High Availability and Disaster Recovery for the Genesys SMS Aggregation Service

Genesys SMS Aggregation Service 2016CD

12/29/2021

Table of Contents

High Availability and Disaster Recovery for the Genesys SMS Aggregation Service

This document explains the High Availability and Disaster Recovery capabilities of the Genesys SMS Aggregation service. The terms *High Availability* and *Disaster Recovery* are often used interchangeably and therefore it's important to first define them:

- High Availability (HA)—HA minimizes service outages from a single datacenter by ensuring that all components are deployed using redundant and/or fault-tolerant components.
- Disaster Recovery (DR)—DR is the ability to provide service in a major event which causes a complete outage at a datacenter.

Hereinafter we will use the abbreviations HA and DR.

Document Overview

This document details the design approaches use to provide these two capabilities.

Datacenters

The SMS Aggregation service runs in the following geographically diverse Datacenters (DCs):

- Ashburn, VA Datacenter (dc3)
- San Jose, CA Datacenter (sv5)

These datacenters are designed to run an active/active configuration and therefore you can use either site. When using these datacenters, note the requirements as defined in HA SMSC , DR SMSC , and DR HTTP .

> **Important**
> If you need additional contact information to access these datacenters, contact your Genesys representative.

High Availability

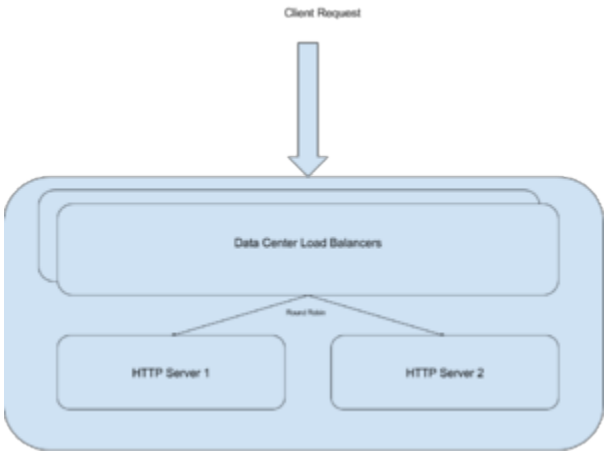
The SMS Aggregation service is made up of several different components. Each of these components is configured for High Availability (HA) using either active/active pairs or active/hot-standby pairs. All client-facing components are made available in an active/active configuration. Any components not configured as active/active have a designated failover policy that is applied in case a component fails. In addition, each component is provisioned with enough capacity to handle the complete system load at any time.

To avoid any interruptions to service during maintenance or outages, follow the recommendations below:

HTTP APIs

The HTTP APIs are available for use at any site where the SMS Aggregation service is located. Details of the URLs are discussed in HTTP API URLs.

The architecture of the SMS Aggregation service is the same for each site and is shown in the following figure.



HTTP API Architecture (click to enlarge)

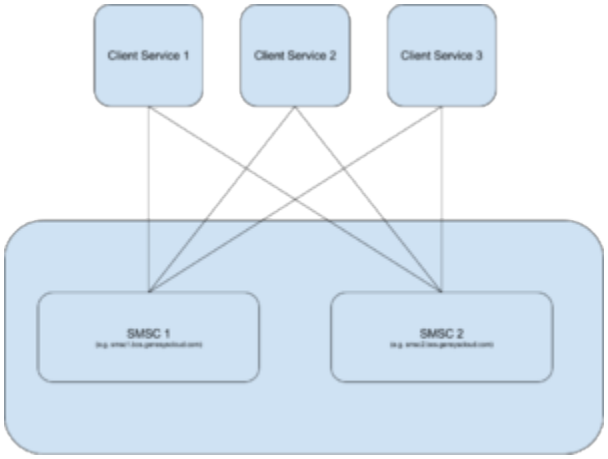
The flow of a client's request is as follows:

1. The client's request is received at the datacenter firewall.
2. The request is forwarded to load-balancing cluster.
3. The load balancer selects the HTTP node using round-robin balancing.

SMSC APIs

The SMSC component provides HA by supporting multiple client connections to each node within the SMSC cluster. The complete list of available URIs are listed here. To take full advantage of the HA SMSC service, a client should comply with the following recommendations:

1. Create one SMPP connection per SMSC node (as shown in the figure below).
2. If the SMSC node becomes unavailable, attempt to rebind in 30 second intervals.
3. Clients should expect MOs and DRs to be sent to ANY active SMPP connection.



SMSC DR Example Configuration (click to enlarge)

Disaster Recovery

This section discusses the Disaster Recovery (DR) plan for the SMS Aggregation service. This service follows the *natively multi-homed design*. With this design, all datacenters are set up as hot sites.

To ensure there is no service impact during any major outage (or maintenance), customers are advised to use all the available sites. However, note the following:

- Clients should support receiving MOs from any site.
- DR for MTs sent to one site may be received at another.

Genesys ensures that no duplicate messages are delivered. The rest of this section discusses the requirements for each component.

HTTP API

There are two strategies you can use for HTTP API DR:

- Single Genesys-provided URL—Genesys maintains short TTLs for external DNS records and, in the event of a failure, updates the DNS information to redirect them to the failover site.
- Client-side failover—customers can configure all site URLs to use as failover sites.

The following section discusses these two approaches.

Genesys-Provided Failover

The URL to use for this strategy is:

- `https://smsc-api.genesyscloud.com/httpapi/receiver`

In the event of a failure, the DNS resolution of this site is updated to the correct location. Therefore, in order to take full advantage of this feature, ensure that the DNS TTL of the authoritative server is obeyed. However, if this cannot be guaranteed (for example, an application layer continues caching forever), then you should use the alternative customer-side approach described below.

Customer Failover Configuration

As an alternative strategy, you can configure both site URLs to use as potential failover sites in your applications. In this configuration, Genesys recommends that you select one of the sites to be primary and the other site used in the case of failure.

The available URLs are:

- `https://smsc-api.dc3.genesyscloud.com/httpapi/receiver`
- `https://smsc-api.bos.genesyscloud.com/httpapi/receiver`

In the event of a outage or failure at one site, you may experience connection timeouts and/or other abnormal behavior. If such events occur, try the following steps:

1. Attempt to resubmit the message after a short duration, such as 30 seconds.
2. If the same error occurs again, use the alternative URL.
3. If you continue to experience problems, contact your Genesys account manager.

SMSC API

Due to the connection-oriented nature of this protocol, your client should create at least one connection to each available SMSC instance from each of your client applications.

The list of available SMSC instances are:

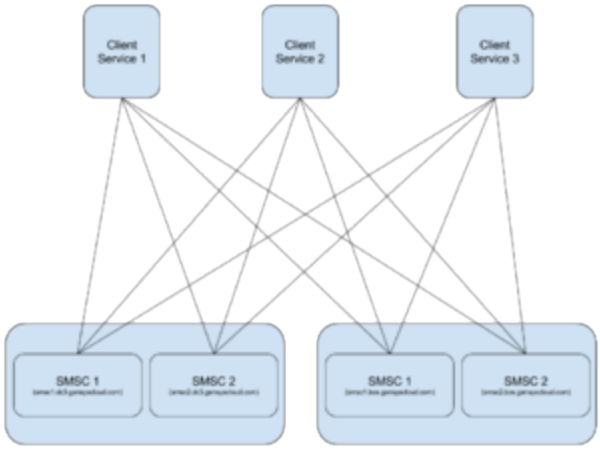
- `smsc1.dc3.genesyscloud.com`
- `smsc2.dc3.genesyscloud.com`
- `smsc1.bos.genesyscloud.com`
- `smsc2.bos.genesyscloud.com`

General Guidelines

In all cases, follow this set of general guidelines:

- A client should bind into all listed SMSCs.
- When an SMSC is unavailable, the client should attempt to rebind. The time between these attempts should be greater than 30 seconds.
- A client should expect messages to be received from all active connections.
- No guarantee is given as to the bind used to send DRs/MOs to the client.
- Clients should not queue messages for inactive binds. Instead, any active binds should be used.

The diagram below shows the desired configuration:



Client Configuration for SMSC DR