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:
     06/21/2010 = Joe, simplify by replacing GMS_TCPSTARTPORT and
GMS_TCPENDPORT with GMS_LISTENER_PORT.GMS_LISTENER_PORT
     06/18/2010 - Joe add content where TBD is in section 4.7
     06/17/2010 - Joe, corrections made during implementation
phase for GMS_LISTENER_PORT and GMS_TCPENDPORT in section 4.5.
Moved these
from group-management-service property to cluster property.
     06/01/2010 - Joe, Add documentation for the gms
configuration.
     05/26/2010 - 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 gf v3.1 module
cluster/gms-adapter.

    3.2. Justification:
     The main theme for Glassfish 3.1 is Clustering and HA. GMS

being am 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:

    Despite specifying v3.1 cluster property **gms-virtual-multicast-uris,** the functionality that implements this is not in scope for gf v3.1.

4.5. Interfaces:

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

GMS configuration in GF 2.1.1 structure in domain.xml

Below is an extraction of relevant GMS configuration related xml
document fragments from a GF v2.1.1 domain.xml.
The example configures a glasfish cluster named *dev-cluster*.

Relevant GMS elements, attributes and properties are in **bold.**

```
<cluster name="dev-cluster"
        config-ref="dev-cluster-config"
        heartbeat-enabled="true",
        heartbeat-address="<legal multicast address>"
        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>

<config name="dev-cluster-config"
       dynamic-reconfiguration-enabled="true">
  ...
    <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 name="failure-detection-tcp-retransmit-timeout"
             value="3000"/>
  </group-management-service>
</config>
```

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.

See link http://wiki.glassfish.java.net/Wiki.jsppage=FaqShoalGMSPropertiesInDomainXML on how to configure properties gms-bind-interface-address and failure-detection-tcp-retransmit-timeout introduced in gf v2.x. See http://appserver.sfbay.sun.com/Wiki.jsp?page=GMSHowTo for documentation on how to configure MEMBER_TYPE added in glassfish v2.x timeframe.

- New gf v3.1 structure in domain.xml

Below is an example illustrating all possible elements/attributes and properties existing in glassfish v3.1 domain.xml for configuring GMS.

```
<cluster
  name="dev-cluster"
  config-ref="dev-cluster-config"
  gms-enabled="true"
  gms-multicast-address="229.9.1.1"
  gms-multicast-port="2299"
  gms-bind-interface-address=
                          "${GMS-BIND-INTERFACE-ADDRESS-dev-cluster}"
  <property name="GMS_LISTENER_PORT"
          value="${GMS_LISTENER_PORT-dev-cluster}"/>
  <property name="gms-list-virtual-multicast-uri"
          value="tcp://ipAddr1:port,tcp://ipAddr2:port,,.."/>
  <property name="gms-member-type"
          value="${GMS-MEMBER-TYPE-dev-cluster}"/>
</cluster>

<config name="dev-cluster-config"
        dynamic-reconfiguration-enabled="true"/>
  ...
  <group-management-service group-discovery-timeout-in-millis="5000">
    <failure-detection
        max-missed-heartbeats="3"
        heartbeat-frequency-in-millis="2000"
        verify-failure-waittime-in-millis="1500"
        verify-failure-connect-timeout-in-millis="10000"/>
  </group-management-service>
```

```
</config>

// If possible, would prefer to advise setting system-property for DAS in
// domain.xml servers.server[@name="server"]. Here is what that
// looks like in v3 today.
<server name="server" config-ref="server-config">
    <resource-ref ref="jdbc/__TimerPool" />
    <resource-ref ref="jdbc/__default" />
     <system-property name="GMS_LISTENER_PORT-dev-cluster"
value="9090"/>
</server>

// in gf v2, we set system-property on "server-config" with other
// system property.
<server config-ref="server-config" name="server">
      <!-- deleted non-essential info to this issue -->
      <system-property
            name="GMS_LISTENER_PORT-dev-cluster"
            value="9090"/>
      <system-property
         name="GMS-BIND-INTERFACE-ADDRESS-dev-cluster"
            value="129.148.71.168"/>
</server>

// setting gms system properties for a clustered instance.
<server config-ref="devtest-cluster-config" name="instance1">
      <!-- deleted non-essential info to this issue -->

     <!-- optional GMS system properties -->
     <system-property
            name="GMS_LISTENER_PORT-dev-cluster"
            value="9121"/>
     <!-- following optional GMS system property specifies that a
       -- clustered instance does not participate in HA when ha is
       -- enabled.
       -->
     <system-property
            name="${GMS-MEMBER-TYPE-dev-cluster}"
            value="SPECTATOR"/>
</server>
```

Summary of xml element/attribute/property changes from
gf v2.x to gf 3.1 domain.xml.

Below are the mapping transformations that gms gf v3.1 upgrade

will perform to upgrade a gf v2 domain.xml to gf v3.1 format.

cluster element changes:

– Renamed 3 attributes under *cluster* :
    From: **heartbeat-enabled**   To: **gms-enabled**
    From: **heartbeat-address**   To: **gms-multicast-address**
    From: **heartbeat-port**   To: **gms-multicast-port**

– Promoted gf v2.1 cluster property **gms-bind-interface-address** to
an attribute of cluster in v3.1.

– Renamed gf v2.1 cluster property **gms-initial-host-list** to
  gms-list-virtual-multicast-uri.

group–management–service element changes:

– Rename an attribute of group–management–service element:
    From: **ping-protocol-timeout-in-millis**
    To: **group-discovery-timeout-in-millis**

– Add a new element under the *group–management–service* element
  called *failure-detection*. Move a number of attributes of
  gf v2.x group–management–service element to be attributes of
  gf 3.1 group–management–service.failure-detection element.

    From: **fd-protocol-timeout-in-millis**
    To: **failure-detection.heartbeat-frequency-in-millis**

    From: **fd-protocol-max-tries**
    To: **failure-detection.max-missed-heartbeats**

    From: **fd-protocol-timeout-in-millis**
    To: **failure-detection.verify-failure-waittime-in-millis**

- Promote gf v2.x group-managment-service property
  failure-detection-tcp-retransmit-timeout to gf v3
  group-management-service.failure-detection attribute
  verify-failure-connect-timeout-in-milis

– Add following properties as predefined cluster property.
  These port ranges must by mutually exclusive for each glassfish
server with gms-enabled running on same machine. (Otherwise, one
will see address in use failures from gms usage of grizzly.)

If there is only one application server running on a machine, then these values need not be set.

   ADD **GMS_LISTENER_PORT** 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.

   Usage described.

   GMS subsystem presets each of these cluster properties to a symbolic token so each clustered instance and DAS can set the values.

   for a cluster named "myFirstCluster", here are the cluster properties.

```
<cluster name="myFirstCluster" ....>
   <property name="GMS_LISTENER_PORT"
             value="${GMS_LISTENER_PORT-myFirstCluster}"/>
</cluster>
```

   Then in server element of a clustered instance, the system property for the cluster is set. If DAS and instance1 and instance2 are all to run on same machine, then all tcp port ranges must be mutually exclusive.

```
<server name="instance1" configref="myFirstCluster-config">
   <system-property name = "${GMS_LISTENER_PORT-
myFirstCluster}"
                    value = "9090"/>
</server>


<server name="instance2" configref="myFirstCluster-config">
   <system-property name = "${GMS_LISTENER_PORT-
myFirstCluster}"
                    value = "9121"/>
</server>

   // DAS.  note that it needs to be able to configure
   // for multiple clusters.
<config name="server-config">
   <system-property name = "${GMS_LISTENER_PORT-
myFirstCluster}"
                    value = "9151"/>
```

```
            <system-property name = "${GMS_LISTENER_PORT-
mySecondCluster}"
                             value = "9181"/>
     </config>
```

Given that the desire is for the need for the above to go away,
GMS_LISTENER_PORT are cluster properties and not attributes.
Since group-management-service can be shared between clusters,
these properties ended up on cluster and not group-management-
service.

- Remove attributes *merge-protocol-max-interval-in-millis* and
merge-protocol-min-interval-in-millis from gf v3.1 element **group-
management-service** since they were never used in gf v2.x 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.

- 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.

****

GMS configuration attribute/property documentation:

     For each GMS configuration attribute/property below, document
whether it is required or optional, whether there exist a default
value, the valid values for the attribute/property, relationship
to other attributes/properties.

TBD: finish documentation below.

- **cluster** attributes

gms-enabled
Its default value is **true**.
If gms-enabled is set to **false**, gms is completely disabled for this glassfish
cluster. No GMS injected services are available. No HA is available.
When **gms-enabled** is set to **false**, no other gms configuration parameters are
processed.

gms-multicast-address

This attribute is optional.
Either this atttribute or cluster property
gms-list-virtual-multicast-uri must be set when **gms-enabled** is set to **true**.
When cluster property gms-list-virtual-multicast-uri is not set to a null or empty
string value, its default value is "229.9.1.1".
Implementation enforced constraints: this value should differ from all other
cluster gms-multicast-address specified in a domain.xml.

gms-multicast-port="2299"
This attribute defaults to the value 2299.

gms-bind-interface-address="<local network interface IP address>"
On a multi-home machine(possessing 2 or more network interfaces),
this attribute allows one to configure which network interface is used for GMS.
This attribute is optional. It does not have a default value. This value
should be a IP address of a local network interface IP address.
Multicast must enabled for this interface UNLESS **gms-list-virtual-multicast-uri**
is set. GMS Implementation must validate that this is a valid local network
interface IP address.
Options to handle invalid value:
1. fail fast - Consider a SEVERE error and do not allow application server to
start. Diagnostic output instructs user on how to configure properly. Plan on
using this approach.
2. print a WARNING that the value is invalid and default to using the
   default algorithm to select which Network Interface is used. Did this in
glassfish v2.x. It is too much of a surprise that this is only a WARNING when it
was set incorrectly.

- **cluster** properties
   name = **GMS_LISTENER_PORT**
   value ="${GMS_LISTENER_PORT-${CLUSTERNAME}}"

Specify a unique GMS port. When multiple glassfish instances are started on one
machine, this value must be unique across the glassfish instances.

   gms-list-virtual-multicast-uri
                value="tcp://ipAddr1:port,tcp://ipAddr2:port,,.."
   The property is optional. It is required to be set when gms-multicast-address
   is not set. This property is not scheduled to be implemented. When this
property is set, **GMS_LISTENER_PORT** must be same value as port listed in this

list for ipAddrX above. Additionally, on a multihome machine, there may be a need to set **cluster** attribute **gms-bind-interface-address** to ipAddrX.

   gms-member-type
            value="${GMS-cluster-name-MEMBER_TYPE}"

Allow any com.sun.enterprise.mgmt.transport.grizzly.GrizzlyConfigConstants properties to be configured in cluster.property. These values enable tuning of the transport. The same GrizzlyConfigConstant occurring in both cluster and group-management-service, the cluster one takes precedence.

* * * * * * * *

- **group-management-service** attributes
  group-discovery-timeout-in-millis

When a GMS client is first joining its group, this is time that it waits for a response back from the Master of the gms group. The default is 5 seconds. After this time is expired, then the instance will make itself the master of the group. Typically, the DAS is started before all other instances in the cluster and it is the defacto Master of GMS group. When the DAS is stopped while the clustered instances are still running, then all remaining gms group members run the same algorithm to select a successor to previous Master of gms group.

- **group-management-service.failure-detection** attributes
     max-missed-heartbeats="3"
     heartbeat-frequency-in-millis="2000"
     verify-failure-waittime-in-millis="1500"
     verify-failure-connect-timeout-in-millis="10000"

```
The existing documentation for GF v2.x group-management-service
attributes at GlassfishWiki GMS FAQ is sufficient for the above
properties. One just needs to map between the GF v2.x names for
these attributes/properties to the GF v3.1 attribute names.
```

group-management-service properties

Allow any com.sun.enterprise.mgmt.transport.grizzly.GrizzlyConfigConstants properties to be configured here. These values enable tuning of the transport.

* * * * * * * * *

```
Logging

Below is an example of gf v2.1 logging in domain.xml.
```

```
<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"/>

        <!-- WORKAROUND in glassfish v2.x to enable Shoal GMS logging.
          --For GF v3.1, module-log-levels.group-management-service
          -- should work. -->
        <property name="ShoalLogger"  value="FINE"/>
    </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.

****

        4.6. Doc Impact:

        Documentation is needed.

        Here was a gf v2 FAQ discussing various group-management-
service attributes. The same basic discussion applies (with
appropriate mapping of names from gf v2 to new gf v3.1 names by
reader) Link: gfv2.1 group management service attributes.

        4.7. Admin/Config Impact:

Admin CLI

Replace deprecated heartbeatport and heartbeataddress with
multicastport and multicastaddress.

Example:

asadmin create-cluster --multicastport 2050 --multicastaddress 230.9.1.1 someName

TBD: Complete admin GUI for GMS.

```
    Admin gui/cli related changes :
    - expose new attributes and attribute name changes
    - property-modified-to-attribute change
    - add ability to provide generic cluster properties in
glassfish v3.1 admin GUI.  In GF v2.1,  this capability was not
provided and only could introduce cluster properties via asadmin
CLI set command. In GF v2.1, there was a capability to add generic
properties for group-management-service and these will still be
needed in GF v3.1.

    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.

ShoalGMSPropertiesInGlassfish_v2_DomainXML

http://appserver.sfbay.sun.com/Wiki.jsp?page=SetGMSIpAddressInAppServer

6. Schedule:
    6.1. Projected Availability:
     // Dates in appropriate precision (quarters, years)
```