Global Server Load Balancing Overview

https://campus.barracuda.com/doc/3539022/

In this article:

Overview

Global Server Load Balancing (GSLB) allows you to coordinate how traffic is processed among multiple data centers. A Barracuda Load Balancer acts as a controller, selecting the location to which traffic is directed based on the parameters that you configure and the health of the data centers. This allows you to allocate the work among multiple data centers and to ensure that if one data center fails then traffic is redirected automatically to a functioning data center.

GSLB Examples

GSLB can be useful when:

- You have a number of server farms that are physically located around the world and you want incoming connections to be directed to the closest healthy server farm.
- You have two data centers and you want one of them to be reserved for use in the event of a disaster. You can assign the first with a high priority and have all traffic directed to it, while the other is used only if the first data center fails.
- You have multiple data centers and each has region-specific content. Depending on the location of the client, requests can be directed to the data center most appropriate for that region.

GSLB Definitions

- A **site** is a network location that hosts data. It may be a Service on a Barracuda Load Balancer with a server farm or one Real Server.
- A **GSLB Controller** is the Barracuda Load Balancer which determines where traffic is directed. It contains configuration information about the sites and it performs health checks on all sites in regular intervals. Only one GSLB Controller is active at a time. It is recommended that you configure one or more backup GSLB Controllers.
- A **region** defines a geographical area, usually composed of one or more countries. You can define custom regions or use the predefined regions.

Site Selection Criteria

The GSLB Service allows you to specify traffic to be directed to a site based on one of three parameters:

- Proximity of the system making the request to a site that can serve the request;
- The region of the system making the request; or
- The priority order of the sites.

How GSLB Works

The GSLB Controller controls which IP address for a sub-domain is given to a client. These steps illustrate the process:

1. A client tries to connect to a domain name such as `www.example.com`. It asks its local DNS server for the IP address of the domain name, and the server issues a DNS request on its behalf.
2. This request is eventually directed to the GSLB Controller (Barracuda Load Balancer) that acts as an authoritative DNS server for the delegated sub-domain www. The GSLB Controller considers the site selection algorithm and the health of the sites and issues a DNS response that contains a list of one or more IP addresses of valid sites.
3. The client tries to connect to the first address in the list.

In the following figure, the selection algorithm is based on the region of the client. The GSLB Controller determines the region where the request originated. The US client is returned the address of the site which handles clients from the US region (207.77.188.166) while the client from Europe is given the address of the site which supports content for the European region (216.129.205.232).

Figure 1. How GSLB Works.
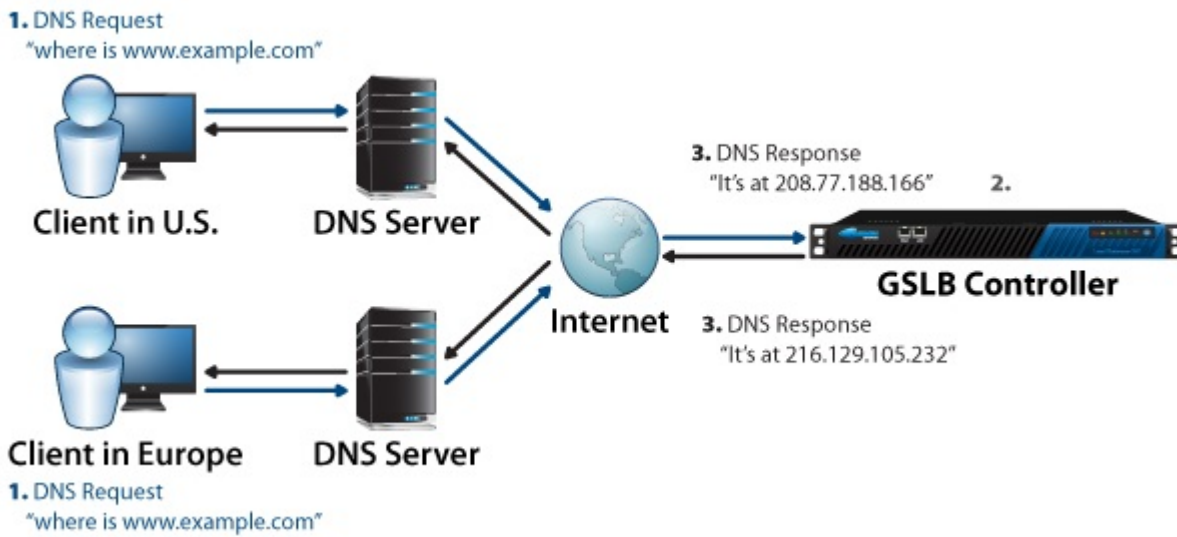
Failover

The record that is returned by the GSLB Controller in response to a DNS query has a time to live (TTL) value of 10 seconds, meaning that the DNS servers across the Internet need to request the IP address of the site again if the record is older than 10 seconds. If a site becomes unavailable, it is removed from the list of returned IP addresses, the caches update quickly, and traffic is directed to a healthy site.

Figures

1. GSLB_crop.png