PABLO D. HAUSER
SECURITY OPERATIONS CENTER, IMPSAT
ph@tectimes.com

LAS REDES DE CONTENIDO

GRANDES

Nunca les surgió la duda de cómo funciona Internet? A mí, particularmente, me ocurrió varias veces, pero terminaba pensando que era, ni más ni menos, que una gran LAN. Esto, hasta cierto punto, es verdad, pero tiene varios otros detalles técnicos y dificultades que no suelen encontrarse en cualquier red de área local.

En esta serie de artículos veremos poco a poco cómo funcionan las grandes redes de datos, cuya máxima expresión es la misma Internet. Para ello, haremos una introducción a los conceptos básicos de la comunicación.

MODELO CLASICO

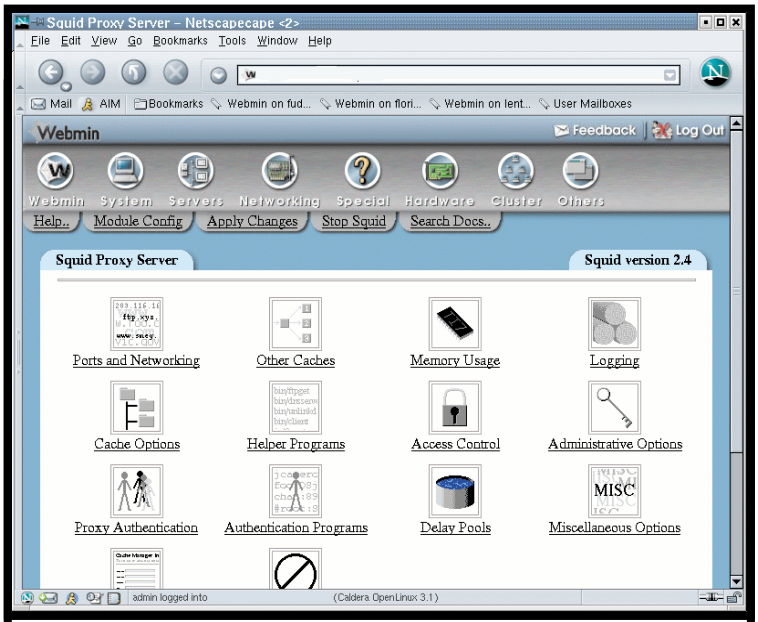
El modelo IP (de **Internet Protocol**) comprende tres capas del modelo OSI: aplicación, transporte y red. El trabajo de la capa de red es llevar los datagramas salto tras salto hasta el host de destino, especificado como dirección IP de destino. IP funciona haciendo su "mejor esfuerzo"; esto significa que no tenemos ninguna garantía de que la entrega de la información (datagramas) se realizará de la manera apropiada.

La capa de transporte provee de un servicio de comunicación end-to-end a las aplicaciones. Actualmente, se encuentran disponibles dos servicios para llevar a cabo esa comunicación: un transporte confiable, en el que la transmisión de bytes es ordenada, implementado por el **Transmission Control Protocol** (TCP); y un transporte no confiable de mensajes, implementado por el **User Datagram Protocol** (UDP).

Por sobre la capa de transporte yace la capa de aplicación, que define los formatos de los mensajes de las aplicaciones y la trayectoria de la comunicación. La Web utiliza un protocolo de aplicación cliente-servidor llamado **Hypertext Transfer Protocol** (HTTP).

Uno de los principios del diseño de la arquitectura de Internet es el "principio end-to-end", el cual indica que toda operación que pueda ser efectuada en el host de destino debe realizarse allí y no en la red en sí. Es por esa razón por la que el servicio IP es tan crudo, y las capas de aplicación y transporte son implementadas únicamente en los hosts particulares y no en todos los caminos de Internet.

Los objetos de aplicación, como páginas web, archivos, etc., se identifican con URLs (en realidad, las URLs identifican "recursos" que pueden ser mapeados a diferentes objetos, llamados "variables"). Las URLs para objetos web tienen la forma **http://host:port/path** (si el puerto no se especifica, por predefinición es el 80); esto significa que el servidor de esa aplicación está ubicado en *host* (también puede indicarse la dirección IP), atiende en el puerto *port* y conoce el objeto bajo el nombre *path* (ruta incluida). Así, las URLs, como su nombre lo indica, nos dicen dónde puede encontrarse el objeto. Para acceder a ese objeto, se abre una conexión TCP hacia el servidor donde corre la aplicación en el *host* y *port* especificado, y se solicita el objeto llamado *path*.



PANTALLA PRINCIPAL DE CONFIGURACION EN MODO GRAFICO DEL PROXY-CACHE SQUID.

INTERNET COMO RED DE CONTENIDOS

Las redes de contenidos buscan brindar acceso a objetos independientemente de su ubicación, en general porque se maneja alguna clase de replicación sobre ellos (muchas veces, esta replicación es dinámica). Sin embargo, desde su diseño, las URLs no fueron creadas para identificar objetos disponibles en varios lugares a la misma vez en la red.

Manejar esta replicación y este acceso independiente de la localización suele implicar la rotura de ese principio "end-to-end" en algún punto. Es entonces cuando la comunicación deja de manejarse end-to-end: elementos de red intermedios que operan en la capa

LAS REDES DE AREA ANCHA (WAN) TIENEN CIERTOS ASPECTOS EN COMUN CON LAS DE AREA LOCAL (LAN), A LAS QUE, QUIZAS, ESTAMOS MAS ACOSTUMBRADOS, PERO TAMBIEN CUENTAN CON ALGUNAS PARTICULARIDADES. EN UNA SERIE DE ARTICULOS TOCAREMOS VARIOS TEMAS QUE INCUMBEN A LAS GRANDES REDES. COMPRENDER COMO FUNCIONAN NOS LLEVARA A ENTENDER SU OBRA CULMINE: INTERNET, LA WAN MAS GRANDE JAMAS CREADA.

REDES

de aplicación (los tipos más conocidos de estos elementos son los proxies) intervienen en la comunicación.

De la misma manera en que los routers IP reenvían datagramas IP ruteándolos hacia su destino de acuerdo con la información de la topología de red que manejan, esos nodos que operan en la capa de aplicación reenvían mensajes de las aplicaciones, utilizando la información de la capa aplicación para decidir dónde mandarlos. Comúnmente, esto recibe el nombre de **ruteo de contenidos**.

Entonces, el éxito de una red de contenidos es manejar la replicación teniendo en cuenta dos tareas distintas: la **distribución**, que asegura la copia y la sincronización de las instancias de un objeto desde el servidor original a varios servidores de replicación; y el **redireccionamiento**, que permite a los usuarios encontrar la instancia de ese objeto que esté más cerca.

Existen varios tipos de redes de contenidos que difieren entre sí por los mecanismos que utilizan, y hay varias maneras de catalogarlas. Aquí utilizaremos la clasificación basada en **quién** es dueño y administra la red de contenidos. Así, encontramos tres tipos de redes: administradas por operadores de red, por proveedores de contenidos y por usuarios.

REDES DE CONTENIDOS DE ISPS

Los operadores de red o ISPs muy frecuentemente instalan proxies que guardan en caché las páginas web para ahorrar ancho de banda. El cliente envía sus solicitudes en busca de objetos hacia el proxy en vez de al servidor original. El proxy mantiene copias en caché de los objetos visitados (en realidad, no copia los objetos, sino las respuestas del servidor) y responde directamente desde allí; si no dispone de la información requerida, obtiene el objeto deseado y entonces sí aloja la copia para futuras respuestas y la manda al cliente.

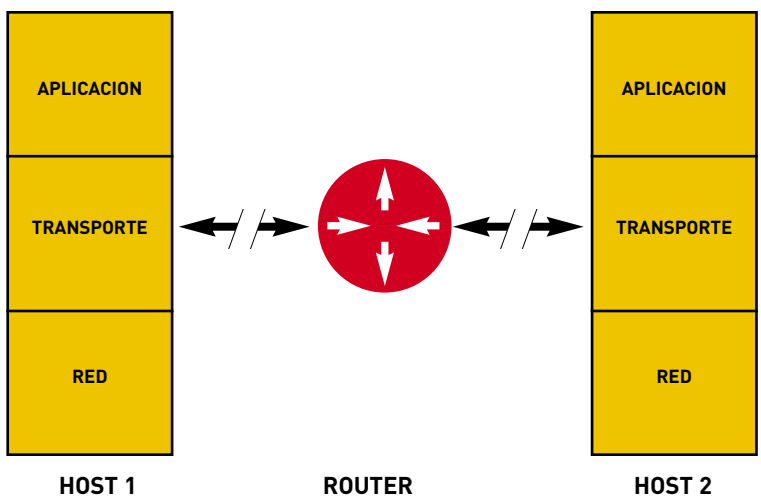
Este esquema de cacheo de proxies puede usarse de manera recursiva; esto es, haciendo que estos proxies contacten proxies "padres" frente a solicitudes que no puedan satisfacer por sí mismos desde su almacenamiento local. Estas jerarquías de proxies conllevan a la construcción de árboles de distribución de contenidos. Esto tiene sentido si la topología de red es en forma de árbol, pero presenta ciertas desventajas: los objetos no demasiado populares (con pocas solicitudes) experimentan demoras, que se incrementan dependiendo de la longitud del árbol, lo que no ocurriría si la solicitud se enviara directamente al servidor original; también sucede que el objeto respuesta originado en el proxy difiere del objeto real alojado en el servidor original, debido a malas configuraciones en la actualización de los proxies.

El proxy **Squid** (www.squid-cache.org), por ejemplo, puede ser configurado para escoger el proxy padre que se consultará mediante solicitudes basadas en el nombre de dominio de la URL requerida, o para obtener el objeto directamente del servidor original. Esto permite la configuración de múltiples árboles lógicos en el grupo de proxies (una forma limitada de ruteo de contenidos).

Los mismos efectos pueden observarse en forma dinámica, utilizando ICP (**Internet Cache Protocol**). ICP permite que un grupo de proxies cooperen intercambiando datos acerca de los objetos que alojan en su caché. De esta manera, un proxy que no posea un objeto puede encontrarlo en otro proxy cercano. Una funcionalidad avanzada de ICP permite seleccionar, entre

LAS REDES DE CONTENIDOS BUSCAN BRINDAR ACCESO A OBJETOS INDEPENDIENTEMENTE DE SU UBICACION, EN GENERAL PORQUE SE MANEJA ALGUNA CLASE DE REPLICACION SOBRE ELLOS. SIN EMBARGO, DESDE SU DISEÑO, LAS URLS NO FUERON CREADAS PARA IDENTIFICAR OBJETOS DISPONIBLES EN VARIOS LUGARES A LA MISMA VEZ EN LA RED.

FIGURA 1

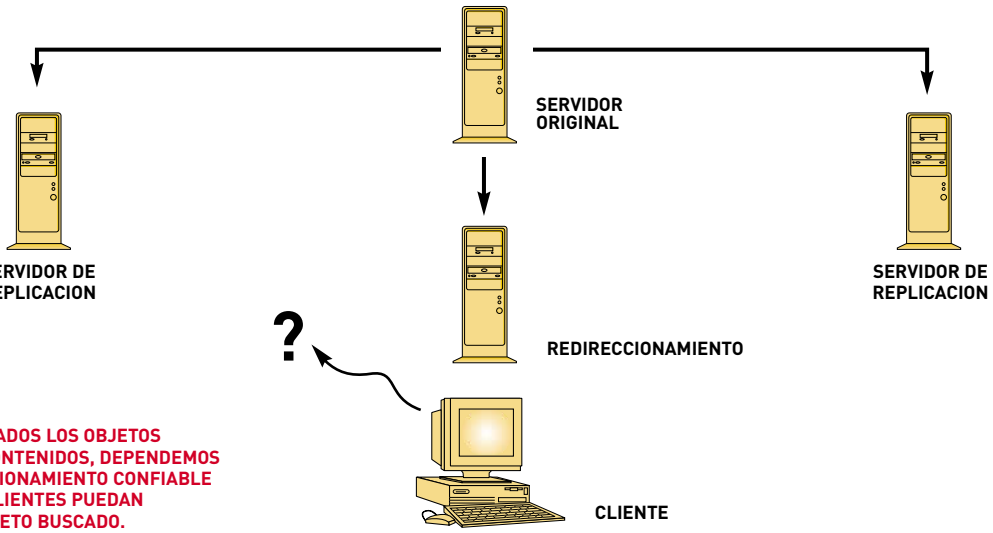


CAPAS DEL MODELO OSI UTILIZADAS POR EL INTERNET PROTOCOL (IP).

LINKS

REDES DE CONTENIDOS	www.eecs.harvard.edu/htk/publication/2002-santa-fe-kung-wu.pdf
INFORMACION SOBRE EL DISEÑO Y EL FUNCIONAMIENTO DEL WEB CACHE COORDINATION PROTOCOL (WCCP)	www.web-cache.com/Writings/Internet-Drafts/draft-forster-wrec-wccp-v1-00.txt www.web-cache.com/Writings/Internet-Drafts/draft-wilson-wrec-wccp-v2-00.txt
INFORMACION SOBRE EL DISEÑO Y EL FUNCIONAMIENTO DEL CACHE ARRAY ROUTING PROTOCOL (CARP)	www.web-cache.com/Writings/Internet-Drafts/draft-vinod-carp-v1-03.txt

FIGURA 2



UNA VEZ REPLICADOS LOS OBJETOS EN LA RED DE CONTENIDOS, DEPENDEMOS DE UN REDIRECCIONAMIENTO CONFIABLE PARA QUE LOS CLIENTES PUEDAN ACCEDER AL OBJETO BUSCADO.

un grupo de proxies, aquél que posea el **Round Trip Time** (RTT) más bajo hacia el servidor original.

Una falla del diseño de ICP es que identifica objetos con URLs. Como dijimos en un principio, una URL, en realidad, identifica un recurso que puede ser mapeado a diversos objetos llamados variables. Así, la información provista por ICP es obsoleta para recursos que tengan múltiples variables. De todos modos, en la práctica, la mayoría de los recursos tienen sólo una variable, por lo que esta falla no complica demasiado las cosas.

Los usuarios suelen configurar los navegadores para que usen un proxy, pero puede darse la configuración automática. Un cliente puede utilizar múltiples proxies por medio de protocolos como el **Cache Array Routing Protocol** (CARP). Para evitar problemas de configuración, los ISPs suelen implementar **proxies de intercepción**: elementos de red tales como routers, corriendo la funcionalidad de Cisco **Web Cache Communication Protocol** (WCCP), redirigen el tráfico HTTP al proxy de forma totalmente transparente para el usuario; el proxy, entonces, responde la solicitud del cliente fingiendo ser el servidor original. Pero si bien esto es atractivo en la teoría, puede generar numerosos problemas... ¿Qué ocurriría si el servidor destino implementara una política de seguridad tal que el solo aceptara servir contenidos a determinadas direcciones IP? Nos veríamos en la disyuntiva de eliminar el proxy, o bien de habilitar la IP del proxy como IP válida en el servidor destino; pero con esta última opción, todo aquel que llegara al proxy y emitiera una solicitud hacia el servidor securizado, obtendría la información que, en realidad, no debería. Otro problema se presenta cuando se pretende que el servidor web entregue información diferente dependiendo del cliente, ya sea por su rango IP, la fecha, la hora o cualquier otra variable enviada en la solicitud. El proxy no reflejará estos detalles y, por ende, la información obtenida tal vez no sea la que se pretendía que obtuviéramos.

Los proxies tienen soporte limitado para asegurar la consistencia de los objetos: o bien el servidor original otorga una fecha de expiración, o el proxy estima el tiempo de vida basado en la última fecha de modificación, utilizando una heurística conocida como TTL (*Time To Live*) adaptante.

REDES DE CONTENIDOS DE PROVEEDORES DE CONTENIDOS

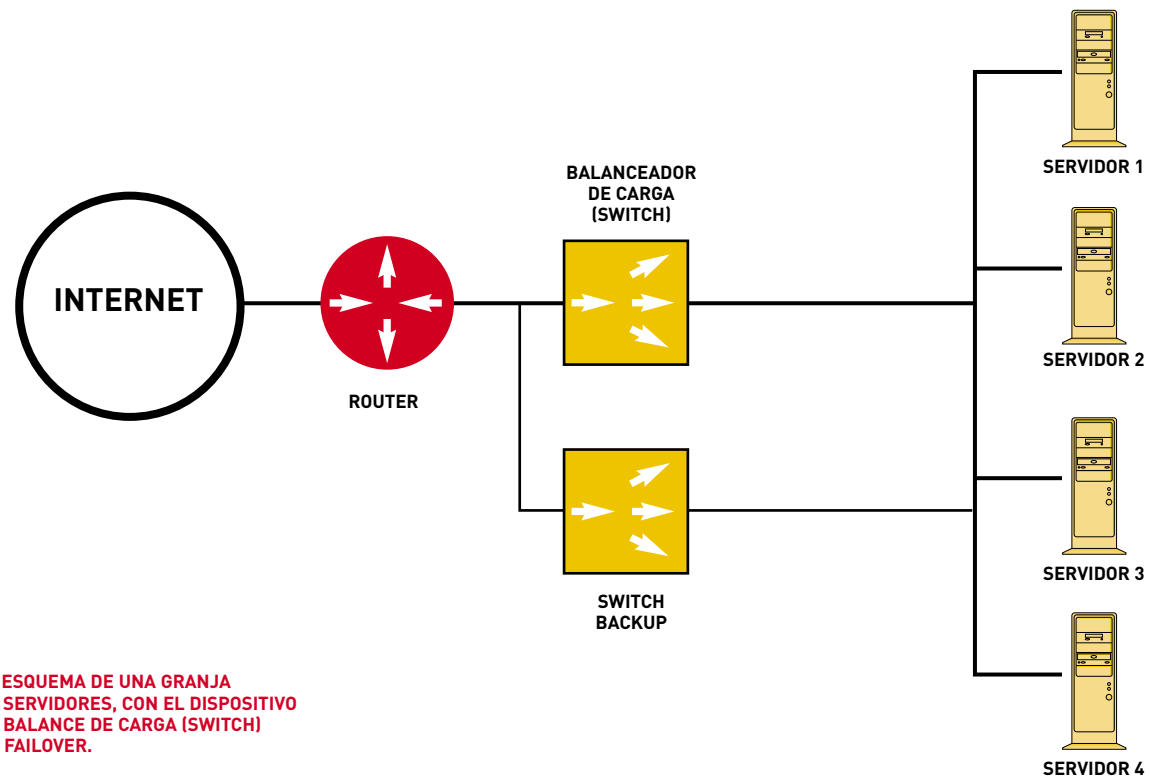
Contrariamente a los ISPs, cuyo objetivo es ahorrar ancho de banda, los proveedores de contenidos desean que aquello que tienen para ofrecer esté ampliamente disponible para los usuarios. Esta categoría se puede subdividir en tres categorías más:

- **Granjas de servidores**: redes de contenidos localmente implementadas que ayudan a proveer mayor capacidad de entrega y alta disponibilidad de los contenidos.
- **Sitios espejados (mirrors)**: redes de contenidos distribuidas que hacen que los contenidos estén disponibles en diversos lugares, de manera que los usuarios pueden elegir obtener el contenido del espejo (*mirror*) más cercano.
- **Content-Delivery Networks (CDNs)**: redes de contenidos "mutualizadas", operadas en beneficio de numerosos proveedores de contenidos, que les permite tener sus contenidos replicados en una enorme cantidad de servidores en todo el mundo, a bajos costos.

GRANJAS DE SERVIDORES

Las granjas de servidores están hechas con un dispositivo balanceador de carga (un switch) que recibe las solicitudes y las despacha hacia una serie de servidores (los servidores físicos). El sistema completo aparece al exterior como un único servidor lógico. El objetivo de la granja de servidores es brindar servicio escalable y de alta disponibilidad. El switch monitorea continuamente los servidores físicos y utiliza varias métricas de carga (algoritmo) para saber dónde realizar el despacho. Como el switch es un único punto frente a una posible falla, suele configurarse un segundo switch en failover para asegurar la alta disponibilidad en todos los puntos. Algunos switches se llaman switches de **capa cuatro** (haciendo referencia al modelo OSI, en el que la capa cuatro es la capa de transporte), lo que significa que van

FIGURA 3



EL ESQUEMA DE UNA GRANJA DE SERVIDORES, CON EL DISPOSITIVO DE BALANCE DE CARGA (SWITCH) EN FAILOVER.

armando una idea de la red que funciona alrededor de ellos a través de la información del primer paquete del intento de conexión, y deciden hacia cuál de los servidores físicos deben dirigir la conexión entrante. Asocian la conexión al servidor físico escogido y lo usan para reenviar todos los paquetes de la conexión. La forma exacta en que los paquetes son enviados hacia el servidor de destino suele variar; generalmente, se realiza cierta clase de manipulación de los encabezados de los paquetes IP y TCP (como hace el NAT, **Network Address Translation**) o encapsulamiento IP. Esta clase de "trucos" no son necesarios si todos los servidores físicos se encuentran en el mismo segmento LAN.
Otros switches más complejos, los de **capa siete** (haciendo referencia al modelo OSI, en el que la capa siete es la capa de aplicación), se fijan en la información de la capa de aplicación de los paquetes, como los encabezados de las solicitudes de URLs y HTTP. En una conexión TCP, los datos de aplicación están disponibles sólo una vez que la conexión se ha establecido. Así, una aplicación proxy que funciona sobre el switch debe aceptar la conexión del cliente, recibir la solicitud y luego abrir otra conexión hacia el servidor físico para reenviársela. Cuando la respuesta vuelve al switch, debe copiar todos los bytes recibidos de la conexión con el servidor hacia la conexión con el cliente.
Este tipo de conexiones consume muchos más recursos en el switch que la simple manipulación que ocurre en los switches de capa cuatro. Pero existen ciertas mejoras aplicadas a los dispositivos que realizan este tipo de análisis; una de ellas es solicitarle al núcleo (kernel) que efectúe el empalme entre las conexiones: después que se envía la solicitud al servidor físico, la aplicación proxy le pide al kernel que se ocupe de empalmar la conexión, y entonces se desentiende de ella. También sería posible fusionar ambas conexiones TCP, es decir, simplemente redireccionar paquetes a nivel de la capa de red para establecer una conexión TCP directa entre el cliente y el servidor. Esto requiere una manipulación compleja de los números de secuencia del paquete TCP (además de las direcciones IP y los puertos) durante el reenvío de los paquetes, ya que ambas conexiones no tendrán los mismos números de secuencia iniciales. El problema es que esta técnica puede volverse extremadamente compleja (e, incluso, imposible) si las opciones TCP difieren en ambas conexiones.

EL OBJETIVO DE LA GRANJA DE SERVIDORES ES BRINDAR SERVICIO ESCALABLE Y DE ALTA DISPONIBILIDAD. EL SWITCH MONITOREA CONTINUAMENTE LOS SERVIDORES FISICOS Y UTILIZA VARIAS METRICAS DE CARGA PARA SABER DONDE REALIZAR EL DESPACHO.

SITIOS ESPEJADOS (MIRRORS)

En algunas redes de contenidos, hay grupos de servidores que están instalados en diversos sitios en Internet, definidos como *mirrors* (espejos) del servidor maestro. La sincronización se realiza casi siempre en forma periódica (seguramente, durante la noche), utilizando herramientas especializadas como **rsync** (http://rsync.samba.org).
La redirección, en la mayoría de los casos, suele ser llevada a cabo directamente por los usuarios. El servidor maestro, al cual el usuario se conecta en primera instancia, muestra una lista de mirrors con información geográfica y la sugerencia de escoger el que se encuentre más cerca de ellos mismos. Este proceso puede automatizarse. Es posible guardar la selección del usuario en una cookie, para que la próxima vez que el usuario se conecte al servidor maestro, se realice una redirección HTTP utilizando la información guardada en la cookie. Otra manera es intentar deducir cuál es el mirror más cercano leyendo información del usuario (por ejemplo, la preferencia del idioma) o la indicada por métricas de red. Estos procedimientos no son muy comunes para sitios espejados, pero sí lo son para sitios de productos comerciales.
En cualquier caso (excepto si la redirección es automática y está basada en **Domain Name Service**, DNS), la URL del objeto cambia según el mirror.

GLOSARIO

CARP: Cache Array Routing Protocol. Protocolo utilizado para routear solicitudes URL hacia cualquier miembro del grupo de proxies establecido en el array (vector). Gracias a esto, se elimina la duplicación de contenidos cacheados.

Datagrama: Entidad independiente que incluye información suficiente para poder ser enrutado desde el equipo de origen al de destino, independientemente de los intercambios anteriores entre ambos y la red de transporte. Es el elemento en el que TCP/IP divide los archivos y otros tipos de contenido antes de enrutarlos a través de una determinada red.

DNS: Domain Name Service. Conjunto de protocolos y servicios que permite a los usuarios utilizar nombres en las URLs, en vez de tener que recordar direcciones IP.

Failover: Término genérico que se usa cuando un nodo debe asumir la responsabilidad de otro, importar sus recursos y levantar el servicio de datos. Se entiende que una situación de failover es una situación excepcional. También se entiende que el servicio de datos sigue levantado, lo cual es el objetivo de la alta disponibilidad.

ICP: Internet Cache Protocol. Protocolo orientado a datagramas para efectuar consulta entre cache proxies.

IP: Internet Protocol. Protocolo que se encarga de realizar la transferencia de los datos.

ISP: Internet Service Provider. Proveedor de servicios de Internet, empresa que brinda acceso a Internet.

Modelo OSI: Una forma amena de entender el modelo OSI se encuentra en **www.geocities.com/SouthBeach/Castle/4775/cursos/tcpip.htm**.

NAT: Network Address Translation. Herramienta que permite utilizar una misma dirección IP para conectar los sistemas de una red a Internet. En esencia, lo que hace NAT es sustituir las direcciones IP internas de los sistemas que forman parte de la red, que no son válidas para su uso en Internet, por la dirección IP externa que le proporciona el ISP (Internet Service Provider o Proveedor de servicios de Internet), que sí es válida para su uso en Internet.

Router: Dispositivo que conecta redes de comunicaciones. Dicho de otra forma, un router es una computadora especializada que resuelve problemas muy concretos de comunicaciones.

TCP: Transfer Control Protocol. Protocolo orientado a la conexión, que se encarga de contabilizar la transmisión de datos entre hosts y registrar si se presentan errores.

TTL: Time To Live. Tiempo de vida asignado a un registro.

UDP: User Datagram Protocol. Protocolo no orientado a la conexión, basado en el intercambio de datagramas.

URL: Uniform Resource Locator. Es una cadena de caracteres con la cual se asigna una dirección única a cada uno de los recursos de información disponibles en Internet. La URL de un recurso de información es su dirección en Internet, la que permite que el navegador la encuentre y la muestre de manera adecuada. Por eso, combina el nombre del servidor que proporciona la información, el directorio donde se encuentra, el nombre del archivo y el protocolo que se debe usar para recuperar los datos.

WCCP: Web Cache Coordination Protocol. Protocolo utilizado para asociar un router con uno o más cachés web (proxies) para propósitos de redireccionamiento transparente de tráfico HTTP, y para permitir que uno de los proxies dictamine cómo el router va a distribuir ese tráfico transparentemente redirigido entre el resto de los proxies asociados.

CDNS

Muchos proveedores de contenidos no pueden afrontar el costo de disponer de un alto número de servidores espejados. Los operadores de CDNs son dueños de una infraestructura de replicación enorme (Akamai, el más grande de ellos, dice poseer ¡15.000 servidores!) y cobran por distribuir los contenidos de terceros. De esta manera, pueden bajar notablemente los costos.
Los servidores CDN no guardan el sitio completo de los proveedores de contenido, sino que almacenan en caché cierta cantidad de contenido según la demanda del cliente. Manejan el almacenamiento en disco de la misma manera en que lo hacen los proxies, y sirven el contenido a los clientes como lo hacen los mirrors.
Dado que el número de estos servidores puede ser muy grande, y acompañado del argumento de que "no se necesita configuración de usuario", las CDNs incluyen complejos sistemas de redireccionamiento que les permiten realizar redirecciones automáticas y totalmente transparentes al usuario. La selección se lleva a cabo sobre la base de métricas de red en cuanto a la carga de cada servidor. El cliente puede ser conectado al servidor a través de redirecciones HTTP o utilizando el sistema de DNS: cuando intenta resolver el hostname de la URL en una dirección IP para conectarse, se le devuelve, en su lugar, la dirección IP del servidor CDN. El uso de DNS asegura que la URL es la misma para todas las copias del objeto. Es en este caso en el que los CDNs realmente transforman las URLs en identificadores independientes de la localización del objeto.

REDES DE CONTENIDOS DE USUARIOS

Las redes de contenidos operadas por usuarios son las tan queridas Peer-to-Peer (P2P). En ellas, la costosa infraestructura de replicación es reemplazada por los usuarios, quienes brindan parte de su capacidad de almacenamiento y procesamiento para que la red esté disponible. De esta manera, no se requiere nada de dinero, y nadie tiene control sobre esta red.
Una de las ventajas de las redes P2P sobre el resto de las redes de contenidos es que, normalmente, están construidas como redes independientes y no conllevan una integración con la Web. De esta manera, se puede construir una libre distribución (algunas de estas redes permiten bajar archivos desde diversos servidores en paralelo) y mecanismos de redireccionamiento e, incluso, se puede utilizar un propio nombre en ese ciberespacio en vez de estar estancado en HTTP y URLs.
Las redes P2P manejan la distribución de una manera directamente proporcional: mientras un archivo sea más popular, más copias de él habrá dispersas, y entonces habrá más copias disponibles. Por supuesto que los mecanismos para hacer esto son mucho más complejos que lo que aquí explicamos, pero ésa es la idea básica.
La parte del redireccionamiento es un tanto más problemática con la mayoría de las redes P2P actuales. Una manera de hacerlo es mediante un servidor centralizado, tal como lo hacía Napster: cada usuario se conecta primero a un servidor central, actualiza el directorio con los objetos disponibles y, luego, busca en el directorio la ubicación de los objetos que los usuarios desean obtener.
Las redes de Gnutella y Freenet, en cambio, tienen una estrategia de búsqueda descentralizada: un nodo consulta a otros vecinos que, a su vez, consultan a otros vecinos, y así sucesivamente, hasta que se encuentra uno que tenga el objeto buscado, o se consumen los recursos para esa búsqueda (por ejemplo, se acaba el tiempo para concretarla).
Aunque actualmente estas redes se utilizan para aplicaciones muy específicas destinadas al intercambio de archivos, las redes P2P aportan nuevos conceptos y técnicas realmente interesantes. Por ejemplo, Edge Delivery Network es una red comercial basada en Freenet. Incluso, algunos proyectos intentan integrar los principios de las redes P2P con la actual arquitectura y los protocolos de la Web (caso BitTorrent, **http://bitconjurer.org/BitTorrent**, y Open Content Network, **www.open-content.net**).

FINALIZANDO

El tema de las redes de contenidos es realmente muy amplio, mucho más de lo que podemos plasmar en estas páginas. Si desean profundizar en este campo, les recomendamos leer los documentos a los que hacemos referencia en la sección de links, y quedamos a la espera de todas las inquietudes que puedan surgirles. Hasta la próxima.■