



BALANCEADOR DE CARGA

/ ficha técnica

gigas

BALANCEADOR DE CARGA

El Balanceador de Carga de Gigas es un dispositivo virtual que se pone al frente de un conjunto de servidores para asignar o balancear las solicitudes que llegan de los clientes. El balanceador incorpora un mecanismo de revisión periódico de cada uno de los servidores, que permite excluir del pool aquellos nodos (o servidores) que están caídos, de una manera totalmente transparente para el usuario.

El balanceador de Gigas es una solución desarrollada por nosotros y por tanto no tiene una licencia comercial, lo cual nos permite ofrecer un producto tecnológicamente avanzado a un precio muy competitivo. Para su correcta instalación es necesaria la contratación de una VLAN. El cliente puede fácilmente incorporar ambos productos a su Cloud Datacenter en cuestión de minutos a través del panel de control de Gigas. Al tratarse de un dispositivo virtual, el usuario puede disfrutar del mismo de forma exclusiva, siendo posible por ejemplo, su reinicio en cualquier momento, así como la posibilidad de aumentar los recursos según sus necesidades.

PRINCIPALES VENTAJAS

Aumenta el volumen de carga del sistema

El balanceador permite aumentar la capacidad de los sistemas entre los que se está distribuyendo la carga. Sea cual sea el límite de conexiones web de un servidor, en el momento en que se instala un segundo servidor y se balancea, pasaremos a disponer de dos servidores que trabajaran en paralelo, haciendo que el sistema soporte el doble de conexiones concurrentes. De la misma forma, al sumar varios frontales, aumenta también la velocidad de conexión.

Mayor seguridad

El balanceador es el único elemento que está expuesto a internet y que tiene direccionamiento público. Los servidores a los cuales el balanceador distribuye la carga se encuentran conectados con direccionamiento privado. De esta manera, si alguien intentara atacar el sistema, solo tendría acceso al balanceador y no podría acceder directamente a los servidores web.

Alta disponibilidad de servicio

Todos los servidores virtuales desplegados por Gigas tienen un sistema de failover. Esto significa que si por ejemplo el servidor (equipo físico donde está instalado ese servidor virtual) tuviera algún problema, el sistema sacaría al servidor virtual de ese

servidor físico, se lo llevaría a otro y lo levantaría. Este proceso de failover o de alta disponibilidad, puede generar un corte de servicio de minutos hasta que se complete el movimiento del servidor. Para aquellos clientes que no pueden soportar ningún corte de servicio, la solución pasa por la instalación de un balanceador. Éste, constantemente está distribuyendo las peticiones entre dos servidores, y chequeando la salud de ambos. Antes de enviar la petición, el balanceador comprueba que el servidor funciona correctamente, y caso que hubiera un problema, en lugar de dirigir esa petición a un servidor (que marca como fallo), se redirigiría al otro. Este entorno de sistemas distribuidos permite que los fallos de alguno de los servidores sean transparentes para los clientes.

Clúster en alta disponibilidad

Las soluciones de clúster garantizan una disponibilidad total ante ciertas situaciones como por ejemplo, un fallo del balanceador. Al estar en clúster, el balanceador se compone y funciona con dos dispositivos (los cuales se monitorizan entre ellos). Cuando la petición llega al balanceador, éste funciona como un ente o clúster en el cual la petición se redirige hacia los servidores web, pero si el balanceador tuviera algún

PUNTOS CLAVES

- Dispositivo virtual dedicado a cada cliente
- Aumenta la capacidad de carga y velocidad de conexión
- Mayor seguridad, protege al servidor
- Sistema de failover transparente para el usuario
- Alta disponibilidad, solución en clúster
- Balanceo configurable según necesidades del cliente
- Información exhaustiva con estadísticas avanzadas
- Direcciones IP virtuales
- Función de auto-escalado

problema, ese otro equipo de balanceo en clúster, asumiría el servicio y de forma transparente seguiría redirigiendo las peticiones. De esa forma aseguramos un entorno de alto rendimiento y disponibilidad para los clientes. La alta disponibilidad del balanceador no añade complejidad en su gestión, ya que el sistema se encarga de mantener sincronizada la configuración entre ambos.

Funcionalidades

CONFIGURACIÓN DEL BALANCEO

El Balanceador de Carga de Gigas ofrece a sus clientes tres tipos de balanceo:

1. **Round Robin o balanceo aleatorio:** para clientes que quieren que cuando llegue una petición al balanceador, estas se redirijan a los servidores que hay por debajo en formato aleatorio, de
2. **Least conexión:** para clientes que prefieren añadir inteligencia a ese balanceo y que las peticiones se distribuyan donde haya menos carga, de manera que se asegure un nivel consistente de carga entre los servidores que se encuentran balanceados.
3. **Balanceo por fuente de origen/mantenimiento de sesión:** ba-

manera que unas veces vaya a uno, y otras veces vaya a otro.



lanceo en el cual siempre se mapea un mismo cliente a un mismo servidor. Hay aplicativos específicos en los que por ejemplo se requiere que hasta que un servidor no se haya caído, todas las peticiones que se originen desde un mismo cliente vayan siempre al mismo servidor. Es el ejemplo claro de los paneles de control en los cuales el usuario para acceder necesita logarse. Si la siguiente vez que el usuario visita la página el balanceador le redirigiera a otro servidor, éste tendría que volver a logarse. Para evitar eso, el cliente puede elegir esta opción de balanceo, que asegura que todas las peticiones del mismo cliente vayan redirigidas siempre al mismo servidor, y se mantenga así la sesión.

AUTO-ESCALADO

Esta funcionalidad permite que el balanceador se encargue de monitorizar el estado de los servidores que están por debajo, de manera que cuando se cumpla algunas de las reglas que el cliente haya definido de antemano, se aumente de forma automática el número de servidores entre los que se está distribuyendo el servicio. Útil para cubrir los picos de demanda o tráfico de manera autónoma, sin la necesidad de que haya un técnico monitorizando el sistema.

Esta funcionalidad avanzada desarrollada en Gigas, se contrata con las opciones de balanceador y VLAN.

Los drivers o reglas que puede configurar el usuario a la hora de que el balanceador vaya ajustando los recursos son básicamente 3:

- **Planificada:** el cliente define la hora exacta para aumentar recursos.
- **Por máximo de CPU** que puede estar utilizando un servidor durante un tiempo determinado: el cliente establece una regla en la que cuando un servidor esté a más de un determinado % de CPU durante el tiempo que se estipule, se entienda que existe un problema de saturación, y automáticamente el balanceador cree nuevos servidores dentro del entorno de balanceo.
- **Por nº máximo de sesiones:** el cliente suele tener muy claro el número de sesiones que su servidor puede soportar, de manera que le puede decir al balanceador que cuando se acerque a ese tope máximo, automáticamente incremente el rendimiento global del servicio añadiendo nuevos servidores.

ESTADÍSTICAS DE BALANCEO

El usuario puede controlar a través del propio panel, las estadísticas del balanceador. Puede ver como se está comportando el balanceador cada vez que le llega una petición, a cuantos servidores está balanceando, a que servidores está redirigiendo cada petición, cual es el estado de esos servidores, si están vivos, si existe algún

problema, etc. Todo esto se puede ver desde las estadísticas que están directamente incluidas en el balanceador.

CONFIGURACIÓN AVANZADA

Existen otras opciones avanzadas de balanceo como por ejemplo:

1. **Hash-type consistence:** permite una asignación estática de las peticiones a un servidor, evitando por ejemplo que al añadir o quitar servidores del grupo de balanceo, se vea afectada la asignación existente.
2. **Número mínimo/máximo:** se establece el número mínimo o máximo de conexiones que el cliente quiere que tenga un determinado servidor, de manera que cuando este servidor llegue a un número máximo de conexiones, la siguiente petición se redirija directamente al otro servidor.
3. **Prioridades entre servidores:** se establece sobre los servidores a los que se está balanceando un nivel de prioridad a la hora de redireccionar peticiones.
4. **Cookie Jsession:** permite que mediante la programación basada en este tipo de cookies, se puedan establecer asignaciones estáticas entre la petición que llega al balanceador y el servidor al que se dirige.

DIRECCIONES IP VIRTUALES (VIPS)

El balanceador tiene asociado dos tipos de IP. Una IP física, y una IP privada que es la que utiliza para conectarse directamente con los servidores a los cuales está balanceando el tráfico. Adicionalmente a estas dos IPs, existe una tercera llamada IP virtual de balanceo (IP sobre la que el cliente configura su dominio). Esta IP virtual es una IP pública que lo que hace es flotar entre dos balanceadores, y apuntará a uno u otro dependiendo de su correcto funcionamiento, de manera que puede estar asociado al primero, y caso que hubiera un problema y fallara, automáticamente se asociaría al segundo, que sigue recibiendo todas las peticiones, y hace que el fallo sea transparente para el usuario.

ADDONS

El Balanceador de Carga de Gigas es un producto independiente, y dispone de diferentes addons a añadir:

- Opción clúster, para garantizar un servicio ininterrumpido en alta disponibilidad.
- Incorporación de IPs virtuales, de forma que un mismo balanceador puede balancear múltiples instancias web, múltiples servicios, o servicios web con distintas IP's. Esto en determinadas configuraciones es necesario, y en nuestro caso el cliente puede ir añadiendo de forma puntual en base a sus necesidades.

CARACTERÍSTICAS TÉCNICAS

Soporta solución en alta disponibilidad

Throughput: hasta 200 Mbits por segundo por balanceador

Hasta 15000 conexiones recurrentes por balanceador

Posibilidad de configurar distintas VIPs de servicio (Virtual IPs)

Distintos sistemas de balanceo y chequeos de salud de los servidores balanceados

Estadísticas en tiempo real

Habilita soluciones de auto-escalado