Universidad Técnica Federico Santa María

Bases para la implementación de una red descentralizada de comunicaciones

Nombre: Cristóbal Troncoso
Rol: 2473031-K

Última revisión: 04/08/2008

Introducción

Los usuarios de Internet estamos sometidos a futuras regulaciones de acceso y contenido de la Red. Como los ISPs son compañias privadas, dueños del equipo físico necesario para enrutar el tráfico en Internet, tienen la libertad de crear sus propias reglas en el momento en que ellos vean que la utilización de Internet va en contra de sus intereses. A pesar de que la Internet esté implementada como descentralizada a un nivel de protocolo, los grandes ISPs controlan gran parte de la infraestructura física de la red.

Otro punto a tomar en cuenta es el de la privacidad de los usuarios de Internet. Si los ISPs así lo quisieran, podrían registrar la actividad total de cualquier persona. [i]

Tenemos entonces tres libertades o derechos que es necesario proteger: Libertad de acceso, libertad de contenido y privacidad de los usuarios; Neutralidad en la Red[ii], redes Tor/Onion[iii], implementaciones de Darknet[iv] y encriptación de datos son algunos de los esfuerzos para lograr conservar estos tres "pilares" fundamentales en una red libre.

Se propone una aproximación distinta al problema: crear una internet totalmente descentralizada, que no dependa de ISPs ni servidores dedicados, utilizando solo equipos clientes de los usuarios.

A través del siguiente se investigará la feasibilidad de la implementación de este tipo de red, enfrentando problemas tales como almacenamiento de los datos, enrutamiento de tráfico y encriptación.

Como queremos solo utilizar computadores personales, de preferencia portátiles, para crear una red sin ningun tipo de servidor central, veremos que tecnologías existentes tenemos para esbozar lo que nos hace falta.

Redes inalámbricas ad hoc[v]

Una red inalámbrica ad hoc es una red inalámbrica descentralizada en la que cada nodo tiene la capacidad de reenviar datos a otros nodos. Cada nodo reenvía o no los datos que le llegan de acuerdo a la conectividad de la red en el momento presente, en contraste a redes 'normales' en que los routers hacen la tarea del routeo. Dentro de lo que se considera como red inalámbrica ad hoc existe el Mesh Networking.

Topología en Malla (Mesh Networking)[vi]

La Topología en Malla es una forma de enrutar tráfico de red a través de nodos. Permite conexiones continuas y reconfiguraciones dinámicas.
La diferencia principal con otros tipos de topología de red es que es auto regenerativa (self-healing), es decir, la red puede seguir operando y enrutando datos en el caso de que un nodo se 'caiga' o no funcione. Además, la red es auto formante (self-forming). Como cada nodo es igual al otro en términos de operación, bastan solo dos nodos para que comienze a formarse la malla.
Los nodos actúan como routers para transmitir los datos de nodos cercanos a nodos destino que pueden estar demasiado alejados para ser alcanzados en un solo salto. Estas propiedades aseguran que una internet, a gran escala, formada por nodos clientes pueda 'vivir' sin importar el mal funcionamiento de unos cuantos nodos.

De manera similar a como opera Internet, los datos saltan de nodo en nodo hasta que alcanzan su destino. Es necesario implementar algoritmos de enrutamiento

en cada dispositivo-nodo para permitir el funcionamiento de la red. Cada dispositivo determina de manera independiente que hacer con los datos que recibe – pasarlos, conservarlos o desecharlos – según sea el caso. Lo ideal es que el algoritmo de ruteo pueda descubrir el camino más corto para llegar al destino.

Como se puede apreciar, una Mesh Network es lo que queremos implementar. Falta elegir un protocolo de ruteo ad hoc y ya se puede tener la fundación de lo que sería una internet descentralizada a nivel físico y a nivel de protocolos.

Ejemplo de protocolo de ruteo: Ad hoc On-Demand Distance Vector Routing (AODV)[vii]

Desarrollado en conjunto por el Centro de Investigaciones de Nokia en la Universidad de California, Santa Barbara y la Universidad de Cincinnati, AODV es un protocolo de ruteo cuya diferencia con otros es que es un protocolo de ruteo reactivo, lo que significa que establece una ruta a su destino solamente cuando se requiere (on demand). Por otro lado, este protocolo intenta disminuir el número de mensajes enviados para conservar la capacidad de la red, evitando repetir requerimientos de datos. También es importante notar que los requerimientos de ruta tienen un TTL (Time to Live), que limita el número en que estos pueden ser retransmitidos.

Teniendo una forma de establecer una arquitectura 'física' de red, y un protocolo de ruteo, es necesario fijar nuestra atención en otro problema. Si queremos utilizar nuestra internet para ver sitios web almacenados en los nodos, necesitamos tener alguna forma de DNS distribuido a través de los nodos. Afortunadamente, eso es lo que quiere lograr Thomas Bocek con su proyecto de tesis en la Universidad de Zurich.

Distributed DNS (DDNS)[viii]

El Domain name Service (DNS) es una base de datos distribuida a nivel mundial, que tiene elementos centralizados, utilizada para traducir hostnames en direcciones IP. DDNS apunta a crear un DNS distribuido y descentralizado, para lo cual utiliza una tabla hash (estructura de datos que asocia claves con valores) distribuida por medio de tecnologías peer to peer.
Algunos de los problemas a resolver por DDNS son: estabilidad, mantener el tráfico de red bajo, y por supuesto, debe seguir funcionando si se cae un nodo.
Este proyecto aún está en vías de desarrollo, a pesar de haber sido inicialmente desarrollado en el 2004.

Otro problema que surge a raíz de tener una red descentralizada es el almacenamiento de datos. ¿Que ocurre cuando necesitamos cierta información, pero el nodo que la contiene esta caído?, o ya que estamos hablando de usuarios comunes y corrientes, supongamos que el nodo simplemente está apagado: Será imposible acceder a la informacion requerida, y obviamente, no se puede forzar a los usuarios a tener sus computadores prendidos (además alejandonos del punto de una Mesh Network), ni menos almacenar el contenido de toda nuestra internet en cada computador. Exploraremos entonces algunos métodos de almacenamiento distribuido de datos.

The Google File System (GFS): Almacenamiento distribuido propietario de Google.[ix]

GFS es un sistema distribuido de almacenamiento escalable diseñado específicamente para aplicaciones donde se hacen muchos requerimientos de datos por segundo. Provee tolerancia a fallos (servidores caidos) a la vez que corre en hardware 'normal'. Un sistema que corre sobre GFS es Bigtable. Bigtable es un sistema de almacenamiento distribuido para datos estructurados en una escala muy grande. Por lo general son petabytes de datos esparcidos a través de

servidores comunes. Muchos proyectos en Google utilizan Bigtable, como el mismo buscador Google, y Google Earth.

Mnet[x]

Mnet es un serivicio de almacenamiento distribuido P2P, sin patrocinamiento comercial. La única aplicación escrita que utiliza el framework Mnet es un software para compartir archivos, similar a Bittorrent en su objetivo.
Durante el proceso de publicación de archivos, los datos son encriptados y divididos en pequeños bloques redundantes. Utilizando un algoritmo de dispersión de información, los datos son subidos a otros nodos que estuvieses corriendo la aplicación.

Podemos ver entonces, que para tener una internet, cada usuario debe destinar un trozo, quizás grande, de espacio de su disco duro. Probablemente también existan conexiones constantes para poder mantener la estructura de red, DNS, y archivos.

Surge la necesidad de implementar algoritmos de encriptación de los datos que viajarán a través de las rutas establecidas, debido a que toda la información de nuestra red pasará a traves de computadores personales. Sería una imprudencia dejar esta información a merced de cualquier persona que quiera espiar los datos que están siendo transmitidos.

Encripción de datos: GnuPG[xi]

A pesar de que GnuPG está orientado a encriptar mensajes de correo electrónico, es posible utilizarlo para la encriptación de datos.
GnuPG encripta mensajes utilizando pares de llaves asimétricas generadas por usuarios de GnuPG. Las llaves públicas pueden ser intercambiadas por los usuarios a través de un método seguro. Los datos son entonces encriptados

usando claves, por ejemplo, de 128 bits. El gran número de operaciones [2^128] requeridas para probar todas las posibles llaves de 128 bits utilizando un ataque de fuerza bruta está fuera de alcance, al menos para el futuro predecible.

Conclusiones

Finalmente, ¿Es posible crear una internet tipo Internet solamente con nodos independientes? Al parecer la respuesta es sí. Utilizando solo computadoras personales es posible crear Mesh Networks. Los demás problemas, como la posibilidad real de un Distributed DNS finalmente quedan como problemas de software. Complejos modelos matemáticos y de estadística son necesarios para implementar estos algoritmos. Así mismo ocurre con los problemas del almacenamiento de archivos. Google ha logrado con éxito distribuir muchos de sus datos a través de servidores, pero en nuestro caso estamos hablando de computadores personales de mediano rendimiento, con poder de cómputo mucho mas humilde.

Si llegara a funcionar un 'Internet' paralelo basado solamente en laptops, probablemente su desempeño no sería el ideal. Quizás, no podamos ver videos en streaming como ya estamos acostumbrados, pero el objetivo no es ese. Como se mencionó al comienzo, se quiere tener una red de comunicación de emergencia, o red libre, en el caso de que las restricciones que impongan los gobiernos o los ISPs nos dejen con una Internet que se parezca mas a un televisor. Probablemente, a medida que se desarrollen nuevas tecnologías de transmisión de datos de largo alcance, como WiMax para computadores portátiles, desarrolladores de software trabajarán en estos problemas de una manera profesional, con patrocinadores comerciales como corresponde.

i http://www.reddit.com/comments/6r6an/the_nsa_records_all_traffic_on_the_internet/
ii http://www.google.com/help/netneutrality.html
iii http://www.torproject.org/
iv http://en.wikipedia.org/wiki/Darknet
v http://en.wikipedia.org/wiki/Wireless_ad-hoc_network
vi http://en.wikipedia.org/wiki/Mesh_networking
vii http://moment.cs.ucsb.edu/AODV/
viii http://distributeddns.sourceforge.net/
ix http://labs.google.com/papers/gfs.html
x http://en.wikipedia.org/wiki/Mnet
xi http://www.gnupg.org/