

Universidad Politécnica de Madrid  
Escuela Técnica Superior de Ingenieros de Telecomunicación



**Desarrollo de escenarios virtuales didácticos  
para experimentar con técnicas de balanceo de  
carga para el acceso a aplicaciones en centros  
de datos**

**TRABAJO FIN DE MÁSTER**

**Santiago Vinicio Mora Tinoco**

2018



Universidad Politécnica de Madrid  
Escuela Técnica Superior de Ingenieros de Telecomunicación

**Máster Universitario en  
Ingeniería de Redes y Servicios Telemáticos**

**TRABAJO FIN DE MÁSTER**

**Desarrollo de escenarios virtuales didácticos  
para experimentar con técnicas de balanceo de  
carga para el acceso a aplicaciones en centros  
de datos**

Autor

**Santiago Vinicio Mora Tinoco**

Director

**David Fernández Cambronero**

Departamento de Ingeniería de Sistemas Telemáticos

2018

## Resumen

En la actualidad, nos enfrentamos a una importante demanda tecnológica y de servicios de internet lo que conlleva a la transferencia de la información de todo tipo, y de servicios web extremadamente alta, por lo que es incuestionable que los sistemas informáticos deben funcionar de forma ininterrumpida y sin errores, debido a esto es que la mayoría de empresas busca maneras de mejorar su infraestructura y otorgar un mejor servicio a sus clientes de forma inmediata, segura y eficaz, teniendo en cuenta siempre la alta disponibilidad, velocidad y seguridad en sus equipos de trabajo.

Por lo antes mencionado es necesario contar con herramientas que nos permitan optimizar el rendimiento del tráfico generado en la red hacia los distintos servidores que almacenan diferentes tipos de servicios, es a raíz de todo esto que surgen las tecnologías de balanceo de carga.

En el trabajo se hará un estudio de las técnicas utilizadas para balancear el tráfico de las aplicaciones en la nube en su acceso a los centros de datos que las soportan. Se estudiarán tanto las técnicas de balanceo de carga entre centros de datos redundantes como entre los diferentes servidores dentro de cada centro de datos, centrándose principalmente en las técnicas basadas en DNS. Adicionalmente, se crearán escenarios virtuales didácticos que permitan experimentar con estas técnicas para afianzar los conocimientos sobre ellas.

Con el objetivo de facilitar el entendimiento de las tecnologías de balanceo de tráfico se propuso el diseño de dos escenarios virtualizados en VNX con lo cual se tiene una mayor flexibilidad y fácil despliegue, el primero consiste en el desarrollo de balanceo de carga realizado entre servidores dentro de un centro de datos esto mediante la utilización de la herramienta HAProxy, mientras que para el segundo escenario se plantea el balanceo de carga entre 2 centros de datos mediante el uso de geolocalización dada por DNS con el uso de la herramienta Bind9.

## Abstract

Currently, we are faced with an important technological and internet services demand which leads to the transfer of information of all kinds, and extremely high web services, so it is unquestionable that computer systems must operate without interruption and without errors, due to this is that most companies look for ways to improve their infrastructure and provide a better service to their customers immediately, safely and effectively, always taking into account the high availability, speed and safety of their equipment. job.

Therefore, it is necessary to have tools that allow us to optimize the performance of the traffic generated in the network to the different servers that store different types of services, it is as a result of all this that the load balancing technologies arise.

In the work, a study will be made of the techniques used to balance the traffic of the applications in the cloud in their access to the data centers that support them. Both the load balancing techniques between redundant data centers and between the different servers within each data center will be studied, focusing mainly on DNS-based techniques. Additionally, didactic virtual scenarios will be created that allow experimenting with these techniques to strengthen knowledge about them.

In order to facilitate the understanding of traffic balancing technologies, the design of two virtualized scenarios in VNX was proposed, with which it has greater flexibility and easy deployment. The first consists of the development of load balancing performed between servers within of a data center this through the use of the HAProxy tool, while for the second scenario the load balancing between 2 data centers is proposed through the use of geolocation given by DNS with the use of the Bind9 tool.

# Índice general

## Contenido

Resumen .....	i
Abstract.....	ii
Índice General .....	iii
Índice de figuras .....	vi
Siglas.....	viii
<b>1</b> Introducción .....	<b>1</b>
1.1 Motivación.....	2
1.2 Objetivo General .....	3
1.3 Objetivos Específicos.....	3
<b>2</b> Centro de Datos .....	<b>4</b>
2.1 Definición .....	4
2.2 Infraestructuras.....	5
2.2.1 Infraestructura Arquitectónica .....	5
2.2.2 Infraestructura Mecánica .....	5
2.2.3 Infraestructura Eléctrica.....	8
2.2.4 Infraestructura de Telecomunicaciones .....	10
2.3 Tipos de Centros de Datos.....	13
2.4 Requerimientos.....	14
2.5 Ventajas y Desventajas.....	16
2.6 Virtualización.....	18
2.6.1 Ventajas .....	19
2.6.2 Desventajas .....	20
2.7 Computación en la Nube.....	21
2.7.1 Ventajas .....	22
2.7.2 Riesgos .....	23
<b>3</b> Balanceo de Carga .....	<b>24</b>

3.1	Definición .....	24
3.2	Funcionalidades .....	24
3.3	Tipos de Balanceos de Carga.....	26
3.4	Beneficios .....	28
3.5	Técnicas de Balanceo de Carga .....	30
3.5.1	Basado en DNS .....	30
3.5.2	Encaminamiento Directo.....	32
3.5.3	NAT .....	34
3.5.4	Túneles IP .....	35
3.6	Algoritmos de Balanceo de Carga .....	37
3.6.1	Round Robin.....	38
3.6.2	Round Robin Ponderado .....	39
3.6.3	Menos Conexiones Activas .....	40
3.6.4	Menos Conexiones Activas Ponderado.....	41
3.6.5	Tablas hash por origen y destino .....	42
3.7	Balanceo de carga por geolocalización DNS .....	43
3.7.1	Definición.....	43
3.8	Software para Balanceo de Carga en Servidores Linux .....	45
3.8.1	Linux Virtual Server.....	46
3.8.2	HAPROXY.....	47
3.8.3	PEN .....	48
3.8.4	POUND .....	48
3.8.5	NGINX .....	49
3.9	Software para Balanceo de Carga con Geolocalización DNS en Linux .....	50
3.9.1	POWER DNS .....	50
3.9.2	BIND9 .....	52
4	Desarrollo de Escenarios Virtuales Didácticos.....	54
4.1	Balanceo de Carga de Servidores en un Centro de Datos .....	55
4.2	Balanceo de Carga entre Centros de Datos por Geolocalización a través de DNS.....	63
5	Conclusiones y Líneas Futuras de Investigación .....	69

5.1	Conclusiones .....	69
5.2	Líneas Futuras de Investigación .....	70
	Bibliografía.....	72



## Índice de figuras

Figura 2.1. Centro de Datos Facebook Suecia.....	4
Figura 2.2. Infraestructura de un Centro de Datos .....	5
Figura 2.3. Enfriamiento de un Centro de Datos .....	6
Figura 2.4. Sistemas de protección contra incendios .....	7
Figura 2.5. Sistemas de control de seguridad y gestión.....	7
Figura 2.6. Infraestructura eléctrica de un Centro de Datos.....	8
Figura 2.7. Infraestructura de Telecomunicaciones de un Centro de Datos .....	10
Figura 2.8. Topologías de Red de un Centro de Datos .....	11
Figura 2.9. Esquema de Virtualización.....	18
Figura 2.10. Computación en la Nube .....	21
Figura 3.11. Balanceo de Carga Basado en DNS. ....	31
Figura 3.12. Balanceo de Carga Basado en Encaminamiento Directo. [16] .....	33
Figura 3.13. Balanceo de Carga Basado en NAT. [16].....	34
Figura 3.14. Envío de unidades de datos en un Balanceo por Túneles IP.....	36
Figura 3.15. Balanceo de Carga Basado en Túneles IP [7].....	28
Figura 3.16. Algoritmo Round Robin. [12] .....	39
Figura 3.17. Algoritmo Round Robin Ponderado. [7].....	40
Figura 3.18. Algoritmo con Menos Conexiones Activas. [7].....	41
Figura 3.19. Algoritmo con Menos Conexiones Activas Ponderado.....	42
Figura 3.20. Algoritmo con Tablas hash por origen y destino. ....	43
Figura 3.21. Balanceo de carga por geolocalización DNS.....	44
Figura 4.22. Esquema de comunicación a los Centros de Datos.....	54
Figura 4.23. Escenario de red del Centro de Datos 1.....	57
Figura 4.24. Escenario de red del Centro de Datos 2 .....	58
Figura 4.25. Configuración HAProxy Centro de Datos 1 .....	58
Figura 4.26. Configuración HAProxy Centro de Datos 2.....	59
Figura 4.27. Ingreso Pagina web Centro de Datos 1 .....	61
Figura 4.28. Estadísticas del Balanceo en el Centro de Datos 1 .....	61
Figura 4.29. Ingreso Pagina web Centro de Datos 2.....	62
Figura 4.30. Estadísticas del Balanceo en el Centro de Datos 2.....	63
Figura 4.31. Escenario de dos Centros de Datos con Geolocalización DNS .....	63
Figura 4.32. Archivo named.conf del DNS Bind. ....	64
Figura 4.33. Archivo de resolución de nombres del DNS Bind .....	65
Figura 4.34. Archivo de Zonas DNS Bind para LAN1.....	65

Figura 4.35. Archivo de Zonas DNS Bind para LAN2. ....	66
Figura 4.36. Ingreso desde la lan1 a la página web .....	68
Figura 4.37. Ingreso desde la lan2 a la página web. ....	68

## Siglas

ACL	Access Control List
ARP	Address Resolution Protocol
BIND	Berkeley Internet Name Domain
BSD	Berkeley Software Distribution
CDN	Content Delivery Network
DNS	Domain Name System
DoS	Denial of Service
DSR	Dynamic Source Routing
FQDN	Fully Qualified Domain Name
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transport Protocol Secure
IMAP	Internet Message Access Protocol
IP	Internet Protocol
IPV4	Internet Protocol version 4
IPV6	Internet Protocol version 6
IPVS	IP Virtual Server
LVS	Linux Virtual Server
MAC	Media Access Control
NAT	Network Address Translation
OSI	Open System Interconnection
POP3	Post Office Protocol versión 3
SMTP	Simple Mail Transfer Protocol
SSL	Secure Sockets Layer

TCP	Transmission Control Protocol
UDP	User Datagram Protocol
URI	Uniform Resource Identifier
UTP	Unshielded Twisted Pair
VNX	Virtual Networks over Linux
VPN	Virtual Private Network
WWW	World Wide Web
XML	eXtensible Markup Language

# 1 Introducción

En la actualidad la utilización masiva de Internet para todo tipo de actividades diarias desde redes sociales ,correo electrónico almacenamiento de información en la nube y diversas transacciones, que van desde una cotidiana búsqueda de información a la compra diaria de diversos productos en línea , lleva consigo una carga creciente de los servidores alojados en diversos centros de datos , sobre todo en los servidores web los cuales son los responsables de la disponibilidad de tiendas electrónicas, portales informativos y páginas corporativas y con ello crecen las expectativas de los usuarios, que esperan, especialmente en el área de servicios, un funcionamiento rápido y seguro de las aplicaciones web. Es así como la disponibilidad de una página repercute en que un usuario interesado se convierta en un cliente.

Las comunicaciones generadas a través de la internet sean convertidas en un factor determinante e importante en el auge informático, viéndose reflejado en las numerosas transacciones que viajan a través de esta gran red. Las instituciones brindan datos al alcance de todos los usuarios, manejándose de acuerdo con los estándares y nivel de seguridad. Cada vez son mayores las peticiones de usuarios dispuestos a satisfacer sus necesidades, y como resultados mayores son los procesos que deben de atender los servidores, los cuales permiten satisfacer dicha demanda.

Una de las problemáticas que más se tratan de evitar hoy en día es los cuellos de botella y así aligerar el tráfico en las comunicaciones, este ha sido un problema constante en las comunicaciones, por lo que se han venido planteando distintas alternativas como medidas para solucionar el incremento acelerado del uso del internet, ya que consigo trae un problema que afecta el servicio que brinda una institución. Al existir mayor cantidad de usuarios, también aumentará el tráfico en la red, y con ello se tendrá una carga de trabajo mayor la cual será soportada por los servidores.

Por ello es que a través de las técnicas de administración de redes se tiene como función organizar, supervisar y controlar cada uno de los elementos de comunicación para garantizar un nivel de servicio optimo a unos costos razonables, de ahí que el objetivo principal de la administración es mejorar la disponibilidad e incrementar la efectividad de la red, debido a todo esto es que se plantea el uso de la tecnología conocida como balanceo de carga la cual a través de diferentes técnicas y algoritmos permite que los servidores atiendan las peticiones o requerimientos de los usuarios. La implementación de este procedimiento ayuda a mejorar el tiempo de acceso debido a

que se trata de evitar la saturación en los recursos de un servidor haciendo que las peticiones sean gestionadas de manera distribuida entre los diferentes equipos administrados en un centro de datos, además de evitar fallos inminentes que puedan sufrir en un momento determinado un servidor, que ante un colapso este no sufrirá una caída total del sistema y con ellos obtener un buen uso de los recursos con que cuenta la organización o entidad y aumentar el rendimiento de la red.

## 1.1 Motivación

Debido al flujo ininterrumpido de innovaciones en el campo de las tecnologías de la información las empresas han considerado la importancia del manejo de esta, así como también de la red que la comunica y la forma en que ésta la trasmite. En consecuencia, son muchas las funciones que se deben llevar a cabo para garantizar el flujo de dicha información, de las cuales la que más destaca es la gestión de la red y garantizar el buen funcionamiento de los dispositivos que la componen. Pero con el rotundo crecimiento de dispositivos conectados a la red y la enorme demanda de recursos en línea ha ido aumentando de manera considerable, esto lleva consigo a que un servidor de internet puede presentar un procesamiento lento de las peticiones que se realizan debido a la congestión presentada en la red, ante esto se pondría en manifiesto que la solución más obvia sería incrementar la memoria, el disco duro o actualizar el procesador; aunque se debe considerar que el tráfico de internet está en constante crecimiento por lo que lo anteriormente expuesto es sólo una solución a corto plazo, además de ser una opción que incluye una inversión generalmente alta, en servidores de mayor rendimiento. La forma más adecuada y que brinde escalabilidad y conserve el servicio a largo plazo, es la de configurar más servidores que gestionen los mismos servicios y repartir las peticiones de los clientes entre ellos. Esto aumenta la velocidad de acceso del usuario al servidor, mejora la fiabilidad del sistema y la tolerancia a fallos, permitiendo reparar o mantener cualquier servidor en línea sin que afecte al resto de los servicios.

Ante la necesidad de brindar una mejora sustancial en el rendimiento de los servicios a través de las redes de datos, nos conduce a la motivación principal para realizar este trabajo puesto que un servidor sobrecargado puede ser perjudicial para cualquier proyecto web, debido a que puede frenar e incluso paralizar algunas áreas de cualquier negocio. Por lo que es necesario contar con una distribución eficiente del flujo de datos, es de esta manera como entre varios servidores web es posible controlar picos de tráfico en épocas de gran demanda y disminuir el riesgo de fallos de un servidor por ello es que optaremos por dicho estudio, y una forma de hacer factible aquello es gracias a la utilización del procedimiento de balanceo de carga, la cual nos permitirá a

través de diferentes herramientas repartir la carga entre diferentes servidores, con ello se puede reducir el posible impacto de una saturación hacia un servidor específico y con ello mejorar la planificación inteligente de los recursos del sistema, permitirá sacar el máximo provecho de nuestros recursos y mejorar las prestaciones.

Así mismo es de gran interés conocer las diferentes metodologías implementadas a raíz de esta tecnología como lo es actualmente la utilización del servidor DNS, para gestionar balanceo de carga a través de la geolocalización de los usuarios que efectúan las solicitudes hacia las diferentes aplicaciones almacenadas en los servidores dentro de los respectivos centros de datos.

## **1.2 Objetivo General**

Como objetivo general se propone que con el desarrollo de este trabajo se pueda entender de mejor manera los conceptos básicos de esta tecnología y así evaluar las funcionalidades de los balanceadores de carga con los cuales se pueden efectuar técnicas avanzadas en la gestión de la red que conlleven a la optimización de la comunicación hacia los centros de datos.

## **1.3 Objetivos Específicos**

- Proporcionar los conocimientos adecuados sobre el funcionamiento y arquitectura de los centros de datos.
- Dar una clara idea del funcionamiento de los procedimientos de balanceo de carga, así como los algoritmos y técnicas que se pueden emplear sobre esta tecnología para la gestión de tráfico en centros de datos.
- Realizar el diseño de un escenario didáctico que permitan explicar de manera más apropiada las técnicas de balanceo de carga sobre servidores dentro de un centro de datos.
- Realizar el diseño de un escenario didáctico que permitan explicar de manera más apropiada el balanceo de carga entre centros de datos por geolocalización a través de DNS.

## 2 Centros de Datos

En el siguiente capítulo se abordará la base teórica con lo referente a los centros de datos, dando a conocer la definición de estos, su infraestructura, características y ventajas y desventajas que tiene dichos sistemas de comunicación.

### 2.1 Definición

Los centros de datos son instalaciones que concentran todos los equipos y elementos necesarios para poder realizar la administración y procesamiento de la información de una organización y que circula a través de las distintas redes de datos.

Este entorno especializado es de vital importancia para cualquier empresa debido a que salvaguarda la información más crítica de la misma que puede ir desde contenido de emails , repositorios de contenido multimedia, sitios web , incluso transacciones bancarias, todas estas influyen de manera importante en la privacidad de los datos de los clientes y así mismo generan un impacto importante en la continuidad de la empresa, debido a esto se ha convertido en la infraestructura de la internet más relevante es por ello que constituye el área de más inversión en las organizaciones.

Después de lo antes mencionado es clara la importancia de un centro de datos dentro los sistemas de telecomunicación, por ello es de suma importancia que estos brinden la suficiente fiabilidad, confiabilidad y además que puedan adaptarse perfectamente al crecimiento masivo que está presentando la internet y se solicita que los mismos sean escalables, un centro de datos se puede observar en la Figura 2.1.



Figura 2.1. Centro de Datos Facebook Suecia.



## 2.2 Infraestructuras

En esta sección se definirán las distintas infraestructuras que componen a los centros de datos, que son necesarias para el cumplimiento de los requerimientos y especificaciones técnicas, que permitan el correcto funcionamiento de estas instalaciones como se puede visualizar en la Figura.2.2. [1].

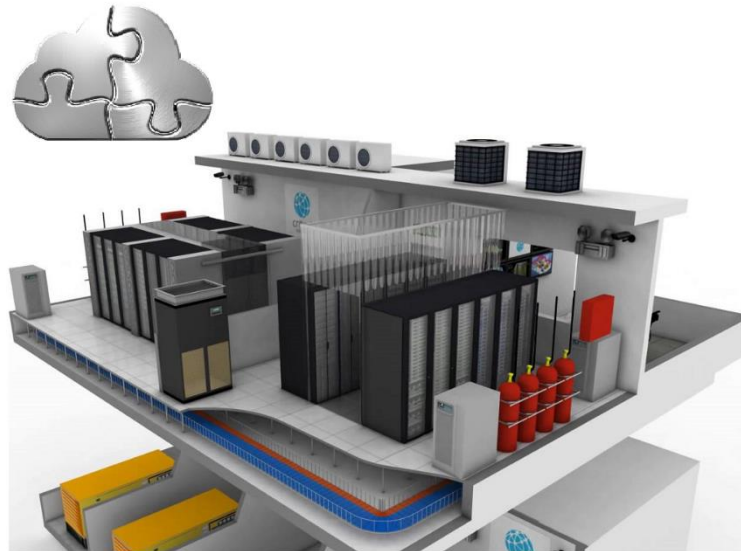


Figura 2.2. Infraestructura de un Centro de Datos

### 2.1.1 Infraestructura Arquitectónica

Se basa principalmente en una serie de normas que debe cumplir el espacio físico donde se acogerán a los equipos entre estas, que no existiese interferencia electromagnética, niveles de ruido demasiado altas, la construcción de estas instalaciones debe estar en una zona que brinde una seguridad física es decir que tenga el menor riesgo posible de catástrofes naturales y del tema de polución y adicional que la ubicación donde se encuentren estos centros de datos tenga un área libre para una expansión de este a futuro.

### 2.1.2 Infraestructura Mecánica.

En este tipo de infraestructura se detallan las especificaciones internas que deben cumplir los centros de datos tanto en el tema de ventilación o enfriamiento de los equipos, los distintos sistemas de protección contra incendios, sistemas de control de

seguridad y así también las consideraciones estructurales apropiadas.

- **Enfriamiento:** Se trata de los diferentes equipos que regulan la temperatura del ambiente dentro de un centro de los datos los cuales se manejan en su mayoría a través de aires acondicionados que a través de ductos distribuyen el flujo de aire frio hacia los diferentes equipos , el control de temperatura depende en gran media de las características eléctricas con las que funcionan los diferentes componentes ya que algunos requieren trabajar con ciertos criterios de temperatura interna , dado esto lo que se quiere evitar es el sobrecalentamientos de los mismos y que provoque en ellos algún tipo de fallo esto se ve representado en la Figura 2.3.

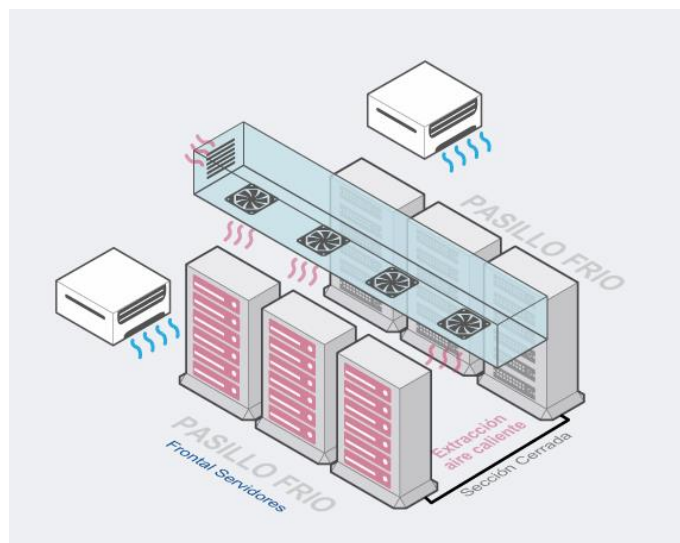


Figura 2.3. Enfriamiento de un Centro de Datos

- **Sistemas de protección contra incendios:** Este tipo de sistemas incluyen todos los dispositivos necesarios para mitigar la propagación del fuego dentro de las instalaciones de los centros de datos y así evitar que el humo y las llamas afecten la integridad de los equipos que componen dicho entorno. Entre los medios que permiten ejecutar estas acciones tenemos los extintores, supresores de gas, dispositivos de alarma y detectores de humo entre estos tenemos 2 nuevas tecnologías implementadas para los mismos como lo son los detectores fotoeléctricos y los iónicos. Estos dispositivos pueden ser observados de mejor manera en la Figura 2.4.

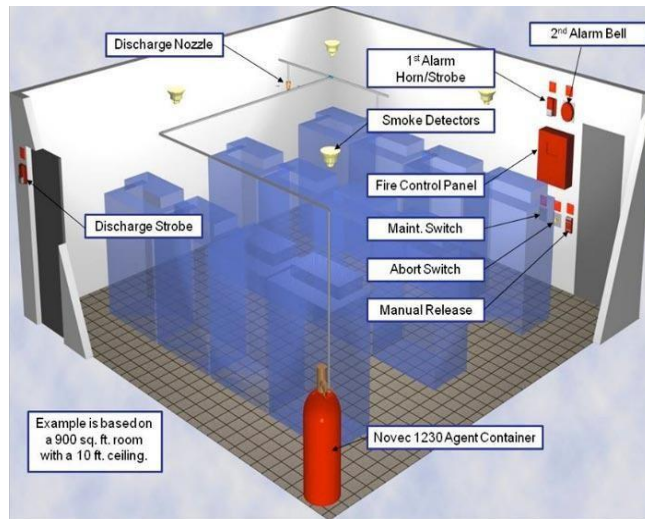


Figura 2.4. Sistemas de protección contra incendios.

- **Sistemas de control de seguridad y gestión:** Este tipo de sistemas es indispensable para proteger tanto la integridad de los equipos como la de la información almacenada en ellos entre los dispositivos que nos permiten tener un mejor control de la seguridad de acceso al centro de datos tenemos los biométricos, llaves para ingreso al rack de comunicación y cámaras de seguridad. Un aspecto importante para tener en cuenta es el manejo de sensores que determinan el estado actual del centro de datos tanto en el aspecto de temperatura como en el caso de desconexiones o fallos eléctricos, esto permitirá tener un mejor control de este y actuar de manera más ágil ante la presencia de cualquier eventualidad como se observa en la Figura 2.5.

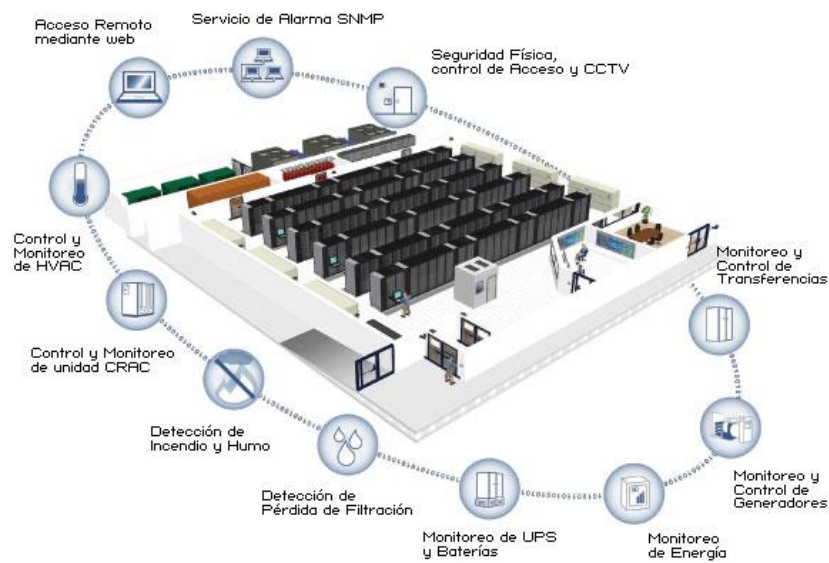


Figura 2.5. Sistemas de control de seguridad y gestión.

- **Consideraciones estructurales apropiadas:** Adicional a los requerimientos ya planteados anteriormente es necesario tener en cuenta las siguientes consideraciones en la estructura de un centro de datos como el que las instalaciones cuenten con un piso falso ya que debajo de los mismos se realizara el recorrido del cableado esto es de suma importancia debido a que a los cables de red tanto de voz como de datos son sensibles a la interferencia electromagnética y al ruido, además de salvaguardar la seguridad de los equipos en posibles inundaciones.

El techo falso es un elemento que es necesario tanto para tener un mejor manejo del cableado eléctrico y el mismo brinde la seguridad apropiada al no estar a la vista y al alcance del personal que ingresa al centro además de lucir un mejor diseño al centro de datos, se debe tener en cuenta que ambos tanto el piso como el techo falso deben contar con los respectivos gabinetes que permiten una mejor distribución del cableado.

### 2.1.3 Infraestructura Eléctrica

Aquí se definen los requerimientos de cada fuente de energía eléctrica requerida para el funcionamiento de todos los elementos que componen a un centro de datos, esta infraestructura debe ser capaz de soportar las necesidades de energía cuando los dispositivos estén operando al máximo de su capacidad, así también se debe controlar los niveles de voltaje apropiados y principalmente poder brindar un flujo eléctrico continuo que proporcione continuidad en el funcionamiento de los equipos. Este sistema eléctrico se puede visualizar mejor en la Figura 2.6. [4].

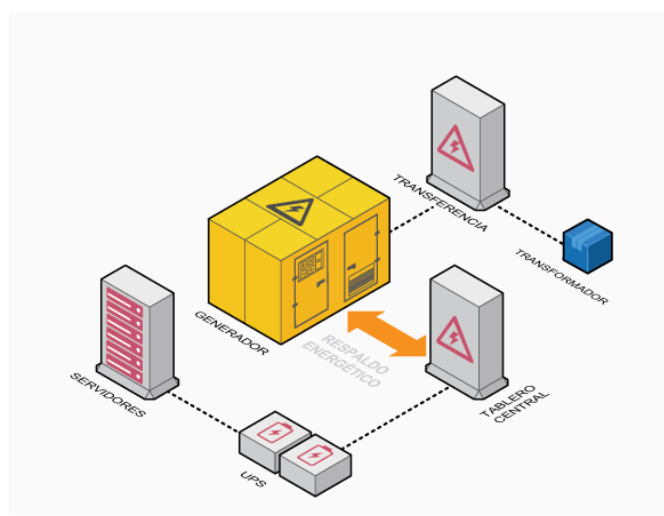


Figura 2.6. Infraestructura eléctrica de un Centro de Datos.

Dentro de la parte eléctrica de un centro de datos se cuenta los siguientes componentes que permiten un funcionamiento adecuado de estos.

- **Generadores:** Son equipos capaces de transformar la energía mecánica en eléctrica permitiendo distribuir la misma a diversas áreas específicas, son un elemento fundamental ya que al presentarse una falla o corte eléctrico externo estos proveerán la energía necesaria para que los equipos del centro de datos puedan seguir funcionando hasta el restablecimiento total de la energía externa.
- **Sistemas de transferencia automática:** Estos sistemas utilizados generalmente en las acometidas eléctricas permitiendo tener una supervisión sobre la distribución y cese de esta hacia un área específica, esto contribuye a que en caso de presentarse alguna caída del suministro eléctrico esta sea capaz de transferir el consumo realizado por los equipos hacia otra fuente de energía como lo son los generadores.
- **Tableros eléctricos de distribución o centrales:** La función principal de este medio es el poder dar una protección máxima de todos los equipos conectados a los diferentes puntos de electricidad puesto que permite el control de las acometidas eléctricas que proveen de energía a cada área del centro de datos, con lo cual se supervisa la distribución de cada una de estas evitando los cortocircuitos y daños en los equipos y sistemas eléctricos.
- **Fuente de energía ininterrumpida (UPS):** Este componente se encarga de almacenar la energía necesaria para poder suministrar de carga eléctrica a los equipos en caso de algún corte se suministró eléctrico externo, este dispositivo es indispensable puesto que evita los picos o variaciones de voltaje presentados cuando existe un corte de energía y esta se restablece, bien desde la red eléctrica alterna o externa impidiendo que pueden impactar en gran medida a los equipos , con esto se mejora la calidad de energía eléctrica otorgada a los elementos del centro de datos.
- **Distribuidores de carga (PDU):** Son dispositivos que son utilizados principalmente para controlar las variaciones de voltaje que afectan a los equipos brindando confiabilidad en la conexión eléctrica, debido a que cuenta con características de transformación, protección y distribución del voltaje, estos por lo regular van colocados en los gabinetes o dentro del rack de comunicación de un centro de datos.

- **Cables eléctricos:** Son conductores metálicos encargados de distribuir la carga eléctrica desde las fuentes de energía hacia los equipos eléctricos dentro de los datos estos tipos de cables deben cumplir con normativas que dependen en gran medida de los valores de corriente y temperatura que se presente en el entorno a trabajar.

#### 2.1.4 Infraestructura de Telecomunicaciones

Esta infraestructura trata primordialmente de los equipos de telecomunicación, cableado estructurado y los diversos componentes que permiten la comunicación tanto interna del centro datos como externa hacia los proveedores y la red de datos de la organización y representados de mejor manera en la Figura 2.7.

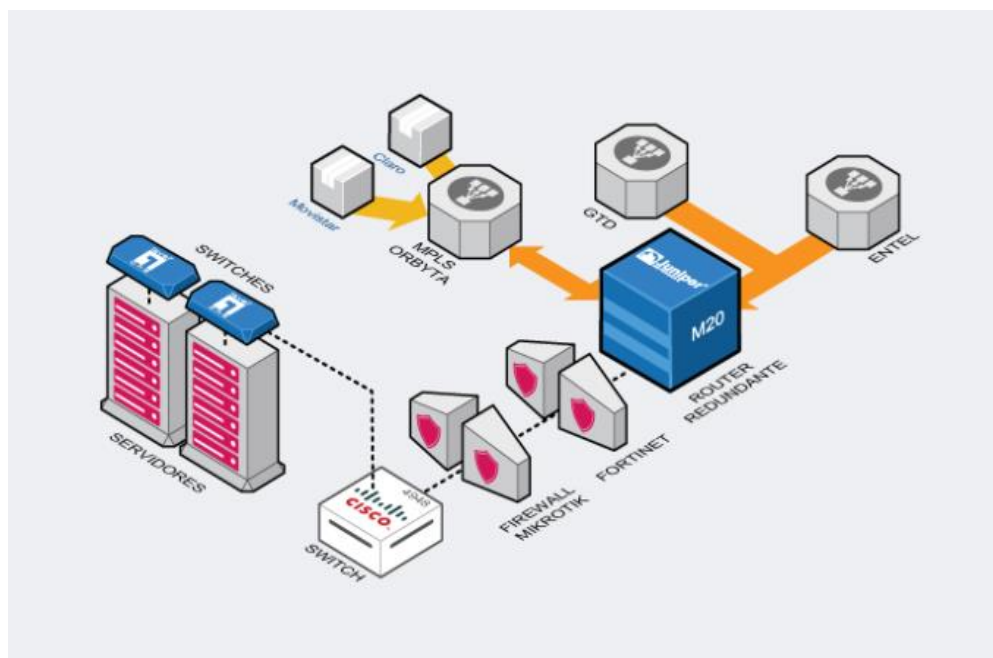


Figura 2.7. Infraestructura de Telecomunicaciones de un Centro de Datos.

A continuación, se detallarán los diferentes componentes que forman parte de este sistema:

- 1) **Topología de Red:** Representan las diferentes configuraciones e interconexiones entre los equipos de telecomunicación, estas pueden ser establecidas tanto a nivel físico como lógico:

- **Topología Lógica:** En esta topología se representa la forma en que los diferentes equipos de comunicación realizan el intercambio de información entre ellos a través de conexiones virtuales que facilitan la administración del acceso a la red, entre las más utilizadas tenemos los bus, estrella y punto a punto.
- **Topología Física:** Representa la disposición de los medios de comunicación para la interconexión física entre los equipos, las topologías más usadas tenemos: anillo, bus, estrella, malla.

Estas topologías de red se pueden observar de mejor manera en la Figura 2.8.

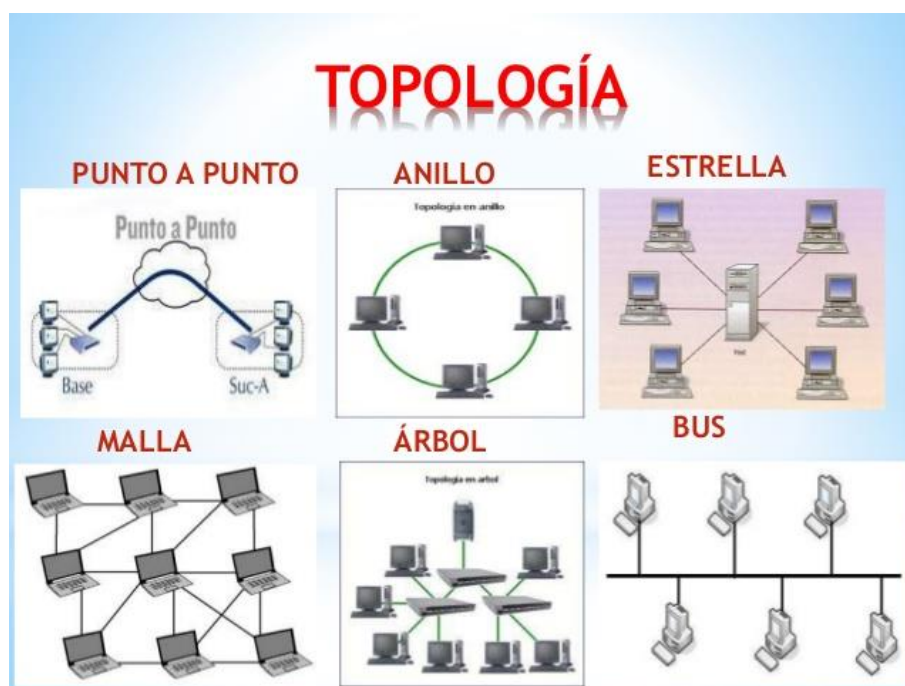


Figura 2.8. Topologías de Red de un Centro de Datos.

**2) Red interna:** La red de datos es el componente más crítico para la prestación de servicios, las transferencias entre diferentes sistemas y para el almacenamiento, por eso la importancia de definir los medios necesarios que permitan que el centro de datos tenga un óptimo desempeño, debido a esto se definen diferentes tipos de jerarquías para definir el diseño de red que viene dado por los niveles núcleo, distribución y acceso. A continuación, se detallan los elementos principales que forman parte de la red interna de un centro de datos.

- **Routers:** Son equipos que permite la interconexión entre varias redes diferentes para que entre estas se pueda realizar el



intercambio de paquetes, incluso poder determinar la ruta que debe recorrer un paquete basado en una dirección de destino. Para poder efectuar esta funcionalidad estos dispositivos manejan diferentes tipos de encaminamiento que le permiten tanto la comunicación interna como externa entre las diferentes redes de datos, entre estos tenemos RIP, OSPF, EIGRP y BGP.

- **Firewall:** Es un equipo que permite realizar un filtrado de los paquetes que son recibidos por parte de redes externas hacia la red interna de la organización y viceversa, esto nos permite tener un control sobre el flujo de paquetes que es recibido y enviado entre las diferentes redes de datos y poder determinar las comunicaciones que son permitidas y rechazadas hacia mis equipos evitando así accesos no autorizados, esto a través de la definición de las ACL respectivas, es de suma importancia contar con este tipo de componentes puesto a que ayuda a bloquear el tráfico considerado malicioso y que puede resultar crítico para mis equipos.
- **Switch:** Es un equipo capaz de poder comunicar uno o más segmentos de red, pudiendo interconectar varios dispositivos dentro de una misma red interna, realizando la transmisión de paquetes entre ellos basándose en la dirección MAC de destino incorporada en las tramas de red. Existen 2 modelos de este tipo de dispositivos los de capa 2 que solo son capaces de realizar funcionalidades de segmentación de la red y los de capa 3 pudiendo realizar tanto segmentación como gestiones de encaminamiento.
- **Servidores:** Son los equipos encargados de ofrecer los distintos servicios esenciales para el funcionamiento de la organización que pueden ir desde correo electrónico, páginas web, base de datos, DNS, etc.
- **Medios de transmisión:** Representan los conectores que permitirán la comunicación entre los diferentes sistemas de telecomunicación tanto de manera local y remota. Estos deben cumplir con ciertos estándares que permitan la óptima transmisión de señal entre el equipo emisor y el receptor. Por lo tanto, su principal objetivo es proveer un sistema total de transporte de información a través de un mismo tipo de cable. Dentro de las



conexiones de un centro estos pueden ser de dos tipos [5]:

- **Par trenzado UTP:** Es el medio de comunicación más utilizado para las conexiones de redes LAN, maneja conectores RJ45 para conectarse a sus equipos finales este tipo de cable provee tasas de transmisión que van de los 100 Mbps hasta los 10 Gbps. Este tipo de cable es muy resistente por ellos su uso típico en las conexiones internas de un centro de datos. Se debe considerar que este tipo de medio maneja diferentes configuraciones para la conexión entre las cuales se podría mencionar que para equipos similares se da un tipo de cable directo y para equipos distintos un cable de tipo cruzado, aunque existen muchas más configuraciones establecidas.
- **Fibra Óptica:** Es uno de los medios incorporados recientemente en los sistemas de telecomunicación esto se debe a que proporciona óptimos resultados al enviar una cantidad masiva de datos a grandes distancias y por brindar conexiones con tasas de transferencia muy elevadas, además de evitar de manera excelente las interferencias electromagnéticas esto se debe a que su composición es realizada a base de hilos de vidrio. Este tipo de medio envía las señales a través de pulsos de luz por ello es por lo que se obtiene conexiones de alta velocidad y permite una fácil detección de fallos, a pesar de todo esto cabe destacar que al ser su composición muy frágil es un cable muy sensible por presentar fallos con la más mínima alteración de este.

**3) Direccionamiento IP:** Es el identificador lógico de los dispositivos de red que es asignado a las interfaces de red de estos, permitiendo la comunicación entre los diferentes componentes de comunicación. Existen actualmente 2 tipos de direccionamiento IP el IPV4 y el IPV6.

## 2.3 Tipos de Centros de Datos

En esta sección se dará a conocer los 2 tipos de datos con los que se puede contar de acuerdo con el número de clientes con el que se desee trabajar.

- ❖ **Centro de datos Corporativo:** Este tipo de centros son administrados por las propias empresas es decir los servicios de comunicación que pueden ser de telefonía e internet, el alojamiento de los equipos como los servidores que manejan los servicios proveídos por las empresas, y el control de la información para beneficio propio convirtiéndose en el núcleo principal para el funcionamiento de la organización.
  
- ❖ **Centro de datos Hosting:** Este centro de datos a diferencia del anterior es administrado netamente por el proveedor de telecomunicaciones, es decir por un tercero independiente de la empresa los cuales comercializan una serie de servicios hacia varias empresas entre estos el internet, conexión de enlace tanto en capa 3 como en capa 2, alojamiento de equipos, páginas web e incluso el tema de VPN. Cabe mencionar que el control sobre los equipos y administración o monitoreo de estos puede ser pactado en las negociaciones respectivas por lo que la empresa y el proveedor definen como se manejan los aspectos antes mencionados.

## 2.4 Requerimientos

En este subcapítulo nos centraremos en presentar las disposiciones necesarias para el correcto funcionamiento de un centro de datos.

1. **Simplicidad:** Es esencial que los centros de datos presenten en un diseño fácil de entender y administrar con el propósito que las personas encargadas de su buen funcionamiento realicen tareas de gestión más óptimas , entre las características que se puede cumplir para mejorar lo mencionado esta tener un etiquetado adecuado de los equipos , de las conexiones realizadas a cada uno de ellos que otorguen una verificación más rápida en caso del algún inconveniente, adicionalmente se podría tener estructura adecuada del control de energía , enfriamiento a través de paneles de control con especificaciones claras.
  
2. **Modularidad:** Puede presentarse que los centros de datos puedan llegar a ser sumamente complejos lo que provoca que estos sean muy difíciles de gestionar y se vuelva tedioso la identificación de problemas y de realización de actividades de mantenimiento sobre estos por lo que es necesario realizar acciones de regulación sobre estos sistemas segmentándolos en partes más

pequeñas que facilitan mucho la administración , la identificación de inconvenientes e incluso los trabajos de réplica y pruebas de contingencia .

3. **Flexibilidad:** Debido a crecimiento continuo de la tecnología sobre todo en el área de las telecomunicaciones, es por esto que sucede un cambio continuo con nuevas innovaciones lo que provoca que los centros de datos deban adaptarse aquello, es decir sean flexibles para poder ejecutar cualquier actualización que se pueda presentar esto radica principalmente en sacar el mayor provecho posible en la durabilidad del mismo.
4. **Escalabilidad:** Este aspecto es muy importante debido a que la estructura de un centro de datos debe ser capaz de acoplarse al crecimiento paulatino de la empresa cubriendo las necesidades de estas, debe poder definirse un crecimiento en tecnología e infraestructura a un plazo definido y sin que estos cambios interfieran con el rendimiento del centro de datos y afecte los servicios proporcionados por las organizaciones.
5. **Estandarización:** Es necesario por definir los modelos de los equipos y la infraestructura que compone a un centro de datos esto permite una manera más precisa el proceso de mantenimiento, así como la identificación de problemas al tener claro la funcionabilidad de cada elemento del centro de datos.
6. **Perfiles de la disponibilidad del sistema:** Es de vital importancia que los centros de datos brinden un servicio continuo ya que esto permite que las empresas puedan proveer un servicio estable hacia sus clientes impidiendo que se provoque pérdidas económicas y que impacten en la continuidad del negocio por ende se establecen los siguientes sistemas de redundancia:
  - **Redundancia de Dispositivos:** Se debe poseer equipos de respaldo que permitan en caso de algún fallo de alguno de los equipos en funcionamiento estos tomen el control del flujo de información impidiendo la pérdida de comunicación, además de aquello permite el mantenimiento de los equipos, debido a que no se afecta el funcionamiento del centro de datos.
  - **Redundancia Eléctrica:** Es indispensable que los equipos continúen en funcionamiento por ello se debe evitar que estos sufran afectaciones energéticas para esto se debe tener fuentes alternas de energía cuando surja un corte o fallo del suministro externo de

energía , por ellos se debe contar con generadores y dispositivos ups que permitan tanto la generación como el almacenaje de energía, adicional a todo aquello los equipos deben poder contar con por lo menos con dos tomas de energía para los equipos de mayor importancia cada uno de los cuales estarán en diferentes fuentes de energía.

- **Redundancia de Enfriamiento:** Es necesario contar con sistemas de contingencia para el enfriamiento del centro de cómputo que actúen en caso de algún inconveniente con los aires acondicionados o medios de ventilación, debido a que el no contar con esta alta disponibilidad podría afectar la integridad de los equipos si llegase a surgir sobrecalentamiento de estos por el aumento de temperatura.
7. **Seguridades Físicas y lógicas:** Los centros de datos deben contar con las medidas de seguridad apropiada y el ingreso solo debe ser permitido a personal autorizado al ser un entorno de suma importancia para el funcionamiento de los servicios de la organización , se debe otorgar un monitoreo continuo de la correcta operación de los equipos , adicional es imprescindible que a pesar de que ingreso el personal autorizado solo puedan acceder a las configuraciones de los equipos personas capacitadas para la administración de estos por ello es requerido manejar niveles de autenticación para poder realizar modificaciones , actualizaciones y correcciones en caso de ser requerido.

## 2.5 Ventajas y Desventajas

En esta sección definiremos las diferentes ventajas y desventajas que nos puede proveer un centro de datos para las redes de telecomunicación [2].

- ❖ **Ventajas:** Entre las principales ventajas que provee un centro de datos tenemos las siguientes:
  1. Reduce en gran medida los costos operacionales ya que se tiene mejor mantenimiento y control de los equipos al tenerlos establecidos en un sitio.
  2. Permite un mejor manejo de los recursos de red esenciales de la empresa, debido a que la administración es centralizada.

3. Otorga el manejo de gran información pudiendo dar alta disponibilidad a esta, a través de diferentes tecnologías brindando estabilidad a los servicios ofrecidos por la organización.
4. Provee de un entorno óptimo con las condiciones adecuadas para el correcto funcionamiento de los diferentes equipos de comunicación, aumentando el tiempo de vida de los mismo y un mejor desempeño.
5. Ofrece una seguridad de la información crítica de la empresa ya que cumple las normas adecuadas para salvaguardar esta ante cualquier eventualidad.
6. Estas instalaciones combinadas con lo antes mencionado provocan que sea un entorno que posibilite el procesamiento eficiente de los datos, estos se transportan de manera segura y viabilidad para futuros crecimiento del negocio.

❖ **Desventajas:** Entre las principales desventajas que provee un centro de datos tenemos las siguientes:

1. El tener toda mi información y mis equipos críticos en una sola ubicación hace que sea propenso a que en caso de cualquier catástrofe sea una pérdida importante para la empresa incluso pudiéndola llevar a la banca rota.
2. Un riesgo que se toma al tener una centralización de todo el flujo de información es que sea más vulnerable a que los hackers concreten sus ataques a estos entornos y es un gran riesgo lógico, debido a que si llegan a filtrar en los datos tendrán un fácil acceso a toda la información crítica de la empresa.
3. El error humano es un factor que considerar, debido a que una manipulación inadecuada, no solo que afectaría a un equipo o información determinada, sino que podría provocar un efecto en cadena y afectar al resto debido a que las redes de telecomunicación es una serie de comunicaciones establecidas.
4. Los costos operacionales podrían llegar a ser elevados si no se tiene una administración correcta de los sistemas eléctricos y de enfriamiento.

## 2.6 Virtualización

Debido a que los requerimientos de recursos en las comunicaciones tales como procesamiento , memoria y almacenamiento son más demandantes cada día , puede producir un crecimiento exhaustivo de los equipos necesarios para cubrir dicha demanda esto genera un conflicto en cuanto a la infraestructura y el aumento considerable de los costos de operación, es a raíz de aquello que surge la virtualización como una tecnología que permite crear múltiples entornos simulados o administrar recursos desde un solo hardware , es decir permite que a través de una maquina física funcione como si fuese múltiples maquinas trabajando sobre este , esto brinda un ahorro significativo tanto de espacio físico , ahorro financiero debido a que ya no se debe adquirir varios equipos para cada requerimiento solicitado sino que todos los servidores y recursos requeridos pueden ser administrados desde un solo equipo físico.

La virtualización trabaja a través de un software denominado Hipervisor el cual es conectado directamente a la maquina física permitiendo la creación de varios entornos virtuales conocidos como máquinas virtuales, permitiendo la separación de los recursos que posee el hardware y pudiéndolos distribuir de manera adecuada, administrando los recursos necesarios para la creación de un servidor esto se observa en la Figura 2. 9.. Lo antes mencionado es manejado por lo regular a través de una consola de administración que por lo general es visualizada a través de la web.

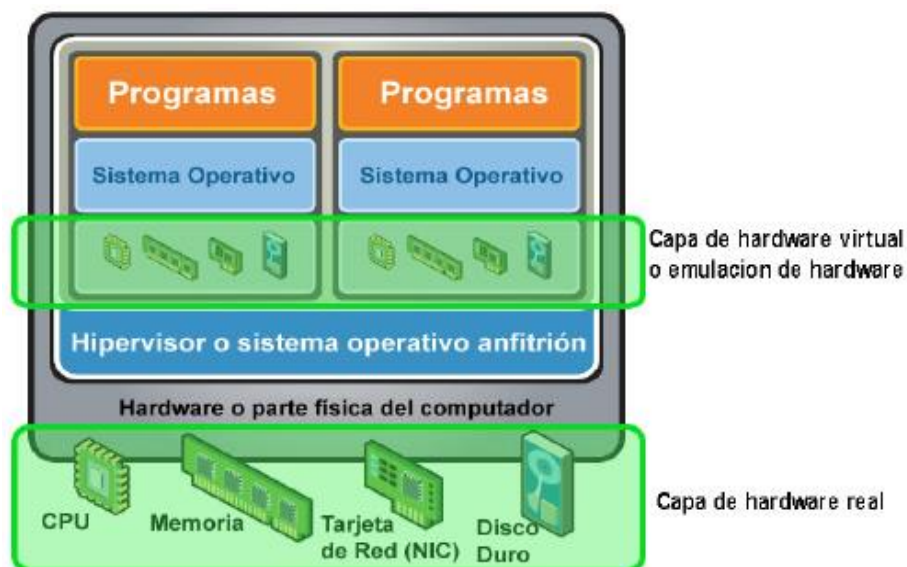


Figura 2.9. Esquema de Virtualización

### 2.6.1 Ventajas

Aquí se definen las diferentes ventajas que nos provee el uso de esta tecnología en el desarrollo de las comunicaciones entre las que tenemos las siguientes:

- Instalación y el manejo se varios servidores incluso con distintos sistemas operativos permitiendo que entre estos no se interfieran uno con el otro impidiendo que se afecte el desempeño de alguno de ellos.
- La recuperación de información es muy fácil debido a que posee herramientas en las que se puede generar clonaciones y copias de respaldo de las máquinas virtuales, incluso realizar este tipo de tareas de manera automática para que se ejecuten cada cierto tiempo.
- Se reduce de manera significativa los costos administrativos, así como un ahorro importante en el aspecto económico debido a que no se debe adquirir una cantidad considerable de equipos de hardware, disminuyendo el costo de energía eléctrica.
- Este tipo de esquema de virtualización es muy flexible al tema de asignación de recursos según requerimientos establecidos de las distintas máquinas virtuales, adicional se puede manejar programas o sistemas operativos ambiguos ya que no presenta problemas de compatibilidad con estos.
- Se tiene un aprovechamiento adecuado de los recursos incluso permitiendo un manejo rápido antes cambios de demanda posible, debido a que se puede dar un aumento en la característica de un servidor ya sea de procesamiento, memoria o almacenamiento en caso de ser necesario sin afectar la funcionabilidad de ese sistema.
- La escalabilidad es un aspecto importante que considerar ya que es de vital importancia hoy en día poder cumplir con las demandas de los clientes de las organizaciones y esta tecnología permite aumentar los recursos de la maquina física según los requerimientos de la empresa
- Cuenta con portabilidad, es decir permite realizar una copia de todas las configuraciones y ficheros de una máquina virtual esto facilita enormemente el hecho de trasladar esta a otro equipo físico. Esto lleva consigo que, en caso de fallo de un sistema físico, los sistemas lógicos implementados ahí pueden distribuirse dinámicamente a otros sistemas.

- Al poseer una consola de administración de todos los recursos virtualizados agiliza el tema de mantenimiento y control de los equipos que brindan el servicio de la empresa.
- El tema de seguridad es de tipo de tecnología es óptimo debido a que el acceso a cada equipo virtual es independiente manejando distintas credenciales, por tanto, en caso de algún ataque solo se afectará a una sola máquina.

### 2.6.2 Desventajas

Aunque es indiscutible los grandes beneficios de contar con la virtualización se deben considerar algunas desventajas que trae consigo aquello en las que se mencionan las siguientes:

- Se presenta una pérdida del rendimiento esto se debe a que el procesamiento de información de un sistema virtual es menor al generado a través de un servidor físico.
- Otro punto que considerar es que todo el funcionamiento de muchas máquinas virtuales radica en un solo equipo físico por lo que en caso de presentarse algún inconveniente sobre este afectará el servicio de todas estas, generando un problema de gran magnitud que, si solo se tuviera con una máquina, ya que casi todos los servicios ofrecidos por la organización quedarían sin funcionar.
- Las máquinas virtuales son más propensas a sufrir afectaciones por errores de configuración o algún tipo de apagado brusco o inadecuado, pudiendo producir errores en el funcionamiento de estas, este tipo de situaciones es menos probable en los equipos físicos ya que poseen una infraestructura más adecuada para su funcionamiento y su tiempo de vida por es aún mayor que un equipo virtual manejado con un solo hardware.



## 2.7 Computación en la Nube

Se trata de una tecnología capaz de permitir el acceso a sistemas informáticos en que los datos y aplicaciones de estos se encuentran hospedados en la internet o centros de datos remotos. Esto permite que los usuarios y empresas que puedan gestionar sus recursos como programas y archivos sin necesidad de tener aquello instalado en sus equipos físicos sino solo contar con un acceso a la red que les permita tener conexión a la internet o a dichos centros [3].

La información es contenida de manera permanente en servidores en Internet o en centros de datos remotos y se envía a cachés temporales de cliente que se conectan por dispositivos tales como laptops, Tablet o celulares como se puede visualizar en la Figura 2.10.

Principalmente este modelo se plantea el uso de servicios de software utilizando la internet como herramienta brindando una confianza al uso de este para poder cumplir las necesidades solicitadas por los usuarios.

Esta tecnología viene siendo influenciada mucho por el tema antes de mencionado de virtualización, además de la automatización y los procesamientos que se ejecutan en paralelo y por el sin número de aplicaciones que ya no se establecen en un solo equipo físico, sino que son compartidas por varios equipos para su funcionamiento. Actualmente se tiene muchas empresas que ofrecen este tipo de servicios entre las que destacan Amazon, Google, Microsoft.

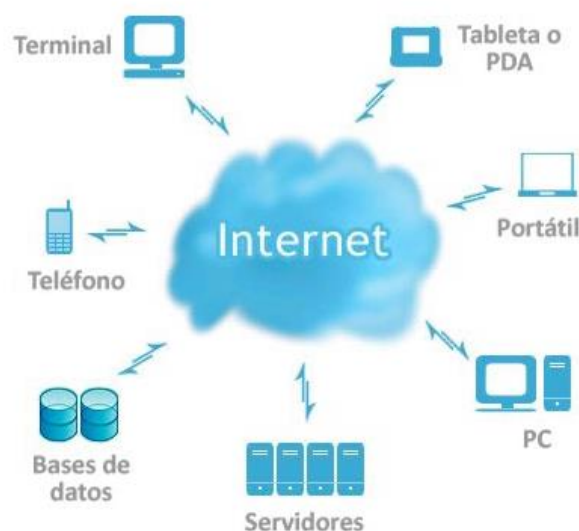


Figura 2.10. Computación en la Nube

## 2.7.1 Ventajas

Entre las ventajas que nos otorga esta tecnología se hará mención de las siguientes:

- Mayor facilidad de acceso hacia los recursos de la infraestructura al ser rápido y de bajo costo esto se debe en gran medida a que no se requiere un equipo específico simplemente una conexión a internet que permita el acceso desde cualquier ubicación.
- Permite la creación de una organización más ágil y rentable basada en integración y automatización. Esto debido a que permite saltarse la fase de adquisición de hardware y el consiguiente gasto, por lo cual es perfecta para la creación de empresas.
- La contratación de este tipo de servicios es mucho más baja que el hecho de adquirir equipos manejar un centro de datos propio, así como también el coste operativo de esta gestión. Además, que La computación en la nube se puede aplicar en casi cualquier entorno: desde el pequeño comerciante que necesita un sitio Web de comercio electrónico, de forma rápida y barata, hasta las grandes empresas, que desean disminuir los gastos y evitar las dificultades de administrar un centro de cómputo complejo.
- Brinda la posibilidad de adaptarse rápidamente a negocios en crecimiento o de picos estacionales, ya que está diseñado para hacer frente a fuertes demandas de carga de trabajo. Esto aumenta la agilidad de respuesta, disminuye los riesgos y los costos operacionales, porque sólo se va escalando según el crecimiento de la empresa o usuario y se paga sólo lo que se usa.
- Permite la compartición de archivos y recursos entre usuarios haciendo eficiente el uso de los recursos informáticos y un flujo óptimo de la información.
- La utilización redundante de múltiples sitios hace posible poder realizar recuperaciones de información en caso de fallo, dando lugar a una disponibilidad optimizada de servicios entregados otorgando una computación ágil y de misión crítica.
- Al tener una centralización de datos y al aumento de la seguridad centrada en los recursos, conlleva a la obtención de niveles más altos de seguridad contra las amenazas emergentes, mediante una mayor visibilidad y prevención. La

seguridad es a menudo tan buena o mejor que los sistemas tradicionales, en parte porque los proveedores están en condiciones de dedicar recursos a la solución de los problemas de seguridad que muchos clientes no pueden pagar.

### 2.7.2 Riesgos

Entre los riesgos que se puede presentar en esta tecnología cabe destacar los siguientes:

- Aunque se contempla que la computación en la nube maneja múltiples recursos en alta disponibilidad, la mayoría de los principales servicios de esta tecnología han sufrido cortes, y los administradores de la infraestructura son capaces de hacer muy poco cuando se ven afectados.
- Hay que tener en cuenta que al depender concretamente de la internet se debe poseer una conexión estable y con un ancho de banda que sea lo suficiente como tener una administración adecuada de los recursos ofrecidos en la nube, en caso de contar con aquello puede ser muy tedioso el trabajo además de poder causar algún tipo de error no deseado.
- Provoca una percepción de inseguridad debido a que al ser la internet una red con acceso a una cantidad considerable de información es mucho más susceptible a los ataques informativos y sobre todo que el control de esta seguridad solo está a cargo del proveedor y no como tal del usuario o la empresa lo que provoca en estas desconfianzas sobre las medidas de seguridad adoptadas y que puedan poner en riesgo la información vital de la empresa que perjudique su continuidad en el negocio.
- El tema de integración es un aspecto que considerar, debido a que no en todos los entornos es factible asociarse de manera fácil a la infraestructura de cómputo en la nube con sistemas desarrollados tradicionalmente, por lo que el cliente debe tener claro qué tan viable resulta implementar una solución basada en la nube dentro de su organización.

### **3 Balanceo de Carga**

Las redes de datos están en una constante evolución y crecimiento, tanto a nivel de tráfico como de complejidad, debido a esto es necesario optimizar las mismas en la mayor medida posible es por ello por lo que en esta sección se tratara de una tecnología importante para conseguir aquello como lo es el balanceo de carga.

#### **3.1 Definición**

El balanceo de carga es una tecnología encargada de permitir la distribución del tráfico o flujo de la red a través de múltiples recursos de computación que ofrezcan los mismos servicios, estos equipos pueden ser de características distintas o semejantes por lo que se determina la disponibilidad de procesamiento y los recursos de cada equipo en un sistema computacional [11].

Este pretende maximizar la utilización de los recursos, posibilitando el mejor desempeño del sistema, esto a través de dividir la carga de trabajo de los procesos u otros recursos que están siendo ejecutados por un equipo, que son generadas por las distintas peticiones provenientes de la red dando un equilibrio a dichos procesos y utilizando de mejor manera los recursos de los componentes que proveen los servicios de la empresa.

El principal objetivo planteado es conseguir que todos los elementos que llevan a cabo la misma tarea estén igualmente cargados con el fin de aumentar la calidad y disponibilidad del servicio, todo el proceso de balanceo de carga es transparente para el usuario final.

#### **3.2 Funcionalidades**

En este subcapítulo se explicará el funcionamiento del balanceo de carga y las diferentes características que posee esta para realizar su propósito.

La manera en que esta tecnología realiza dicha distribución de tráfico es a través de un dispositivo de red denominado balanceador de carga, el cual intercepta este tráfico redireccionando todos los pedidos dirigidos a un servicio entre los distintos servidores que lo ofrecen para asegurar un flujo continuo y uniforme consiguiendo de esta manera la máxima velocidad de ejecución.

Entre las distintas funciones realizadas por este balanceador de carga son las siguientes:

- Receptar el flujo de red que es dirigido a un destino específico.
- Se encarga de clasificar el tráfico en peticiones individuales para luego determinar a qué servidor destinar esta petición.
- Este realiza un monitoreo sobre los servidores que se desea balancear para verificar el estado en el que se encuentran los mismos y si se encuentran disponibles, es decir que están respondiendo al tráfico en caso de no ser así son separados de los servidores que están siendo parte del balanceo hasta que este vuelva a funcionar.
- Nos proporciona redundancia ya que se cuenta con varios escenarios a prueba de fallos, por tanto, es capaz de en caso de caída o pérdida de servicio de un equipo reasignar el mismo a otro equipo disponible permitiendo la continuidad de los servicios ofrecidos por la organización.
- Provee de una distribución de contenidos realizando cosas como el hecho de leer urls, interceptar cookies y analizar código XML.
- Son capaces de tomar decisiones basadas en el tráfico tradicional de información de OSI capa 2 y 3.
- Los balanceadores de carga más avanzados, pueden tomar decisiones inteligentes de gestión del tráfico basadas en información de la capa 4-7 incorporadas en las solicitudes emitidas por los clientes.
- Los clientes de las aplicaciones realizan las peticiones directamente a través de una conexión con una dirección IP virtual en el balanceador de carga, en lugar de disponer de una dirección IP de servidor individual. Este luego transmite la solicitud del cliente al servidor de aplicaciones adecuado. Toda esta operación es transparente para el cliente, para quien parece que se conectan directamente al servidor de aplicaciones.
- Estos equipos manejan escenarios de redundancia aplicados en caso de fallos el primero de ellos es un estado activo-pasivo en el cual, en caso de caída de uno de los servidores, el equipo en estado pasivo sin intervenir en la carga de trabajo pasa a entrar en funcionamiento e incorporarse en el grupo de

equipos balanceados hasta que el servidor que presento el inconveniente recupere su funcionalidad. El segundo estado es activo-activo en el cual todos los equipos que dispone el centro de datos están en estado de funcionamiento y en caso de que uno de ellos falle el resto toman el control de la carga de trabajo de ese equipo.

### 3.3 Tipos de Balanceo de Carga

En esta sección se dará a conocer los distintos tipos de balanceo de carga que se pueden aplicar.

Los balanceos de carga pueden ser determinados por la ejecución de procesos entre estos tenemos [6]:

- **Balanceo de Carga Estático:** En este método, el rendimiento de los procesadores se determina en el comienzo de la ejecución, después dependiendo de su desempeño la carga de trabajo se distribuye en el inicio por el procesador maestro. los procesadores esclavos calculan el trabajo asignado y envían su resultado al maestro. Una tarea siempre se ejecuta en el procesador al que se asigna los métodos de equilibrio de carga estáticos por lo que no es preventivo.

El objetivo del método de equilibrio de carga estática es reducir el tiempo de ejecución de un programa concurrente mientras se minimiza los retrasos de comunicación. Este tipo de balanceo presente algunos inconvenientes como:

- 1) Es muy difícil predecir de forma exacta el tiempo de ejecución de todas los componentes en las que segmenta un programa antes de ejecutarlo.
- 2) Determinados sistemas pueden tener retardos en las comunicaciones que pueden variar bajo diferentes circunstancias, lo que hace imposible determinar el valor de esta variable.
- 3) Regularmente los problemas necesitan un número indeterminado de pasos computacionales para alcanzar la solución. Un caso sobre esto es cuando los algoritmos de búsqueda normalmente atraviesan un grafo buscando la solución, y ante esto no se sabe cuántos caminos hay que probar, independientemente de que la programación sea secuencial o paralela.

4) Una de las principales desventajas que presenta casi todos los esquemas estáticos es que el final de la selección de un host para la asignación del proceso esto se realiza cuando se crea el proceso. Por lo que no se puede cambiar durante la ejecución del proceso para realizar cambios.

➤ **Balanceo de Carga Dinámico:** Este balanceo toma las decisiones durante el equilibrio de carga del sistema distribuido es decir es realizado durante la ejecución de los procesos. Las tareas pueden moverse de un sistema sobrecargado a un sistema bajo de carga para recibir un servicio más rápido. El enfoque dinámico es mejor ya que considera el estado del sistema en equilibrar la carga.

En este balanceo la división de la carga computacional depende de las tareas que se están ejecutando y no de la estimación del tiempo que pueden tardar en ejecutarse. Aunque este lleva consigo una cierta sobrecarga durante la ejecución del programa, resulta una alternativa mucho más eficiente que el balanceo de carga estático. Las tareas son repartidas entre los procesadores durante la ejecución del programa.

Dependiendo de dónde y cómo se almacenen y repartan las tareas este se clasifica en dos tipos:

- **Distribuido:** En este sistema se ejecuta el algoritmo de equilibrio de carga por todos los nodos en el sistema y la responsabilidad de equilibrio de carga es compartido por todos los nodos. Por lo regular tienden a producir más mensajes esto se debe a la metodología distribuida en la que cada nodo necesita comunicarse entre sí para tomar decisiones.

Un gran inconveniente que presentarse es que el nodo maestro únicamente puede repartir una tarea a la vez, y después de que haya enviado las tareas iniciales sólo podrá responder a nuevas peticiones de una en una. Por tanto, se pueden producir colisiones si varios esclavos solicitan peticiones de tareas de manera simultánea por esta razón es más recomendable su uso cuando se manejan pocos esclavos y las tareas de computación son intensivas.

- **Centralizado:** en este sistema, los nodos no son responsables de la carga. En cambio, se elige un maestro que balancee la carga para todo el sistema, es decir manejan un funcionamiento Maestro/Esclavo.

Una vez que un nodo esclavo finaliza una tarea, este solicita una nueva al maestro. Esta técnica también se basa en una programación por demanda y no sólo es aplicable a problemas que tengan tareas de una misma cantidad de trabajo. Un uso adecuado de esta técnica sería para problemas donde el número de tareas pueda variar durante la ejecución. Una problemática que presentarse es que si llegase a fallar el equipo maestro toda la comunicación se corta no se tiene redundancia de comunicación.

### 3.4 Beneficios

En esta sección se dará a conocer una serie de puntos importantes con los que el balanceo de carga contribuye a mejorar en gran medida los temas de comunicación en una red de datos y la importancia del uso de esta tecnología en el ámbito de las telecomunicaciones.

1. Nos permite asegurar que un servidor no se vea sobrecargado por una gran afluencia de tráfico y que esto provoque una caída del sistema evitando su acceso para los usuarios. Este procedimiento es de vital importancia debido a que es difícil predecir la cantidad de solicitudes que se enviarán a un servidor, por ello los balanceadores de carga cuando detectan que un servidor está muy saturado de solicitudes dirigen este flujo de datos hacia otro equipo que cuente con mejor capacidad de respuesta, sin lo antes mencionado se puede generar una pérdida en el servicio evitando la continuidad del negocio.
2. Contribuye a maximizar el throughput y además se mejora el rendimiento de los sistemas, esto se debe en gran medida a una mejor administración del tráfico que es generado hacia los diferentes equipos que ofrecen el servicio en un centro de datos, esto se logra equilibrando la carga de trabajo a través de múltiples dispositivos, esto ayuda a que se optimice el uso de recursos que ofrece cada componente computacional.
3. Al otorgarnos una distribución del tráfico en diversos servidores, esto nos



permite una reducción en los tiempos de respuesta, ya que el tráfico de un servidor lento es transferido automáticamente a otro más rápido incluso si se reciben muchas peticiones a la vez.

4. Nos brinda confiabilidad ya que evita que cualquier servidor se convierta en un punto único de fallo ya que los balanceadores de carga monitorean la disponibilidad de los servidores de aplicaciones para evitar la posibilidad de enviar las solicitudes de cliente a un recurso de servidor que es incapaz de responder, es decir si se produce una caída de un equipo esto no influiría en el funcionamiento de todo el sistema ya que las funciones que ejercía el dispositivo que presento el inconveniente son asumidas por el resto de los equipos disponibles que forman parte del sistema de balanceo con esto se obtiene una alta disponibilidad del servicio ofrecido por las empresas.
5. Los balanceadores de carga tienen una característica importante que es la persistencia, esto permite que se pueda manejar todas las peticiones de un cliente específico en un solo servidor en lugar de distribuir las solicitudes entre todo el sistema, esto se realiza principalmente en circunstancias en las que se requiere conservar el estado de conexión de un cliente por ejemplo en los casos en que estos realicen compras electrónicas se pueda asegurar que cada vez que un usuario agrega un elemento a su carrito de compras la información se enviará al servidor específico de la gestión de la cesta de la compra.
6. Nos ofrece seguridad debido a que al impedir la congestión de red y con ello evitar que un servidor se sature y deje de funcionar, se reduce en gran medida la probabilidad de recibir ataques de negación de servicios DoS donde su característica principal es hacer que una máquina o recurso de red quede inutilizada o no disponible para sus usuarios esto a través del envío de cantidades gigantescas de peticiones a servidores específicos.
7. Facilita la realización del mantenimiento en los servidores tales como actualizaciones y configuraciones del este pueden tener lugar durante su funcionamiento sin que se produzca una remarcable pérdida del rendimiento del sistema. ya que se reconoce el estado de mantenimiento de este equipo y se desvía las peticiones de manera apropiada a otros equipos, esto es factible ya que se cuenta con una alta disponibilidad brindada por los otros servidores que componen este sistema, esto hace del balanceo de carga una solución muy flexible.

8. La escalabilidad es un aspecto importante que considerar es que una de las enormes ventajas ofrecidas, ya que actualmente la demanda en los servicios ofrecidos es cada vez mayor lo que conlleva que se requiere un aumento de los recursos, en esta tecnología los nuevos servidores pueden ser fácilmente incorporados a la infraestructura del centro de datos y el balanceador de carga iniciará de inmediato el envío de tráfico a los nuevos servidores.

## 3.5 Técnicas de Balanceo de Carga

En este subcapítulo se detallarán las diferentes técnicas que se realizan en un balanceo de carga así mismo como los pros y contras que radican en cada una de ellas.

### 3.5.1 Basado en DNS

Este tipo de técnica se sustenta principalmente es que a un nombre de dominio específico se le asignen varios registros de host con las direcciones IP de los servidores que están brindando el servicio, esto es usado principalmente en servidores web. Cuando el cliente realiza una petición al servidor DNS este va rotando entre las diferentes direcciones manejadas por ese dominio utilizando un algoritmo Round Robin, es decir manejando las consultas de manera cíclica entre uno u otro registro [16].

Hay que tener en cuenta que cuando el usuario realiza una consulta el cache del explorador de internet almacenara la dirección del servidor que atendió en ese momento su petición esto producirá que la próxima vez que el cliente realice una petición a este dominio sea redireccionado a la IP del servidor que atendió su última consulta.

Este tipo de balanceo si cliente realiza una petición y el servidor no es capaz de responder a su solicitud porque el mismo este caído o presente algún inconveniente este probara con la siguiente dirección de manera cíclica y así sucesivamente hasta recibir la respuesta de un servidor otorgándole acceso al servicio. Como se maneja el esta técnica se observa en la Figura 3.11.

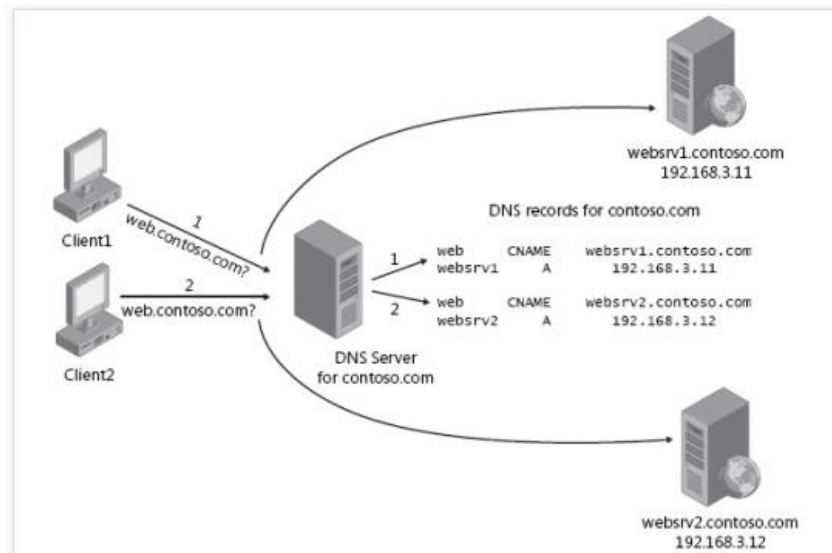


Figura 3.11. Balanceo de Carga Basado en DNS

### ❖ Ventajas

1. Permite el trabajo con sistemas heterogéneos es decir poder manejar este tipo de balanceo entre equipos con diferentes sistemas operativos y arquitectura.
2. Nos brinda el balanceo de grandes redes sin necesidad de contar con un equipo dedicado que brinde este tipo de servicio lo que nos proporciona un ahorro económico.
3. Es muy sencillo de implementar esta técnica y además de proporcionar un balanceo sencillo sin dificultades de administración ya que se distribuye la carga de manera aleatoria.

### ❖ Desventajas

1. En caso de la caída de un servidor los clientes que estaban enviando peticiones a este equipo seguirán realizándolo por lo que estarán sin servicio hasta que la cache de su explorador se actualice y sea dirigidos a un servidor que este operativo.
2. No se cuenta con una distribución apropiada de carga ya que esta es manejada de manera cíclica por ello un servidor puede estar recibiendo peticiones más pesadas que otros lo que imposibilita un

manejo uniforme de dicha carga y que puede llevar consigo inconvenientes de saturación en un servidor específico poniendo lento el funcionamiento del sistema o incluso provocando que este deje de funcionar.

3. No se toma en cuenta el tiempo de respuesta otorgada por cada servidor, así como la congestión de red que ese pueda estar produciendo.
4. No hay manera de asegurar la conexión al mismo servidor dos veces, si es necesario como el caso del comercio electrónico en la realización de una compra.

A continuación, se nombrarán tres técnicas que manejan como tal un nodo director, es decir un equipo que es el receptor de tráfico de los paquetes el reenvió de estos hacia los servidores que forman parte del sistema se los realiza a través de las siguientes técnicas:

### 3.5.2 Encaminamiento Directo

Este tipo de técnica se caracteriza principalmente porque todos los equipos conforman parte del sistema de balanceo pertenecen a un mismo segmento de red y todos estos manejan una misma dirección pública. El balanceador recibe todo el tráfico y distribuye a los diferentes nodos una vez recibida las peticiones esos servidores son los encargados de procesar estas consultas y comunicarse directamente con el usuario, con esto se reduce en gran medida la carga de trabajo que manejaría el balanceador al disminuir el procesamiento ya que solo recibiría los paquetes entrantes. Como se puede visualizar en la Figura 3.12. [10]

#### ▪ **Ventajas:**

- 1) Este balanceo nos permite sobre todo impedir las cuestiones de congestión de tráfico, sobrecarga y encolamiento de paquetes en el balanceador de carga que conllevan a la formación de un cuello de botella, ya que existe respuestas directas a los clientes dada por los servidores.
- 2) No otorga eficiencia ya que al manejarse en el mismo segmento

físico de red no existe un cambio en las cabeceras de red, transporte o aplicación.

▪ **Desventajas:**

- 1) El hecho de manejar un segmento de red similar para el sistema de balanceo de servidores provoca que solo se posea una conexión a nivel de capa 2.
- 2) En este tipo de técnica se debe tratar de garantizar que los equipos no respondan a comandos ARP para evitar conflictos con otros protocolos logrando a través de esto que los servidores respondan con la misma dirección IP, pero con diferente dirección MAC. Con esto el balanceador redirige el tráfico a un servidor a través de su MAC.

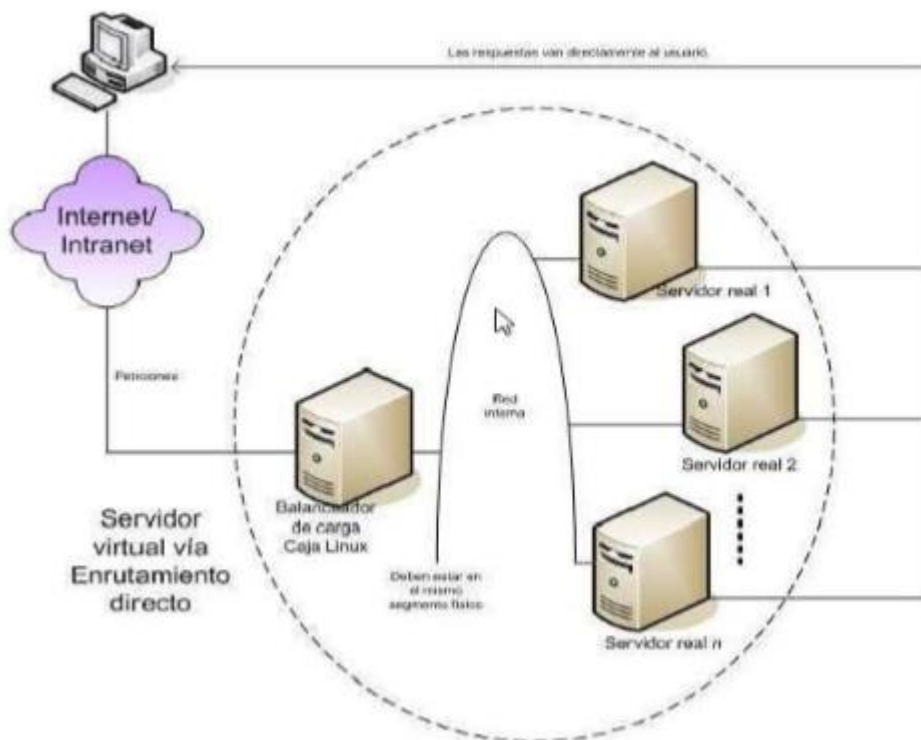


Figura 3.12. Balanceo de Carga Basado en Encaminamiento Directo

### 3.5.3 NAT

Este tipo de técnica se basa principalmente en el uso de traducciones de direcciones IP tanto de origen como de destino de los paquetes que son receptados por el balanceador de carga el cual los envía hacia las direcciones reales de los servidores que forman parte del sistema de balanceo para el servicio solicitado, una vez que estos procesan la petición realizada reenvían la respuesta al balanceador para que este rescriba su dirección origen y sea transmitido al cliente con la dirección IP virtual del equipo balanceador. Por ello en esta técnica el balanceador es el encargado tanto de recibir como de enviar el tráfico de red generado a los servidores, es decir todo el flujo de red tanto de entrada como de salida es administrado por el balanceador. Esto se puede observar mejor en la Figura 3.13 [20].

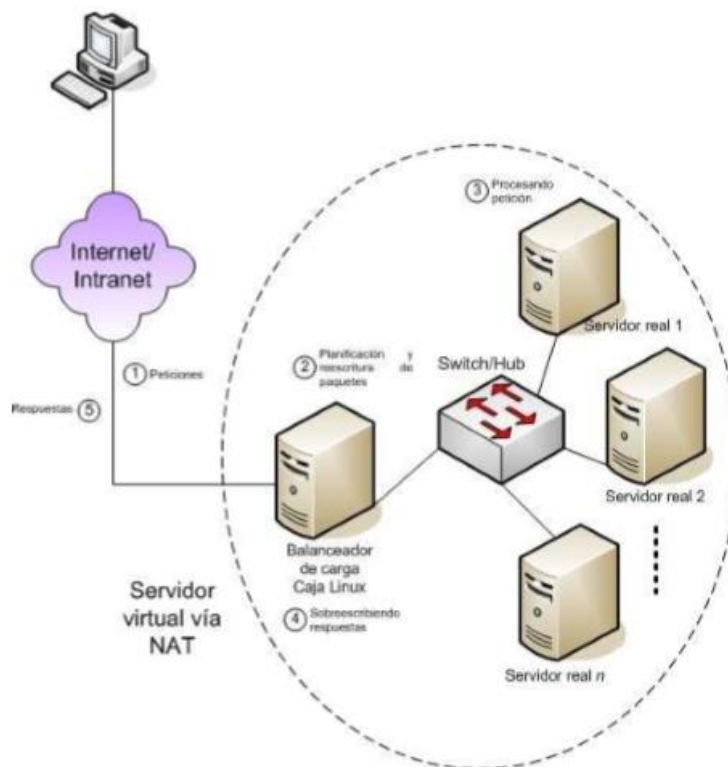


Figura 3.13. Balanceo de Carga Basado en NAT

#### ✓ Ventajas

1. Este tipo de técnica es su facilidad, rapidez y pocos requerimientos de hardware para dicha implementación, ya que se puede utilizar cualquier servidor para el procesamiento TCP/IP.

2. Permite un mejor nivel de seguridad debido a que se usa servidores reales remotos ya que se camufla las direcciones privadas de los servidores reduciendo así el riesgo de generación de ataques.
3. Nos permite mantener sesiones establecidas a un servidor específico debido a que el balanceador administra tanto el tráfico entrante como saliente y permite almacenar en cache las conexiones de un determinado cliente a un servidor esto es muy útil en el tema de comercio electrónico donde es necesario mantener una sesión a un determinado servidor durante todo el proceso de compra.

#### ✓ **Desventajas**

1. En el hecho de que todo el flujo de datos sea administrado por el balanceador puede producir sobrecarga, congestión y encolamiento de paquetes debido a que se es muy propenso a producir un cuello de botella sobre este equipo.
2. El manejo de traducciones de dirección IP muchas veces provoca conflictos en la transmisión de paquetes ya que muchas de estas suelen ser ineficientes y difíciles de interpretar para ciertos equipos.
3. Con esta técnica se pierde escalabilidad y rendimiento debido a que todo el tránsito de recepción y envío de paquetes pasa por un único equipo lo que provoca lentitud en el procesamiento de las solicitudes dadas por los clientes.

#### 3.5.4 **Túneles IP**

Actúa de manera similar a un balanceo por encaminamiento directo, pero con la pequeña diferencia que en este tipo de técnica tanto el balanceador como los servidores reales pueden estar en redes remotas encapsulando los paquetes transmitidos a través de túneles por medio de tramas IP con esto se logra que no necesariamente pertenezcan a un mismo segmento de red, esto siendo observado de mejor manera en la Figura 3.15.

El funcionamiento de este tipo de técnica se basa en la recepción de las peticiones por parte del balanceador el cual encapsula la unidad de datos en otra unidad de datos

manteniendo las direcciones origen y destino enviadas por el cliente sin modificaciones, y además conteniendo la IP origen al balanceador y como destino el servidor real disponibles que pueda ofrecer el servicio solicitado, una vez recibido por el servidor el requerimiento este desencapsula la unidad de datos la procesa y envía la respuesta al cliente directamente utilizando la IP pública del sistema, es por ellos necesario que los servidores posean la misma dirección pública que el balanceador para poder aceptar la unidad de datos original enviada por el cliente y responder ante el requerimiento. Este tipo de funcionamiento se puede visualizar mejor en la Figura 3.14 [7].



Figura 3.14. Envío de unidades de datos en un Balanceo por Túneles IP

### ➤ Ventajas

- 1) Al igual que el encaminamiento directo se evita la formación de cuellos de botella con lo que no se sobrecarga al balanceador debido a que este solo recibe el flujo de entrada y cada servidor es capaz de responder a las peticiones enviadas por el balanceador directamente a los clientes, pero esta vez a través del uso de túneles IP unidireccionales.
- 2) Nos otorga escalabilidad ya que a medida que el sistema lo vaya requiriendo, se podrán ir añadiendo nuevos equipos a la granja de servidores sin ningún tipo de problema.
- 3) Permite implementar sistemas con alta disponibilidad y balanceo de carga muchos más óptimos, sin tener que depender de la ubicación física de los equipos, debido a que tanto el balanceador como el servidor puede estar en segmentos de red diferentes y conectados de manera remota.



### ➤ Desventajas

- 1) Se puede presentar inconvenientes con algunos equipos que no tengan soporte a túneles IP.
- 2) Al realizar el encapsulamiento se forman nuevas unidades de datos que podrían estar con cabeceras sobrecargadas pudiendo generar lentitud en el procesamiento de estas solicitudes.

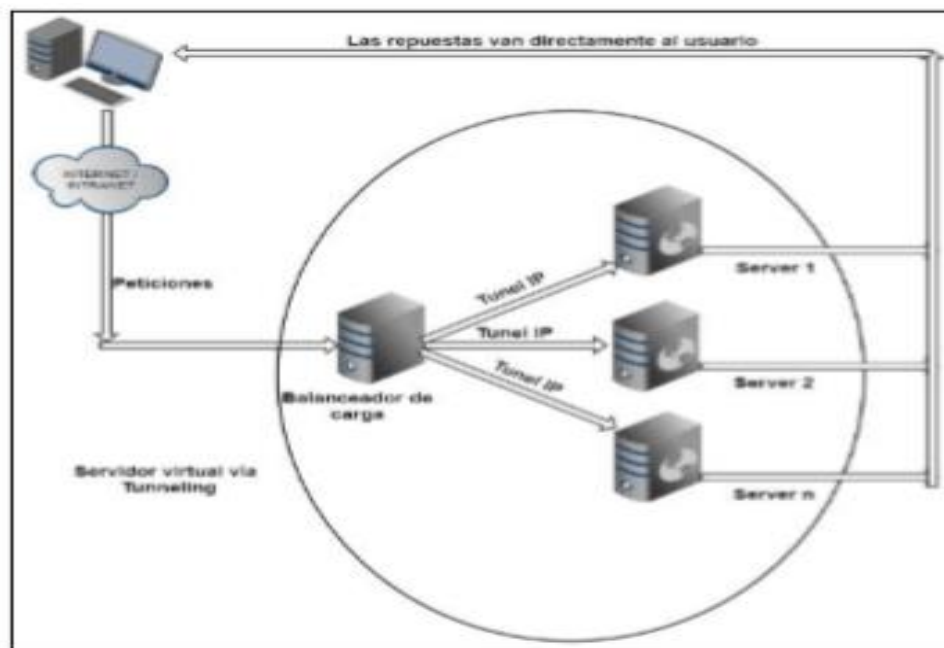


Figura 3.15. Balanceo de Carga Basado en Túneles IP

## 3.6 Algoritmos de Balanceo de Carga

En esta sección se analizará los diferentes algoritmos que son implementados para la distribución de la carga entre los distintos recursos de computación que ayudan a mejorar el rendimiento de los centros de datos según la infraestructura de red con la que se cuenta y las necesidades presentadas por la organización. Estos algoritmos son manejados a través de dos políticas de balanceo que es indispensable tenerlas claras las cuales son:

- **Adaptativas:** Este criterio toma en cuenta el ejemplo de algoritmos que utilizan información en tiempo real, es decir cuando se están ejecutando las

solicitudes al sistema de balanceo que les permite elegir el servidor más adecuado para cada petición en concreto y optimizar la distribución de la carga de trabajo [19].

- **No Adaptativas:** Este se caracteriza por la utilización de algoritmos que no se ajustan al estado actual del sistema, es decir no se analiza los criterios de estado de los servidores que ofrecen el servicio, además que para este todos los requerimientos usan esta política que se aplica por el lapso de peticiones del cliente. Aquí se contienen algoritmos tales como los Random o Round Robin [14].

Entre los algoritmos que se establecen para la distribución de la carga entre los diferentes recursos computacionales tenemos:

### 3.6.1 Round Robin

Este tipo de algoritmo el balanceador se encarga de distribuir la carga de manera secuencial entre los diferentes servidores que forman parte del sistema, enviando las solicitudes de forma cíclica empezando por el primer servidor, el segundo y así sucesivamente hasta que se recorra todos los equipos y después retomando esta secuencia nuevamente hasta que se haya distribuido todo el tráfico recibido. Esto se puede visualizar mejor a través de la Figura 3.16 [12].

Todo el flujo de datos de la red que es receptado es tratado con la misma importancia independientemente de la urgencia o criticidad de la solicitud o de la carga que este genere a un servidor, esto puede conllevar a que servidores reciban peticiones de carga mucho mayor, mientras que otros apenas se encuentren utilizando recursos.

También se debe considerar que no considera la cantidad de conexiones entrantes o del tiempo de respuesta que experimenta cada servidor ya que muchas solicitudes pueden requerir diferentes tiempos de procesamiento con lo cual podríamos estar saturando a ciertos equipos enviándoles solicitudes de mayor procesamiento y causando lentitud en la respuesta hacia el cliente. Principalmente este tipo de algoritmo recomienda cuando se tienen servidores con iguales o similares capacidades de procesamiento, memoria y almacenamiento, ya que se le asigna la misma cantidad de pedidos para intentar realizar una distribución equitativa de la carga.

Una limitación que es importante tomar en cuenta es que esta demora en el

reconocimiento de problemas presentados en el sistema por lo que se podría estar enviando solicitudes a un servidor que presente lentitud o que no se encuentre disponible en un momento determinado.

Es el algoritmo más utilizado. Es simple y fácil de implementar. Ofrece el mejor rendimiento para algunas aplicaciones generalmente se usa en servidores web donde las solicitudes HTTP son de naturaleza similar y, por lo tanto, se distribuirán por igual.

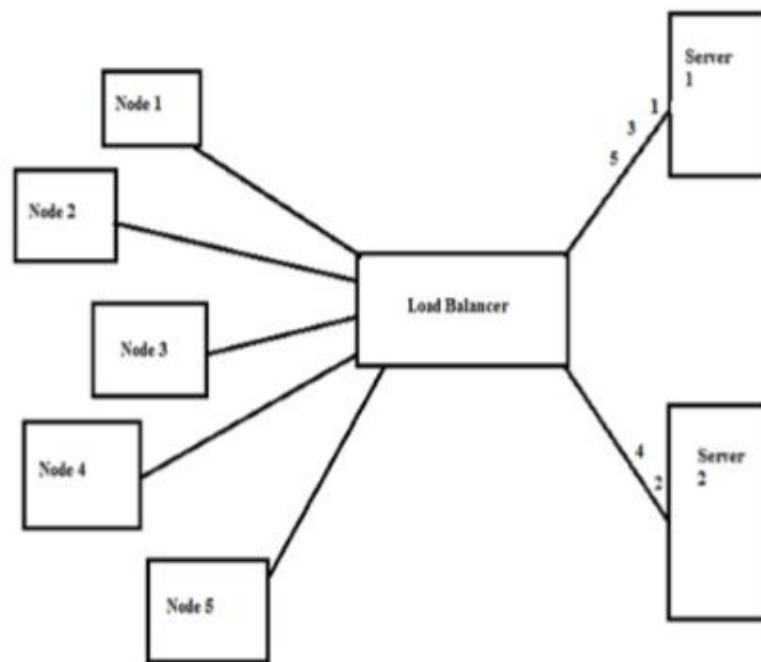


Figura 3.16. Algoritmo Round Robin

### 3.6.2 Round Robin Ponderado

Este algoritmo se maneja de manera similar al round robin solo en que este se establece una distribución con mayor carga a ciertos servidores que cuenten con mejor capacidad de procesamiento, memoria y almacenamiento que el resto, esto se gestiona a través de un indicativo otorgado por factor de peso para cada servidor que es establecido en los archivos de configuración del balanceador [9].

Los servidores que son capaces de manejar una gran carga se le asignan más solicitudes por parte del balanceador cuando recepta el tráfico de la red es recomendable utilizar este tipo de algoritmo cuando los recursos computacionales no

son homogéneos en sus características, esto se puede observar en la Figura 3.17.

Su uso es muy apropiado si hay diferencias representativas en la capacidad de los servidores que otorgan el servicio en el sistema. No obstante, si la carga de peticiones varía de forma drástica, los equipos con una capacidad mayor responderán a más peticiones de las que deberían e igual podría producir una distribución inapropiada y como consecuencia una saturación de los equipos que cuenten con un factor de peso muy elevado.

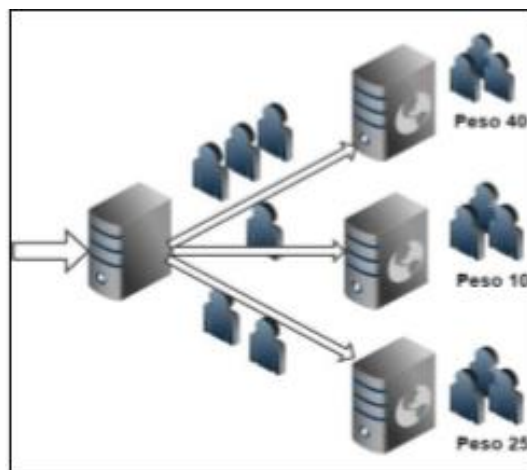


Figura 3.17. Algoritmo Round Robin Ponderado

### 3.6.3 Menos Conexiones Activas

Se maneja una programación dinámica debido que se tiene un mecanismo de consulta continua que nos permite comprobar el número de conexiones establecidas activas en cada servidor que se encuentra atendiendo los requerimientos del cliente. Al poseer un registro de las conexiones activas para cada servidor y ante la recepción de una nueva petición se asigna la misma al servidor con menos conexiones en ese momento. Esto es apreciable en la Figura 3.18 [13] [15].

Para los balanceadores de carga que está administrando una colección de servidores con un rendimiento similar, la programación de menos conexión es buena para distribuir el flujo de tráfico sin problemas cuando hay un alto grado de variación en la carga de las solicitudes receptadas.

Este algoritmo toma en cuenta el problema que surge debido a que un servidor se sobrecargará más rápido que otro porque algunos clientes permanecen conectados a

ese servidor mucho más tiempo que el otro o cuando la carga de un cierto tipo de solicitud toma un mayor tiempo de procesamiento a un servidor.

Asume que las capacidades de procesamiento de todos los servidores son las mismas y asigna la solicitud recién llegada al servidor con menor número de conexiones establecidas. Sin embargo, el rendimiento del sistema no es el apropiado cuando las capacidades de procesamiento de los servidores son diferentes ya que no se estaría aprovechando la capacidad de estos equipos, además de que podría haber retrasos en las respuestas a las solicitudes. es por eso que es recomendable tener un sistema con equipos homogéneos que permitan tener una distribución equitativa de la carga.

Nos permite evitar el problema presentado en los dos algoritmos anteriores los cuales, Al repartir las peticiones no consideran cuántas conexiones debe mantener un servidor por un tiempo determinado lo que podría conllevar a una sobrecarga de uno de ellos si se acumularan varias conexiones activas a un mismo equipo.

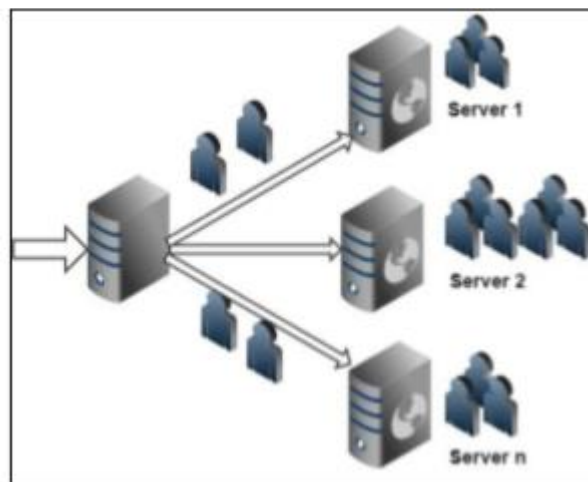


Figura 3.18. Algoritmo con Menos Conexiones Activas

### 3.6.4 Menos Conexiones Activas Ponderado

Este algoritmo es manejado de igual manera que las menos conexiones activas, pero en este caso, además de considerar el servidor con menos conexiones establecidas se determina la capacidad de procesamiento que es capaz de aceptar el servidor y de esta manera asignar una carga de trabajo apropiada tomando en cuenta estos dos criterios para realizar la distribución del tráfico. Esto se puede verificar de mejor manera en la Figura 3.19.

Se distribuye más solicitudes a los servidores con menos conexiones activas en relación con sus capacidades, la misma es indicada y ajustada por la información de carga dinámica, la agregación de un parámetro de capacidad hace que este algoritmo sea idóneo cuando la infraestructura tiene servidores reales con capacidades de hardware diferentes [17].

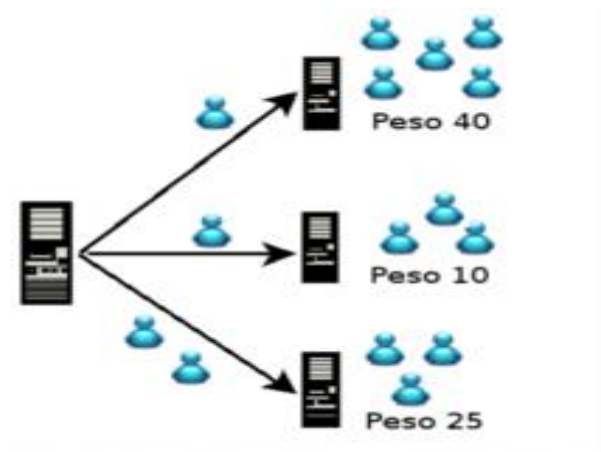


Figura 3.19. Algoritmo con Menos Conexiones Activas Ponderado

### 3.6.5 Tablas hash por origen y destino

Este algoritmo se basa principalmente en la utilización de una tabla de asignaciones fijas la cual contiene las direcciones IP origen y destino, a través de estas se indica que servidor deberá atender la petición. balanceador compara las direcciones de las tramas TCP/IP que receipta de estas tablas y actúa en consecuencia para realizar la distribución de la carga [8].

Se genera una clave hash única y que es iniciada por la fuente de origen de la petición y la dirección IP de destino del servidor y del cliente, este ha sido asignado a un servidor particular con la ayuda de esta clave. Las solicitudes se distribuyen a los servidores en base al clave hash. Por lo tanto, esta se usa para determinar qué servidor recibirá la solicitud. Si un servidor no funciona dejando de brindar el servicio la distribución realizada cambia. Esto se puede observar de manera más apropiada en la Figura 3.20.

Este algoritmo asegura que el mismo cliente con la misma dirección IP llegará al mismo servidor siempre y cuando esté en funcionamiento, si el resultado hash cambia

debido al número de servidores en ejecución, muchos clientes serán dirigidos a un servidor diferente. Este algoritmo es generalmente usado en modo TCP donde no se puede insertar ninguna cookie.

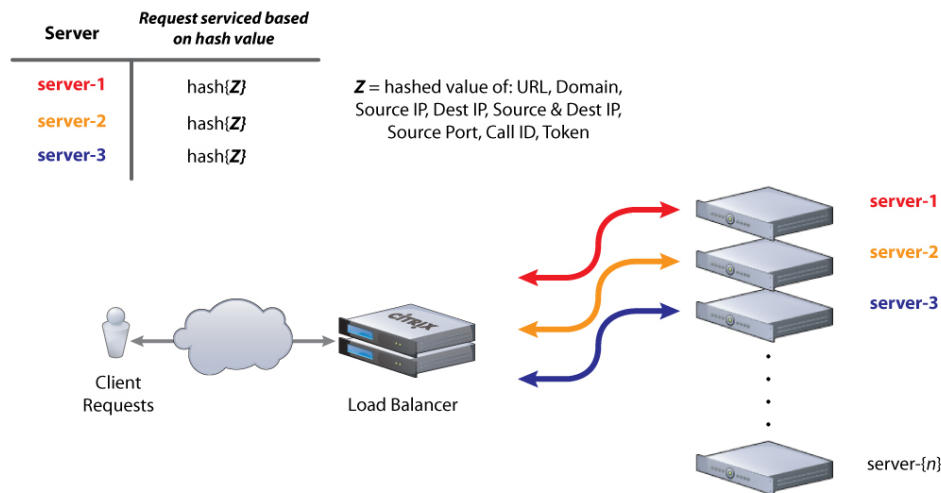


Figura 3.20. Algoritmo con Tablas hash por origen y destino

### 3.7 Balanceo de carga por geolocalización DNS

En esta parte del capítulo se explicará lo referente a la geolocalización DNS, así como ciertas características y ventajas que presenta este tipo de tecnología a la hora de efectuar el balanceo de carga entre medios computacionales.

#### 3.7.1 Definición

Este tipo de tecnología consiste en que cuando un cliente realiza una consulta a su servidor DNS, el servidor DNS busca la ubicación del cliente en función de su dirección IP pública en el paquete de consulta DNS. El servidor DNS encuentra el servidor de archivos más cercano a esta ubicación y devuelve la dirección IP de este servidor de archivos en la respuesta DNS. Esto le permite enviar a los usuarios a un servidor que sea más apropiado para ellos, que podría ser uno más cercano o uno que ofrezca contenido en otro idioma [23][24].

Po ello una solicitud de DNS entrante, realiza la búsqueda geográfica en el momento

de la solicitud, y devolver resultados diferentes en función de la dirección IP que genero dicha petición, el DNS actuará como un punto de conexión a través del cual el tráfico se transmite en todo el mundo en función del servidor de nombres DNS utilizado por el usuario o la dirección IP del usuario en función de la geolocalización del usuario final. Esto se puede observar de mejor manera en la Figura 3.20.

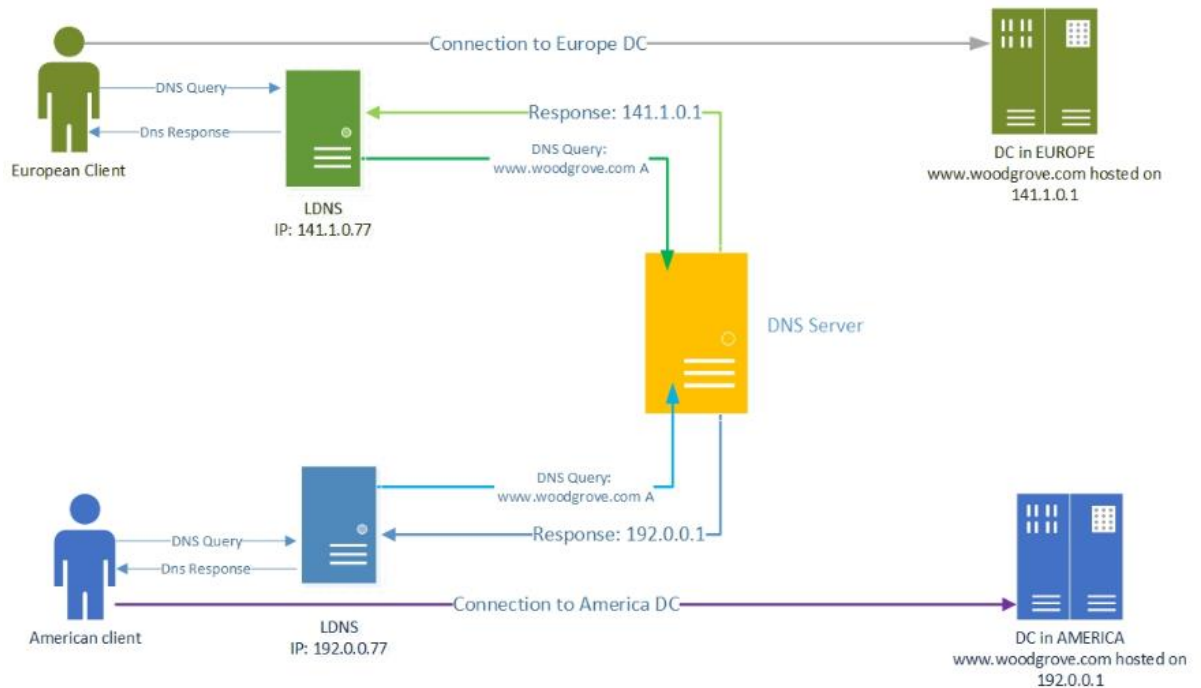


Figura 3.21. Balanceo de carga por geolocalización DNS

Esto ayuda a distribuir las cargas del sistema a través de toda su infraestructura de replicación distribuida geográficamente. Este mecanismo es independiente del protocolo, es universal y más simple para que los usuarios finales manejen aquello de manera transparente.

La geolocalización DNS da la capacidad básica de hacer un equilibrio de carga simple y de gran escala sin la necesidad de efectuar cambios a clientes o protocolos que personalizados.

Entre los grandes beneficios que nos otorga esta tecnología es la de permitir tiempos de carga de páginas web más rápidos en diferentes regiones debido a que el servidor más cercano se encarga de todo el contenido. También permite ejecutar diferentes sitios web para el mismo dominio en diferentes países, ya que se ejecutarán en



diferentes servidores.

Después de lo antes mencionado el balanceo de carga de DNS por geolocalización permite a los administradores de red maximizar el uso de la infraestructura geográficamente dispersa, aumentando el rendimiento al minimizar la distancia entre los centros de datos y los solicitantes. Este se basa en una metodología simple de balanceo de carga Round Robin. Esto significa que el balanceo de carga ejecuta una lista de servidores, enviando una conexión a cada servidor y comenzando en la parte superior de la lista cuando llega al final.

Adicional a los claros beneficios en cuanto a rendimiento, el balanceo de carga geográfica permite la entrega de contenido y servicios específicos de la ubicación, aprovechando la tecnología CDN existente.

Una característica a tener en cuenta es a poder realizar una distribución geográfica de la carga los administradores de un sitio web puede ofrecer contenido específico de ubicación, personalizar los contenidos, dirigir publicidad y ofertas que son relevantes para un cliente de una región específica, además de aquello permite la detección de fraudes, analizar el tráfico, lograr cumplimiento, gestionar digitalmente derechos e implementar funciones de direccionamiento y cercado geográfico.

Incluso el administrador de alguna aplicación realiza una fácil gestión de las actualizaciones de su aplicativo ya que podría lanzar una nueva versión solo a un conjunto selecto de clientes, para evitar la interrupción de toda la base de usuarios.

### **3.8 Software para Balanceo de Carga en Servidores Linux**

En esta sección se hará referencia a los distintos softwares que nos permiten implementar un balanceador de carga sobre un equipo con sistema operativo Linux el cual nos permita realizar las diferentes técnicas y algoritmos de balanceo mencionados en capítulos anteriores sin necesidad de contar con equipos de red especializados para balanceo de carga como es el caso de los Fortinet, Cisco y demás marcas que proveen de este tipo de equipos.

Entre algunos de los softwares disponibles para poder administrar balanceadores de carga sobre sistemas operativos Linux tenemos los siguientes:

### 3.8.1 Linux Virtual Server

Este software es considerado la principal alternativa de código abierto existente en el mercado siendo un proyecto que nos brinda la información y los componentes necesarios para montar un servidor virtual en una granja de servidores con Linux, además de proporcionar las herramientas necesarias para construir un sistema de balanceo de carga sobre un equipo con sistema operativo Linux, permite la distribución de carga de servicios de red tales como servidores web y de correo, utilizando conmutación en capa 4 en el kernel de Linux esto con lleva se pueda balancear la carga entre sesiones TCP y UDP entre múltiples servidores reales. Por lo tanto, proporciona una forma de escalar los servicios de Internet más allá de un solo host. Es muy rápido y permite que dichos servicios se escalen para dar servicio miles de conexiones simultáneas.

Pues bien, esta conmutación en capa 4 funciona al multiplexar las conexiones TCP / IP entrantes y los datagramas UDP / IP a servidores reales. Los paquetes son recibidos por un director de Linux y se toma una decisión sobre a qué servidor real enviar el paquete. Una vez que se toma esta decisión, los paquetes subsiguientes para la misma conexión se enviarán al mismo servidor real. Por lo tanto, se mantiene la integridad de la conexión.

Este software consiste en plantear un balanceador de carga, también conocido como director, que será la máquina que será accesible directamente para los clientes y luego tendremos los servidores que serán aquellos que receptorán las peticiones de los usuarios vía el balanceador de carga, y responderán a las peticiones que el funcionamiento típico de la mayoría de estos equipos.

A pesar de que LVS se ejecuta sobre Linux es capaz de balancear las conexiones de carga de los usuarios finales que ejecutan cualquier sistema operativo a servidores reales que ejecutan de igual manera cualquier sistema operativo, pero siempre que estas conexiones utilicen TCP o UDP. Un detalle a tener en cuenta es que los servidores podrán estar o bien en la misma red física o en redes diferentes lo que permitirá el tener servidores en sistemas distribuidas geográficamente.

Permite las técnicas y algoritmos de balanceo excepto el uso de tablas hash todo esto mencionado en capítulos anteriores. Una enorme ventaja de este tipo de software es la cantidad de aplicaciones que se enlazan para optimizar las cualidades de este entre algunas de estas aplicaciones como hearbeat, keepalived, entre otras.

### 3.8.2 HAPROXY

Es un balanceador de carga a nivel de aplicación basado en eventos, además es de código abierto y puede balancear la carga de cualquier servicio TCP/HTTP de alto desempeño y confiable, que ofrece servicios de proxy y balanceador de carga brindando alta disponibilidad para aplicaciones basadas en estos servicios, diseñado con el objetivo de reenviar los datos de manera rápida, ejecutando la menor cantidad de operaciones posibles [18].

Nos permite distribuir la carga de conexiones TCP y HTTP que llegan como peticiones por parte del cliente. En el modo TCP, las decisiones de equilibrio de carga se toman para para tipo de conexión. En el modo HTTP, las decisiones se toman por solicitud. Es especialmente para sitios web que deben receptor cargas muy altas de tráfico y, al mismo tiempo, requieren persistencia de sesión o procesamiento de capa 7.

Su arquitectura está optimizada para mover datos tan rápido como posible con las menores operaciones posibles implementa una capa Modelo que ofrece mecanismos de derivación en cada nivel para garantizar que los datos no lleguen a niveles más altos niveles cuando no es necesario este software otorga grandes ventajas al no consumir mucho procesamiento en su ejecución ayudando al kernel hacer su trabajo más rápido esto debido a que solo requiere el ejecutable HAProxy y un archivo de configuración para ejecutarse [21].

Se ha convertido en el balanceador de carga de opensource de mayor uso que se aplica con la mayoría de las distribuciones de Linux, y a menudo se implementa de manera predeterminada en las plataformas de la nube., aunque su desarrollo es de carácter personal, a sido en gran parte avalado por diferentes empresas que incluso lo incluyen en productos comerciales.

Nos otorga un soporte específico para servicios como HTTP, SMTP, MySQL, entre otros, pero el protocolo HTTP es el más desarrollado y sobre este se pueden añadir, quitar o modificar cabeceras HTTP, tanto en la petición como en la respuesta por lo que su configuración es muy potente y permite establecer arquitecturas complejas.

Este no es un balanceador de carga basado en paquetes debido a esto no verá paquetes IP ni datagramas UDP, no realizará NAT o incluso menos DSR. Estas son tareas para capas inferiores. Algunos componentes basados en kernel como IPVS ya lo hacen bastante bien y se complementan perfectamente con HAProxy.

### 3.8.3 PEN

Se trata de un balanceador de carga altamente escalable, disponible y robusto para protocolos basados en TCP y UDP, tales como DNS, HTTP o SMTP. Permite que varios servidores aparezcan como uno al exterior y detecta automáticamente los servidores que están caídos y distribuye clientes entre los servidores disponibles.

Este balanceador realiza un seguimiento de los clientes e intentará enviarlos de vuelta al servidor que visitaron la última vez. La tabla del cliente tiene un número de ranuras (por defecto 2048, configurable mediante argumentos de línea de comandos) cuando esta tabla se llena, la ranura menos usada recientemente será arrojada para dejar espacio para la nueva.

Cuando este software detecta que un servidor no está disponible, escanea uno que comience con la tarea recientemente establecido por el servidor que quedo fuera de servicio de esta forma obtenemos balanceo de carga y failover.

Admite el filtrado GeoIP y se puede configurar para la terminación SSL, el retorno directo del servidor y el proxy inverso transparente.

Correctamente configurado, nos permite garantizar que un conjunto de servidores esté siempre disponible, incluso cuando los servidores individuales se cierran para mantenimiento o reconfiguración, además funciona en una larga lista de sistemas operativos, incluidos, entre otros, FreeBSD, Linux, HP-UX, Windows y Solaris.

Unas desventajas a tomar en cuenta es que no tiene interfaz gráfica y brinda estadísticas poco claras acerca de que máquina atendió la petición.

### 3.8.4 POUND

Es un reverse proxy, balanceador de carga y un servidor web front-end HTTPS., fue desarrollado exclusivamente para permitir la distribución de carga entre varios servidores web y permitir usar el protocolo de seguridad SSL para aquellos servidores web que no lo ofrecen de forma nativa está distribuido bajo la licencia GPL.

El balanceador de carga eventualmente actúa como un proxy inverso que distribuye uniformemente el tráfico entrante / solicitudes recibidas del puerto 80 o 443 (o ambos)

en dos o más servidores web back-end. Este distribuirá las solicitudes desde los navegadores del cliente entre varios servidores de fondo, mientras mantiene la información de la sesión. Es capaz de descifrar las solicitudes HTTPS de los navegadores cliente y las pasará como HTTP simple a los servidores back-end, además verificará las solicitudes de corrección y solo aceptará las bien formadas

Dado que su uso es bastante concreto, su instalación y configuración es realmente sencilla, permitiéndonos tener un balanceador de carga operativo en un instante. Es un programa muy pequeño, auditado fácilmente por problemas de seguridad, no se realiza el almacenamiento en caché por lo que cada solicitud se pasa sin cambios a un servidor, presenta algunos problemas con la implementación SSL.

### 3.8.5 NGINX

Es un servidor HTTP, reverse proxy y servidor proxy IMAP/POP3 es lanzado el código fuente bajo una licencia estilo BSD. Es conocido por su estabilidad, gran conjunto de características, configuración simple, y bajo consumo de recursos, este tiene versiones tanto gratuitas con ciertas limitaciones como de pago. Entre las características HTTP nos encontramos: manejo de ficheros estáticos, índices y auto índices, reverse proxy acelerado sin caché y con balanceo de carga y tolerancia a fallos, FastCGI, arquitectura modular y soporte SSL.

Se puede configurar como un equilibrador de carga simple pero potente para mejorar la disponibilidad y eficiencia de sus servidores. En una configuración de equilibrio de carga, nginx actúa como punto de entrada único para una aplicación web distribuida que trabaja en múltiples servidores separados.

Es un balanceador de carga HTTP muy eficiente para distribuir el tráfico a varios servidores de aplicaciones y mejorar el rendimiento, la escalabilidad y la confiabilidad de las aplicaciones web con nginx. Entre los algoritmos de balanceo de carga que se pueden utilizar son round robin, menos conexiones activas y tablas hash

Una de las cualidades principales es la de trabajar junto con diversas tecnologías de desarrollo y lenguajes. La asincronía es su característica fundamental, junto con su rapidez, ya que es un servidor web ligero. Realiza esto sin reducir el amplio rango de funcionalidades del servidor: avanzadas y, sobre todo, muy flexibles, además capaz de adaptarse a muchas situaciones distintas y fue pensado desde el primer día para

atender grandes necesidades de tráfico

### **3.9 Software para Balanceo de Carga por Geolocalización DNS en Linux.**

En esta sección se mencionarán algunos softwares que nos permiten implementar un balanceador de carga sobre un equipo con sistema operativo Linux que hará la función de DNS, el cual nos brindará la posibilidad de balancear la carga entre diferentes centros de datos a través de la geolocalización del usuario que realiza la petición y direccionar esta solicitud al servidor del centro de datos más cercano geográficamente al cliente.

Entre los softwares que nos permiten realizar la función antes mencionada tenemos los siguientes:

#### **3.9.1 Power DNS**

Es principalmente un servidor DNS autoritativo escrito en C++ y distribuido bajo los términos de la licencia GNU/GPLv2 creado por la compañía danesa PowerDNS.com BV, con colaboración de la comunidad del software libre. Funciona en todos los sistemas operativos Linux y los derivados de Unix. Incluye soporte para utilizar diversos motores de bases de datos, desde los sencillos archivos tipo BIND hasta las bases de datos relacionales.

Cuenta tanto con un servidor autoritativo y el repetidor. Si bien la mayoría de los servidores de nombres combinan completamente estas funciones, este las ofrece por separado, pero puede combinar el uso autoritativo y recursivo sin problemas. El servidor autoritativo responde a las preguntas sobre los dominios que conoce, pero no se difundirá en la red para resolver consultas sobre otros dominios una vez que este responde una pregunta, sale de la base de datos y se puede confiar en que tiene autoridad. No hay forma de contaminar el caché o confundir al daemon. El encargado de resolver los nombres, por el contrario, de forma predeterminada no tiene conocimiento de los dominios en sí, pero siempre consultará a otros servidores autorizados para responder a las preguntas que se le hagan.

Entre sus beneficios es que se puede satisfacer las necesidades de las organizaciones pequeñas al ser fácil de instalar, así como para servir grandes volúmenes de consultas en grandes cantidades de dominios. Además, mediante el uso de técnicas de programación inteligente nos ofrece un rendimiento de resolución de dominio muy alto. Un aspecto a considerar es la seguridad que la proporciona mediante el uso de funciones de lenguaje, el código fuente de esta es razonablemente pequeño, lo que facilita la auditoría. De la misma manera, las características de la biblioteca se han utilizado para mitigar los riesgos de desbordamientos de búfer.

Nos otorga muchas estadísticas sobre su funcionamiento, lo que es útil tanto para determinar la escalabilidad de una instalación como para detectar problemas para ellos cuenta como una herramienta conocida como PowerAdmin con la cual se tiene administración de las funciones de este software, así como todos los tipos de zonas y registros DNS.

En las funcionalidades que brinda de geolocalización y poder realizar una distribución de la carga a través de la localización de los clientes lo realiza cuando se recibe una consulta a un servidor de nombre DNS autoritativo, este conoce la dirección del servidor de almacenamiento en caché recursivo que lo consultó y, en función de esta información, puede devolver una respuesta diferente según la dirección de origen. Esto se conoce como GeoDNS o DNS basado en GeoIP, y se usa a menudo para devolver la dirección de un recurso que está más cercano (en red) a la resolución del usuario, esto lo permite gracias al uso de un back-end GeoIP de Aki Tuomi para PowerDNS que hace justamente eso. La información de ubicación en cada país se puede obtener de la base de datos GeoLite de MaxMind esto bajo pago, o se puede optar por crear la suya propia.

El uso del backend Geo se puede utilizar para distribuir consultas de forma global utilizando una tabla de direcciones IP / mapeo de país, muchas de las cuales están disponibles de forma gratuita en línea o pueden adquirirse por una pequeña tarifa. Esto permite que los visitantes se envíen a un servidor cerca de ellos, sin demoras apreciables, ya que de lo contrario se incurriría con un redireccionamiento de nivel de protocolo. Además, de que se puede usar para proporcionar servicio en varios clusters, cualquiera de los cuales puede dejarse de usar fácilmente, por ejemplo, para fines de mantenimiento. Para esta la geolocalización es necesario solo configurar el back-end con unas pocas directivas, básicamente proporcionamos la ruta a los datos de ubicación y a un archivo de configuración YAML que define las zonas geo-habilitadas que ofrecemos en el servidor.

### 3.9.2 BIND9

Es un software de código abierto que permite publicar su información de Sistema de nombres de dominio (DNS) en Internet y resolver consultas DNS para sus usuarios. El nombre BIND significa "Dominio de nombre de Internet Berkeley", porque el software se originó a principios de la década de 1980 en la Universidad de California en Berkeley. Siendo el software DNS más utilizado en Internet, proporcionando una plataforma robusta y estable sobre la cual las organizaciones pueden construir sistemas informáticos distribuidos sabiendo que esos sistemas cumplen totalmente con los estándares DNS publicados [22].

Provee características principales de los componentes del Sistema de Nombres de Dominio, entre los cuales incluyen: Un servidor de sistema de nombres de dominio (named), biblioteca resolutoria de sistema de nombres de dominio y herramientas para verificar la operación adecuada del servidor DNS (bind-utils).

BIND implementa los protocolos DNS. Los protocolos DNS son parte de los estándares básicos de Internet. Especifican el proceso mediante el cual una computadora puede encontrar otra computadora en función de su nombre. La distribución del software BIND contiene todo el software necesario para hacer y responder preguntas del servicio de nombres.

Este software se distribuye en 3 partes importantes que se menciona a continuación:

- 1. Resolver el nombre de dominio:** Es un programa que resuelve preguntas acerca de nombres enviando esas preguntas a servidores apropiados y respondiendo apropiadamente a las respuestas de los servidores para aquello un navegador web utiliza una biblioteca local de resolución en la misma computadora para buscar nombres en el DNS. La resolución dada normalmente reenviará las consultas a una resolución de caché, un servidor o un grupo de servidores en la red dedicada a los servicios de DNS. Estos enviarán consultas a uno o varios servidores autorizados para encontrar la dirección IP para ese nombre DNS.
- 2. Servidor de autoridad de nombre de dominio:** No es más que un servidor DNS autoritativo que responde a las solicitudes de los resolutores, utilizando información sobre los nombres de dominio para los que es autorizada. Puede proporcionar servicios de DNS en Internet instalando este



software en un servidor y proporcionándole información sobre sus nombres de dominio.

- 3. Herramientas:** Se incluyen algunas herramientas de diagnóstico y operativas. Entre ellas la popular herramienta DIG, que no es específica de BIND y se pueden usar con cualquier servidor DNS.

Este software posee características que vale la pena mencionar y que nos ayudan a optar por este entre las que tenemos entre otras las siguientes:

- Es una fuente abierta transparente. Si la organización requiere de alguna funcionalidad que no esté en BIND, se puede modificarla y contribuir con la nueva función a la comunidad enviando esta nueva fuente.
- Ha ido evolucionado continuamente para ser un sistema DNS muy flexible y completo independientemente de cuál sea su aplicación, contando con la mayoría de las características requeridas por las organizaciones.
- Siendo la primera solución, la más antigua y la que se implementa más comúnmente, hay más ingenieros de red que ya están familiarizados con BIND que con cualquier otro sistema. La ayuda está disponible a través de nuestra lista de correo de la comunidad, o puede suscribirse para soporte experto y confidencial a cualquier hora y día.
- Es compatible con la mayoría de distribuciones de Linux entre las que están CentOS, Red Hat Enterprise Linux, Debian, Fedora, FreeBSD, Solaris, Ubuntu, además de también ser manejable en Windows. Las versiones más actualizadas de BIND están siempre disponibles desde el sitio web y servidor ftp. Muchos de los sistemas operativos también ofrecen paquetes BIND para sus usuarios. Se pueden construir con un conjunto diferente de valores predeterminados que la distribución estándar y algunos de ellos agregan un número de versión propio que no se corresponde exactamente con la versión BIND. Para obtener asistencia con la configuración y comprender en general cómo usar BIND, el Manual de referencia administrativa (ARM) de BIND es la herramienta principal, mientras que los usuarios de Windows pueden encontrar útil la explicación de las versiones disponibles para Windows, además que en toda la internet la información sobre esta herramienta es extensa pudiendo encontrar varios libros excelentes en BIND.

## 4 Desarrollo de Escenarios Virtuales Didácticos

En este capítulo se tratará de brindar un mejor entendimiento de la teoría mencionada en este trabajo por medio del desarrollo de escenarios virtuales didácticos, pero principalmente de la tecnología de balanceo de carga y la implementación de esta para beneficio de las comunicaciones hacia los centros de datos.

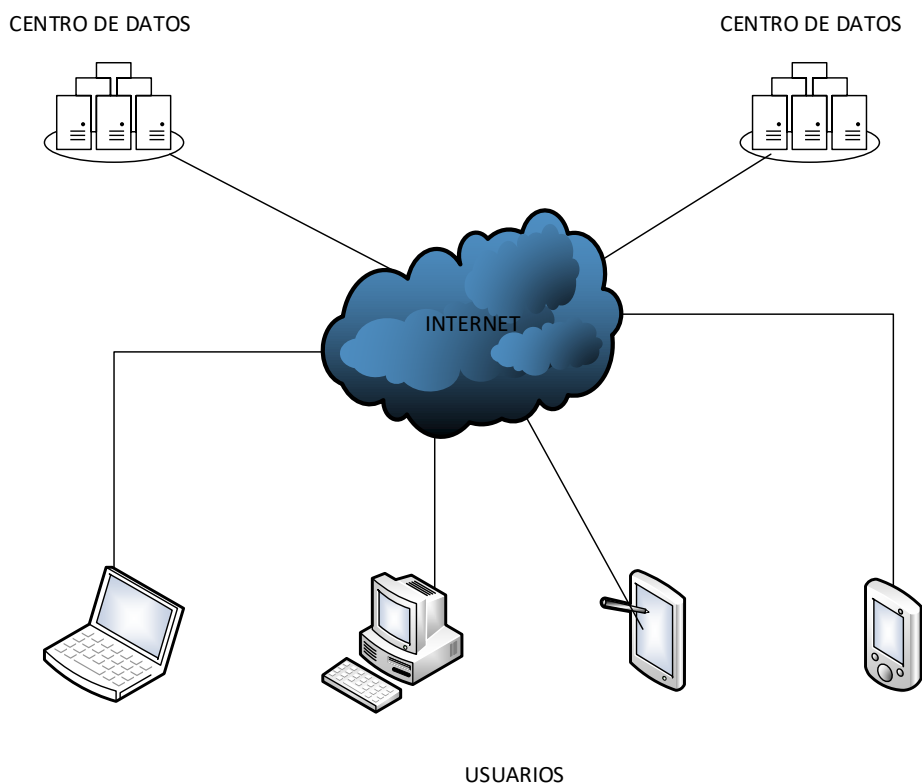


Figura 4.22. Esquema de comunicación a los Centros de Datos

Uno de los principales problemas que surgen es es como mejorar la comunicación hacia los diferentes centros de datos como se visualiza en la Figura 4.22., teniendo como consigna el constante crecimiento de las telecomunicaciones y la gran demanda que los usuarios generan hacia los servicios ofrecidos en la internet , por lo que se desea optimizar el rendimiento en la transmisión y recepción de los datos generada hacia los centros de datos con el fin de poder brindar un mejor servicio, es debido a esto que se plantean dos mecanismos de balanceo de carga que serán detallados en las siguientes secciones los cuales nos permitirán comprender de mejor manera algunas de las soluciones planteadas para solventar esta problemática.

Los escenarios serán construidos con la utilización de la herramienta VNX, gracias a que esta nos otorga la facilidad de desplegar grandes escenarios de redes virtuales dentro de un entorno virtual, además de contar con una simple configuración y

despliegue a través de comandos en un terminal Linux que se encuentran en detalle en la web de esta herramienta, una vez realizado aquello permite la interacción con el escenario mediante la ejecución de comandos dentro de la máquinas virtuales desplegadas.

## 4.1 Balanceo de Carga de Servidores en un Centro de Datos

En esta sección se realizará el diseño de dos escenarios de red virtuales que nos permita comprender la distribución del tráfico entre los distintos servidores que brindan un servicio específico dentro de un centro datos.

Para el desarrollo de los dos escenarios mencionados nos basamos en la comunicación que se genera entre un usuario y un servidor web cuando este intenta ingresar a una página web específica, los centros de datos disponen de 2 servidores web y un balanceador de carga con el cual podemos distribuir el tráfico que surge cuando el usuario realiza peticiones a dicha página. Estos escenarios serán representados por dos centros de datos distintos que estarán constituidos con los mismos elementos, pero ambos se manejarán con distintos algoritmos de balanceo de carga. Para poder gestionar estos escenarios generados a través de la herramienta VNX se requiere ejecutar el siguiente comando dentro de la ruta donde se encuentra el archivo de configuración XML que para el caso del primer centro de datos es CD1.xml mientras que para el segundo es CD2.xml.

### Arranque y Creación del escenario:

```
# sudo vnx -f "nombre del archivo".xml -v -create
```

Para brindar el servicio de la página web se cuenta con 2 servidores web apache2 en cada centro de datos, para poder arrancar el servicio en estos equipos solo basta con ejecutar el siguiente comando en la terminal de Linux ya teniendo arrancado el escenario en VNX.

### Inicio del servicio apache en los servidores web:

```
# sudo vnx -f "nombre del archivo".xml -x start-www
```

Como se hace referencia en el marco teórico de este trabajo existen varias herramientas de software que nos permiten la implementación de un balanceador de carga a través de un sistema operativo Linux, pero de todas estas se decidió por el uso de HAProxy debido a los siguientes aspectos:

- Es un software libre y gratuito, pero a pesar de aquello cuenta con mucha información obtenida a través de manuales, guías de configuración y uso de esta herramienta, además de contar con una gran comunidad que nos permite resolver un sin número de dudas o problemáticas presentadas.
- Nos brinda confiabilidad ya que a pesar de ser gratuito es avalado por un sin número de grandes organizaciones reconocidas tales como Github, Twitter, Tuenti entre otras, en algunas de ellas incluido en productos comerciales como es el caso de Azure y Amazon Web Services, además de todo aquello para en constante actualización otorgando nuevas mejoras en cada una de ellas.
- Nos permite el balanceo sobre cualquier servicio que se ejecute sobre el protocolo TCP y especializado sobre todo en servicios HTTP siendo el más desarrollado por esta herramienta y disponer de soporte específico para este protocolo, debido a que nuestro desarrollo se basa en servicios web es una opción apropiada como balanceador.
- Cuenta con una configuración fácil de comprender y muy detallada sobre las distintas implementaciones que se puede realizar y a pesar de esto cuenta con propiedades muy potentes en dicha configuración que permite establecer arquitecturas de redes complejas.
- Es una herramienta rápida debido a que no requiere grandes requisitos de hardware para su implementación, además de que su arquitectura está optimizada para mover datos lo más rápido posible con las menores operaciones posibles. Como tal, implementa un modelo en capas que ofrece mecanismos de derivación en cada nivel para garantizar que los datos no alcancen niveles más altos cuando no se necesitan. La mayor parte del procesamiento se realiza en el kernel, y HAProxy hace todo lo posible para ayudar al núcleo a hacer el trabajo lo más rápido posible.
- Nos brinda seguridad al poderse implementar SSL en las conexiones establecidas al balanceador, además realiza una inspección de los protocolos

HTTP, SSL o de carga útil si los elementos de protocolo no válidos son rechazados, a menos que se les indique que los acepten de todos modos.

- Es una herramienta muy flexible ya que es de los pocos softwares para balanceo de carga que tiene características más desarrolladas para acomodar su funcionamiento a las necesidades que se requieran para la distribución de tráfico ya que se puede utilizar la mayoría de algoritmos de balanceo de carga que fueron detallados en este trabajo.
- Un aspecto importante de considerar es que monitoriza el tráfico y muestra las estadísticas y el estado de las instancias asociadas en una página web. El estado de los servidores se monitorea continuamente utilizando parámetros por servidor. Esto garantiza que la ruta al servidor sea operativa para el tráfico regular.

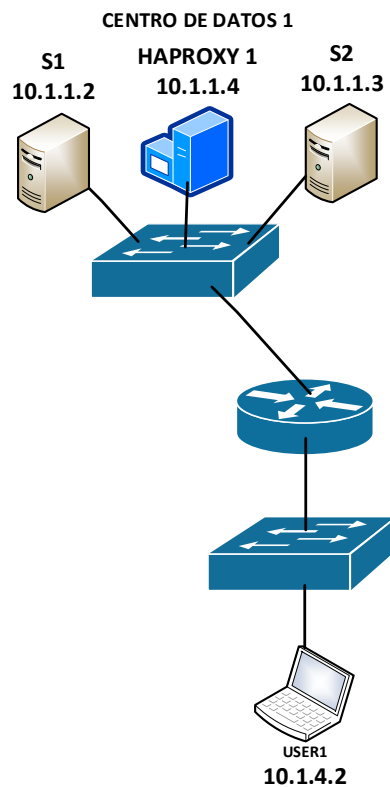


Figura 4.23. Escenario de red del Centro de Datos 1

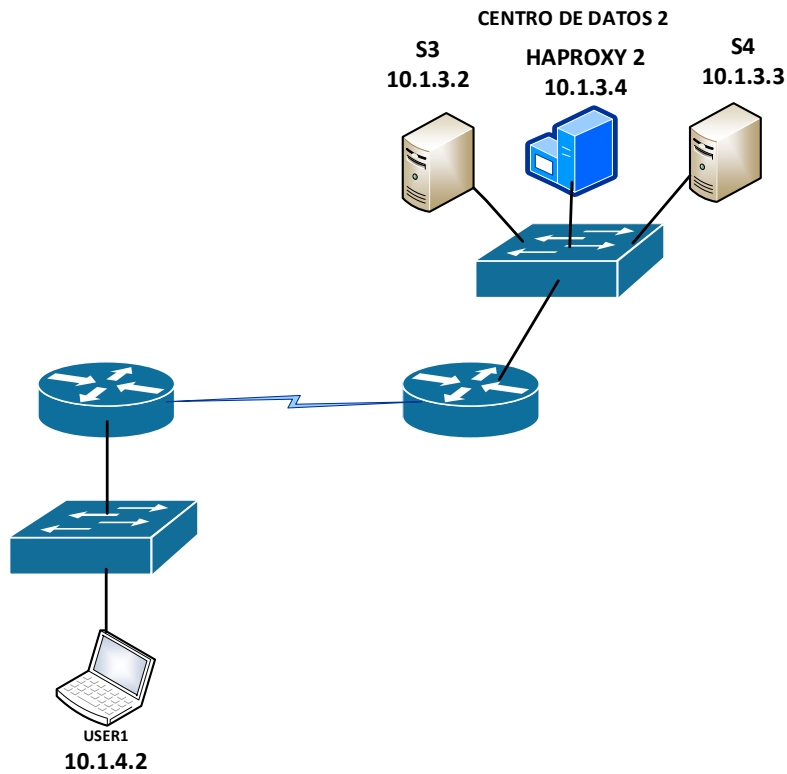


Figura 4.24. Escenario de red del Centro de Datos 2

Al definir la utilización de HAProxy como balanceador de carga, se procede con la instalación de este, una vez realizado aquello se genera un archivo configuración haproxy.cfg ubicado en la ruta /etc/haproxy que nos permitirá establecer los diferentes parámetros para el funcionamiento de esta herramienta, para este escenario se definió la configuración que se visualiza en la Figura 4.25. y en la Figura 4.26, del cual tenemos que considerar y tener claro los siguientes aspectos importantes:

```

GNU nano 2.9.3                                haproxy.cfg
log      global
mode     http
option   httplog
option   dontlognull
timeout  connect 5000
timeout  client 50000
timeout  server 50000
frontend Accesos_HTTP
bind :80
option   forwardfor
stats   uri /haproxy?stats
default_backend Granja_puerto_80

backend Granja_puerto_80
mode     http
balance roundrobin
server S1 10.1.1.2:80 maxconn 400 check
server S2 10.1.1.3:80 maxconn 400 check

```

Figura 4.25. Configuración HAProxy Centro de Datos 1

```
GNU nano 2.9.3 haproxy.cfg

log global
mode http
option httplog
option dontlognull
timeout connect 5000
timeout client 50000
timeout server 50000
frontend Accesos_HTTP
bind :80
option forwardfor
stats uri /haproxy?stats
default_backend Granja_puerto_80

backend Granja_puerto_80
mode http
balance roundrobin
server S3 10.1.3.2:80 maxconn 400 check weight 2
server S4 10.1.3.3:80 maxconn 400 check weight 1
```

Figura 4.26. Configuración HAProxy Centro de Datos 2

El bloque denominado frontend se define una serie sockets en escucha que aceptan peticiones de clientes, así como diversos parámetros relacionados con el tratamiento de las solicitudes entrantes, entre los que se establecen los siguientes:

- **bind:** En este parámetro se detalla la dirección IP del equilibrador de carga con su respectivo puerto asociado como en este caso se manejará la misma IP local del equipo solo se define el puerto que en este caso es el 80 por el cual se receptara el trafico generado por los clientes.
- **stats:** Se define la dirección uri donde se puede ingresar a través del navegador web para comprobar las estadísticas del tráfico generado hacia los servidores que forman parte del sistema, así como el estado de los mismos.
- **default\_backend:** Se direcciona el trafico recibido por el balanceador hacia el sistema de servidores definido en el backend.

El bloque backend se definen los servidores encargados de responder las peticiones del cliente y permitirá analizar la solicitud del cliente para lo cual se destacan los siguientes parámetros:

- **mode:** Se especifica el tipo de tráfico que será analizado por el balanceador de carga y direccionado hacia los servidores para nuestro caso se establece http, con cual se permitirá analizar la solicitud del cliente a nivel de aplicación con lo que se consigue examinar el contenido

de la petición y poder lograr persistencia de sesión en las solicitudes del usuario.

- **balance:** En este parámetro se puede define el algoritmo de balanceo de carga que es utilizado para la distribución de tráfico hacia los servidores que brindan el servicio web. Para el caso del primer centro de datos de la Figura 4.23 se establece un balanceo round robin mientras que para del segundo centro de datos de la Figura 4.24 se usa un algortimo de balanceo de carga round robin ponderado.
- **server:** En este se establecen la IP o nombre de dominio de los servidores del centro de datos asociados con su respectivo puerto a los cuales se les distribuirá el tráfico generado por los usuarios cuando realicen peticiones a la página web.
- **maxconn:** Determina el máximo número de conexiones simultaneas que se pueden establecer hacia cada servidor.
- **check:** necesario para que el balanceador pueda realizar un chequeo constante del estado del servidor.
- **weight:** Se establece el peso o ponderación que se le otorga a un servidor específico para distribuir una carga de trabajo determinada sobre este. A través del uso de este parámetro es factible la utilización de algoritmos basados en peso como el round robin ponderado y menos conexiones activas ponderado.

Para poder iniciar el servicio de HAProxy así como la configuración ya establecida de estos escenarios solo es necesario ejecutar el siguiente comando en la terminal de Linux:

#### **Inicio del servicio HAProxy en los balanceadores de carga:**

```
# sudo vnx -f "nombre del archivo".xml -x haproxy
```

Una vez ya teniendo iniciados los servicios respectivos en cada escenario, procederemos a comprobar el correcto funcionamiento del balanceo de carga en cada centro de datos.



➤ **Prueba de funcionamiento en el Centro de datos 1**

Para comprobar el correcto funcionamiento del balanceo en el primer centro de datos ingresaremos a la pagina web ofrecida por los servidores del primer centro de datos como se observa en la Figura 4.27 y comprobaremos que la distribución del tráfico se está gestionando mediante el algoritmo de balanceo round robin.



Figura 4.27. Ingreso Pagina web Centro de Datos 1

Por medio de la dirección IP del balanceador y la uri establecida en la configuración de HAProxy, podemos verificar el estado de la distribución de tráfico establecido para los servidores que gestiona el balanceador de carga.

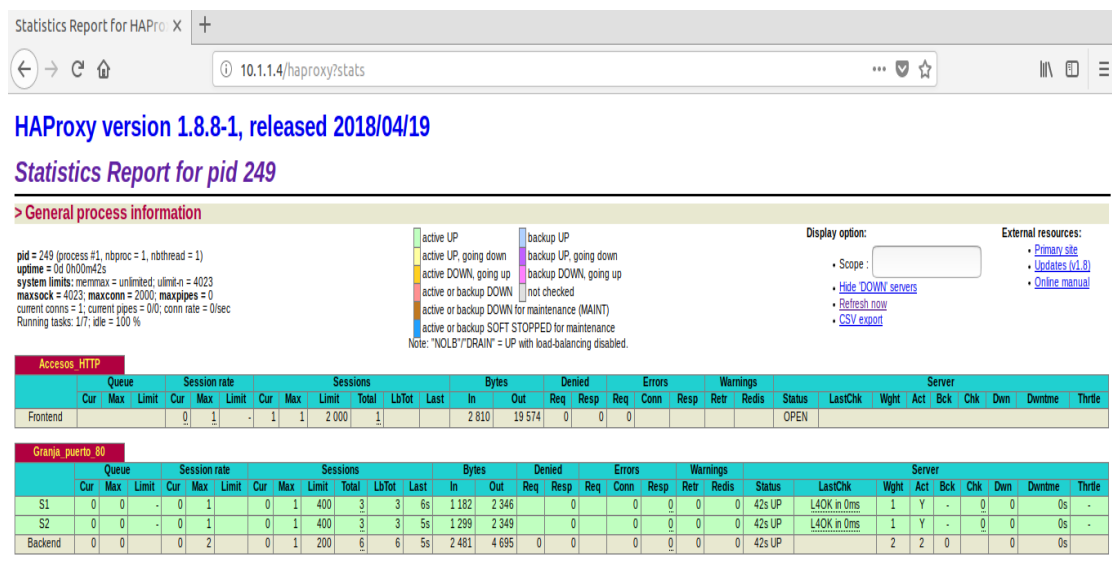


Figura 4.28. Estadísticas del Balanceo en el Centro de Datos 1

Como visualiza en la Figura 4.28. se comprueba que el balanceo de Carga está funcionando según los parámetros establecidos al distribuir la carga entre uno y otro servidor según el algoritmo round robin de manera secuencial.

### ➤ Prueba de funcionamiento en el Centro de datos 2

Al igual que el primer centro de datos en este se comprobará el funcionamiento correcto del balanceo al ingresar a la página web de la Figura 4.29., pero con un algoritmo basado en round robin ponderado.



Figura 4.29. Ingreso Pagina web Centro de Datos 2

De manera similar se ingresa a las estadísticas del balanceo del segundo centro de datos modificando la dirección IP del balanceador de carga en este caso y manteniendo la misma uri como se observa en la Figura 4.30. la misma que nos permite concluir que la distribución de carga definida para este centro de datos está funcionando según lo requerido por el algoritmo round robin ponderado dando el doble de carga a un servidor que a otro.

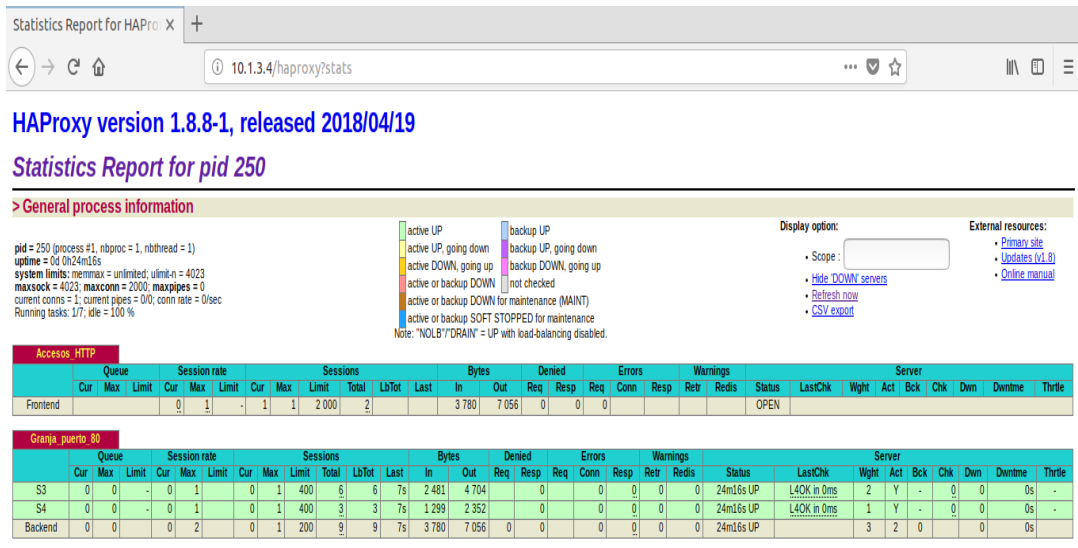


Figura 4.30. Estadísticas del Balanceo en el Centro de Datos 2

## 4.2 Balanceo de Carga entre Centros de Datos por Geolocalización a través de DNS

En este subcapítulo procederemos al desarrollo de un escenario virtual en cual se realizará el balanceo de carga entre los dos centros de datos que ya fueron definidos en la sección anterior, esto se realizara a través del mecanismo de geolocalización por DNS como se visualiza en la Figura 4.31. por medio del cual realizaremos la distribución del tráfico hacia uno u otro centro de datos dependiendo de cual se encuentre más cercano a la localización del usuario que realiza la solicitud a la página web que está almacenada en los servidores web de ambos centros de datos.

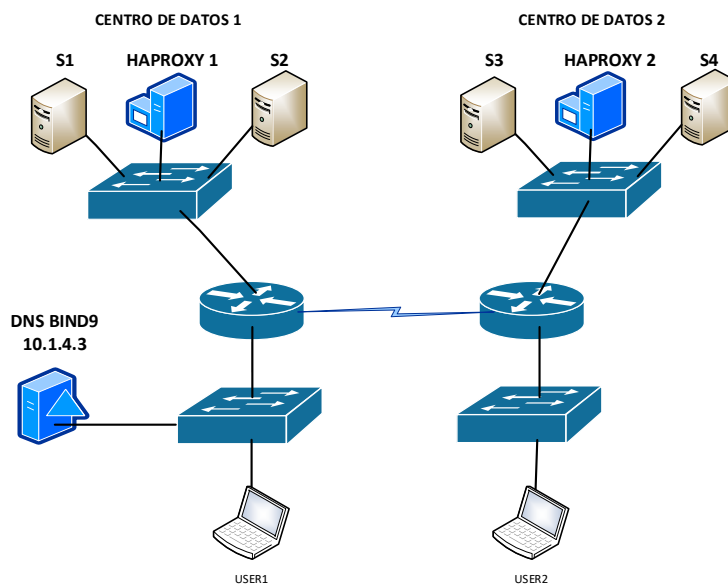


Figura 4.31. Escenario de dos Centros de Datos con Geolocalización DNS

Como se detalla en la teoría de este trabajo existen pocas herramientas de software que nos permitan la implementación de realizar un balance de carga entre centros de datos basándonos en la geolocalización a través de DNS el cual es instalado en un sistema operativo Linux, debido a que de las dos opciones mencionadas POWER DNS no ofrece un soporte adecuado ni una información adecuada y detallada de esta información caso contrario que sucede con BIND9 que es el software DNS más utilizado en Internet, proporcionando una plataforma robusta y estable con información muy amplia y detallada de las funcionalidades que la misma ofrece independientemente de cuál sea nuestra aplicación, es probable que BIND tenga las características requeridas, además de que se encuentra con una comunidad muy amplia para resolución de cualquier duda o problemática encontrada es por eso que opto por la utilización de esta para la realización de este escenario virtual.

Una vez definido el software para la gestión de DNS se procede con la instalación de este y adicionalmente se inician los servicios de apache2 y HAProxy indicados en la sección anterior, ya realizado aquello se debe tener claro la configuración que se establecio en los distintos archivos que se generan cuando se instala Bind9 con el fin de poder poner en funcionamiento el balanceo de carga entre estos centros a través de la ubicación del usuario que realizada las solicitudes, por ellos a continuación se detalla la configuración establecida en estos archivos los cuales se encuentran en la ruta /etc/bind, así como una breve descripción de los parámetros mas importantes a tener en cuenta cuando se realiza este tipo de mecanismo de balanceo.

A screenshot of a terminal window titled "DNS - console #1". The window shows the GNU nano 2.9.3 editor editing the file named.conf. The content of the file is as follows:

```
include "/etc/bind/named.conf.default-zones";
acl "lan1" {
  10.1.4.0/24;
  localhost;
};
acl "lan2" {
  10.1.5.0/24;
  localhost;
};
```

Figura 4.32. Archivo named.conf del DNS Bind

En el archivo de configuración named.conf establecemos las ACL respectivas en las cuales definimos las redes de la lista de clientes a los que permitiremos efectuar consultas DNS sobre nuestros servidores web, adicional indicamos el directorio en el cual serán requeridas estas ACL, todo esto se puede observar mejor en la Figura 4.32.

```

GNU nano 2.9.3 db.prueba.net.lanone
;
; BIND data file for local loopback interface
;
$TTL 604800
@ IN SOA prueba.net. root.prueba.net. (
    2 ; Serial
    604800 ; Refresh
    86400 ; Retry
    2419200 ; Expire
    604800 ) ; Negative Cache TTL
;
@ IN NS prueba.net.
prueba.net. IN A 10.1.4.3
DNS.prueba.net. IN A 10.1.4.3
host.prueba.net. IN A 10.1.4.2
www.prueba.net. IN A 10.1.1.4
www.prueba.net. IN A 10.1.3.4

```

Figura 4.33. Archivo de resolución de nombres del DNS Bind

Luego debemos proceder a crear los archivos que definirán las zonas que estarán dedicadas a la resolución de nombres es decir tendrán el conjunto de entradas de registros de recursos (RR), en nuestro para las dos redes creadas se contara con los archivos db.pruebas.net.lanone y db.pruebas.net.lantwo., ambos archivos contaran con la misma configuración debido a que además del balanceo entre los centro de datos también se desea que se disponga de alta disponibilidad para que el servicio sea continuo para usuarios que gestionen sus peticiones a un centro de datos el cual tengas sus servidores web no operativos. Como se visualiza en la Figura 4.33 se determinan el nombre de mi dominio que para este escenario se estableció en prueba.net, así como también las FQDN para los distintos registros de host el propósito de este archivo de zona es principalmente asignar nombres de host y servicios a direcciones IP específicas, para nuestra página web se define el FQDN www.prueba.net el cual estará registrado con las direcciones IP de los balanceadores de carga de ambos centros de datos como se detalla en la Figura 4.33.

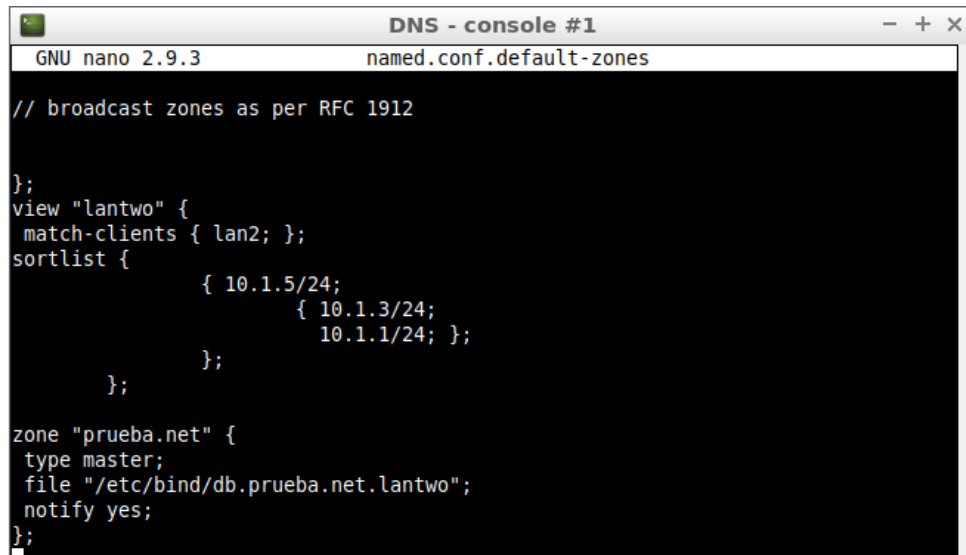
```

GNU nano 2.9.3 named.conf.default-zones
// prime the server with knowledge of the root servers
view "lanone" {
    match-clients { lan1; };
    sortlist {
        { 10.1.4/24;
            { 10.1.1/24;
                10.1.3/24; };
        };
    };
};

zone "prueba.net" {
    type master;
    file "/etc/bind/db.prueba.net.lanone";
    notify yes;
};

```

Figura 4.34. Archivo de Zonas DNS Bind para LAN1



```
GNU nano 2.9.3          named.conf.default-zones
// broadcast zones as per RFC 1912
};
view "lantwo" {
  match-clients { lan2; };
  sortlist {
    { 10.1.5/24;
      { 10.1.3/24;
        10.1.1/24; };
    };
  };
};
zone "prueba.net" {
  type master;
  file "/etc/bind/db.prueba.net.lantwo";
  notify yes;
};
```

Figura 4.35. Archivo de Zonas DNS Bind para LAN2

En el archivo `named.conf.default-zones` se define el funcionamiento de nuestro servidor DNS en donde detallaremos los diversos parámetros y configuración realizada para hacer posible el balanceo de carga por geolocalización a través de este equipo, como se visualiza en la Figura 4.35 y Figura 4.36. Es necesario considerar los siguientes aspectos importantes que se establecieron en este archivo:

- **view:** Es la característica perfecta de Bind9 para implementar el mecanismo descrito anteriormente. Las vistas nos permiten definir diferentes configuraciones "virtuales" dentro del mismo servidor, y también para especificar quién debe ver qué configuración, la clave de la utilización de esta es que el servidor usa la dirección IP de origen de la consulta para seleccionar cuál de las vistas definidas debe consultarse para responder. Esto significa que debe de asociar una vista con una o más direcciones IP de origen por lo que cuando un cliente envía una consulta, se utiliza la vista que coincide con su dirección IP. Como el caso definido en nuestro escenario se manejan 2 redes para los usuarios que tendrán acceso a los servicios web se procedió con la creación de dos vistas. La sentencia `view` empaqueta una lista de acceso que controla quién ve la vista y algunas opciones que se aplican a todas las zonas en la vista [25].
- **match-clients:** En este parámetro se ingresan el conjunto de direcciones IP que podrán tener acceso a las zonas definidas dentro de la vista para nuestro caso este conjunto de direcciones fueron establecidas en las ACL descritas con anterioridad con los nombres de `lan1` y `lan2`.
- **sortlist:** En este se define el orden el que serán usados los registros de recursos

definidos en los archivos de zona creados para la resolución directa de nombres , esto nos permite definir un orden de prioridad que para nuestro caso es de gran ayuda, debido a que al contar con dos centros de datos podemos designar una alta disponibilidad entre ambos esto provocara que cuando un usuario realice la consulta al servidor web del centro de datos mas cercano y si los equipos que brindan el servicio no se encuentran funcionando se asignara al siguiente centro de datos que se encuentre operativo registrado en el archivo de zonas directas.

- **Zone:** Define una zona de registro de recursos, aquí se ingresa la ruta donde se encuentra el archivo de las zonas de resolución directa de nombres.

Para poder iniciar el servicio de Bind9, así como la configuración ya establecida de este escenario solo es necesario ejecutar el siguiente comando en la terminal de Linux:

### **Inicio del servicio Bind9 en el servidor DNS:**

```
# sudo vnx -f "nombre del archivo".xml -x bind
```

Para el funcionamiento correcto de este escenario es necesario realizar el registro tanto del nombre de dominio como de la IP del servidor DNS en el usuario y en el mismo servidor por lo que para aquello solo es necesario la ejecución del siguiente comando:

### **Registro de nombre de dominio e IP del DNS en los equipos:**

```
# sudo vnx -f "nombre del archivo".xml -x dns
```

Después del inicio de todos los servicios necesarios para el funcionamiento de este escenario lo siguiente es verificar que esté funcionando de manera adecuada según los requerimientos dados. Para comprobar el correcto funcionamiento del balanceo de carga ingresaremos a la dirección de la página web ofrecida por los servidores de ambos centros de datos, las solicitudes serán generadas por dos usuarios que pertenecen tanto a la lan1 (coit) como a la lan2 (PC2), así comprobaremos que la distribución del tráfico se está gestionando mediante el mecanismo planteado.

```
coit@coit-vm: ~
File Edit Tabs Help
coit@coit-vm:~$
coit@coit-vm:~$
coit@coit-vm:~$ curl -l www.prueba.net
<html>
<font face="Arial" color="blue">
<center>
<h3>Welcome to</h3>
<h1>S1 virtual machine web server CD1</h1>
<br>
<pre>
000000      0000 00000      000 0000000  00000
`888.      .8' `888b.      `8' `8888      d8'
 `888.      .8'  8 `88b.      8      Y888..8P
  `888.      .8'  8 `88b.      8      `8888'
   `888.8'      8      `88b.8      .8PY888.
    `888'      8      `888      d8' `888b
     `8'      o8o      `8      o888o  o88888o
</pre>
<h3>Virtual Network over Linux</h3>
```

Figura 4.37. Ingreso desde la lan1 a la página web.

```
PC2 - console #1
root@PC2:~#
root@PC2:~#
root@PC2:~# curl -l www.prueba.net
<html>
<font face="Arial" color="blue">
<center>
<h3>Welcome to</h3>
<h1>S3 virtual machine web server CD2 </h1>
<br>
<pre>
000000      0000 00000      000 0000000  00000
`888.      .8' `888b.      `8' `8888      d8'
 `888.      .8'  8 `88b.      8      Y888..8P
  `888.      .8'  8 `88b.      8      `8888'
   `888.8'      8      `88b.8      .8PY888.
    `888'      8      `888      d8' `888b
     `8'      o8o      `8      o888o  o88888o
</pre>
```

Figura 4.38. Ingreso desde la lan2 a la página web.

Como se observa en la Figura 4.37 y Figura 4.38. el balanceo de carga entre los dos centros de datos por geolocalización DNS está funcionando según lo establecido ya que tanto el usuario de la lan1 como de la lan2 están siendo direccionados hacia el servidor web del centro de datos mas cercanos a la localización de los mismos.



## 5 Conclusiones y Líneas Futuras de Investigación

Después de haber efectuado la descripción, el desarrollo de escenarios y el respectivo análisis de los temas abarcados en las secciones anteriores, en el presente capítulo se explican las conclusiones obtenidas a través del estudio y realización de escenarios virtuales didácticos, sobre los conocimientos de centros de datos y balanceo de carga presentadas en este trabajo. Además, se plantean posibles líneas de investigación futuras, donde se pueda profundizar en ciertos puntos que quedan pendientes después de la realización del trabajo.

### 5.1 Conclusiones

Este trabajo se ha finalizado cumpliendo con los objetivos planteados inicialmente, dando a conocer las funcionalidades y beneficios que nos otorga la tecnología de balanceo carga para la gestión del tráfico en centros de datos que hoy en día es de gran importancia tomando en cuenta el enorme crecimiento de las redes de comunicación producido por la enorme demanda de los servicios ofrecidos por la internet.

A través del uso de la herramienta VNX se facilitó el desarrollo de los escenarios propuestos durante este trabajo, debido a que no fue necesario de enormes requerimientos de hardware como de software, además de que nos permite implementar soluciones que respondan a problemas reales, todo esto nos hace destacar la gran importancia de la virtualización en el tema de gestión de redes de datos.

Por medio de la realización del primer escenario el cual se basa en el balanceo de carga de servidores en un centro de datos se logro que se comprenda de mejor manera la técnicas y algoritmos de distribución de tráfico que se pueden emplear según los requerimientos planteados en cada organización, mejorando el rendimiento de estas puesto que se realiza una mejor administración de la carga de trabajo con lo que se optimiza el uso de los recursos que ofrece cada componente computacional el cual brinda los distintos servicios en un centro de datos, además de otorgar aspectos de seguridad muy importantes como lo es la prevención de ataques DoS.

Con la realización del segundo escenario el cual se basa en el balanceo de carga entre centros de datos por medio de geolocalización DNS se dio a conocer como a través de

este mecanismo se maximiza significativamente el uso de la infraestructura geográficamente dispersa, aumentando el rendimiento al minimizar la distancia entre los centros de datos y los solicitantes con lo que se consigue que los tiempos de carga de páginas web sean más rápidos en diferentes regiones debido a que el servidor más cercano es el que atiende los requerimientos, incluso a través de este mecanismo se puede realizar gestiones específicas (actualización, personalizar los contenidos, publicidad, etc.) para un conjunto selecto de clientes.

Para el desarrollo de estos escenarios virtuales decidió por la implementación de un sistema automatizado con la Herramienta VNX con lo cual a través de una simple ejecución de ciertos comandos se despliega y se pone en funcionamiento los mecanismos de balanceo propuestos en este trabajo con lo cual se logra que las personas interesadas en conocer más sobre el uso de estas tecnologías puedan gestionar de manera sencilla los recursos computacionales centrándose básicamente en comprender los aspectos más importantes de la misma y no realizando las gestiones tediosas de proceso de instalación y configuración de los componentes que forman parte de los escenarios planteados.

Durante la realización de este trabajo surgieron una serie de problemáticas entre las que se destacan la falta de herramientas de software libre en Linux que permitan realizar balanceo a través de geolocalización DNS, por lo que esto conlleva a que solo se plantee el uso de Bind9 para esta gestión además de que la mayoría de información relevante de como poder realizar la configuración para poner en funcionamiento este mecanismo se basa principalmente en este software.

Dentro de la realización de los escenarios virtuales el aspecto que más conlleva tiempo fue el uso de la herramienta VNX, debido a que no se estaba familiarizado con esta por lo que se procedió con la lectura de documentación y práctica con ejemplos, además se presentaron algunos problemas con la administración de directorios necesarios para la ejecución automática de los servicios utilizados para la implementación de los mecanismos de balanceo de carga planteados.

## **5.2 Líneas Futuras de Investigación**

En el proceso de realización de este trabajo surge un tema de estudio adicional muy interesante como es el caso de los cluster DNS que no es más que un grupo de

servidores DNS que comparten registros entre sí, con los se puede realizar funcionalidades mucho mas completas en los centros de datos al utilizarlos en conjunto las técnicas y algoritmos de balanceo de carga presentadas en este trabajo debido a que a nuestros escenarios propuestos no solo que les podemos brindar una distribución correcta de la carga de trabajo sino que se pueden implementar mecanismos de redundancia muy fiables lo que con lleva a poseer un sistema con alta disponibilidad lo que mejora rotundamente las prestaciones en las redes de comunicación brindando de esta manera una mejor experiencia en los servicios ofrecidos al usuario y las organizaciones.

Una de las alternativas a realizar a futuro es el robustecimiento de estos escenarios aumentando el numero de usuarios que realizan solicitudes a los servidores, así como también realizando el balanceo de carga de otro tipo de servicios como FTP o sistemas de base datos, esto nos permitirá tener un mejor panorama del uso algoritmos de distribución de trafico que no fueron realizados en los escenarios planteados tales como menos conexiones activas y tablas hash los cuales son mencionados en el marco teórico de este trabajo.

Un aspecto importante a considerar para trabajos futuros es el de implementar diferentes pruebas de rendimiento sobre estos escenarios que nos permitan tener una clara perspectiva de los algoritmos y técnicas de balanceo de carga más apropiados para los diferentes escenarios que se deseen plantear más adelante.

## Bibliografía

- [1] Edwin Homero Yaselga, “Diseño Del Centro De Datos Para Petroecuador En El Edificio Matriz En Base Al Estándar TIA-942-2,” Junio 2013.
- [2] Jorge Maldonado Mahuad, “Diseño Del Centro De Datos Para Basado En Estándares. Caso Práctico: Diseño Del Centro De Datos Del Colegio Latinoamericano,” Mayo 2010
- [3] Jorge Lastras Hernansanz, Javier Lázaro Requejo y Jonatan David Mirón García, “Arquitecturas De Red Para Servicios En Cloud Computing,” 2008/2009
- [4] Hugo Paltán Orellana, “El Desarrollo De Estándares Y Procedimientos Para La Creación De Un Data Center En La UPSE,” Octubre 2013.
- [5] Liliana Raquel Castillo Devoto, “Método Para El Manejo Del Balanceo De Carga En Sistemas De Cómputo Distribuido De Alto Desempeño,” Noviembre 2008.
- [6] Andrea Mesa Múnera, “Diseño De Infraestructura De Telecomunicaciones Para Un Data Center,” 2009.
- [7] Gabriel Genaro Vera Mite, “Diseño e Implementación De Un Portal Web Para El Consejo Nacional Electoral Con El Fin De Ayudar En La Capacitación A Los Ciudadanos Del Ecuador Acerca Del Código De La Democracia, Enfocado En La Infraestructura Tecnológica, Implementando El Esquema De Red Y Servicios De Balanceo De Carga Usando Herramientas Open Source Para Los Ambientes De Desarrollo Y Producción Del Portal Web Abc De La Democracia,” 2016.
- [8] Bustos Burbano Andrés Gustavo, “Configuración De Un Cluster De Alta Disponibilidad Y Balanceo De Carga En Linux Para Satisfacer Gran Demanda Web Y Servicios De Resolución De Nombres,” Marzo 2007.
- [9] Christian Kuster, Joaquín Oldán, “Balanceo de carga en redes IPv4, un enfoque de Redes Definidas por Software,” 2015.
- [10] Daniel Antonio Eugenin Morales, “Integración De Aplicaciones De Software Libre Aplicado A Un Cluster De Alta Disponibilidad Y Balanceo De Carga De Servidores Proxy,” 2005.
- [11] Franco Bocchio, “Modelo Cloud Computing como Alternativa para Escalabilidad y Recuperación de Desastres,” 2014.
- [12] Anupama, “Load Balancing The Network Traffic In The Nth Mode Of Iptables,” 2016.
- [13] YU SHENGSHENG, YANG LIHUI, LU SONG, ZHOU JINGLI, “Least-

- Connection Algorithm based on variable weight for multimedia transmission,” pp. 1–5.
- [14] Yenedys Yurys Diaz Perez, Hector Andres Hoyos Millan, “Técnicas Para Administrar Y Gestionar Redes,” 2008.
  - [15] Dr. Mustafa ElGili Mustafa, “Load Balancing Algorithms Round-Robin (Rr), Leastconnection, And Least Loaded Efficiency,” , no.1, pp. 1–5, 2017.
  - [16] Carlos Polivio Zumba Vásquez, “Diseño Y Desarrollo De Un Prototipo De Un Cluster En Linux De Alta Disponibilidad Para Satisfacer La Demanda De Acceso Web En La Carrera De Ingeniería En Sistemas Computacionales Y El Balanceo De Carga De Los Servicios,” 2011.
  - [17] Argha Sen, “Improving Performance Of Clusters Using Load Balancing Algorithms,” May 2011.
  - [18] Abraham Jaramillo Garófalo, “Configuración, Optimización Y Evaluación De Un Servidor De Alta Disponibilidad Con Equilibrado De Carga,” 2015 - 2016.
  - [19] Virginia Martín-Rubio Pascual, Antonio Fuentes Bermejo, “Estudio E Implantación De Un Sistema De Alta Disponibilidad En Rediris,” pp. 1–10, Marzo 2009.
  - [20] Félix Ortigosa Martínez, “Implementación De Un Servidor Web Virtual Balanceador De Carga Basado En Linux,” Agosto 2004.
  - [21] “HAProxy Starter Guide”, [En línea] <https://cbonte.github.io/haproxy-dconv/1.9/intro.html>.
  - [22] “El sistema de nombres de dominio: Bind 9.2.1”, [En línea] <https://blyx.com/public/docs/bind.html>.
  - [23] John Hawley, “GeoDNS—Geographically-aware, protocol-agnostic load balancing at the DNS level,” pp. 1–10, July 13th–17th, 2009.
  - [24] Juha Tawaststjerna, “Content Distribution Networks and GeoDNS load balancing,” Toukokuu 2014.
  - [25] Waldner, “Geolocation-aware DNS with Bind”, [En línea] <https://blyx.com/public/docs/bind.html>, February 2010
  - [26] “VNX”, [En línea] [https://web.dit.upm.es/vnxwiki/index.php/Main\\_Page](https://web.dit.upm.es/vnxwiki/index.php/Main_Page)
  - [27] Repositorio de los archivos de configuración de los escenarios virtuales creados en VNX para este trabajo, [En línea] <https://github.com/sanmor16/Mecanismos-de-Balanceo-de-Carga-Vnx>