Wiki.jsp?page=GMSHowTo> for documentation on how to configure cluster and group-management-service properties added in glassfish v2.1 timeframe.

Logging

```

<log-service alarms="false" file="{com.sun.aas.instanceRoot}/
logs/server.log" log-rotation-limit-in-bytes="2000000" log-
rotation-timelimit-in-minutes="0" log-to-console="false" retain-
error-statistics-for-hours="5" use-system-logging="false">
  <module-log-levels admin="INFO" classloader="INFO"
  cmp="INFO" cmp-container="INFO" configuration="INFO"
  connector="INFO" corba="INFO" deployment="INFO" ejb-
  container="INFO" group-management-service="INFO" javamail="INFO"
  jaxr="INFO" jaxrpc="INFO" jdo="INFO" jms="INFO" jta="INFO"
  jts="INFO" management-event="INFO" mdb-container="INFO"
  naming="INFO" node-agent="INFO" resource-adapter="INFO"
  root="INFO" saaj="INFO" security="INFO" self-management="INFO"
  server="INFO" synchronization="INFO" util="INFO" verifier="INFO"
  web-container="INFO"/>
  </log-service>

```

Need to make sure that the above way of specifying the log level works for GMS in v3.1. Need to eliminate the need to specify the ShoalLogger property.

- New 3.1 structure in domain.xml

Changes in gms config elements in DTD from v2 to v3.1.

Add following properties to group-management-service element.

ADD TCPSTARTPORT property which has an integer value.

ADD TCPENDPORT property which has an integer value.

Notes: initial GMS over Grizzly implementation needs these values. These properties define the range between which grizzly will select a tcp port for listening to. These will remain properties since we hope to not have to set them explicitly in final product.

ADD MEMBER_TYPE, IS_BOOTSTRAP_SEED

3 properties introduced : *member-type*, *is-bootstrap-seed*, *list-virtual-multicast-uri*. *member-type* defines whether the instance is a CORE member or a SPECTATOR member. *is-bootstrap-seed* defines if this node will be a bootstrapping host for other members to use for discovery purposes. *list-virtual-multicast-uri* is a comma separated list of initial bootstrapping tcp addresses. This address list must be specified on all members of the cluster through this property.

Add a new element under the *group-management-service* element called *failure-detection*. This element holds all the failure-detection related attributes. Also the gf 2.1 group-managment-service property *failure-detection-tcp-retransmit-timeout* has been renamed to the attribute of group-management-service.*failure-detection.verify-failure-connect-timeout-in-millis*.

Renamed 2 attributes under *cluster* :

From: heartbeat-address	To: gms-multicast-address
From: heartbeat-port	To: gms-multicast-port

Remove attributes *merge-protocol-max-interval-in-millis* and *merge-protocol-min-interval-in-millis* element group-management-service since they were never used in gf v2 and will not have any meaning in gf v3.1.

Need to specify a symbolic value that can be replaced per instance for cluster properties *gms-bind-interface-address* and *gms-member-*

type. The symbolic token value is set per clustered server instance via `system-property`.

Essentially the structure should look as follows (after changing the names of some attributes) :

```
<cluster config-ref="dev-cluster-config"
  name="dev-cluster"
  gms-enabled="true"
  gms-multicast-address="228.8.20.94"
  gms-multicast-port="17227"
  <property name="gms-list-virtual-multicast-uri"
    value="tcp://ipAddr1:port,tcp://ipAddr2:port,..." />
>
</cluster>
```

```
<config dynamic-reconfiguration-enabled="true" name="dev-cluster-
config"/>
```

```
...
<group-management-service
  group-discovery-timeout-in-millis="5000"
  bind-interface-address="228.234.54.55"
  >

  <failure-detection max-missed-heartbeats="3"
    heartbeat-frequency-in-millis="2000"
    verify-failure-waittime-in-millis="1500"
    tcp-retransmit-timeout="3000"/>
  <property name="is-bootstrap-seed" value="false"/>
  <property name="TCPSTARTPORT"
    value = "<integer larger than 8024 and less than
tcpEndPort property>"/>
  <property name="TCPENDPORT" value="<integer>"/>
  <property name="MEMBER_TYPE"
    value="${GMS_CLUSTER_NAME_MEMBER_TYPE}/>
  </group-management-service>
</config>
```

- Property name changes

fd-protocol-max-tries changed to *failure-detection.max-missed-heartbeats*

fd-protocol-timeout-in-millis changed to *failure-detection.heartbeat-frequency-in-millis*

ping-protocol-timeout-in-millis changed to *group-discovery-timeout-in-millis*

vs-protocol-timeout-in-millis changed to *failure-detection.verify-failure-waittime-in-millis*

failure-detection-tcp-retransmit-timeout changed to *failure-*

detection.verify-failure-connect-tinmeout-in-millis.

In the *cluster* element, the following attribute names need to be changed:

heartbeat-address changed to *gms-multicast-address*

heartbeat-enabled changed to *gms-enabled*

heartbeat-port changed to *gms-multicast-port*

- backward compatibility for the property name changes

The server mbeans are auto-generated. So some amount of work needs to be manually done to make sure that the domain.xml is backward compatible.

4.6. Doc Impact:

Documentation will need to incorporate the changes suggested above.

4.7. Admin/Config Impact:

Admin gui/cli related changes :

- expose new attributes and attribute name changes

- property-modified-to-attribute change

4.8. HA Impact:

The config changes will only affect how GMS is started.

4.9. I18N/L10N Impact:

No.

4.10. Packaging & Delivery:

No impact.

4.11. Security Impact:

No impact.

4.12. Compatibility Impact

If older attribute/property names are used, then the class/interface needs to provide a solution to deal with backward compatibility.

// List any requirements on upgrade tool and migration tool.

4.13. Dependencies:

5. Reference Documents:

```
// List of related documents, if any (BugID's, RFP's,
papers).
// Explain how/where to obtain the documents, and what each
// contains, not just their titles.
http://appserver.sfbay.sun.com/Wiki.jsp?
page=SetGMSIPAddressInAppServer
```

6. Schedule:

6.1. Projected Availability:

```
// Dates in appropriate precision (quarters, years)
```