

Single-Endpoint Client Configuration: Using External Load Balancer (LBR) with Oracle Global Data Services (GDS)

Simplifying Oracle GDS with External Load Balancers for Seamless Connectivity

REV 1 , Version **[1.0]**
Copyright © 2025, Oracle and/or its affiliates
Public

Oracle Global Data Services (GDS) provides high availability, load balancing, and seamless failover for distributed and replicated databases, with Global Service Managers (GSMs) dynamically managing workloads based on real-time performance and availability. While GDS optimizes database connectivity, some organizations choose to integrate external load balancers to further abstract the infrastructure and provide a durable, single endpoint configuration that can transparently withstand GSM setup and configuration changes. This white paper explores the use of the Oracle Cloud Infrastructure (OCI) Network Load Balancer and the F5 BIG-IP Local Traffic Manager (LTM) with GDS, detailing the configuration steps and best practices.

Introduction

Oracle Global Data Services (GDS) is a powerful framework designed to manage replicated and distributed databases, providing high availability, load balancing, and seamless failover capabilities. At the heart of Oracle GDS are the Global Service Managers (GSMs), which act as intelligent traffic directors, dynamically managing connections to the databases based on workload, availability, and defined policies, and ensuring continuous database connectivity and optimal performance.

In certain scenarios, customers may choose to deploy an external load balancer in front of GSMs within their Oracle GDS configurations. Such an external load balancer may enable a single endpoint configuration and allow for making GSM setup changes transparent to the application. This approach may be particularly relevant in environments with stringent high-availability requirements or need to abstract GSM setup changes from the application. An external load balancer can transparently distribute traffic across multiple GSM instances, ensure instant redundancy in case of a GSM failure, and provide centralized traffic control across hybrid and multi-cloud architectures.

This blog post explores the use of external load balancers with Oracle Global Data Services, providing a detailed overview of the available options, their features, and implementation best practices. We will discuss two popular load balancer choices: Oracle Cloud Infrastructure (OCI) Network Load Balancer (NLB) and F5 BIG-IP Local Traffic Manager (LTM). We will outline the technical steps for setting up these external load balancers with Oracle Global Data Services. Additionally, we will list best practices for ensuring robust, secure, and high-performing configurations.

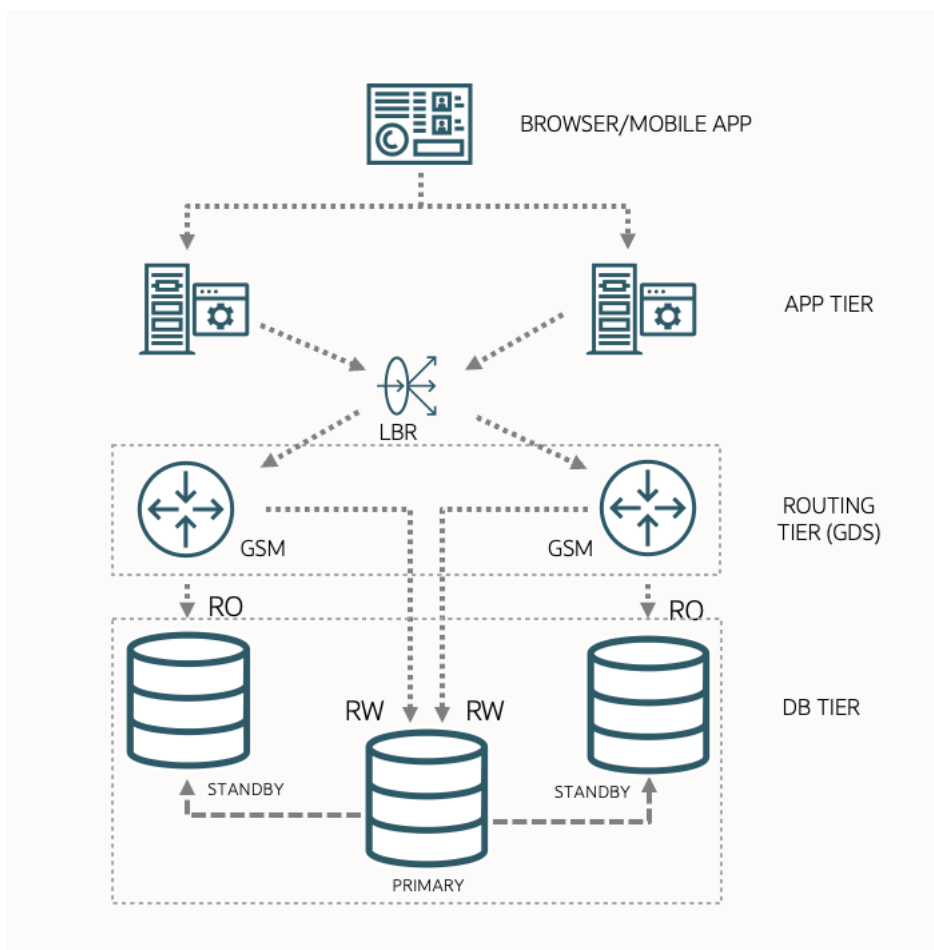


Figure 1. Using an external load balancer with Oracle Global Data Services

Setting Up External Load Balancers with Oracle GDS

Prerequisites

In a GDS configuration, the client application initially connects to a Global Service Manager (GSM) node. The GSM then provides a TNS redirect, enabling the client to establish a direct connection to the database hosting the requested service. The GSM does not act as a proxy or provide NAT (Network Address Translation) services. Similarly, the external load balancer configuration applies only to this initial client-to-GSM connection.

To successfully configure an external load balancer with Oracle Global Data Services (GDS), the following pre-requisites must be met:

1. System Configuration Without Load Balancer

The system should first be fully configured and operational without the load balancer. The load balancer's virtual IP (VIP) is not required in any GSM configuration steps. Clients should be able to connect directly to GSMs using GSM node addresses.

2. Direct Connectivity Between Client and GSM Nodes

The client application must be able to connect to the GSM Listener and the ONS servers on the GSM nodes. By default, these operate on ports 1522 (Listener) and 6234 (ONS), but these may vary based on configuration. The addresses of all GSM nodes must be resolvable and accessible by the client application.

3. Direct Connectivity Between Client and Database Nodes

The client application must also have the ability to connect directly to the database nodes. This requires SQL connections to the Database Listeners, which typically use port 1521, though the exact port may vary. The addresses of all database nodes must be resolvable and accessible by the client application.

By ensuring these prerequisites are met, you establish the foundational network and connectivity setup required for the successful integration of an external load balancer into an existing GDS environment.

Supported Load Balancers

The load balancer used for handling the initial client request to the GSM must operate at Layer 4, as no packet inspection is required or desirable in this configuration. The primary role of the load balancer is to provide a Virtual IP (VIP) and transparently forward traffic to one of the GSM nodes.

This section outlines configuration details for two recommended Layer 4 load balancers: Oracle OCI Network Load Balancer and F5 BIG-IP Virtual Edition (VE).

Oracle OCI Network Load Balancer

Oracle Cloud Infrastructure (OCI) offers two types of load balancers: the Oracle Load Balancer, which supports Layer 7 load balancing, and the OCI Network Load Balancer, which provides simpler Layer 4 functionality. For this use case, the OCI Network Load Balancer is recommended due to its straightforward configuration and compatibility with GSM requirements. While the Oracle Load Balancer can also be used, the Network Load Balancer is ideal for handling Layer 4 traffic in this scenario.

F5 BIG-IP Virtual Edition (VE)

For Layer 4 load balancing, we used the F5 BIG-IP Virtual Edition (VE). Our testing specifically used F5 BIG-IP VE Release 17.1, but other recent releases of F5 BIG-IP should also provide the required functionality. This load balancer offers robust and reliable Layer 4 traffic handling, ensuring seamless connectivity between clients and GSM nodes.

Topology

The topology below shows the updated GDS deployment architecture with an external Load Balancer, in this case F5 (but it can also be the OCI Network Load Balancer), within the network.

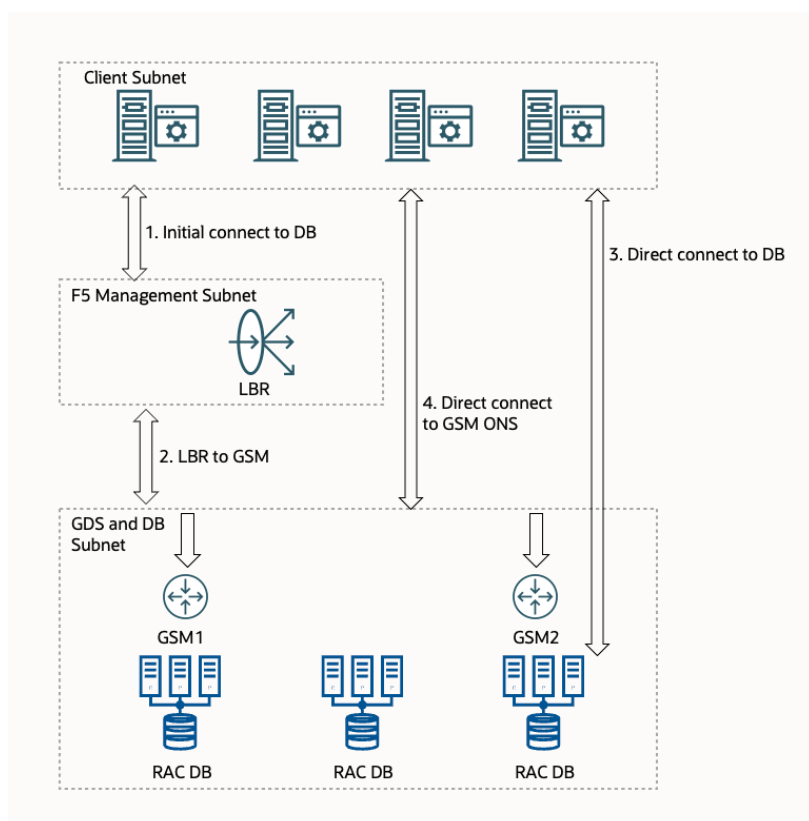


Figure 2. Request flow in an Oracle GDS setup with an external load balancer

Request Flow

1. The client initiates a connection to the Load Balancer (LBR), which transparently forwards the request to a Global Service Manager (GSM) via the Virtual IP (VIP).
2. The GSM redirects the client to the appropriate database where the requested service is offered.
3. Once the redirection is complete, the client establishes a direct connection to the database, bypassing the LBR for subsequent database interactions.
4. Upon successfully connecting to the database, the client receives Oracle Notification Service (ONS) information from the database node via the Auto-ONS process and subscribes to the GSM ONS server for real-time event notifications.
5. After the initial connection, the client communicates directly with both GSM and database nodes.
6. The LBR does not function as a NAT gateway and does not act as a SQL proxy—it is only used for the initial connection to the GSM.

GSM Configuration

In this setup, Global Service Managers (GSMs) function as Remote Listeners, managing and directing client connections to the appropriate database services. The Local Listeners reside on one or more database nodes, handling direct communication with the databases. For high availability, it is recommended to deploy at least two GSMs to ensure redundancy and failover protection. Depending on the expected connection traffic, additional GSMs may be required to maintain optimal performance and scalability.

The configuration steps for GSM and database nodes are not covered in this document, as the initial GSM setup does not impact the integration of a load balancer. The load balancer is simply introduced to manage the initial client connection to GSM.

Configuring the Oracle Cloud Infrastructure (OCI) Network Load Balancer (NLB)

Creating OCI NLB

The OCI Network Load Balancer can be created from the Networking section of the OCI in the OCI UI or with OCI CLI. Here we walk through the Create Screens in the UI.

Add Details

Creating an OCI NLB will also assign a Public or Private IP and place the NLB in a desired subnet. This can be either the same subnet as the GSMs and Databases or, like the F5 configuration, a separate subnet which can access the required ports of the Database subnet.

Create network load balancer

Add details

- Configure listener
- Choose backends
- Review and create

A network load balancer provides automated traffic distribution from one entry point to multiple servers in a backend set. The network load balancer ensures that your services remain available by directing traffic only to healthy servers in the backend set.

Load balancer name: my_nlb

Choose visibility type

- ☒ **Public**
You can use the assigned public IP address as a front end for incoming traffic.
- ☐ **Private**
You can use the assigned private IP address as a front end for internal incoming VCN traffic.

☐ Allow IPv6 address assignment
Enables a dual-stack IPv4/IPv6 implementation for your load balancer. Learn more about [IPv6 addresses](#).

Assign a public IP address

- ☒ **Ephemeral IPv4 address**
Automatically assign an IPv4 address from the Oracle pool.
- ☐ **Reserved IPv4 address**
Select an existing reserved IPv4 address or create a new one from one of your IP pools.

Choose networking

Virtual cloud network in **Ravi, GDS** ([Change compartment](#))
rnet

Subnet in **Ravi, GDS** ([Change compartment](#))
rnetsubnet1

☐ Use network security groups to control traffic

[Show advanced options](#)

[Next](#) [Cancel](#)

Configure Listener

Give the Listener a name.

Specify TCP as the Protocol and 1522 as the Ingress Traffic Port. This will be the Port at which the NLB receives requests from clients.

Create network load balancer

Configure listener

- Add details
- Choose backends
- Review and create

A listener is a logical entity that checks for incoming traffic on the network load balancer's IP address. You must configure at least one listener for each traffic type in order to handle UDP/TCP or UDP/TCP traffic. You can configure additional listeners after you create your network load balancer.

Listener name: my_listener

Specify the type of traffic your listener handles

UDP **TCP** UDP/TCP

Ingress traffic port

- ☐ Use any port
This will use a wildcard, or 0, as the port.
- ☒ **Specify the port**
1522

This port will be used for the backend ports, if any, selected in the following step.

Proxy protocol

- ☐ Enable proxy protocol V2
Sends additional connection information, such as the source and destination IP addresses and ports, to their backends across layers of network address translation (NAT) or TCP proxies.

[Previous](#) [Next](#) [Cancel](#)

Choose Backends

Give the Backend Set a name. This will be the pool of GSMs to direct Traffic to.

Click on 'Add Backends'. On the popup screen, add each of the GSM nodes, with a port of 1522. This must match the Port used by the GSM Listeners.

Click 'Add Backends' when finished.

On the original screen, ensure that 'Preserve Source IP' is unchecked.

For the Health Check Policy, enter the following information:

- For Protocol: Choose TCP.
- For Port: Choose 1522, the port used by the GSM Listeners
- For Interval: Leave as default or choose a value that is equal to the `TRANSPORT_CONNECT_TIMEOUT`

- For Timeout: Leave as default or choose a value that is at least 3x the Interval
- Choose Next to complete this page

Create network load balancer

Add details

Configure listener

Choose backends

Review and create

A network load balancer distributes traffic to backends within a backend set. A backend set consists of one or more instances.

Backend set name
my-backendset

Select backends

Add by compute instance or IP addresses. Optional

No backends selected. Click add backends to select resources from a list of available resources.

Add backends

☒ Preserve source IP

Specify health check policy

A health check is a test to confirm the availability of backends. A health check can be configured with various parameters.

Protocol
TCP

Interval in MS. Optional
10000

Number of retries. Optional
3

Request data. Optional
"GET / HTTP/1.1"

This will be base64 encoded

Response data. Optional
"HTTP/1.1 200 OK"

This will be base64 decoded

☐ Fail open
If selected, the network load balancer continues to route traffic to backends in this backend set if some are unhealthy.

☐ Enable instant failover
When enabled, existing flows will be redirected to a healthy backend when the existing backend becomes unhealthy.

Show advanced options

Add backends

Choose how to add backends by selecting compute instances or by entering IP addresses.

Backend type

☒ Compute instances☐ IP addresses

Source preservation must be disabled in order to add IP addresses. Learn more about source preservation.

Backend IPv4 compute instances

Instance in Raw_GDS	IP address	Availability domain
No data available	No data available	No data available

Port 1502Weight 1X

+ Another backend

[illegible]

Review and Create

Review the information and click 'Create Network Load Balancer' when ready.

Please Review the 'Additional Notes on the TNS Monitor from the Oracle Networking team' posted below.

Configuring the Health Monitor

Notes on the TCP Monitor

The TCP Health Monitor was created when creating the NLB.

This should be sufficient but please consult the Section in F5 configuration in this document called 'Suppressing Monitor Logging' which also applies here.

Configuring a TNS Monitor

A more complete health check for Oracle Listeners is a TNS Ping. This may be more appropriate in Enterprise environments. This check contacts the Listener with a SQL NET Ping request and waits for a reply. Oracle has a utility called *tnsping* that does just this, but we can emulate what *tnsping* does by allowing the load balancer to send and receive the equivalent.

For OCI NLB, we cannot enter the required Send/Receive responses in the UI as the UI only accepts Text which it will encode into base64, and we need to send non-ASCII header information as well.

We can accomplish this with OCI CLI, however. This assumes you have already created a TCP Monitor as above.

First, ensure you have the backend-set-name and the Network Load Balancer OCID. We will send the command '(CONNECT_DATA=(COMMAND=ping)(CONNECTION_ID=loadbalancer))' and look for the response 'ERR=0' in the Reply text.

```
$ oci nlb health-checker update --backend-set-name "my_backendset" --
network-load-balancer-id $NETWORK_LOAD_BALANCER_ID
--request-data
"AIAAAAEAAAABQAEsAAAgAP//fwgAAAEAADkASgAAAAAACa8AAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAoQ09OTkVDVF9EQVRBP
ShDT01NQU5EPXBpbmcpKENPTk5FQ1RJT05fSUQ9bG9hZGJhbGFuY2VyKSsk="
--response-data "RVJSPTA="
```

The response data is merely the desired reply text 'ERR=0' but base64 encoded. If a different reply text is desired, then the text must first be encoded as base64.

To alter the Send string if required:

1. Convert the Send string (e.g. (CONNECT_DATA=(COMMAND=ping))) first to Hex
2. Concatenate the Headers (in Hex), provided below to the Send string in Hex.
3. Take the concatenated Hex string and convert it to base64 using an available conversion tool
4. Use the base64 string in the OCICLI command above.

The required header Hex string referenced above is:

```
00830000010000000140012c00002000ffff7f08000001000039004a0000000000
0026bc0000000000000000000000000000000000000000000000000000000000002000002000
00000000000000000000
```

LBR High Availability

The OCI Network Load Balancer comes with High Availability built in. A standby load balancer is created and, in the case of a failure, the LBR IP is automatically migrated.

More information is available here: https://docs.oracle.com/en-us/iaas/Content/Balance/Concepts/load_balancer_types.htm

Installing and configuring F5 Virtual Edition (VE)

Installing F5 VE

The process of installing F5 Virtual Edition (VE) is documented in the official F5 documentation and is not covered here. F5 documentation also covers the installation of F5 VE in Oracle Cloud Infrastructure (OCI).

https://clouddocs.f5.com/cloud/public/v1/oracle/oracle_deploy.html

All the steps outlined in the F5 documentation should be followed, up until before the step "Create a Virtual Server" which will be covered here.

At this point, we assume that F5 VE has been successfully installed, activated, and the user is able to access the F5 web-based configuration utility.

Once F5 VE is installed, it must be configured to distribute traffic to the GSM nodes in the GDS setup. This involves creating GSM nodes, defining a load-balancing pool, and configuring a virtual server.

Creating the GSM nodes

The first step is to register the GSM nodes in the F5 configuration.

On the web-based configuration utility:

1. On the Main tab, expand Local Traffic, and then click Nodes.
2. Click the Create button.
3. In the Address box, specify the IP Address of one of your GSM nodes.
4. In the Name box, specify a name for the node. This is only used as a reference within F5 and can be the short host name, for example.
5. We will add the Health Monitors later.
6. Click Repeat.
7. Repeat this procedure for all the GSM nodes. GSM1 and GSM2 in our example topology.

Creating the Pool

Once the nodes have been created, group them into a pool to enable load-balancing.

1. On the Main tab, expand Local Traffic, and then click Pools.
2. Click the Create button. The New Pool screen opens.
3. From the Configuration list, select Advanced.
4. In the Name box, enter a name for your pool.
5. From the Load Balancing Method list, choose 'Round Robin'
6. In the Address box, specify the IP Address of one of the GSMs.
7. In the Service Port box, specify the GSM Listener Port. This port must match the port number of the Listener on the GSM Host. For example, 1522.
8. Click the Add button.
9. Repeat steps 6-8 for each of the GSM nodes.
10. Click Finished.

Note: Optionally, TCP Profiles can be created for greater control and to not interfere with SQL NET configuration. For a default SQL*NET configuration, TCP Profiles should not be necessary. This is especially true since our connection is not a long-running connection but a quick connection to the GSM Listener. We can instead rely on the Health Monitor settings in F5 to stop sending connections to GSMs that are congested or failed.

Creating the Virtual Server

A Virtual Server must be created to serve as the primary entry point for client connections. This Virtual server will use the secondary private IP address that's associated with the external network interface. This IP Address will be the one that Clients connect to in their connect strings.

Assigning a Secondary Private IP in OCI

1. In the OCI Console, navigate to the F5 Compute instance.
2. Go to Attached VNICs, select the F5 External VNIC.
3. Assign a Secondary Private IP Address.
4. This IP will serve as the connection endpoint for clients in their TNS connect strings.

Configuring the Virtual Server in F5

1. In the F5 Configuration Utility, navigate to: Local Traffic → Virtual Servers
2. Click Create.
3. Complete the following information:

Attribute	Value	Comments
Name	Any unique name for this Virtual Server such as 'GSM Region1'	
Destination Address/Mask	The secondary IP Address on the External NIC	This must be an already available Virtual IP that is attached to the VM
Service Port	1522	This is the port that the F5 will receive requests on. It does not have to match the GSM port but it is good practice to do so.
Source Address Translation	Auto Map	
Default Pool	Select the GSM pool created earlier.	
Default Persistence Profile	select dest_addr	
Fallback Persistence Profile	select source_addr	

Configure any other settings as needed, and then click Finished.

Creating and Configuring Health Monitors

A Health Monitor is used by the F5 Load Balancer to detect the health of the GSM Listeners and to remove from the pool any GSM that is non-responsive.

What we are interested in monitoring is the health of the GSM Listener. Note that we are not interested in monitoring the health of the back-end databases. Database events are still monitored by FAN events and GSM will also relocate DB Services as needed.

F5 has also de-supported the Oracle Health Monitor for recent releases of Oracle

Database: <https://my.f5.com/manage/s/article/K40226145> Thus, it is not recommended to use F5's Built-in Oracle Health Monitor.

Within F5, we have two options for monitoring the health of the GSM Listeners:

TCP Monitor

A TCP Monitor will check that something is listening on the intended port. This is sufficient for most cases.

Among the monitors F5 provides is a TCP Half-Open Monitor. This Monitor forgoes the client acknowledgement (step 3 in the 3-way handshake) and thus is lighter on the server.

If a TCP Half-Open Monitor is available, then this is recommended over a TCP Monitor.

Configuring the TCP Monitor

To configure this in F5:

1. Go to Local Traffic → Monitors → Create
2. Choose a Name and choose TCP Half Open from the Type dropdown.
3. For the Interval, choose a value that is equal to the `TRANSPORT_CONNECT_TIMEOUT` or leave as default if this is not known.
4. Set Timeout to `Interval*3+1`. Or leave as default if Interval was defaulted.
5. Set the Alias Service Port as 1522 (or the listen port of GSM)

Suppressing Monitor logging

TCP Monitors will often cause unnecessary logging by GSM as it is not a recognized SQL NET connection.

These logs can be disabled by setting the parameter `log_suppress_nodes` in the `gsm.ora` file. This file can be found in `$GSM_HOME/network/admin`.

This parameter takes a list of addresses (IP or hostname) of nodes that should not be logged.

An example is below where we have specified the IP address of F5 load balancers that will be performing health checks on this Listener.

```
LOG_SUPPRESS_NODES=(10.0.2.1, 10.0.2.2)
```

TNS Ping Monitor

For a more comprehensive health check of Oracle Listeners, a TNS Ping test can be utilized, particularly in enterprise environments. This method verifies listener availability by sending a SQL*Net ping request and awaiting a response. While Oracle provides the `tnsping` utility for this purpose, the same functionality can be replicated by configuring the load balancer to send and interpret equivalent requests.

Configuring a TNS Ping Monitor on F5

To set up a TNS ping monitor on F5, follow these steps:

1. Create a TCP Monitor
 - Navigate to Local Traffic → Monitors → Create.
 - Assign a Name and select TCP as the Monitor Type.
2. Set the Send String
 - The `tnsping` utility sends a request using a string like:
(CONNECT_DATA=(COMMAND=ping)(CONNECTION_ID=loadbalancer))
 - The `CONNECTION_ID` is optional and can be any value for tracking requests.
Since the request requires additional header data without an ASCII representation, enter the header in Hex format in the F5 Send String field. Hex is entered in the Send string box by preceding it with a `'\x'`
 - The final Send String should look like this:

```
\x00\x83\x00\x00\x01\x00\x00\x00\x01\x40\x01\x2c\x00\x00\x20\x00\xff\xff\x7f\x08\x00\x00\x01\x00\x00\x39\x00\x4a\x00\x00\x00\x00\x00\x00\x26\xbc\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x20\x00\x00\x20\x00\x00\x00\x00\x00\x00\x00\x00\x00(CONNECT_DATA=(COMMAND=ping)(CONNECTION_ID=loadbalancer))
```

3. Set the Receive String

- The GSM Listener will reply with a response like: (DESCRIPTION=(TMP=))(VSNNUM=0)(ERR=0)(ALIAS=GSMA))
- The Receive String box should contain a recognizable portion of this response to confirm a successful connection.
- Enter ERR=0 (without quotes) in the Receive String field to validate the listener's reply.

In addition, the following should be noted:

- The Interval and Timeout can be set as the TCP Monitor above.
- The *tnsping* request check will work on *tnsping* response timeout. The response timeout will have to be appropriately configured to avoid false positives which will depend on listener load and protocol in use.
- The TNS check will not work on a TLS Listener. (See general note on TLS at the end of this document) In the TLS case a simpler TCP Check is recommended.
- The Listener has a RATE_LIMIT feature which if enabled, limits the number of connections it processes per second. A *tnsping* based health check cannot be used on a rate limited port.

Applying the Health Monitors

To apply either of the Health Monitors in the F5 environment:

1. Select a Node from Local Traffic→ Nodes
2. For Health Monitor, choose 'Node Specific'
3. Choose the Health Monitor created earlier and click 'Update'
4. Do this for all Nodes

Load Balancer High Availability

To ensure a highly available environment, a secondary F5 Load Balancer (LBR) can be configured as a failover counterpart. Detailed instructions for setting up a secondary F5 LBR in Oracle Cloud Infrastructure (OCI) are available in F5 documentation. See, https://clouddocs.f5.com/cloud/public/v1/oracle/oracle_deploy.html#deploy-big-ip-b

Setting up a Second LBR and Configuring Synchronization

Setting up a failover-ready LBR involves establishing trust between the two load balancers and configuring a synchronization (sync) group. Below are key considerations and troubleshooting tips when following the instructions above:

Check	Details
NTP Servers	The two LBRs must be time-synchronized. Configure Network Time Protocol (NTP) by adding an NTP server in the F5 UI under System → Configuration → Device → NTP. Do this for both load balancers.
SelfIPs	When following the steps above: • Verify that Self IP addresses are configured correctly, ensuring the correct VLAN assignment, as this may be the source of miscommunications. • Ensure that Port Lockdown settings are not too restrictive and allow communications between the loadbalancers

Check	Details
Firewalls and Security Lists	Ensure that both load balancers can communicate over essential ports. This includes 443 (HTTPS), 4353 (F5 Config Sync), 1026 (UDP failover), and 22 (SSH).

Automating the Failover of the OCI Virtual IP

While F5 devices monitor each other and automatically initiate failover, F5 Virtual Edition (VE) cannot directly alter the VIP configuration on the host VM. In OCI, this is handled via an automation script available on GitHub.

See, <https://github.com/f5devcentral/f5-oci-failover>

Upon failover, these scripts trigger the migration of the Virtual IP to the new active F5 instance. The setup can be tested by forcing a failover event and verifying that client connections seamlessly transition to the new active F5 instance.

Configuring the Client Applications

Once the failover setup is complete, client applications should be configured to send all requests to the Load Balancer endpoint. Below is an example TNS connection string following Oracle Maximum Availability Architecture (MAA) best practices:

```
my_global_svc =
(DESCRIPTION =
(ADDRESS = (PROTOCOL = TCP) (HOST = gsm-lbr.example.com) (PORT = 1522))
(CONNECT_TIMEOUT=90) (RETRY_COUNT=10)
(RETRY_DELAY=5) (TRANSPORT_CONNECT_TIMEOUT=3)
(CONNECT_DATA = (SERVICE_NAME = my_global_svc.region.oradbcloud)))
```

To verify the connection using SQL*Plus:

```
$ sqlplus user/pwd@my_global_svc
```

ONS Configuration in Oracle GDS

Oracle Notification Service (ONS) is a critical component in enabling Fast Application Notification (FAN), which informs clients about database events such as service availability, node failures, or recoveries. Typically, ONS servers run on database nodes (default port 6200) and broadcast FAN events to subscribed clients.

However, in an Oracle Global Data Services (GDS) setup, ONS services are centralized on the GSM servers, operating on default port 6234.

- GSM servers aggregate FAN events from all database ONS servers in the GDS-managed pool.
- Clients subscribe to the GSM ONS network rather than individual database ONS servers.
- ONS notifications remain consistent, even if services are relocated to different databases.

The ONS architecture used in GDS setups ensures reliable FAN event propagation, simplifies client-side ONS configuration, and supports dynamic database relocation without additional client modifications.

Simplifying Client Configuration with Load Balancers

Integrating an external Load Balancer (LBR) within a region further reduces client configuration complexity while maintaining high availability at all levels.

With an LBR in place, clients only need to connect to a single LBR endpoint, making them agnostic to backend changes, including:

- The number and addresses of Global Service Managers (GSMs)
- The number and locations of database nodes providing a service
- Changes in underlying database configurations (RAC, Single Instance, Data Guard, GoldenGate, etc.)

This approach enhances resilience, scalability, and ease of maintenance, ensuring seamless connectivity without requiring manual client-side updates.

Application Security

Global Service Managers reside in the same network as the target databases, often the same subnet. This ensures not only a performant solution but also that GSM nodes benefit from the same network security provided to database network configurations.

The use of TLS listeners is supported for Global Data Services although the additional steps and caveats are not listed in this document.

High availability at each level in an Oracle GDS setup

Database High Availability

High availability at the database level is achieved through Oracle Real Application Clusters (RAC) and Oracle Data Guard, ensuring continuous service availability and resilience against failures.

- **Node-Level Failure:** In a RAC environment, failure of an individual node triggers RAC failover and service migration, ensuring uninterrupted access.
- **Service Routing:** GSM dynamically routes connection requests to active service instances.

- **Cluster-Level Failure:** If an entire RAC cluster fails, Data Guard automatically triggers a failover to a standby database, maintaining service continuity.
- **Role Transition Management:** GSM handles automatic role transitions, ensuring that services remain available without manual intervention.
- **Fast Connection Failover (FCF) Support:** GSM also facilitates Oracle Notification Services (ONS), allowing clients to receive real-time outage notifications and failover seamlessly.

Global Service Manager (GSM) High Availability

To maintain high availability of the Global Service Manager (GSM), multiple redundant GSM nodes should be deployed.

- **Stateless Architecture:** GSMs are lightweight and stateless, meaning any GSM instance can handle incoming requests.
- **ONS Network Redundancy:** GSMs form an ONS network, ensuring notifications continue even if a GSM node fails.
- **Failover Handling:** If a GSM fails, clients automatically reconnect to an available GSM, ensuring uninterrupted operation.

Load Balancer High Availability

A highly available load balancer (LBR) setup ensures that client connections remain resilient to LBR failures.

- **Standby Load Balancers:** The deployment includes redundant standby load balancers to handle failures.
- **Automatic Failover:** In the event of a primary LBR failure, traffic is seamlessly rerouted to a standby LBR, ensuring continuity.
- **Consistent Client Connectivity:** Since the client connect address remains unchanged, applications remain unaffected during LBR failovers.

Global Traffic Management (GTM) Considerations

The setup described primarily focuses on a Local Traffic Management (LTM) solution, where traffic is managed within a single region.

- **Regional GSM:** Within each region, local GSM instances manage database services.
- **Cross-Region Traffic Management:** In a multi-region deployment, a Global Traffic Manager (GTM) is required to distribute client connections across regions.

- GTM Solutions: Enterprises can leverage Oracle Cloud Infrastructure (OCI) Traffic Manager or F5 GTM DNS services to intelligently route traffic across geographically distributed database environments.

Conclusion

Oracle Global Data Services (GDS) provides a powerful framework for managing distributed and replicated databases, ensuring high availability, intelligent connection routing and load balancing, and seamless failover. Integrating external load balancers, such as Oracle Cloud Infrastructure (OCI) Network Load Balancer or F5 BIG-IP LTM, can further enhance GDS setups by providing a single endpoint configuration and GSM failover handling.

This white paper has outlined some of the best practices, configuration steps, and deployment strategies to successfully integrate external load balancers with Oracle GDS. As businesses continue to scale across multi-cloud and hybrid environments, leveraging Oracle GDS with external load balancers can allow for seamless application connectivity and business continuity.

Connect with us

Call **+1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

Copyright © 0000, Oracle and/or its affiliates. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle, Java, MySQL, and NetSuite are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.