OPERACIÓN CONTINUA EN LAS REDES (LOAD BALANCING: CONCEPTO, GESTION Y MONITOREO)

DORADO RUIZ, HECTOR JOAQUIN
HUERTAS FIGUEROA, PEDRO LUIS

UNIVERSIDAD TECNOLOGICA DE BOLIVAR
FACULTAD DE INGENIERIA DE SISTEMAS
CARTAGENA DE INDIAS
2008

OPERACIÓN CONTINUA EN LAS REDES (LOAD BALANCING: CONCEPTO, GESTION Y MONITOREO)

DORADO RUIZ, HECTOR JOAQUIN

HUERTAS FIGUEROA, PEDRO LUIS

MONOGRAFIA PRESENTADA PARA OPTAR POR EL TITULO DE INGENIERO DE SISTEMAS

Director

ISAAC ZUÑIGA SILGADO

Ingeniero de Sistemas

UNIVERSIDAD TECNOLOGOCA DE BOLIVAR

FACULTAD DE INGENIERIA DE SISTEMAS

CARTAGENA

2008.

Nota de aceptación

_____

_____

_____

_____

_____

_____

Firma Presidente del Jurado

_____

Firma del Jurado

_____

Firma del jurado

Cartagena de Indias, D. T. y C., Julio de 2008

Cartagena de Indias, D. T. y C., Julio 31 de 2008

Señores:
Comité Facultad Ingeniería de Sistemas
Universidad Tecnológica de Bolívar
La ciudad.

De la manera mas atenta, nos permitimos presentar a su consideración y aprobación, el trabajo de grado titulado: "**OPERACIÓN CONTINUA EN LAS REDES (LOAD BALANCING: CONCEPTO, GESTION Y MONITOREO)".**

Esperamos que el presente trabajo se ajuste a las expectativas y criterios evaluativos de la Universidad para los trabajos de grado.

Agradeciendo de antemano su colaboración.

Cordialmente,

HECTOR JOAQUIN DORADO RUIZ	**PEDRO LUIS HUERTAS FIGUEROA**
CC: 92.532.917 de Sincelejo	**CC: 1.050.947.699 de Turbaco**

Cartagena de Indias, D. T. y C., Julio 31 de 2008

Señores:
Comité Facultad Ingeniería de Sistemas
Universidad Tecnológica de Bolívar
La ciudad.

A través de la presente me permito entregarle la monografía titulada: "**OPERACIÓN CONTINUA EN LAS REDES (LOAD BALANCING: CONCEPTO, GESTION Y MONITOREO)"**, para su estudio y evaluación la cual fue elaborada por los estudiantes **HECTOR JOAQUIN DORADO RUIZ y PEDRO LUIS HUERTAS FIGUEROA** de los cuales acepto ser su director.

Atentamente,

_____
Ing. ISAAC ZUÑIGA SILGADO

Cartagena de Indias, D. T. y C., Julio 31 de 2008

Señores:
Comité Facultad Ingeniería de Sistemas
Comité de Evaluación de Proyectos
La ciudad.

Estimados señores:

Con el mayor agrado me dirijo a ustedes para poner a consideración el trabajo final titulado: "**OPERACIÓN CONTINUA EN LAS REDES (LOAD BALANCING: CONCEPTO, GESTION Y MONITOREO)".** El cual fue llevado a cabo por los estudiantes **HECTOR JOAQUIN DORADO RUIZ y PEDRO LUIS HUERTAS FIGUEROA**, bajo mi orientación como asesor.

Agradeciendo su amable atención.

Cordialmente,

_____
Ing. ISAAC ZUÑIGA SILGADO

AUTORIZACIÓN

Cartagena de Indias, D. T: y C., Julio 31 de 2008

Yo, **HECTOR JOAQUIN DORADO RUIZ**, Identificado con el numero de cedula 92.532.917 de Sincelejo, autorizo a la Universidad Tecnológica de Bolívar para hacer uso de mi trabajo de monografía y publicarlo en el catalogo Online de la biblioteca.

_____

HECTOR JOAQUIN DORADO RUIZ

AUTORIZACIÓN

Cartagena de Indias, D. T: y C., Julio 31 de 2008

Yo, **PEDRO LUIS HUERTAS FIGUEROA**, Identificado con el numero de cedula 1.050.947.699 de Turbaco, autorizo a la Universidad Tecnológica de Bolívar para hacer uso de mi trabajo de monografía y publicarlo en el catalogo Online de la biblioteca.

_____

PEDRO LUIS HUERTAS FIGUEROA

A Dios por darme la vida y ser mi guía en todo este proceso que sirve como cimiento para terminar de forjar mi futuro.

A mis padres Mary y Ever, quienes con mucho esfuerzo pusieron todo su empeño para poder cumplir este sueño que se esta haciendo realidad.

A mi familia, amigos y compañeros de estudio de los cuales tuve un muy buen apoyo y me brindaron consejos para poder seguir adelante y surcar todos los obstáculos que se colocaban en mi camino.

Héctor J. Dorado Ruiz

Quiero dedicar este logro principalmente a Dios
Por haberme acompañado siempre durante toda esta carrera

A mis padres Carlos y Fanny, por apoyarme y confiar en mí en todo momento, y a toda mi familia por darme su apoyo siempre

A los amigos que encontré durante todo este proceso y los que he tenido desde siempre por brindarme su apoyo siempre que los necesite.

PEDRO LUIS HUERTAS FIGUEROA

AGRADECIMIENTOS

Agradecemos principalmente LA UNIVERSIDAD TECNOLOGICA DE BOLIVAR y a todo el cuerpo de docentes en general, por brindarnos las herramientas de apoyo necesario en nuestro desarrollo académico e investigativo.

Especialmente a nuestro asesor ISAAC ZUÑIGA SILGADO, por la orientación brindada durante la investigación y por sus incentivos a profundizar la temática y obtener de esta manera las bases necesarias para llevar a cabo esta monografía.

Por ultimo queremos agradecer a la FUNDACION UNIVERSITARIA TECNOLOGICO COMFENALCO y todo su cuerpo de docentes por estar presente desde los inicios y hasta el final de esta carrera.

CONTENIDO

CAPÍTULO 2

CAPÍTULO 3

RESUMEN

El presente documento es el resultado de una amplia investigación a cerca de la operación continua de redes, principalmente utilizando el balanceo de cargas. Además de otro tipo de herramientas y aspectos de igual importancia como lo son la redundancia y loas aspectos físicos de la instalación de la red.

En el primer capitulo básicamente se definen, el concepto de balanceo de cargas y sus características, así como también se describen cada uno de los puntos donde se puede aplicar, también allí se encuentran clasificados los diferentes métodos utilizados para realizar el balanceo de cargas tales como: DNS Round robín, LB Round robin, Weigthed LB, Dynamic Weighted LB, Sticky sessions, y la combinación de Dynamic Weighted-Sticky Session.

En el siguiente capitulo se dan a conocer herramientas para hacer balance de cargas; existen software y también dispositivos con estas características, se muestran detalladamente cada una de sus funcionalidades.

En el capitulo 3, se define el concepto de operación continua en las redes, su importancia, se muestra también la mayoría de los aspectos físicos que debemos tener en cuenta en nuestro centro de computo, para que este tenga un buen funcionamiento y lograr su constante operación.

INTRODUCCION

Las aplicaciones de voz, datos y video lideran las cuestiones relacionadas con la operación continua en las redes. Las diferentes compañías deben determinar sus necesidades de red y establecer de manera adecuada los niveles de servicio, los diseños y monitorización de las capacidades para conducir esas necesidades a buen término.

Load Balancing o balanceo de carga, es una herramienta que desde sus inicios ha sido utilizada como solución a la congestión de la red y en primer lugar para garantizar la operación continua de las mismas. Este es un concepto que se puede aplicar a diversos componentes dentro de una red, tales como: el servidor global, servidores generales, firewall, etc.

En este documento se reconocen e identifican las necesidades de hacer balanceo de carga en cada uno de los puntos críticos de una red, lo cual ayuda a mitigar el tráfico entre ellos y garantizar su constante operación. Pero debemos tener en cuenta que el balanceo de carga no es suficiente para mantener una red en continuo funcionamiento, también se señalan otro tipo de conceptos como la redundancia, que se puede aplicar en diversos campos, tales como: redundancia en enlaces, en equipos, y en aplicaciones.

Además se dan a conocer características físicas que se deben tener en cuenta en un centro de cómputo, que acompañen el balanceo de carga, para así brindar un alto rendimiento y disponibilidad los 365 días del año y las 24 horas del día.

Esta monografía esta dirigida a todas aquellas personas o administradoras de red interesadas en el tema, y resulta mas interesante que un libro ya que es el resultado de una investigación profunda y contiene información no solo de una sino de diversas fuentes bibliográficas sobre los temas relacionados con ella.

LISTA DE FIGURAS

LISTA DE TABLAS

CAPÍTULO 1.

LOAD BALANCING: CONCEPTOS BASICOS, APLICACIONES Y CLASIFICACION

1.1 CONCEPTOS DE LOAD BALANCING O BALANCEO DE CARGA

1.1.1 DEFINICIÓN DE LOAD BALANCING

Load Balancing o balanceo de carga es un proceso que se encarga de dividir cierto trabajo entre varias partes, las cuales tienen las mismas características y capaces de realizar las mismas tareas[1]. Es la técnica mediante la cual el tráfico de una red puede ser distribuido hacia el otro lado a diferentes servidores, siendo esto transparente al usuario final, ya que aparentemente responde como si estuviera funcionando como un solo servidor que responde a las peticiones de todos.

El balanceo de carga no es para nada un nuevo concepto en el ámbito de las redes de comunicaciones, pero si ha venido evolucionando, ya que al principio en los inicios del Internet cuando aun este era poco utilizado, debido a que no todos tenían acceso a el, y los sitios Web también eran muy poco utilizados para el

[1] http://www.elserver.com/blog/2008/04/09/dynamic-weighted-sticky-session-load-balancing/

comercio, un solo servidor podía responder a todas las peticiones de clientes sin ningún tipo de problema. Pero cuando mas y mas empresas reconocieron el poder que el Internet podía brindar, esto empezó a cambiar, y un solo servidor ya no era suficiente para responder a la gran demanda de peticiones por parte de los clientes, es aquí cuando surge la necesidad de crear formas ingeniosas como la redundancia y el balanceo de carga. En la figura 1 podemos observar un esquema simple de balaceo de cargas, en el cual intervienen dos balanceadores que se encargan de repartir el trafico entre los Web Server.

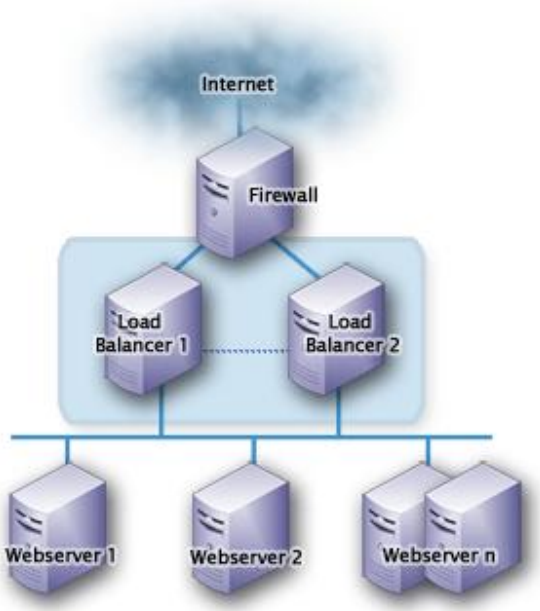


Figura 1. Esquema simple de Load Balancing[2].

En el caso de los servicios Web, los clientes acceden desde Internet y aunque aparentemente todos reciben el servicio de un solo servidor, el tráfico se esta direccionado a servidores diferentes que tienen las mismas capacidades y características, pero todo esto es transparente al usuario final.

[2] Tomado de Internet: http://www.elserver.com/blog/2008/04/09/dynamic-weighted-sticky-session-load-balancing/

El balanceo de carga, además de equilibrar el trabajo entre servidores y mejorar la disponibilidad de servicio, proporciona otras ventajas como la redundancia y la seguridad, además de la escalabilidad sin necesidad de migración de tecnologías, ni interrupción de servicio.

El balanceo de carga proporciona las siguientes funciones[3]:

- Intercepta el trafico de red destinado a un sitio (Como el trafico Web)
- Comparte el tráfico en los pedidos individuales y determina cual servidor recibe el pedido individual.

- Mantiene un reloj sobre los servidores disponibles, asegurando que estén respondiendo al trafico, sino, son sacados de rotación.

- Suministra redundancia dando trabajo a más de una unidad en un escenario fail-over.

- Ofrece la distribución de contenido aware realizando acciones como lectura de URL's e interceptando cookies.

Un balanceador de carga puede mediante una VIP (IP virtual) enviar tráfico a un determinado servidor disponible y luego enviar el tráfico del servidor de vuelta a través de Internet hasta su destino, la parte crítica es que en la mayoría de los casos, el tráfico debe atravesar el balanceador de carga en el camino devuelta hacia Internet.

[3] BOURKE, Tony. Server Load Balancing. United States of America: O'Reilly & Associates, Inc., 2001. P. 14

Existen muchas técnicas para realizar balanceo de carga, pero estas pueden ser clasificadas en dos tipos:

- Balanceo de carga Bridge-Path
- Balanceo de carga Route-Path

Para entender la mecánica de cada uno de estos métodos veamos la siguiente ilustración:

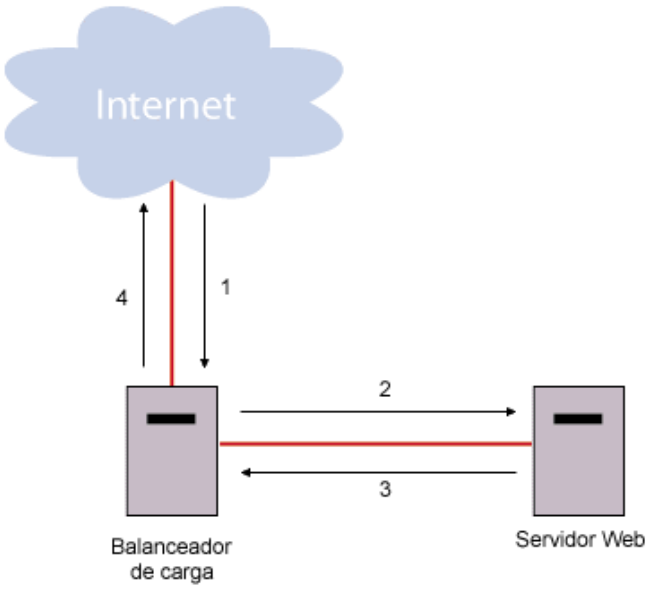


Figura 2. Proceso de Balanceo de carga[4]

Como se puede apreciar en la figura 2, un usuario desde Internet intenta acceder a un servicio Web, cuando este envía la petición (1), es recibida inicialmente por el balanceador de carga y el se encarga de hacer llegar dicha solicitud al servidor Web (2), que a su vez le responde nuevamente al balanceador (3), para que lo haga llegar al usuario (4).

[4] Tomado de Internet: http://www.tejedoresdelweb.com/w/Balance_de_carga

1.1.1.1 BALANCEO DE CARGA BRIDGE-PATH

Como hablamos de Balanceo de carga Bridge-Path, actúa como un puente entre dos redes, es decir, realiza el proceso tal cual como se detalla en la figura 2, este método es factible solo cuando existe un solo camino para llegar al objetivo, ya que se encuentra en la capa 2, lo que quiere decir que presenta conflictos al hacer redundancia, aparte de no ser escalable.

1.1.1.2 BALANCEO DE CARGA ROUTE-PATH

A diferencia del anterior, el Route-Path, lo encontramos en la capa 3, y esta en capacidad de enrutar en un clúster[5] de servidores. Además de ofrecer por sus características mayor redundancia y escalabilidad.

1.1.2 SESIÓN PERSISTENTE

Para entender el concepto de sesión persistente, también llamada sesión pegajosa o sticky-session, primero debemos entender que sucede en una transacción entre un cliente y un servidor Web. Esta transacción puede constar de múltiples conexiones TCP, tal y como lo muestra la figura 3.

[5] Clúster: agrupamiento de servidores con el fin de obtener una mayor tolerancia a fallas y un mayor rendimiento.

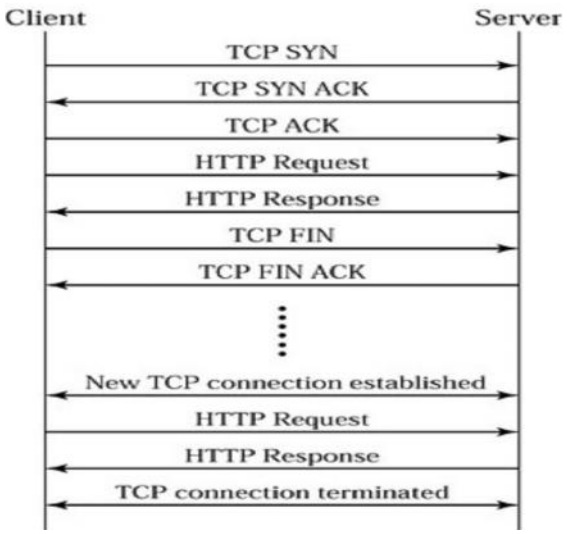


Figura 3. Proceso de Transacción Cliente-Servidor[6]

Inicialmente el navegador del cliente abre una conexión TCP para el sitio WEB, y envía una solicitud http, el servidor envía una respuesta con todo los objetos de la pagina, entonces el navegador del cliente los recibe y los monta en la pagina; cuando el cliente hace clic en alguno de los vínculos se abre otra conexión TCP para enviar la solicitud y recibe como respuesta los objetos que contienen la pagina siguiente. Cuando hay un solo servidor, las peticiones de todos los clientes van a ese servidor.

Cuando existe un clúster de servidores, lo que sucede es que cada vez que el cliente realiza una conexión TCP, es direccionada al servidor basado en la carga que tengan los servidores en ese momento, lo que quiere decir que un cliente es atendido por varios servidores.

Pero existen casos en los que es necesario que el cliente sea atendido por un solo servidor, y es cuando aparece el concepto de sesión persistente. Entonces una

[6] Tomado de: KOPPARAPU, Chandra. Load Balancing Servers, Firewalls and Caches. New York United States: John Wiley & Sons, Inc., 2002. P. 25

sesión persistente es la capacidad de persistir todo el periodo de sesiones para un usuario con el mismo servidor para la duración de una solicitud de transacción. En la figura 4, podemos observar como un balanceador de carga mantiene a un usuario pegado al mismo servidor durante la misma sesión.

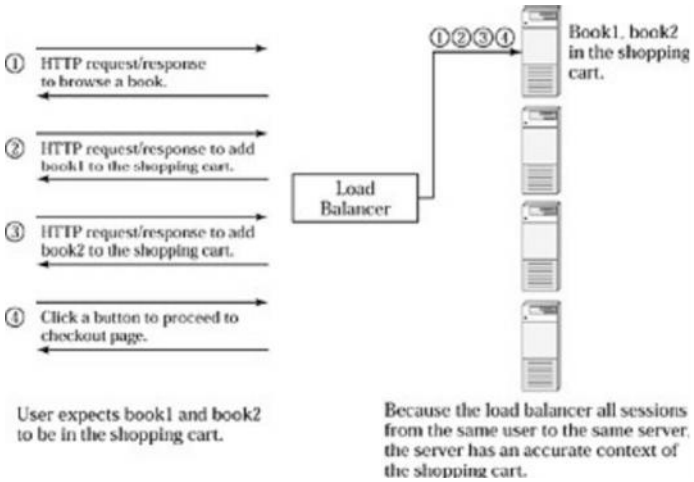


Figura 4. Sesión persistente.

Para lograr que el balanceador de carga pueda mantener sesiones persistentes debe saber como identificar al usuario y cuando empieza o termina una transacción. Para esto existen los siguientes métodos basados en direcciones IP, algunos de estos son:

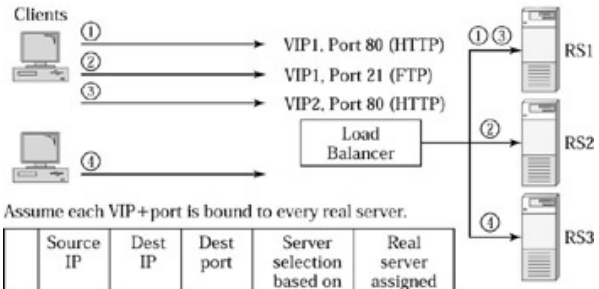
- Fuente IP, Virtual IP y Puerto
- Fuente IP y Virtual IP

1.1.2.2 ALGUNOS TIPOS DE SESION PERSISTENTE

FUENTE IP, VIRTUAL IP Y PUERTO

Al utilizar este método el balanceador de carga tiene en cuenta como base los tres campo de cada paquete TCP SYN: dirección origen, dirección destino y numero de puerto destino, en este paquete la dirección IP destino será la IP (Virtual IP) del balanceador de carga y el puerto destino indica la aplicación visitada por el usuario. El balanceador de carga selecciona un servidor basado en el equilibrio de la carga para la primera conexión recibida de una dirección IP determinada, un Virtual IP específico y un número de puerto, a partir de esto las siguientes conexiones con los mismos valores serán direccionadas al mismo servidor, siempre y cuando el temporizador de sesiones de persistencia no haya expirado.

En caso de que los tres campos con concuerden, así por ejemplo, el cliente quiera acceder a una aplicación diferente, este no será dirigido al mismo servidor, por el contrario será direccionado en base a la carga de los servidores, tal y como lo muestra la figura 5.



Assume each VIP+port is bound to every real server.

	Source IP	Dest IP	Dest port	Server selection based on	Real server assigned
1	C1	VIP1	80	Load balancing	RS1
2	C1	VIP1	21	Load balancing	RS2
3	C1	VIP1	80	Persistence	RS1
4	C2	VIP2	80	Load balancing	RS3

Figura 5. Sesión de persistencia basada en IP, VIP y Puerto[7]

[7] KOPPARAPU, Chandra. Load Balancing Servers, Firewalls and Caches. New York United States: John Wiley & Sons, Inc., 2002. P. 28

FUENTE IP Y VIRTUAL IP

Con este método las aplicaciones que se encuentran en un mismo servidor pueden compartir la información, en este caso solo tenemos en cuenta la IP del cliente y la VIP del Balanceador, así todas las conexiones provenientes de una dirección IP serán dirigidas a un solo servidor sin importar la aplicación o numero de puerto al que se dirige como se puede apreciar en la figura 6.

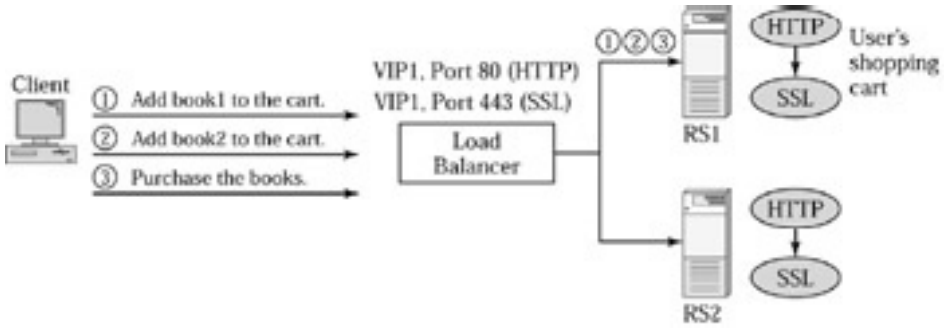


Figura 6: Sesión persistente basado en IP y Virtual IP[8].

1.1.2.3 COOKIE SWITCHING

Una cookie es un objeto que es controlado por los servidores Web, cuando un cliente hace una solicitud, puede enviar una cookie como parte de la respuesta, el navegador almacena la cookie en el cliente, y este envía la cookie en todas las peticiones al servidor.

Existen tres formas diferentes para realizar el Switcheo de Cookies:

- Lectura de Cookies
- Inserción de Cookies

[8] KOPPARAPU, Chandra. Load Balancing Servers, Firewalls and Caches. New York United States: John Wiley & Sons, Inc., 2002. P. 29

- Reescritura de Cookies

LECTURA DE COOKIES

La lectura de cookies funciona de la siguiente manera: cuando un cliente hace una petición por primera vez a un servidor Web, el balanceador de carga verifica si existe una cookie, sino existe lo crea una cookie con el identificador de uno de los servidores y lo dirige hacia el, cuando el cliente intenta realizar una nueva conexión TCP, lleva consigo y en forma transparente la cookie con el identificador del servidor, el balanceador verifica que existe una cookie y lo envía al servidor correspondiente.

INSERCION DE COOKIES

Con este método, cuando el cliente realiza la primera conexión TCP, es dirigida a uno de los servidores basado en la carga, cuando el servidor responde el balanceador de la carga le inserta una cookie con el identificador de ese servidor, siendo esto transparente a la aplicación. La desventaja de este método es el rendimiento de los gastos generales y el potencial inducido por el balanceador de carga.

REESCRITURA DE COOKIES

Este método toma un poco de los dos anteriores, el mayor problema en la inserción de cookies es que requiere una compleja copia de memoria y lo que es mas importante puede causar que el tamaño máximo del paquete sea excedido, lo que si podemos tener es un marcador en el paquete para el nuevo cookie y todo lo que debe hacer el balanceador de carga es configurar correctamente su valor.

La ventaja de este método es que no crea la enorme sobrecarga en el balanceador de carga, también es mejor que el de lectura de cookies ya que, disminuye el impacto de crear la cookie solo en el servidor.

1.1.2.4 APLICACIONDE DE COOKIE SWITCHING

Aunque inicialmente el cookie switching fue creado para solucionar el problema del mega proxy, puede servir para otros propósitos, así por ejemplo, si tenemos un sitio Web con tres tipos de clientes: Plata, oro y platino, si tenemos servidores de mayor gama que otros, podemos garantizar que los clientes platino sean dirigidos a los servidores de gama alta, y si en un momento dado aumentan las solicitudes de los usuarios plata, no afecta los demás niveles de clientes.

El cookie switching también puede ser utilizado para mejorar la calidad de servicio (QoS). Los tradicionales switches capa 2/3, solo pueden mirar los encabezados IP, mientras que un balanceador de carga puede examinar las capas superiores y fijar la procedencia IP o el tipo de servicio (ToS) o bits que indican la prioridad del paquete.
Al mirar las cookies el balanceador de carga, puede reconocer el cliente, usuario o tipo de tráfico y tomar decisiones inteligentes con los paquetes de datos.

1.1.2.5 CONSIDERACIONES DE COOKIE SWITCHING

Aunque esta es una poderosa herramienta utilizada para equilibrar la carga, persistencia en el periodo de una sesión y proporcionar control de tráfico, existen algunos aspectos que debemos considerar, como por ejemplo la privacidad del usuario, ya que las cookies pueden ser utilizadas en algunos sitios Web para realizar seguimiento de los patrones de uso del Internet. Pero existen dos tipos de cookies: las temporales y las permanentes, las primeras son aquellas que se

alojan en el equipo solo para el periodo de sesiones de navegación y las permanentes son aquellas que pueden quedarse almacenas para siempre en el ordenador del usuario.

Pero para esto la mayoría de los navegadores que utilizamos hoy en día, tienen la funcionalidad de permitir o no permitir las cookies temporales o permanentes por separado.

1.1.3 SERVIDORES DE BALANCEO DE CARGA

Un balanceador de carga es un dispositivo que distribuye la carga entre varias maquinas, como se dijo anteriormente, tiene el efecto de hacer varias maquinas aparecen como una sola. Existen varios componentes de los dispositivos SLB que se examinan a continuación:

VIP's

La IP virtual es la instancia del Load Balancing, el punto al que los navegadores acceden para ir a un sitio determinado. Un VIP tiene una dirección IP que debe ser de acceso público. Por lo general un puerto TCP o UDP esta asociado a una VIP, así como el puerto 80 para el trafico Web.

Un VIP tendrá por lo menos un servidor real que le han sido asignadas, a las que va a distribuir el tráfico, por lo general hay varios servidores reales y el VIP extienda el tráfico entre ellos métodos y métricas.

SERVIDORES

Un servidor es un dispositivo que ejecuta un servicio, comparte la carga con otros dispositivos, tiene una dirección IP y por lo general un puerto TCP/UDP asociado con el y no tiene que ser públicamente direccionable (dependiendo de la topología de la red)

GRUPOS

Cuando hablamos de grupos nos referimos a un conjunto de servidores con carga equilibrada, a lo que se le llama granja de servidores.
USUARIO – NIVELES DE ACCESO

El nivel de acceso de los usuarios se refiere a la cantidad de control del usuario cuando se ha conectado a un balanceador de carga. Y existen de los siguientes niveles:

- Solo lectura: un acceso de solo lectura, es un nivel en el que no se pueden realizar ningún tipo de cambios, en este modo solo se tiene acceso a ver ajustes, configuraciones, pero nunca de realizar un cambio.
- Súper usuario: es e nivel de acceso al que el se le otorga autonomía total del sistema, el súper usuario puede: agregar cuentas, borrar archivos y modificar cualquier parámetro del sistema.
- Otras categorías: existen algunos productos que ofrecen niveles adicionales, que le permiten acceder o modificar parámetros específicamente.

1.1.4 REDUNDANCIA

El concepto de redundancia es muy simple, si un dispositivo falla, otro tomara su lugar o función sin afectar el adecuado funcionamiento de la red. Generalmente

los dispositivos de balanceo de carga tienen esta funcionalidad, se tienen los dos dispositivos implementados y se utiliza un protocolo que comprueba la "salud" de su socio (el otro dispositivo). En algunos escenarios ambos dispositivos están activos y aceptan el tráfico, a este escenario se le conoce como **active-active** (figura 7), mientras que en otros, solo uno de ellos se encuentra funcionando, y el otro a la espera de que este falle, a este escenario se la conoce como **active-standby** (figura 8).

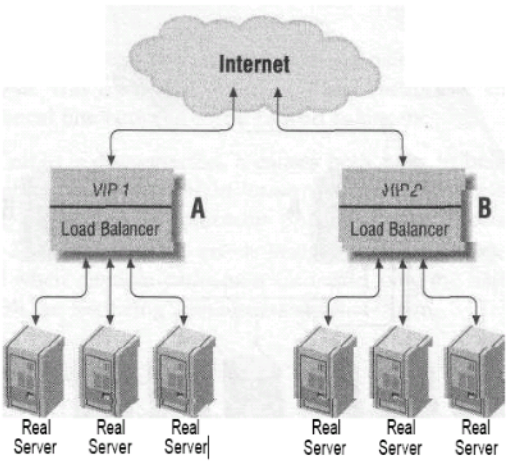


Figura 7. Escenario active-active[9]

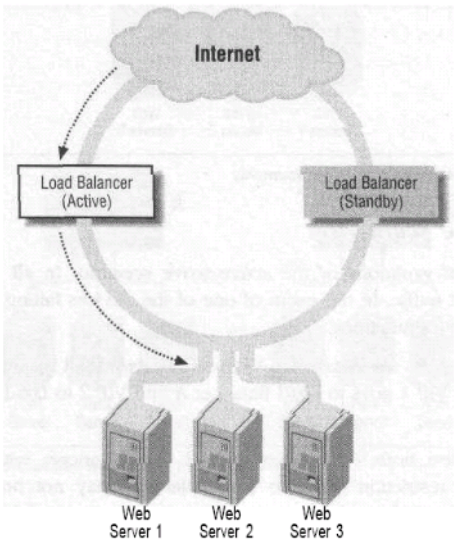


Figura 8. Escenario active- standby[10]

[9] Tomado de: BOURKE, Tony. Server Load Balancing. United States of America: O'Reilly & Associates, Inc., 2001. P. 27

[10] Ibis, P. 30

1.2 APLICACIONES DE LOAD BALANCING

1.2.1 LOAD BALANCING EN SERVIDORES GLOBALES (SGLB)

1.2.1.1 ¿PORQUE HACER SGLB?

Son dos los motivos que impulsan la necesidad de hacer balanceo de carga en los servidores globales: la alta disponibilidad y tiempos de respuesta más cortos[11].

Hablando de la disponibilidad, utilizamos un balanceador de carga alternativo en la red, en caso de que el otro falle. Pero ¿que sucedería si perdemos el poder en el centro de datos en donde la granja de servidores y los balanceadores están ubicados? Ó ¿Qué pasaría si perdemos la conexión a Internet debido a un fallo en el proveedor de servicio (ISP)? ¿O si ocurre un desastre natural como inundaciones o un terremoto y hace que caiga todo el centro de datos?

Si aplicamos SGLB tendríamos varios centros de datos, con diferentes granjas de servidores para garantizar una disponibilidad continua a los usuarios aun si todo un centro de datos esta inactivo.

Un factor que hasta hace poco era difícil de controlar es el retraso en los tiempos de respuesta: tiempo de respuesta del usuario, demora de Internet, y demora del lado del servidor. No se puede controlar la demora por parte del cliente, ya que influyen aspectos como el acceso desde la última milla, el rendimiento del equipo, etc. Pero con el uso de GSLB podemos operar el sitio Web o servidores de

---

[11] *BOURKE, Tony. Server Load Balancing. United States of America: O'Reilly & Associates, Inc., 2001. P. 64*

aplicaciones en múltiples granjas y centros de datos, y con esto podemos atender a los usuarios desde el punto en que se ofrezcan mejores tiempos de respuesta.

1.2.1.2 DNS: VISION GENERAL

Cuando deseamos acceder a un sitio Web, por ejemplo: loadbalancing.com, escribimos la dirección URL en el navegador: www.loadbalancing.com, el navegador debe encontrar primero cual es la dirección IP del sitio y es aquí donde DNS entra en juego.

El domino de Internet tiene la estructura de un árbol invertido tal y como se muestra en la figura 9, En el nivel superior se encuentran varios niveles como: com, gov, edu y cada uno de estos dominios contiene otros dominios y subdominios, entonces loadbalancing.com es un subdominio del dominio com, así mismo loadbalancing.com puede contener otros subdominios como sub.loadbalancing.com; un nombre de dominio dentro de sub.loadbalancing.com podría ser ftp.sub.loadbalancing.com; estos subdominios también puedes ser llamados dentro de la zona de dominio loadbalancing.com.

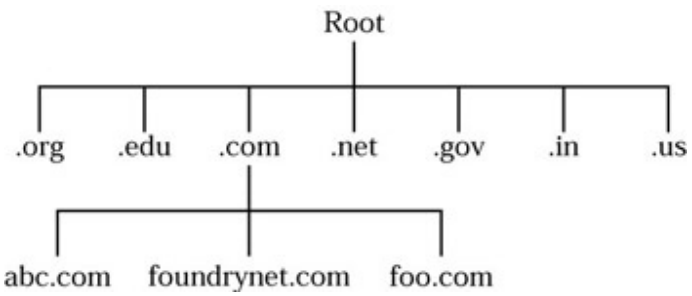


Figura 9. Árbol invertido DNS

Lo que sucede cuando digitamos www.loadbalancing.com en nuestro navegador es lo siguiente: en nuestra LAN tenemos un DNS local, nuestro navegador hace

un llamado a este para resolver http://ww.loadbalancing.com y obtener su IP, este la obtiene una IP y lo dirige hacia el siguiente servidor DNS, y este en primer lugar lo dirige a la raíz del servidor de nombres, el cual le devuelve una lista de servidores para el dominio .com, luego investiga cuales son los servidores con autoridad para el dominio loadbalancing.com y por ultimo el servidor con la autoridad sobre loadbalancing.com devuelve la IP del servidor Web para http://www.loadbalancing.com. Este proceso lo podemos entender mejor con la figura 10.

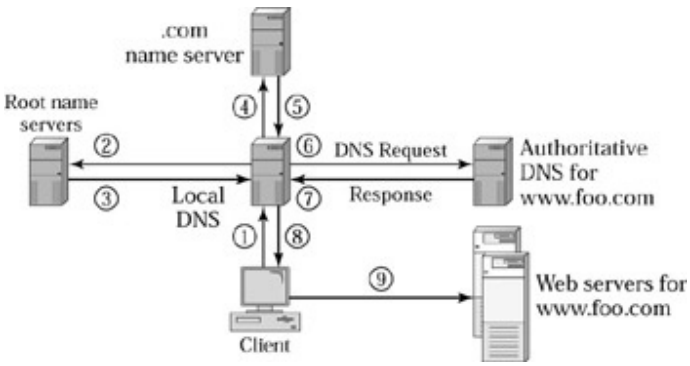


Figura 10. Sistema DNS

1.2.1.3 REDIRECCIONAMIENTO HTTP

Uno de los enfoques que pueden utilizarse sin ninguna modificación del sistema actual de DNS es este método llamado redireccionamiento HTTP, la definición del protocolo HTTP, incluye una forma de que un servidor Web entregue como respuesta HTTP que contienen un error de redireccionamiento y a la vez la nueva URL. Este lo que hace es informar al navegador que debe ir a la nueva URL si desea encontrar la información que requiere.

VENTAJAS

- No necesita cambiar los parámetros o configuración DNS
- Cuando el primer servidor recibe la petición HTTP, sabe cuando la dirección IP del cliente es útil, basado en políticas de selección.

DESVENTAJAS

- Solo funciona como su nombre lo indica en Aplicaciones HTTP
- Aumentan los tiempos de respuesta, ya que el servidor debe resolver nuevamente la segunda URL, debe realizar una nueva conexión TCP, y enviar una nueva solicitud HTTP.

1.2.2 LOAD BALANCING EN FIREWALLS

1.2.2.1 CONCEPTO DE FIREWALL

Un firewall es un dispositivo o software de seguridad que separa la red interna de la red externa. Todo el tráfico que entre o salga de la red interna de fluir a través del firewall en virtud de la topología utilizada como lo muestra la siguiente figura:

Figura 11. Como trabaja un firewall

El firewall o corta fuegos, aplica seguridad basado en políticas de acceso y protege la red interna de usuarios maliciosos. Estos dispositivos analizan los paquetes entrantes o salientes para determinar si dicho paquete o conexión debe permitirse.

1.2.2.2 ¿PORQUE HACER LOAD BALANCING EN EL FIREWALL?

Tenemos que un firewall es un punto de acceso a la red, así mismo un solo punto de fracaso para toda la red, si perdemos el firewall, perdemos la conexión entre la red interna y la red externa. Aunque algunos firewall tienen la funcionalidad de trabajar en grupo, se componen de dos firewalls, en donde uno actúa como reserva de la otra unidad, esto mejora la disponibilidad pero no la escalabilidad.

El balanceo de carga en el firewall permita mejorar la escalabilidad, así como la alta disponibilidad mediante de la distribución de carga a través de múltiples cortafuegos y tolerar el fracaso de un firewall.

1.2.2.3 LOAD BALANCING EN EL FIREWALL

El diseño básico para balancear la carga entre firewalls se conoce como sándwich firewall, consiste en colocar un balanceador de carga en ambos lados de los cortafuegos, debido que el tráfico es originado de un lado al otro, y sin importar su dirección siempre hay que pasar por un balanceador antes que por el firewall como se muestra en la siguiente figura:

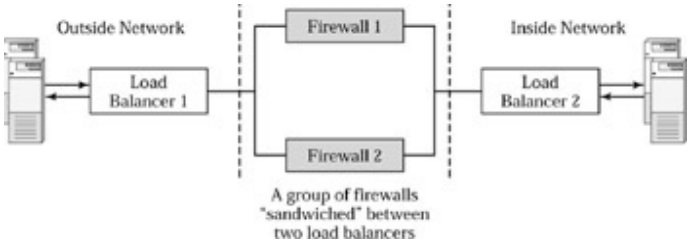


Figura 12. Sandwich firewall

1.2.3 LOAD BALANCING EN CACHÉS

1.2.3.1 CONCEPTO DE CACHÉ

Una cache se encarga de almacenar el contenido de las páginas Web frecuentemente visitadas, esto con el fin de mejorar tiempos de respuesta y ahorrar ancho de banda. Esto se logra de la siguiente forma: cuando un cliente desea acceder por primera vez a una pagina Web, la cache la toma y la almacena en memoria, esto con el fin de que si el mismo usuario desea volver a ingresar a esa pagina, es cargada directamente de la caché sin necesidad de ir nuevamente al servidor de origen.

1.2.3.2 LOAD BALANCING EN CACHÉS

El balanceo de carga de carga a través de cachés es muy diferente al balanceo en servidores. El balanceo en servidores se busca el que tenga menor carga para enviar la siguiente solicitud. En el balanceo de carga con caché tenemos que prestar atención a los contenidos que se encuentran en cada una de las caches para minimizar los tiempos de respuesta. Cuando hacemos una solicitud

determinada a uno de los servidores, esta se guarda en la caché, de esta manera los siguientes clientes que hagan la misma solicitud tendrán mejores tiempos de respuesta.

1.3 METODOS DE LOAD BALANCING

1.3.1 DNS ROUND ROBIN

Con este método no se requiere hardware, ni servidores con características especiales para gestionar el balanceo, consiste en asignar varias direcciones IP a un mismo sitio, cada IP en un servidor diferente, y cada uno de los servidores con el mismo contenido. Una de las principales desventajas de este método es que si uno de los servidores activos presente fallas y quede fuera de servicio todos los usuarios que tengan asignada esa IP perderán el acceso.

1.3.2 LB ROUND ROBIN

Este método es parecido al anterior, la única diferencia es que no se hace una distribución basada en DNS sino que se asignan los servidores al azar.

1.3.3 WEIGHTED LB

Este método se basa en la capacidad de carga de los servidores del grupo de balanceo, la cual debe ser definida por el administrador de la red. Así el administrador debe suministrar al balanceador de carga la información acerca de la capacidad de cada uno de los servidores.

1.3.4 DYNAMIC WEIGHTED LB

Este método al igual que el anterior, se basa en la capacidad de carga de los servidores del grupo de balanceo, pero en esta ocasión el balanceador de carga puede medir y estimar la capacidad de cada uno de los servidores (a diferencia del anterior que debía ser definida por el administrador de la red).

1.3.5 STICKY-SESSIONS O STICKY-USER

Este método es utilizado en ocasiones donde se requiere que el usuario se mantenga por todo el periodo de sesión con un solo servidor, estas son las llamadas sesiones persistentes, tal y como se hablo al principio de es este capitulo. La desventaja de este método es que no se tiene muy en cuenta las características o capacidad de cada uno de los servidores.

1.3.6 UNA NUEVA SOLUCION: DYNAMIC WEIGHTED – STICKY SESSION

Actualmente se empiezan a utilizar estos métodos de forma complementaria, proporcionando un balanceo de cargas teniendo en cuenta la capacidad de cada uno de los servidores para la debida distribución de los usuarios a través de la red y también se encarga de mantenerlos lo mas posible pegados a un mismo servidor, en la medida que no sea mas beneficioso moverlo a otro.

CAPITULO 2

2. SISTEMAS DE GESTION Y MONITOREO DE LOAD BALANCING

2.1 PROGNOSIS Dynamic Load Balancing

2.1.1 *Descripción*

Un sistema ineficiente supone mayores costes operacionales. Mediante la distribución automática de nuevos procesos entre CPUs y discos con la menor carga de trabajo y, eliminando las ineficiencias en procesos y volúmenes swap, el hardware estará preparado para un uso mas efectivo.

Mediante el balanceo de la carga de trabajo entre CPUs y discos, Dynamic Load Balancing mejora sustancialmente los tiempos de respuesta de sus aplicaciones críticas NonStop y asegura que sus recursos hardware funcionen con óptima eficiencia. Puede especificar qué procesos batch, TACLs, procesos transitorios, así como servidores Pathway son optimizados, dándole un control total sobre la distribución de la carga de trabajo sobre su red.

Dynamic Load Balancing proporciona considerables mejora en el rendimiento del sistema. Proporciona un alto rango de opciones de configuración, que pueden usarse para adaptar el producto a sus requerimientos específicos.

Por ejemplo, se pueden excluir tareas (jobs) específicas o usuarios de CPUs concretas o hacer cumplir las prioridades de usuarios de forma que las

Aplicaciones críticas no sean impactadas. Como dato importante, destacar que estos cambios pueden ser implementados mientras el sistema está funcionando.

Dynamic Load Balancing también proporciona ahorro potencial en el hardware. Le ayuda a aplazar o eliminar el coste de compras de hardware incrementando el rendimiento de los recursos existentes, a través de una distribución efectiva de la carga de trabajo.

Figura 13. Diagramas de CPUs y PCB balanceados

Mediante la asignación automática de nuevos procesos y volúmenes swap a los recursos con la menor carga de trabajo Dynamic Load Balancing asegura que las CPUs y discos este corriendo a una eficiencia óptima.

Proporciona de forma notoria una mejora del rendimiento de CPUs mediante el balanceo de la carga de trabajo. Los nuevos procesos se inician en la CPU mas libre.

2.1.2 *Utilidades*

Gestión efectiva de CPU

Nuevos procesos son iniciados en la CPU con menor carga de trabajo. La CPU más apropiada se determina utilizando atributos como longitud de colas para determinar la carga de trabajo de CPU, las páginas libres como indicación de utilización de memoria y disponibilidad, y los PCBs ó TLEs para indicar la utilización de recursos del sistema.

Asignación dinámica de ficheros swap

Dynamic Load Balancing selecciona volúmenes de swap óptimos para nuevos procesos. El resultado es ficheros swap balanceados, creación más rápida de procesos y mejoras en la utilización de discos.

Opciones de configuración sofisticadas

Usted tendrá control complete sobre los recursos que Dynamic Load Balancing controla. Puede ser configurado para excluir objetos particulares y procesos, forzar máximas prioridades para usuarios y restringirlos para CPUs específicas, o especificar parámetros para la selección de CPUs en la distribución de nuevos procesos.

Compatible con todo el software

Puede ser aplicado a software de third-party, Guardián o desarrollado por el usuario.

2.2 BARRACUDA LOAD BALANCER

2.2.1 *Descripción*

Diseñado para lograr objetivos de alta disponibilidad y seguridad, el Barracuda Load Balancer integra el balanceo de carga y la prevención contra la intrusión en la red en un dispositivo económico y fácil de usar. El Barracuda Load Balancer proporciona amplias capacidades de recuperación inmediata después de fallas (failover), en caso de fallas en el servidor, en la distribución del tráfico entre múltiples servidores y en la protección integral contra intrusiones en la red. Sin costos adicionales de licencia por servidor o puerto, el Barracuda Load Balancer ofrece la solución más económica posible para escalar y proteger centros de datos de misión crítica.

Alta disponibilidad y escalabilidad

Para ambientes de mucho tráfico, el Barracuda Load Balancer puede distribuir el tráfico con base en varios algoritmos de programación, incluyendo round robin, cálculo de carga o basado en menos conexiones. Para las aplicaciones que requieren persistencia, el Barracuda Load Balancer puede mantener el estado usando persistencia por IP o cookies.

El Monitor de Servicios integrado al Barracuda Load Balancer asegura que los servidores y sus aplicaciones asociadas siempre estén operativos. En caso de daño o caída de un servidor o aplicación, el Barracuda Load Balancer activa el

recuperador de fallas entre los servidores para asegurar una disponibilidad continua. Para mitigar los riesgos asociados con fallas en los 'load balancers' (Balanceadores de carga) mismos, se pueden instalar dos Barracuda Load Balancers en una configuración activa/pasiva.

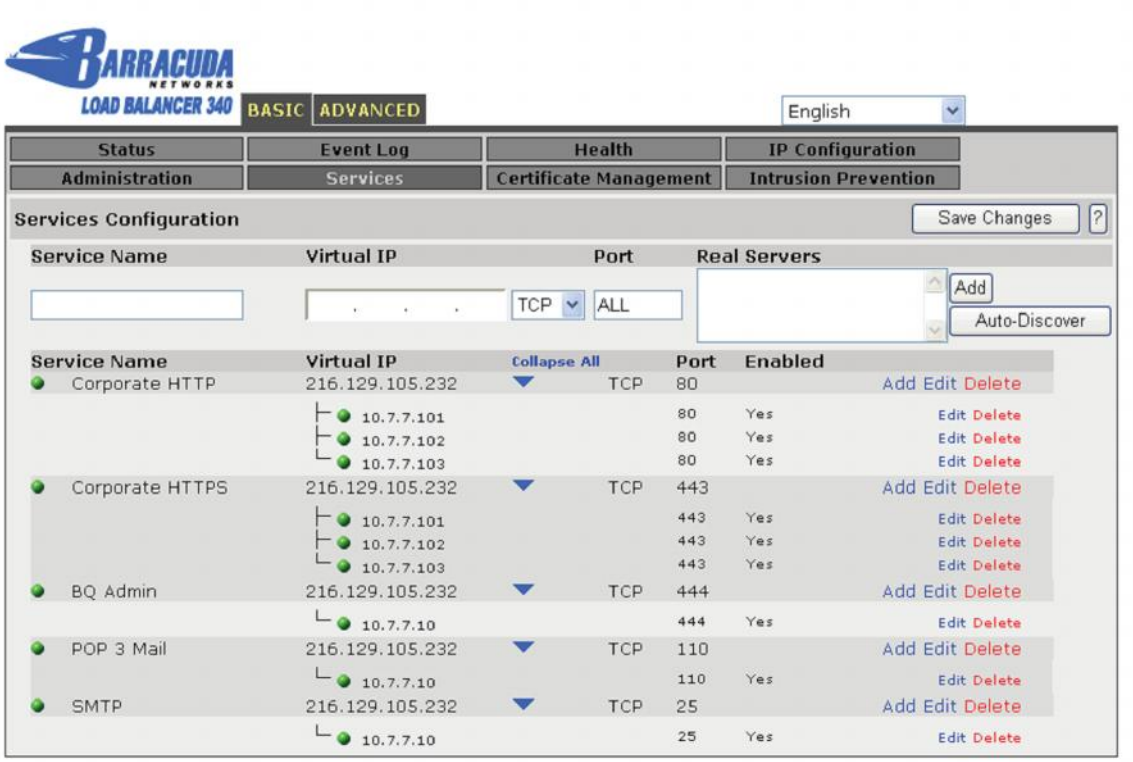


Figura 2. El Barracuda Load Balancer puede distribuir el tráfico para múltiples servicios, incluyendo Web, email, y otras aplicaciones de red.

Facilidad de Uso y Mantenimiento

Para minimizar la administración relacionada con seguridad, el Barracuda Load Balancer recibe actualizaciones Energize automáticamente para lo más reciente en definición de prevención de intrusión y actualizaciones de seguridad. Las actualizaciones Energize son distribuidas cada hora por la Central de Barracuda,

un centro de tecnología avanzada donde los ingenieros continuamente monitorean y alivian las más recientes amenazas de Internet.

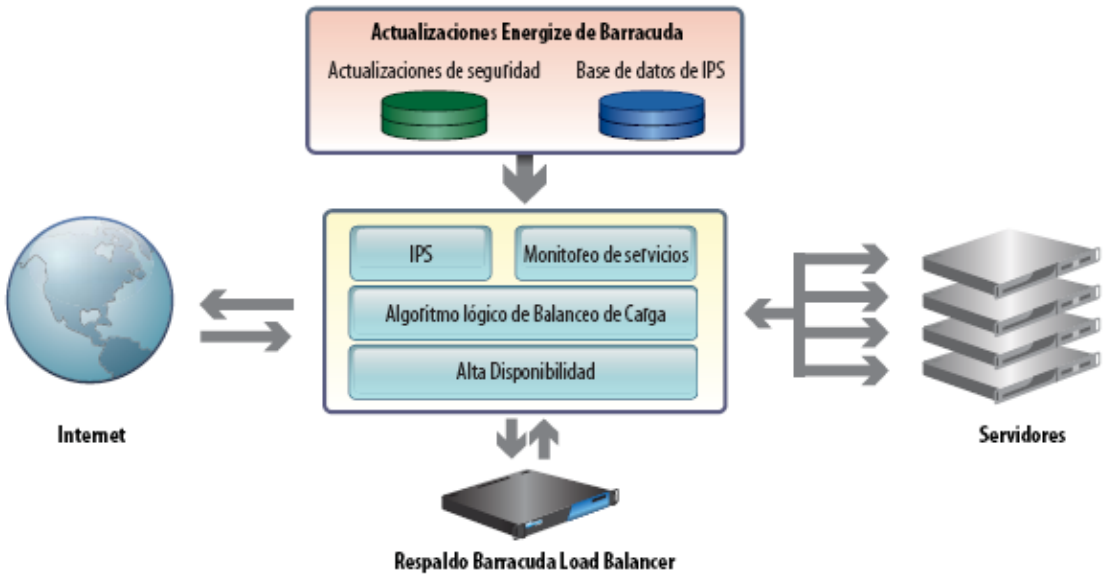


Figura 15. Arquitectura Barracuda Load Balancer

. Figura 16. Típica instalación

2.2.2 Características

BALANCEO DE CARGA

• Balanceo de carga en capa 4 y 7

• Persistencia de sesiones capa 4

• Persistencia de cookies capa 7

• Soporte para Route-Path & Bridge-Path

• Modo Direct Server Return

• Monitoreo de salud del servidor

• Peso de servidores

• Soporte para menos conexiones y round-robin

• Monitoreo adaptable

• Chequeo de salud programable

DISPONIBILIDAD Y PERFORMANCE

• Alta disponibilidad

• Detección de falla de servidores

• Throughput gigabit por el cable

• Añadir y remover servidores con rapidez

• Modo de mantenimiento de servidores

USABILIDAD Y ADMINISTRACION

• Fácil de usar interfaz vía web

• Monitoreo de salud

• Monitoreo de performance

• Soporte para SNMP

• Actualizaciones automatizadas

• Protegido por SSL y ACL

Estadísticas de tráfico

• Configuración de respaldos automatizados

• Modo de auto-descubrimiento

• Sin restricciones por puerto o servidores

• Application Programming Interface (API)

• Actualizaciones de firmware con un solo click

PROTOCOLOS SOPORTADOS

• HTTP

• HTTPS (SSL)

• SSH

• SMTP

• IMAP

• RDP (Citrix/terminal services)

• POP3

• NNTP

• ASP

• Streaming Media

• DNS

• LDAP

• RADIUS

• TFTP

• Otros servicios TCP/UDP

SEGURIDAD

• Protección integrada de IPS y ataques

• Actualizaciones automatizadas de definiciones

• IPS

• Nivel de servicio ACL

• Protección DDoS

2.3 D-LINK DFL-2500

2.3.1 *Descripción*

El DFL-2500 es un firewall VPN de última generación diseñado para empresas corporativas, universidades y organismos que buscan la mejor relación precio/rendimiento. Este equipo es una poderosa solución de seguridad que

entrega firewall integrado, balanceo de carga, sistema de tolerancia a fallas, Zone-Defense, filtro de contenido, autenticación de usuarios, bloque IM/P2P, protección contra negación de servicios (DoS) y soporte para redes virtuales. El DFL- 2500 es rackeable (19”), tiene una altura de 1U, incluye 8 puertas 10/100/1000 Mbps configurables. Excepto la interfaz DMZ, el DFL-2500 puede administrar múltiples segmentos de red para diferentes grupos, departamentos por piso, satisfaciendo completamente los requerimientos de seguridad del cliente. Además, el DFL-2500 posee una nueva interfaz Web GUI más amistosa, una estructura de diseño más limpia entregando al cliente un look más profesional.

2.3.2 *Características*

• PRINCIPALES CARACTERÍSTICAS Y FACILIDADES:

Rendimiento y Capacidad

• Rendimiento firewall: 600 Mbps

• Rendimiento VPN: 300 Mbps

• Sesiones concurrentes: 1.000.000

• Sesiones/seg. 15.000

• Políticas: 4,000

• Usuarios: Sin restricción

MODO DE OPERACIÓN

• Layer 3: Router, NAT

• Layer 2: Modo Transparente

• Network Address Translation (NAT)

• Port Address Translation (PAT)

• Proactive Network Security: Firewall para Switch Zone-Defense

• Configuración de políticas programables en el tiempo

VIRTUAL PRIVATE NETWORK (VPN)

• Encriptación: DES/3DES/AES/Twofish/Blowfish/

CAST-128/NULL

• Servidor VPN: PPTP/L2TP/IPSec

• Túneles VPN dedicados: hasta 2.500

DIRECCIONAMIENTO IP & ROUTING

• Static IP address

• PPPoE para xDSL, cliente PPTP para xDSL, cliente

DHCP para interfaz WAN

• IP Alias

• Static Routes

• OSPF Dynamic Routing

NETWORKING

• Soporte de múltiples enlaces WAN

• Soporte de VLAN IEEE 802.1q: hasta 1024 VLANs

• IP Multicast: IGMP v2*, IGMP snooping*

• Cliente DDNS: DynDNS.org, TZO.com, dhs.org,

BALANCEO DE CARGA DE TRÁFICO

• Balanceo de carga de tráfico de salida*

• Balanceo de carga de servidores

• Algoritmo de balanceo de carga:

Round Robin

Connection Rate

IP Address/Network Stickiness

LOGGING Y MONITOREO

• Capacidad Interna de log: 1.600 registros

• Soporte de servidor log externo: syslog server

• Monitoreo de rendimiento en tiempo real

• Soporte SNMP v1, v2

• SNMP Trap*

• SNMP Standard MIB-II y Custom MIB

CAPITULO 3

3. OPERACIÓN CONTINUA EN LAS REDES

3.1 Introducción

Una Red con operación continua va más allá de la simple disponibilidad. Una definición podría ser la siguiente: " Se trata de una red que llega casi al 100% de disponibilidad, mientras alcanza excelentes niveles de procesamiento y de tiempos de respuesta, sin tener en cuenta cualquier tipo de fallos en cualquier sitio de la red". A esta descripción clásica se une cada vez más la tolerancia a fallos y la capacidad de recuperación frente a desastres.

3.2 Criterios para determinar una efectiva operación continúa de red

La disponibilidad debería situarse entre el 99,9% y el 100% a lo largo de toda la red. Aunque el 99,9% es una ratio de disponibilidad comúnmente aceptada, a menudo sólo refleja la disponibilidad de un solo componente de la red y no la disponibilidad de todo el entorno. A continuación se muestra una tabla de disponibilidades y su equivalencia en tiempo de servicio.

Disponibilidad(%)	Referencia	Tiempo promedio fuera de servicio por año.
90.000	Un-9	876H (36.5D)
99.000	Dos-9s	87.6H (3.65D)
99.900	Tres-9s	8.76H
99.990	Cuatro-9s	52.56M (< 1H)
99.999	Cinco-9s	5.25M

Tabla 1. Disponibilidad de una red

Una disponibilidad de 99.99% se traduce en 52.56 minutos fuera de servicio una red en promedio al año.

La red requiere un muy buen rendimiento. Así, una red que funciona bien pero está tan congestionada que el tráfico no fluye como debería, no goza realmente de una buena disponibilidad.

Los desastres también afectan a la operación continua de las redes o a los servicios que reparten, de manera que debería estar soportada basándose en requisitos de continuidad de negocio.

Dado que la operación continua en la redes tendrá un mayor coste que una red normal, las organizaciones necesitan determinar qué aplicaciones (sólo las críticas o todas ellas) deberían estar soportadas por estas nuevas redes. En este sentido, es conveniente remarcar el hecho de que en entornos de múltiples localizaciones, se establecerán diferentes requisitos de alta disponibilidad para cada una de ellas.

Las organizaciones deben, al mismo tiempo, determinar los criterios de diseño (basados en los costes), los cuales pueden incluir atributos de resistencia y elasticidad, redundancia y revisión de puntos únicos de fallo en la red. Las empresas que definan y diseñen apropiadamente sus redes de alta disponibilidad lograrán los niveles deseados de fiabilidad y seguridad[12].

Existen algunas consideraciones de tipo físicas que hay que tener presentes a la hora del diseño de una red, a continuación expondremos algunas de esas consideraciones.

3.3 Consideraciones físicas del sitio de ubicación de la red

Inicialmente lo que se debe establecer en la planificación de una red es la mejor ubicación del centro del centro de cableado, porque es allí donde se instalan la mayoría de cables y dispositivos de red.

Los Centros de Computo están compuestos de un sistema de comunicaciones de red de alta velocidad y alta demanda capaz de manejar el tráfico para SAN (Storage Area Networks), NAS (Network Attached Storage), granja de servidores de archivos/aplicaciones/Web, y otros componentes que deben ser localizados en un ambiente controlado.

El control de ambiente se refiere a la humedad, inundación, electricidad, temperatura, control de fuego, y por supuesto, acceso físico y seguridad física. Las comunicaciones dentro y fuera del centro de cómputo se proveen por enlaces WAN, CAN/MAN y LAN en una variedad de configuraciones dependiendo de las necesidades particulares de cada centro. Un centro de cómputo diseñado apropiadamente proporcionara disponibilidad, accesibilidad, escalabilidad, y

[12] Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

confiabilidad 24 horas al día, 7 días a la semana, 365 días al año descontando el tiempo fuera de servicio por mantenimiento. Las compañías telefónicas trabajan un 99.999% de disponibilidad y los centros de cómputo no deben ser diferentes.

La planeación de los centros de datos se ha convertido casi en una especialidad para los integradores de tecnología y a las empresas de arquitectura. Algunas firmas cuentan con un grupo interdisciplinario conformado por arquitecto, ingeniero electricista, ingeniero de sistemas e ingeniero electrónico dentro de su personal. El equipamiento que alberga el centro de cómputo es bastante complejo, cada uno con requisitos específicos de calefacción, enfriamiento, presupuesto de consumo eléctrico y consideraciones de espacio. Un centro de cómputo típico contiene los siguientes componentes:

- Infraestructura de cómputo y redes (cableado, fibra, y electrónicos)

- Centro de Operaciones (Network Operation Center – NOC)

- Sistemas eléctricos de distribución, generación y acondicionamiento - UPS, generadores con control ambiental y sistemas HVAC.

- Sistemas de detección y supresión de fuego.

- Seguridad física, control de acceso y monitoreo ambiental.

- Protección de circuitos (protección de iluminación en algunos casos).

- Tierra física.

- Racks y gabinetes para equipo.

- Canalizaciones: Piso falso y bandejas en techo.

- Equipo de Telecomunicaciones

Los centros de cómputo deben ser cuidadosamente planeados ANTES de construirse para asegurar su conformidad con todas las normas y reglamentos

aplicables. Las consideraciones de diseño incluyen selección de sitio, ubicación, espacio, electricidad, capacidad de enfriamiento, carga de piso, acceso, seguridad, limpieza ambiental, prevención de peligros y crecimiento. Para poder calcular las necesidades anteriores, el proveedor de la solución debe conocer los componentes que contendrán el centro de cómputo y sobre todo, el crecimiento esperado o pronosticado.

El grupo TIA TR-42.1.1 tiene la tarea de desarrollar la norma "Telecommunications Infrastructure Standard for Internet Data Centers." "El alcance de este grupo de trabajo incluirá topologías y desempeño para cableado de fibra y cobre, y demás aspectos de la infraestructura IT que permitirán que las instalaciones rápidamente puedan incorporar nuevas tecnologías, tales como redes 10 Gb/s. La TIA/EIA ha adoptado recientemente la propuesta ***TIA/EIA-942 'The Telecommunications Infrastructure Standard for Data Centers'*** (aun en borrador). Entre los requisitos se considerarán necesidades de flexibilidad, escalabilidad, confiabilidad y administración de espacio[13].

Se establece a continuación criterios ideales para establecer la correcta y segura disposición física del centro de cómputo en las organizaciones.

3.3.1 Edificio, área y espacio

Edificio

Es trascendental la ubicación del edificio y su construcción misma para la operación eficiente del centro de cómputo y como primera medida, debe

[13] http://gerenciait.com/links/articulos/ci_links_temacentral_abril_p2.htm

considerarse si se trata de un edificio nuevo de construir o uno ya existente a adecuarse, para ello se mencionan los siguientes puntos:

a) Realizar un estudio de la zona a fin de evitar estar expuestos al peligro por sismos, contaminación, incendio, explosión, inundación, radiaciones, interferencia de radar, vandalismo, disturbios sociales, así como riesgos provocados por las industrias cercanas y todo lo que pueden ocasionar problemas con el equipo de procesamiento de datos y archivos.

b) Seleccionar la parte más segura dentro del edificio para el centro de cómputo y contar con facilidades de energía eléctrica, acometidas telefónicas, aire acondicionado, servicios públicos y salida de emergencia adecuada.

c) Cuando el acceso al centro de cómputo deba efectuarse a través de otros departamentos, será necesario prever el paso de las máquinas a través de diferentes puertas, ventanas, pasillos, montacargas, etc. Los elevadores deberán soportar cuando menos una carga estándar de 1135kg. (2500lbs.) y ser lo suficientemente largos para acomodar los equipos en este.

d) Se deben definir claramente las rutas de acceso del personal para la carga de documentos, respaldos en unidades magnéticas, elaboración de reportes, etc., cuidando que no existan sobre el piso escalones, rampas, cables, etc.

e) La construcción del piso debe soportar el peso de los equipos que serán instalados. Las designaciones típicas de los equipos IBM no rebasan de los 340kg/m2.

f) La puerta de acceso al centro de cómputo debe tener 95cm. de ancho mínimo y abrir hacia afuera.

g) Se deben de usar materiales de construcción no combustible y resistente al fuego.

h) Recubrir las paredes con pintura lavable, con el objeto de que no se desprenda polvo y sea fácil su limpieza.

i) Construir el mínimo de ventanas exteriores (o ninguna) a fin de evitar interferencias.

j) Si el falso plafón se utiliza como pleno para el retorno del aire acondicionado, deberá pintarse el techo real con pintura de aceite o sintética de color claro.

Área y Espacio

Se recomienda que en el área del centro de cómputo existan separadores de aluminio y cristal o cuartos independientes para la instalación de todo el equipo y se debe considerar lo siguiente:

a) La configuración definitiva del sistema a instalar: el procesador, impresoras, estaciones de trabajo, módems, multiplexores y demás periféricos.

b) Para hacer una distribución adecuada se deberá poseer un plano del local elegido en escala 1:50 sobre el que se ubicarán las plantillas de los equipos cuidando sus áreas de servicio y pruebas (espacio adicional al área del equipo para su mantenimiento).

c) Es necesario plantear la secuencia de conexión de los equipos para los direccionamientos de los mismos.

d) Se recomienda la ubicación de la consola del sistema como máximo a 6 metros de distancia del rack del procesador y que sea visible el panel de control del mismo.

e) Por el polvo que desprenden las impresoras y el ruido que hacen al imprimir, se deben instalar en un cuarto independiente junto con una estación de trabajo a un

metro de distancia de la impresora del sistema para facilitar el suministro de los reportes.

f) Se debe tener en cuenta el espacio a ocupar del equipo adicional como son: Comunicaciones, módems, teléfonos, nobreak, un archivo mínimo, cintas de respaldo, una mesa de trabajo, mueble para manuales y papelería, además del espacio para futuro crecimiento.

3.3.2 Canalizaciones de cableado

En un centro de cómputo donde existe gran variedad de cables necesarios para el funcionamiento y comunicación de los procesadores con sus equipos periféricos, tanto por seguridad como por cuidar los acabados en la decoración interior, los ductos son un factor de gran importancia para ocultar los cables de señal.

Aún contando con piso falso en el centro de cómputo se deben distribuir los cables a través de canaletas o ductos especiales para cables dando una apariencia ordenada y facilidad para el mantenimiento. Existen varios tipos de ductos son: PVC, el cual es igual para la canalización aparente, METALICOS, NORYL, POLYCARBONATO, etc., (los últimos dos soportan temperaturas arriba de los 125°C).

El sistema modular de cableado de comunicación permite conducir cables para voz, datos, video, fibra óptica y electricidad en canales independientes y cuenta con toda la gama de conectores RJ-11, RJ-45, F. TWINAX. BNC. TOKEN-RING. RCA, etc., tanto en PLUG, JACK, ADAPTADOR O RECEPTACULO.

Nunca deberá conducir señal y electricidad por la misma tubería o ducto.

Cable Coaxial

El tipo de cable para la conducción de señal de datos coaxial o twinaxial provee un alto rango de inmunidad a las interferencia electromagnéticas y de radiofrecuencia, lo cual es de suma importancia en ambientes contaminados o zonas con interferencias, también alcanza distancias más grandes para la transmisión de señal en comparación con el cable de par trenzado (twisted pair).

Cable Par Trenzado (Twisted Pair)

Hoy en día el sistema de cableado estructurado ha dado las facilidades de convertir un departamento, área o edificio en inteligente, donde cada oficina cuente con los servicios de señal que necesite utilizando el cableado de par trenzado (twisted pair).

Este cableado estructurado consiste en un sistema de distribuidores donde en uno le llega las señales de voz, datos o video de los equipos o procesadores y el otro distribuidor es la concentración de todos los cables que llegan de las oficinas y en este se realiza el patcheo de la señal o servicio requeridos.

Cable de Fibra Óptica

La fibra óptica es el medio de transmisión de hoy y del futuro, es de alto grado de inmunidad a las interferencias electromagnéticas y cumple con el ancho de banda requerido para las aplicaciones de alta velocidad de datos. La fibra óptica es últimamente aplicada como medio de transmisión entre los pisos de un edificio o como BACKBONE.

3.3.3 Temperatura y Humedad

Los fabricantes de los equipos presentan en sus manuales los requerimientos ambientales para la operación de los mismos, aunque estos soportan variación de temperatura, los efectos recaen en sus componentes electrónicos cuando empiezan a degradarse y ocasionan fallas frecuentes que reduce la vida útil de los equipos.

Se requiere que el equipo de aire acondicionado para el centro de cómputo sea independiente por las características especiales como el ciclo de enfriamiento que deberá trabajar día y noche aún en invierno y las condiciones especiales de filtrado.

La alimentación eléctrica para este equipo debe ser independiente por los arranques de sus compresores que no afecten como ruido eléctrico en los equipos.

La determinación de la capacidad del equipo necesario debe estar a cargo de personal competente o técnicos de alguna empresa especializada en aire acondicionado, los que efectuarán el balance térmico correspondiente como es:

1. Para Calor Sensible.

Se determinan ganancias por vidrio, paredes, particiones, techo, plafón falso, piso, personas, iluminación, ventilación, puertas abiertas, calor disipado por las máquinas, etc.

2. Para Calor Latente.

Se determina el número de personas y la ventilación.

La inyección de aire acondicionado debe pasar íntegramente a través de las máquinas y una vez que haya pasado, será necesario que se obtenga en el

ambiente del salón una temperatura de 21ºC +/ 2ºC y una humedad relativa de 45% +/- 5%. Es necesario que el equipo tenga controles automáticos que respondan rápidamente a variaciones de +/- 1ºC y +/- 5% de humedad relativa.

Estas características de diseño también han demostrado ser de un nivel de confort bueno y aceptado par la mayoría de las personas.

Se recomienda mantener las condiciones de temperatura y humedad las 24 horas del día y los 365 días del año.

Debe tenerse en cuenta que una instalación de aire acondicionado debe proveer como mínimo el 15% de aire de renovación por hora, por el número de personas que en forma permanente consumen oxígeno y expelen anhídrido carbónico, si no se considera, al cabo de un tiempo de operación comienzan a manifestarse malestares como dolor de cabeza, cansancio o agotamiento y disminuyen en el rendimiento del personal.

No deben usarse equipos de aire acondicionado de ventana que no regulen la humedad ni filtren el aire, porque los gases de la combustión de motores y polvo son aspirado y enviado al centro de cómputo.

El polvo y gases corrosivos pueden provocar daños en el equipo, una concentración alta de gases tales como dióxido de sulfuro, dióxido de nitrógeno, ozono, gases ácidos como el cloro, asociados con procesos industriales causan corrosión y fallas en los componentes electrónicos.

Este tipo de problemas son usuales en las ciudades muy contaminadas, por lo que se debe tener en cuenta en el diseño del aire acondicionado instalar filtros dobles o de carbón activado de tal manera que forme un doble paso de filtro de aire, con objeto de evitar causarle daño a las máquinas del sistema y degradaciones en sus componentes electrónicos. Todos los filtros que se usen no deberán contener materiales combustibles.

Una alta humedad relativa puede causar alimentación de papel impropio, accionamiento indebido de los detectores de humo e incendio, falta de confort para el operador y condensación sobre ventanas y paredes cuando las temperaturas exteriores son inferiores a las del centro de cómputo.

Una baja humedad relativa crea la facilidad para que con el movimiento de personas, sillas rodantes, papel y mobiliarios generen la electricidad estática.

El mejor método de distribución de aire para el centro de computo es el de usar el piso falso para la salida de aire y el plafón falso para el retorno mismo. Debe preverse una renovación de aire mayor al 15 %.

3.3.4 Prevención de inundaciones

Los centros de cableado tienen que estar libre de cualquier amenaza de inundación. No debe existir riesgo de ingreso de agua, por esto deben estar alejados de regaderas y tuberías, en caso de existir algún peligro de inundación, se debe proporcionar drenaje de piso. De haber regaderas contra incendio, estas deben ser de espuma o de CO_2, con el fin de prevenir el incendio y la integridad de los equipos.

También se deben tener en cuenta las siguientes recomendaciones como prevención:

a) Si el centro de cómputo en la planta baja o en el sótano, es importante que se considere y elimine cualquier posibilidad de inundación.

b) Eleve 20cm. su piso normal y verifique que en el área y sus alrededores haya buen sistema de drenaje y que este funcione adecuadamente.

c) Coloque una protección adicional en las puertas a fin de evitar que se introduzca en el agua, en caso de que ésta subiera arriba de los 20cm o hasta el nivel del piso falso.

3.3.5 Sistema de prevención y supresión de fuego

Estas son algunas consideraciones que se deben tener en cuenta para prevenir y combatir el fuego.

a) La mejor prevención contra incendios consiste en emplear materiales no combustibles o en su defecto, tratarlos con pinturas, impregnaciones u otros que impidan o retarden su inflamación.

b) Debe instalarse un sistema de detección de humo e incendio distribuido por toda el área, tanto debajo del piso falso, en las salidas de aire acondicionado, en el falso plafón como las visibles en el techo. Este sistema de detección debe activar una alarma, la que avisara al personal para efectuar el plan de contingencia ya establecido.

c) Deben emplearse suficientes extintores portátiles de bióxido de carbono. Este es el agente recomendado para el equipo eléctrico (fuego clase "C"). La ubicación de los extinguidores debe estar marcada en el techo y ser accesible a las personas que trabajan en el área. Además, deben poder ser retirados con facilidad en caso de necesidad. Estos extintores deben ser inspeccionados una vez por año como mínimo y las instrucciones para su uso deben ser colocadas al lado de los mismos e incluidas en el programa de seguridad.

d) Es aconsejable colocar una boca de agua con manguera a una distancia efectiva del centro de proceso, como agente extintor secundario para escritorios, sillas, muebles, etc. (fuego clase "A").

3.3.6 Iluminación, piso falso, acústica y tomacorrientes

Iluminación

Es muy importante contar con buena iluminación en toda el área, que facilite la operación de los equipos y para el mantenimiento de los mismos. Para evitar la fatiga de la vista es necesario instalar lámparas fluorescentes blancas compatibles con la luz del día y pintar la oficina con colores tenues y el techo blanco para activar la reflexión.

Debe evitarse que lleguen los rayos directos del sol, para observar con claridad las distintas luces y señales de la consola y tableros indicadores de los equipos. Los circuitos de iluminación no se deben tomar del mismo tablero eléctrico que para alimentar los equipos de cómputo. El nivel de iluminación corresponde a 40 watts por metro cuadrado de superficie de salón, usando lámparas fluorescentes.

Se deberá colocar un interruptor de pared que controle la iluminación principal de la habitación en la parte interna, cerca a la puerta.

Se recomienda el uso de luces de emergencia alimentadas del UPS (Uninterruptible Power Supply) o con baterías, que automáticamente se encienden ante una falta de energía eléctrica comercial.

Piso Falso

El piso falso da la facilidad de distribuir el aire acondicionado de una manera más eficiente para el enfriamiento de los equipos, ocultar el cableado de instalación eléctrica y distribuir el cableado de sentido a las necesidades requeridas así como

sus cambios de posición y mantenimientos. Se pueden mencionar algunas de las ventajas al usar el piso falso:

a) Permite un espacio entre el piso real y el piso falso, que se puede usar como cámara plena para el aire acondicionado, facilita la distribución y salida del mismo donde se requiera.

b) Proveer una superficie uniforme y plana que cubra todos los cables de señal de interconexión, cajas, cables y bocas de alimentación de energía eléctrica, líneas telefónicas y de comunicaciones, etc.

c) Permite cambios de distribución de los equipos o ampliaciones de los mismos con el mínimo de costo y tiempo.

d) Es construido por paneles antiestáticos por una densa barrera termo acústica, envuelto con lámina electro galvanizada, proporcionando solidez para un soporte de cargas óptimo resistente a la humedad y al fuego.

e) La base guarda uniformidad estructural para soportar cargas distribuidas en un área mínima de 40cm cuadrados.

f) Los pisos falsos metálicos, presentan la facilidad de ser conectados a tierra en diferentes puntos, lo cual ayuda a descargar la estática que se produce en las superficies.

El piso falso debe ser de módulos intercambiables de 61x61 cm. Y pueden ser construidos de acero, aluminio, hierro, etc. En el caso de los pisos de madera, la parte inferior de las losas deberá quedar recubierta con una lámina metálica, de tal forma que al descansar sobre los pedestales la placa haga contacto físico y forme un plano de tierra elevado, que facilite la descarga electrostática. Esto implica que los pedestales deberán ser conectados a tierra, lo cual se comprobará previamente a la instalación de sistema.

La carga de algunos equipos en sus puntos de apoyo puede ser de hasta 455 kg (1000 lbs.), por lo que le piso falso debe ser capaz de soportar cargas concentradas de 455 kg en cualquier punto con una máxima deflexión de 2mm.

Si el espacio entre el piso real y el piso falso se usa como cámara plena, es necesario que tanto el piso firme como las paredes que limitan la cámara no desprendan polvo en absoluto y deberá estar sellado lo más herméticamente posible, para evitar fugas de aire o para evitar que entre polvo y basura. Es necesario un escalón o rampa de acceso al centro de cómputo para igualar los niveles de piso, por seguridad el escalón o rampa deberá ser del mismo material del piso falso y estar recubierta con hule estriado perpendicular a la dirección de circulación o acceso, y en caso de la rampa tener una elevación menor de 12º.

Acústica

El total del nivel de ruido en el centro de cómputo, es acumulado por todos los ruidos del salón es afectado por los arranques físicos de los motores de los equipos y los movimientos en la operación. Para proveer una mayor eficiencia y una operación confortable, se recomienda aplicar material acústico en paredes y techos del salón, como son texturas a base de tirol o recubrimientos de enjarres.

Tomacorrientes

Se debe tener como mínimo dos receptáculos para tomacorrientes dúplex de CA, dedicados, no conmutados, ubicados cada uno en circuitos separados. También se debe contar por lo menos con un tomacorriente dúplex ubicado a 15 cm por encima del piso.

3.3.7 Acceso y seguridad física

Consideraciones que se den tener en cuenta para el acceso y la seguridad del centro de cómputo:

a) El centro de cómputo debe tener una sola entrada para controlar el acceso a la instalación. Las puertas adicionales para salida de emergencia sólo podrán ser abiertas desde adentro y deberán estar siempre cerradas. Esta puerta de acceso única, permitirá tener un mejor control del paso al centro de cómputo, tanto del personal como visitantes.

b) Dependiendo de factores tales como el edificio en donde está instalado el centro de cómputo y que este albergue otras funciones, es primordial el hecho de evitar el libre acceso a áreas restringidas. La identificación de las personas deberá ser total, antes de permitirles el paso hacia áreas más críticas.

c) Excepto para el personal de servicio, no se debe permitir que cualquier visitante tenga acceso al centro de cómputo o sus alrededores. Si esto es requerido o necesario, dicho visitante deberá ser acompañado por el personal responsable autorizado o de vigilancia durante su permanencia en área. Tanto el personal de servicio como los visitantes deberán ser llamados para revisión de cualquier objeto de mano que pretendan introducir al área restringida como: maletas, bolsas, portafolios, bultos, etc.

d) El acceso puede ser mejor controlado por medio de cerraduras electromecánicas operadas a control remoto, previa identificación de la persona. Existen cerraduras eléctricas que se pueden abrir con tarjetas magnéticas programables o tableros de control con password (clave de acceso), cuya clave puede ser cambiada periódicamente y es posible registrar automáticamente las

entradas, intentos de violación e inferir cuando se está haciendo mal uso de una clave confidencial.

e) También existe dispositivos de monitoreo a base de cámaras de T.V. en circuito cerrado, de modo que una persona de vigilancia pueda estar checando simultáneamente todas aquellas áreas que son de fácil acceso desde el exterior del edificio y poder notificar oportunamente al vigilante más cercano sobre lo que considera sospechoso y que es necesario interceptar. Una comunicación directa entre todos los puntos de vigilancia mencionados y el puesto de monitoreo, es indispensable.

f) La vigilancia personal es de los mejores medios de seguridad por lo que el personal deberá ser instruido para que vigile a cualquier persona que no conozca y que se encuentre dentro de la instalación y que en adición sepa que no está autorizada para permanecer ahí. Cuando menos una persona de cada turno deberá ser asignada como responsable de la seguridad interna.

3.3.8 Mantenimiento preventivo

Es muy importante saber las condiciones de operación de los equipos y prevenir riesgos y efectos de problemas que puedan afectar la operación de los mismos, por lo que se recomienda que periódicamente se elaboren los calendarios y se realice el mantenimiento preventivo oportunamente. Se recomienda contar con una póliza de mantenimiento de servicio de algún proveedor o con personal altamente capacitado para la realización del mantenimiento preventivo.

a) Se deberá revisar las especificaciones en el manual de operación de cada equipo.

b) Para el piso falso y plafones, se deberán mantener aspirados y limpios sobre todo si se usan como cámara plena de aire acondicionado para que no suelte polvo para los equipos.

e) El aire acondicionado por sus condiciones de uso que es exclusivo para el centro de cómputo y su funcionamiento es de las 24:00 hrs. del día todo el año, se requiere de un mantenimiento preventivo del compresor y filtros de aire de la manejadora, etc.

f) Los detectores de humo e incendio probarlos para que activen el sistema de alarmas y estén en condiciones de operación para cuando se requiera.

g) La instalación eléctrica, UPS´s y reguladores, verificar que proporcionen los voltajes correctos, revisar cables flojos o de falso contacto, interruptores, etc.

3.4 Conexiones eléctricas

La instalación eléctrica es un factor fundamental para la operación y seguridad de los equipos en el que se debe completar el consumo total de corriente, el calibre de los cables, la distribución efectiva de contactos, el balanceo de las cargas eléctricas y una buena tierra física.

Una mala instalación provocaría fallas frecuentes, cortos circuitos y hasta que se quemen los equipos.

La instalación eléctrica para el área de sistemas, debe ser un circuito exclusivo tomado de la sub-estación o acometida desde el punto de entrega de la empresa distribuidora de electricidad, usando cables de un solo tramo, sin amarres o conexiones intermedias. Para el cálculo de la línea se debe tomar un factor de

seguridad de 100% en el calibre de los conductores para una caída máxima de voltaje de 2%.

Se debe construir una tierra física exclusiva para esta área, la cual se conecte a través de un cable con cubierta aislante al centro de carga del área de cómputo.

3.4.1 Tierra física

La conexión a tierra constituye el voltaje de referencia, o sea, cero voltios. Todos los voltajes son mediciones de potenciales eléctricos, desde un punto con respecto a otro; normalmente son con respecto a la tierra.

El propósito de la conexión a tierra es mas por razones de seguridad, que por motivos de referencia de potencial, la idea es impedir que las partes metálicas de los equipos informáticos se carguen con voltaje y lo descarguen a través de una persona o dispositivo electrónico.

Estas son algunas consideraciones que hay que tener en cuenta para una buena conexión de tierra:

a) Se deberá elegir un jardín o lugar en donde exista humedad, en caso contrario es necesario colocar un ducto que aflore a la superficie para poder humedecer el fondo.

b) Hacer un pozo de 3 metros de profundidad y 70 centímetros de diámetro.

c) En el fondo se debe colocar una capa de 40 cm. de carbón mineral sobre la cual descansará una varilla copperwel.

d) Encima del carbón se deberá agregar una capa de sal mineral de 5 cm. y otra de pedacería de aluminio y cobre de 40 cm., cubriéndose después con tierra hasta la superficie.

e) El tablero principal para los dispositivos se debe proveer trifásico y con doble bus de tierra, (5 hilos), uno para el neutro eléctrico y otro para proveer tierra física a las maquinas.

f) Como una medida de seguridad deberá instalarse en un lugar próximo a la puerta un control para cortar la energía a todo el equipo de cómputo en cualquier situación de emergencia, y deberá estar debidamente señalizado.

g) El espacio próximo al control de interruptores debe permanecer libre de obstáculos para su fácil operación.

Todos los conductores eléctricos hacia el centro de carga de la sala deben instalarse bajo tubería metálica rígida y de diámetro adecuado, debidamente conectadas a tierra.

Los circuitos a cada unidad deben estar en tubo metálico flexible, en la proximidad de la maquina que alimentarán, para evitar trasferencia de energía radiante de los mismos, a los cables de señal de los dispositivos y por otra para evitar peligros de incendio.

Todos los interruptores deben estar debidamente rotulados para su rápida operación por parte del personal autorizado.

Para las conexiones de los contactos polarizados 125 VCA 3 hilos, debe utilizarse el código de colores:

FASE: Negro, rojo o azul, NEUTRO: Blanco o gris y TIERRA FÍSICA: Verde

Al efectuar los cálculos de la instalación eléctrica al tablero del equipo, los conductores, reguladores de tensión, interruptores termo magnético, etc., se deben calcular teniendo en cuenta la corriente de arranque de cada máquina, la cual generalmente es superior a la nominal.

Dicha corriente de arranque debe poder ser manejada sin inconvenientes, por todos los elementos constitutivos de la instalación. Se debe considerar una expansión del 50% como mínimo.

3.4.2 Reguladores de voltaje

Es indispensable la instalación de un regulador de voltaje para asegurar que no existan variaciones mayores al ±10% sobre el valor nominal especificado, que dé alta confiabilidad, protección total de la carga y rechace el ruido eléctrico proveniente de la línea comercial contaminada por motores, hornos, etc., éste deberá soportar la corriente de arranque con baja caída de tensión y estar calculado para las necesidad es del sistema y la ampliación futura que se estime necesaria.

La regulación debe ser rápida efectuando la corrección para cualquier variación de voltaje o de carga entre 1 y 6 ciclos.

Las variaciones que soportan los equipos son las siguientes:

Tolerancia de voltaje 115 volts +10% -10%, 208 volts

Tolerancia de frecuencia 60 Hz +-1/2 Hz, +6% -8%

Se requiere instalar un arrancador electromagnético con estación de botones, para proteger los equipos que no estén soportados por el UPS, de sobretensiones al momento de cortes de energía momentáneos y que estén únicamente con regulador de voltaje, el cual al momento de cualquier corte eléctrico, des-energizará los equipos y cuando regrese la corriente eléctrica, no entrará de lleno a los mismos si no hasta que una persona active el botón de arranque.

3.4.3 Utilización de UPS

Para proteger de fallas de energía eléctrica comercial y evitar pérdida de información y tiempo en los procesos de cómputo de los equipos, se requiere de una UPS la cual abastezca eléctricamente como mínimo a los equipos del backbone.

El uso de una fuente interrumpida de energía evita fallas en los sistemas de cómputo entregando una tensión:

a) De amplitud y frecuencia controlada.

b) Sin picos ni ciclos faltantes.

c) En fase y redundante con la línea externa, independiente del comportamiento de la red comercial.

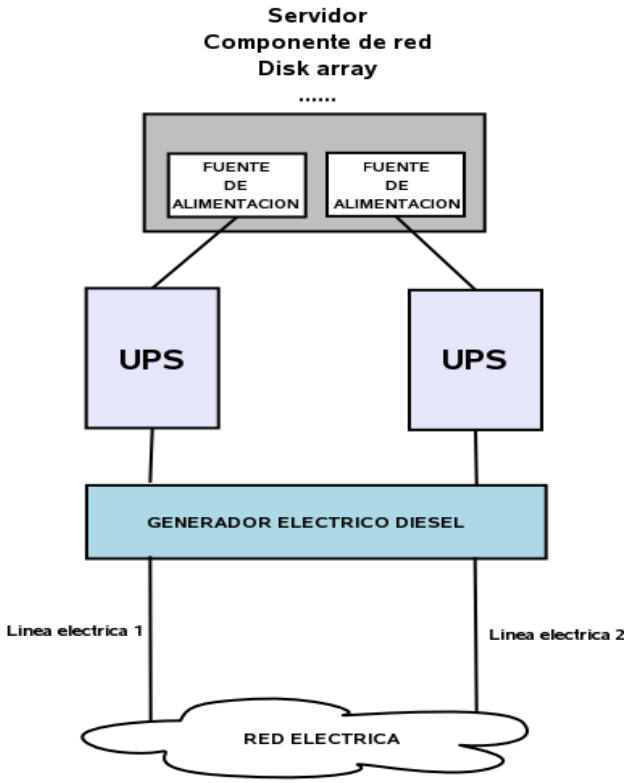


Figura 17. Redundancia eléctrica[14]

La UPS en condiciones normales de energía comercial funciona como un regulador de voltaje, y en una baja o corte de energía, entra la carga de las baterías (Battery Backup) de un modo sincronizado que le es transparente al funcionamiento de los equipos.

Una vez restablecida la energía, las baterías se recargan automáticamente.

La figura 17 Muestra redundancia en la energía eléctrica ya que tiene dos fuentes independientes cada una conectada a una UPS, para garantizar el fluido eléctrico en la red.

3.4.4 Dispositivos Pararrayos

Los rayos representan una de las fuentes primarias de los disturbios, es por eso que se hace necesario que toda la red posea este tipo de protección.

Los tipos de pararrayos mas comunes son los de tipo punta Franklin, Ionizados y Radioactivos. Estos últimos se encuentran prohibidos por leyes de protección ambiental.

Los pararrayos Ionizados operan ionizando el medio circundante, facilitando de esta forma la descarga de rayos con voltajes que la conexión a tierra pueda manejar. Por esta razón se recomiendan su uso por encima que los punta Franklin a pesar de ser mas costosos.

[14] Tomado de: http://www.linux-es.org/node/211

3.5 Redundancia en la red

De nada sirve tener servidores con componentes duplicados y redundantes y un suministro eléctrico constante y equilibrado si algunos de los componentes de la red fallan y no podemos acceder al servidor.

Los componentes más normales en una red son:

- **Routers (enrutador)**: Es un dispositivo que interconecta segmentos de red o redes enteras
- **Switch (Conmutador)**: Es un dispositivo que interconecta dos o más segmentos de red
- **Tarjeta de red o NIC**: Es un dispositivo electrónico que permite a una DTE (Data Terminal Equipment), ordenador o impresora, acceder a una red y compartir recursos
- **Cables de red**: Para interconectar los diferentes componentes, existen muchos y variados tipos, siendo los mas comunes el cable de par trenzado y el de fibra óptica
- **Líneas de conexión**: a la red de área amplia, WAN (por ejemplo Internet)

Cualquiera de estos componentes puede fallar, dejando al sistema incomunicado. Pero existen técnicas para evitar que esto ocurra, lo que se suele hacer es configurar la red, para que al menos existan 2 caminos diferentes entre dos componentes A y B. En la Figura 18 se tiene un esquema, en el que se puede ver como configurar una red con redundancia doble desde el servidor hasta Internet. De esta manera se puede estropear un router, un switch y una tarjeta de red a la

vez sin que perdamos conectividad. El mismo esquema se podría ampliar para tener redundancia triple o cuádruple de los componentes.

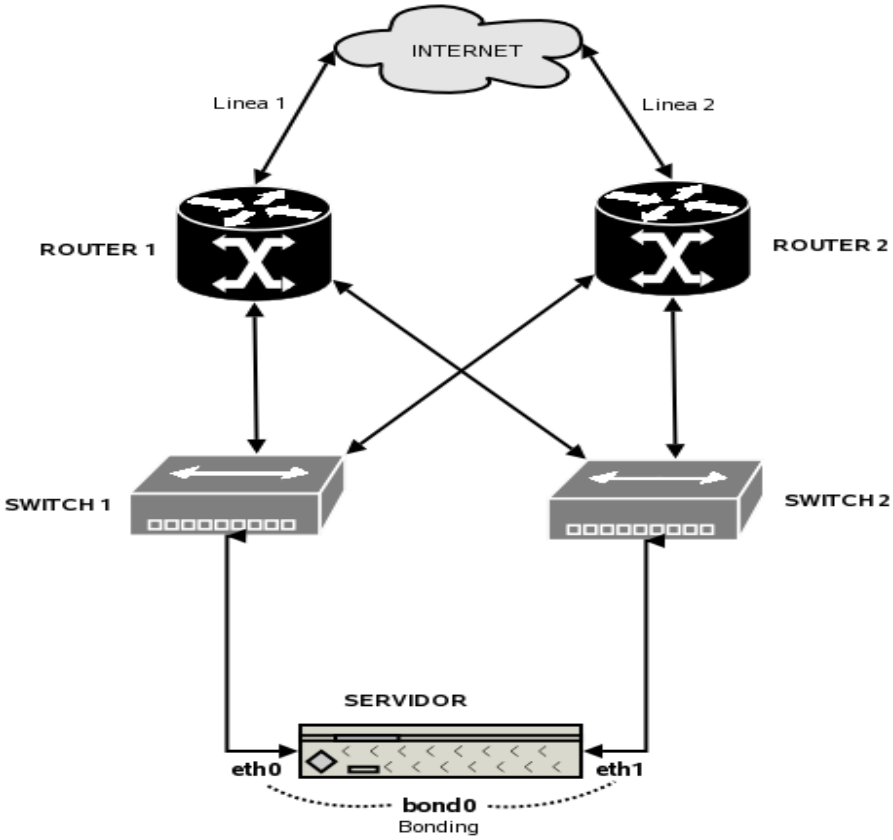


Figura 18. Red redundante[15]

3.5.1 Redundancia de equipos

Tener equipos redundantes significa poseer una copia en standby de los equipos que conforman la red, para ser reemplazado en caso de que alguno falle. Los equipos que normalmente son redundados son los que conforman el core (núcleo) de la red, como servidores, switches y routers.

[15] Tomado de: http://www.linux-es.org/node/211

Se debe poseer un stock de equipos en el sitio de trabajo, esto ofrece la ventaja de disponibilidad inmediata en caso en caso necesario de cambiar el equipo defectuoso.

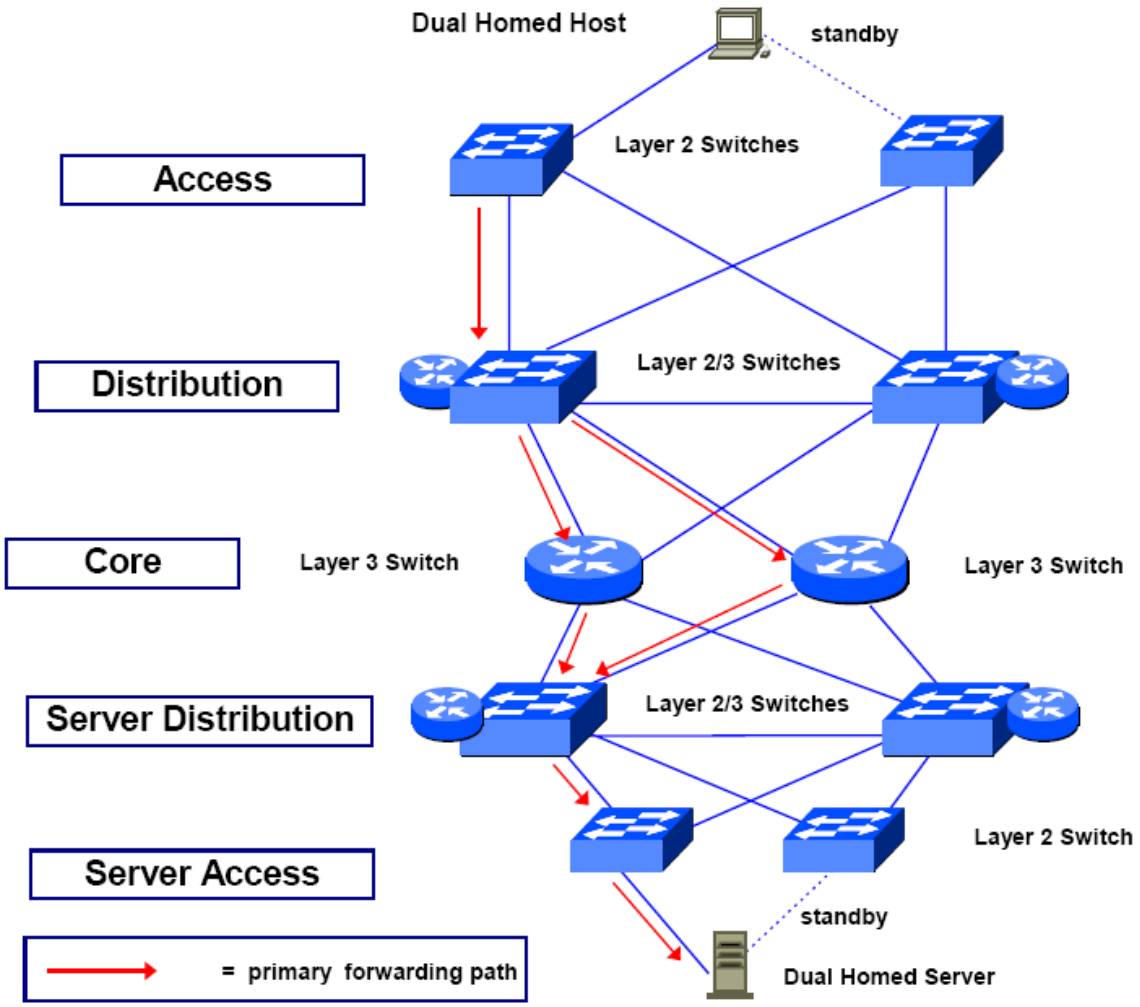


Figura 19. Red redundante sin puntos de falla[16]

[16] Tomado de: http://www.cisco.com/warp/public/779/largeent/learn/technologies/campuslan.pdf

En la figura 19 los elementos de red que proporcionan redundancia no necesitan ser colocados con los elementos de red primaria. Esto reduce la probabilidad que los problemas con el ambiente físico interrumpan el servicio.

Problemas con software virus o actualizaciones o configuración de errores y cambios puede ser tratado a menudo por separado dentro de la trayectoria primaria y las trayectorias secundarias de la ruta sin interrumpir totalmente el servicio. Por lo tanto, la redundancia del nivel de red puede también reducir el impacto de los mecanismos de falta del no-hardware.

Con las características apropiadas del resiliency, más diseño y la configuración cuidadosos, la carga del tráfico entre las capas respectivas de la topología de red (Ej. Capa de Acceso o Capa de distribución, Figura 3) puede ser compartida entre las trayectorias primarias y secundarias de la ruta. Por lo tanto, la redundancia del nivel de red puede también proporcionar funcionamiento y capacidad de creciente, qué alternadamente ayuda a reducir el costo incremental de una red redundante.

Las redes redundantes se pueden configurar automáticamente al Fail-over de primario a las instalaciones secundarias sin la intervención del operador. La duración de la interrupción del servicio es igual al tiempo que toma para que ocurra el Fail-over. Los tiempos de Fail-over más bajos que algunos segundos son posibles.

3.5.2 Redundancia de enlaces y los caminos de datos

Involucra dos formas de redundancia que dependen la una de la otra. La redundancia en los enlaces de red se refiere a la redundancia en la conectividad de los dispositivos de la red, y la redundancia en los caminos de los datos se refiere a los protocolos que determinan como los datos viajaran, sobre los enlaces físicos.

La redundancia en los enlaces se logra estableciendo más de un enlace de cada componente con el resto de la red, la forma de uso de estos enlaces esta determinada por protocolos de software: el STP (Spanning Tree Protocol)[17] y el protocolo de enrutamiento utilizado en la red.

La redundancia en el camino de los datos determina como serán enviados los datos sobre múltiples enlaces. Este tipo de redundancia tiene dos formas: resguardo en caliente, que significa que el enlace físico existe y esta listo para activarse a través del Spanning Tree Protocol; y paralelismo, el cual usa el protocolo de enrutamiento presente en la red (RIP, OSPF, etc.)

[17] El STP (Spanning Tree Protocol) es un estándar utilizado en la administración de redes, basado en el algoritmo de Árbol Abarcador, para describir como los puentes y conmutadores puedes comunicarse para evitar bucles en la red.

CONCLUSIÓN

Hoy en día, existe la necesidad de utilizar redes de comunicación para optimizar las tareas o procesos en cualquier empresa, tanto que no se puede tolerar que el funcionamiento de estas pierda su continuidad, en esto influyen diversos aspectos y técnicas que nos ayudan a mantener las redes en constante operación.

El balanceo de cargas, además de garantizar el funcionamiento continuo de las redes, también nos permite solucionar problemas de congestión en la red, ya nos que ayuda a mitigar el trafico que fluye por la red, a través de diversos métodos y algoritmos, que hacen una distribución eficiente del trafico a diferentes dispositivos, todo esto totalmente transparente al usuario final.

La redundancia es un concepto al que llegamos al hacer load balancing, esta consiste en tener fuentes o caminos alternativos que funcionan o se activan cuando los elementos principales fallan, brindándonos de esta manera una continua operación y una alta disponibilidad en los servicios.

Por ultimo y no menos importante debemos tener en cuenta las características físicas del sitio donde se encuentra nuestro centro de cómputo, tal como conexiones eléctricas, estática, cableado, condiciones de humedad, etc., ya que esta parte es la raíz de nuestras redes o sistemas de información

Se concluye que el uso de todas estas herramientas y elementos antes mencionados, nos dan como resultado: un alto rendimiento en la red, mejor ancho de nada, aumentan el buen uso de los recursos de la red, y por supuesto una **operación continúa en las redes**.

RECOMENDACIONES

Como valor agregado a esta investigación, se sugiere a averiguar e indagar en fuentes distintas a las que se proveen en este documento, pues el tema de Operación continua es muy amplio, dado por la importancia que hoy por hoy se les da a las redes en general.

Teniendo en cuenta que las aplicaciones de voz, datos y video en la actualidad lideran las cuestiones relacionadas con la operación continúa de redes, y balanceo de cargas es recomendable profundizar mucho más en estos temas por separado para tener bases mas amplias, y tener un continuo seguimiento ya que evolucionan muy rápido, debido a las exigencias que se demandan en el mercado empresarial.

BIBLIOGRAFIA

KOPPARAPU, Chandra. Load Balancing Servers, Firewalls and Caches. New York United States: John Wiley & Sons, Inc., 2002. 117p.

BOURKE, Tony. Server Load Balancing. United States of America: O'Reilly & Associates, Inc., 2001. 175p.

SHIMONSKI, Robert. Windows® Server 2003 Clustering & Load Balancing. United States of America: The McGraw-Hill Companies, Inc., 2003. 381p.

CASTRO JIMENEZ, Katerine, FERRER GOMEZ, Reymundo. Monografía Load Balancing: Balanceo de carga. Concepto, Estado del arte y Aplicabilidad en Linux y Windows. Cartagena Colombia: Universidad Tecnológica de Bolívar, 2004.

MSc. Ing. Lilliam Pajés Mora. Conferencia: 8 elementos a considerar en el diseño de redes. UMSS Cochabanba .Marzo 2007.

El Server [En línea]

<http://www.elserver.com/blog/2008/04/09/dynamic-weighted-sticky-session-load-balancing/.> Dynamic-Weighted Sticky-Session Load Balancing. Por Joel Chornik el 9 Abr 2008. (Consulta: 12 de Julio de 2008)

Tejedores del WEB [En línea]

<http://www.tejedoresdelweb.com/w/Balance_de_carga> (Consulta: 15 de Julio de 2008)

Scribd [En línea]

<http://www.scribd.com/doc/3081866/Administracion-de-Centros-de-Computo> (Consulta: Julio 11 de 2008)

OreillyNet [En línea]

http://www.oreillynet.com/pub/a/oreilly/networking/news/bourke_1100.html. (Consulta: 15 de Julio de 2008)

GerenciaIT

http://gerenciait.com/links/articulos/ci_links_temacentral_abril_p2.htm (Consulta: 12 de Julio de 2008)

Linux-es

http://www.linux-es.org/node/211 (Consulta: 26 de Julio de 2008)

Cisco

<http://www.cisco.com/warp/public/779/largeent/learn/technologies/campuslan.pdf > (Consulta: Julio 26 de 2008)

Barracuda [En Línea]

<www.barracuda.com> (Consulta: Julio 15 de 2008)

RedIris

http://www.rediris.es/mmedia/gt/gt2003_2/distvideo.ppt (Consulta Julio 29 de 2008)

Totemguard

http://www.totemguard.com/soporte/files/Continuidad_de_Negocio.ppt (Consulta Julio 29 de 2008)