

ARQUITECTURAS DESCENTRALIZADAS

Freddy L. Abad Leon



TABLA HASH DISTRIBUIDA

Una tabla de hash distribuida almacena pares una clave y un valor, estas permiten consultar el valor asociado a una clave, en las que los datos se almacenan de forma distribuida en una serie de nodos y proveen un servicio eficiente de búsqueda que encuentra el valor asociado a una clave. Mediante un sistema de enrutado que permite encontrar de forma eficiente el nodo en el cual está almacenada la información que se necesita. La responsabilidad de mantener el mapeo de las claves a los valores está distribuida entre los nodos, de forma que un cambio en el conjunto de participantes causa una cantidad mínima de interrupción. Esto da a los DHTs el poder de escalar a cantidades de nodos extremadamente grandes, y que puedan manejar constantes errores, llegadas y caídas de nodos. Estas tablas forman una infraestructura que puede ser usada para construir servicios más complejos, como sistemas de archivos distribuidos, distribución de archivos peer-to-peer, sistemas de distribución de contenido, caché web cooperativo, multicast, anycast, servicios de DNS, y mensajería instantánea. Ejemplos de redes distribuidas que usan las tablas:

- BitTorrent
- Codeen
- Coral Content Distribution Network
- Freenet
- Deluge
- Dijjer
- eMule
- FAROO
- GUNet
- JXTA
- KTorrent
- LimeWire
- NEOnet
- OneSwarm
- Overnet
- The Circle
- Transmission
- μ Torrent
- Vuze

- Warez P2P
- YaCy
- Ares Galaxy
- Twister
- GNU Ring

FUNCIONAMIENTO

En un sistema basado en una Tabla Hash Distribuida, a los elementos de datos se les asigna una llave aleatoria a partir de un identificador grande (por ejemplo, un identificador de 128 o de 160 bits). Así mismo a los nodos del sistema también se les asigna un número aleatorio a partir del propio espacio identificador. La esencia de todo sistema basado en una DHT es implementar un esquema eficiente y determinista que sólo mapee la llave de un elemento de datos hacia el identificador de un nodo, basándose en cierta distancia métrica. Al buscar un elemento de datos, la dirección de red del nodo responsable de ese elemento de datos es devuelta, mediante un enrutamiento de una petición de un elemento de datos hacia el nodo responsable.

El sistema Chord por ejemplo sus nodos están organizados lógicamente en un anillo de tal forma que un elemento de datos con llave k se mapea hacia el nodo con el identificador más pequeño $id \geq k$. El nodo se conoce como sucesor de la llave k , y se denota como $succ(k)$. Para buscar realmente un elemento de datos, una aplicación que se ejecute en un nodo cualquiera llamaría entonces a la función $LOOKUP(k)$, la cual posteriormente devolvería la dirección de la red de $succ(k)$. En ese punto, la aplicación puede contactar al nodo para obtener una copia del elemento de datos. La organización de los nodos en una red sobrepuesta, o su gestión de membresía funciona de manera que para buscar una llave cada nodo establecerá atajos hacia otros nodos, de tal manera que las búsquedas pueden realizarse en cierto número de pasos $O(\log(N))$, donde N es la cantidad de nodos que participan en la superposición. Se crea un espacio de identificadores. Mediante una función hash a cada recurso y nodo se le asigna un id. Cada nodo se encarga de un segmento del espacio, asignado según su id. Cada nodo debe saber dónde están los recursos de su segmento. Cada nodo conoce solo un subconjunto de otros nodos. Donde cada uno de los nodos hace de representante de un área del espacio. Se enrutan búsquedas a los vecinos, según el área al que pertenezca la clave de búsqueda

NAPSTER

Es un servicio de distribución de archivos de música. Fue la primera gran red P2P de intercambio creado por Sean Parker y Shawn Fanning.

FUNCIONAMIENTO

Combina tres funciones en una, las cuales son:

- Un buscador – Dedicado a encontrar solo archivos MP3.

- Compartir archivos – La habilidad de negociar con archivos MP3 directamente, sin tener que utilizar un servidor centralizado para almacenamiento.
- Un servicio IRC – Una manera de encontrar y chatear con otros usuarios interesados en archivos de este tipo.

Napster comparte archivos de manera distinta. En lugar de almacenar los ficheros en un ordenador central, las canciones son mantenidas en las máquinas de los usuarios. Esto se llama peer-to peer (P2P). Cuando se descarga una canción usando Napster, se descarga del ordenador de otra persona (cualquier persona en cualquier lugar del mundo)

GNUTELLA

Es un proyecto de software distribuido para crear un protocolo de red de distribución de archivos entre pares, sin un servidor central. Fue la primera red descentralizada de intercambio de archivos P2P. Utilizando un cliente Gnutella, los usuarios pueden buscar, descargar y cargar archivos a través de Internet. La red Gnutella es una alternativa descentralizada, a sistemas semi-centralizados como Napster.

Los clientes populares de Gnutella aun activos son: Shareaza , Zultrax P2P , y WireShare (para Windows), Apollon (para Linux). Gtk-gnutella (Windows, macOS y Linux).

Gnutella no es un sitio web, es una red descentralizada en la que los usuarios pueden ver los archivos de un pequeño número de otros miembros de la red Gnutella y, a su vez, pueden ver los archivos de otros miembros de la red, en una especie de efecto en cadena. Después de instalar y lanzar Gnutella, el equipo del usuario se convierte en un cliente y un servidor en la red, que se llama GnutellaNet. Gnutella permite a los miembros de la red compartir cualquier tipo de archivo, mientras que Napster está limitado a archivos de música MP3 .

FUNCIONAMIENTO

Gnutella es una red P2P pura. Esto es, todos los nodos tienen la misma función, peso e importancia dentro de la red. El funcionamiento de la red pasa por tres fases:

1. **Entrada.** En esta fase un nuevo nodo se conecta a otro que ya esté dentro de la red. Cómo se encuentra un nodo ya conectado está fuera del protocolo, pero normalmente los clientes Gnutella se distribuyen con una lista de nodos que se espera estén siempre conectados y se escoge alguno al azar. Un nodo cualquiera puede estar conectado a varios nodos, y recibir conexiones de nuevos nodos formando una malla aleatoria no estructurada.
2. **Búsquedas.** Cuando un nodo desea buscar un fichero, le envía un mensaje a todos los nodos a los que está conectado. Estos buscan localmente si lo ofrecen, y a la vez reenvían la búsqueda a todos los nodos a los que ellos están conectados. Esta estrategia de difusión

se llama inundación de la red, y existen mecanismos para evitar reenvíos infinitos y bucles. Cuando una petición llega a un nodo que ofrece el fichero, se contesta directamente al nodo que inició la búsqueda.

3. **Descargas.** La descarga se realiza directamente desde los nodos que contestaron a la búsqueda del fichero. Los ficheros pueden partirse en varios trozos servidos por diferentes nodos, y los clientes suelen incluir un sistema de comprobación final de la integridad del fichero.

FREENET

Es una red de distribución de información descentralizada y resistente a la censura. Freenet tiene por objeto proporcionar libertad de expresión a través de las redes de pares mediante una fuerte protección del anonimato; como parte del apoyo a la libertad de sus usuarios, Freenet es software libre.

Freenet trabaja por medio de la puesta en común del ancho de banda y el espacio de almacenamiento de los ordenadores que componen la red, permitiendo a sus usuarios publicar o recuperar distintos tipos de información anónimamente.

El funcionamiento de FreeNet es muy similar al de I2P en el sentido de que permite que sus usuarios establezcan conexiones seguras al interior de una gran red de nodos que son utilizados para enrutar tráfico entre cada uno de los participantes de la red, brindando un funcionamiento muy similar al de una VPN

cada nodo en FreeNet reserva un pequeño espacio en disco para almacenar ficheros relacionados con la red, esta zona de almacenamiento es conocida como “DataStore” cada uno de estos espacios de almacenamiento tienen asociada una clave la cual es empleada para recuperar dicho espacio para leer y/o escribir, esta zona de almacenamiento se encuentra distribuida entre sobre todos los nodos conectados en la red de Freenet

los ficheros almacenados en el DataStore se cifran y solamente pueden ser accedidos por medio de una clave para descifrarlos, los usuarios de cada instancia de FreeNet tienen poco control sobre lo que se almacena en el DataStore, los cuales pueden ser conservados o eliminados de forma automática en función de que tan populares son. De esta forma, para que algo sea borrado de FreeNet, es necesario no buscarlo y esperar que nadie más lo haga, lo cual desde un punto de vista práctico es poco viable, es por esta razón que es “resistente a la censura”.

BITTORRENT

Es un protocolo o un conjunto de normas, que establece las bases para un intercambio de archivos basado en la filosofía peer to peer. Para BitTorrent, cada ordenador es destino y fuente. Así los

usuarios no sólo descargan el archivo desde la fuente original, sino que van compartiendo partes ya descargadas entre todos. Ésta es la clave del protocolo BitTorrent y de la tecnología P2P.

Un tracker (servidor especial conocido como rastreador) se encarga de organizar la red de peers (conocida como enjambre) y distribuye a cada cliente para que se conecte con otro cliente. El tracker en esencia es el organizador de toda la distribución que contiene toda la información necesaria para que los peers se conecten entre sí utilizando el protocolo BitTorrent. El tracker así es el único punto de encuentro al que los clientes deben conectarse obligatoriamente para poder iniciar una descarga.

Cuando un cliente comienza a descargar un archivo, pide más información al tracker. Éste le proporciona una lista inicial de peers escogida al azar de entre todos los usuarios que están descargando o sirviendo el archivo. Esta lista se irá aumentando con nuevos clientes, así como también perderá a aquellos que se hayan desconectado.

Cada cliente contacta con el tracker cada cierto tiempo o cuando se desconecta para actualizar su estado, indicando la cantidad de bytes que ha descargado y subido al enjambre. Si un peer que está descargando (conocido como leecher) completa el archivo, entonces pasará a ser un seed. En general, cuantos más seeds y leechers haya en el enjambre, mayores velocidades de transferencia se tienen. El proceso tiene distintas etapas:

Primero, es necesario fraccionar el contenido del fichero en pequeños fragmentos (entre 64KB y 4MB y conocidos como piezas), después se crea un hash o identificador para cada fragmento del fichero y, finalmente, se crea un archivo con extensión .torrent que especifica la localización del tracker y la información sobre los fragmentos de los que hablábamos antes. Aunque largo, este proceso se simplifica mediante el uso de clientes, ya que es el propio software el que se encarga de todo.

El archivo .torrent se sube a un servidor web, de tal forma que esté disponible para todos los usuarios interesados en descargarlo. Para comenzar la transferencia del archivo original, sería tan sencillo como descargar el fichero *.torrent asociado y abrirlo desde algún programa o cliente preparado para tal efecto.

Con este archivo se abre con un cliente que admita este tipo de ficheros (el más conocido uTorrent). Además de descargar, permite especificar multitud de variables como la velocidad de descarga máxima o el puerto a utilizar, además de visualizar el número de usuarios conectados, los que están descargando y los que están sirviendo el archivo.

GLOBULE

Fue una red de distribución de contenido colaborativo de código abierto. Se implementa como un módulo de terceros para el Servidor HTTP Apache que permite a cualquier servidor replicar sus documentos a otros servidores de Globule. Esto puede mejorar el rendimiento del sitio, mantener

el sitio disponible para sus clientes, incluso si algunos servidores están inactivos, y en cierta medida ayudar a resistir a las multitudes de flash y el efecto Slashdot. Esta red de distribución esta descontinuada en la actualidad.

Globule se encarga de mantener la coherencia entre las réplicas, supervisar los servidores y re direccionar automáticamente a los clientes a una de las réplicas disponibles. Globule también admite la replicación de documentos PHP que acceden a bases de datos MySQL. Se ejecuta en sistemas Unix y Windows.

Funciona mediante evitar cargas elevadas en nodos individuales a través de una abstracción de indexación o tabla de hash desordenada distribuida (DSHT); Los DSHTs crean grupos de nodos autoorganizables que recogen información entre sí para evitar comunicarse con servidores más distantes o cargados.

La tabla hash *descuidada* se refiere al hecho de que Coral estaba compuesto de anillos concéntricos de tablas de hash distribuidas, cada anillo que representa una gama geográfica más amplia. Las DHT están compuestas de nodos, todos dentro de una latencia unos de otros (por ejemplo, un anillo de nodos dentro de los 20 milisegundos entre sí). Evita los puntos calientes continuando a consultar anillos progresivamente más grandes si no están sobrecargados. En otras palabras, si los dos anillos de la parte superior están experimentando demasiado tráfico, un nodo sólo ping más cerca: cuando un nodo que está sobrecargado se alcanza, la progresión hacia arriba se detiene. Esto minimiza la aparición de puntos calientes, con la desventaja de que se reduce el conocimiento del sistema como un todo.

BIBLIOGRAFIA

- I. Anónimo, “Redes entre pares”
<https://www.yumpu.com/es/document/view/15668845/tratabilidad-y-np-completitud-gsync/7>
- II. Colaboradores O&P, “¿Qué es Napster y cómo funcionaba antes?”
<http://www.ordenadores-y-portatiles.com/napster.html>
- III. Colaboradores Wikipedia, “Napster”
<https://es.wikipedia.org/wiki/Napster>
- IV. Colaboradores Wikipedia, “Gnutella”
<https://es.wikipedia.org/wiki/Gnutella>
- V. Mitchell Bradley, “Gnutella P2P Free File Sharing and Download Network”
<https://www.lifewire.com/definition-of-gnutella-818024>
- VI. Adastra, “Preservando el Anonimato y Extendiendo su Uso – Conceptos Básicos, Funcionamiento y Claves en FreeNet”
<https://thehackerway.com/2012/01/25/preservando-el-anonimato-y-extendiendo-su-uso-conceptos-basicos-funcionamiento-y-claves-en-freenet-parte-xxxvi/>
- VII. Colaboradores Wikipedia, “Globule (CDN)”

- [https://en.wikipedia.org/wiki/Globule \(CDN\)](https://en.wikipedia.org/wiki/Globule_(CDN))
- VIII. González María “Que es y cómo funciona el protocolo BitTorrent”
<https://www.xatakamovil.com/conectividad/que-es-y-como-funciona-el-protocolo-bittorrent>
- IX. Colaboradores Wikipedia, “Coral Content Distribution Network”
https://en.wikipedia.org/wiki/Coral_Content_Distribution_Network