

22 de Diciembre de 2017

Informe de Proyecto ELO-323: "BIRD Internet Routing Daemon"

Carlos Antinopai Araya 201103003-6

José Rojel Díaz 201123008-6

INTRODUCCIÓN

Nuestra motivación parte en nuestro interés por el servicio de *streaming* multimedia Netflix, y cómo abordan los problemas de diseñar y administrar una red de distribución de contenidos que sirve a más de 89 millones de clientes en todo el mundo con más de 125 millones de horas de contenido en *streaming*.

Open Connect es el nombre de esta red creada por Netflix, la cual en pocas palabras administra servidores con catálogo pre-cargado cerca de la ubicación final de sus clientes, mediante una asociación comercial y a veces tecnológica con los proveedores de servicios de internet.

BIRD es el demonio (programa) de ruteo en internet que utilizan estos servidores para adaptarse a la siempre creciente y móvil gama de dispositivos que ejecutan la aplicación Netflix, y es de nuestro interés mostrar en una simulación cómo opera.

Índice

1. Netflix
 - 1.1 Crecimiento de Netflix
 - 1.2 Open Connect
 - 1.3 Open Connect Appliances
2. BIRD
 - 2.1 ¿Cómo funciona BIRD?
 - 2.2. Simulación

1. Netflix

1.1 Crecimiento de Netflix

Netflix, Inc. Es una empresa estadounidense de entretenimiento que proporciona streaming multimedia (películas, series de televisión, documentales, etc.) a cambio de una tarifa plana mensual. Inicó con un modelo de negocio basado en el alquiler de títulos en DVD y VHS, los cuales mandaba por correo físico en territorio nacional. En la década del 2000, introdujo un servicio de streaming online, el cual a día de hoy compone la mayoría de su modelo de negocio e infraestructura. Como podemos apreciar en la Figura 1, la cantidad de tráfico que su modelo de distribución de contenidos tenía que manejar tuvo hasta la fecha un crecimiento acelerado, lo que hizo notar las falencias de un modelo no pensado para una cantidad masiva de clientes a nivel mundial: El tráfico de streaming multimedia comenzó a ser una porción importante del tráfico total de los ISPs (proveedores de servicios de internet) de sus clientes, lo que tenía un impacto en la calidad de servicio entregado; la calidad estándar de vídeo y audio subió en las últimas décadas, requiriendo mejorar la resolución, número de canales, ancho de banda y diversidad de formatos de compresión.

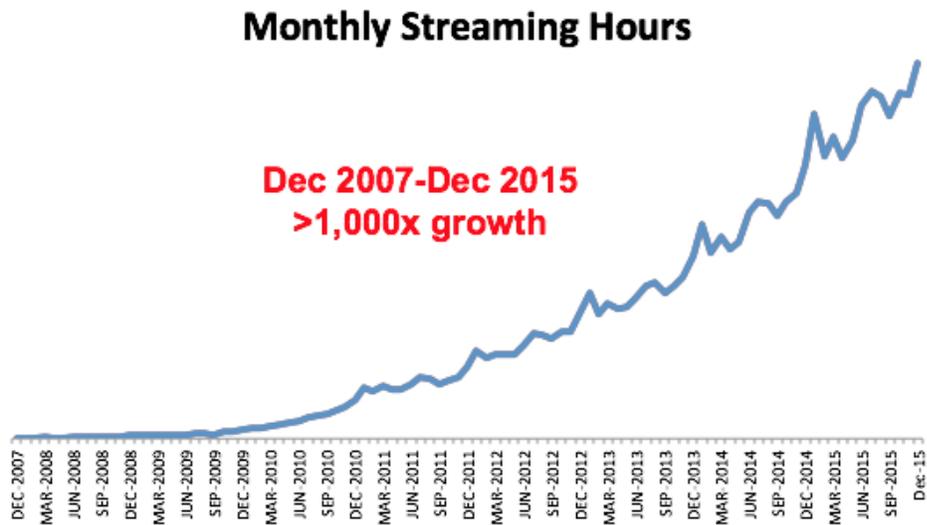


Figura 1. Horas de streaming traficadas por Netflix en promedio mensualmente.

1.2 Open Connect

Para adaptarse a la creciente e inmensa llegada de usuarios, la empresa optó por rediseñar su modelo de distribución de contenidos. De esta forma, nació Open Connect, su red de distribución de contenidos a nivel global. La idea detrás de esto es garantizar un tráfico estable de streaming a sus consumidores a la vez que tener la capacidad de ser escalable, interactivo y monitoreable.

La red se puede dividir en 2 fases principales:

A) Intercambio de tráfico con ISPs: Netflix ofrece ser quien administre el tráfico de sus clientes en lugar del ISP, transmitiendo el streaming de datos directamente desde servidores convenientemente dispuestos cerca de éstos. La Figura 2.1 muestra las 2 modalidades que ofrece Netflix a los ISPs para el intercambio de tráfico (peering). El ISP puede directamente redirigir el tráfico correspondiente a los clientes de Netflix a un determinado punto de intercambio (IX), desde donde operan servidores con contenido precargado para transmitir y actualizados periódicamente. La otra opción es instalar sin costo un servidor de Netflix directamente en los centros de datos del ISP, el cual es actualizado y cargado desde un IX.

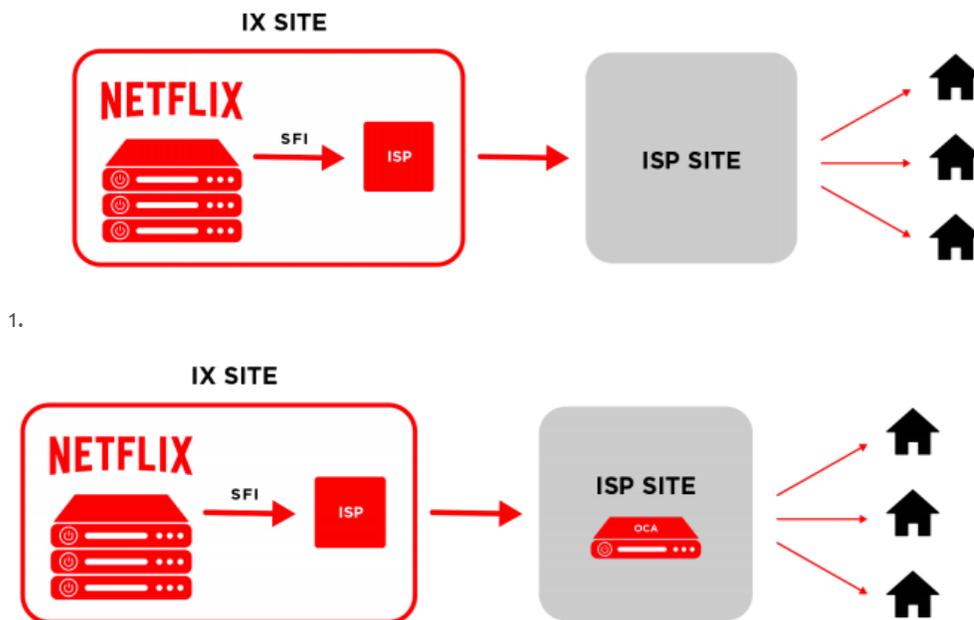


Figura 2.1. Diagrama de relación entre puntos de intercambio (IX) y proveedores de servicios de internet (ISPs).

B) Streaming y control entre usuario y servidor: La empresa utiliza los mencionados servidores sólo para la transmisión del contenido multimedia equivalente al catálogo. Todo lo que ocurre desde la aplicación Netflix del usuario entre que la inicia, navega por el catálogo y presiona el botón *reproducir* ocurre mediante una nube de Amazon Web Services (AWS). Ésta escoge el servidor adecuado para reproducir el contenido y mantiene la información del usuario. Podemos apreciar un resumen de este proceso en la Figura 2.2.

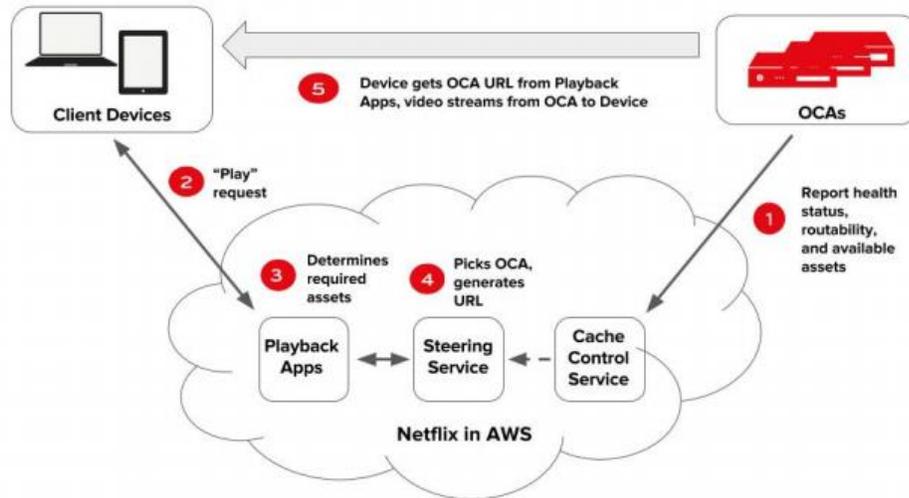


Figura 2.2. Proceso entre usuario y streaming del contenido escogido.

1.3. Open Connect Appliances

Los servidores (instalados en los IX o centros de datos de cada ISP) antes mencionados son denominados por Netflix como OCA (Open Connect Appliance) y fueron diseñados con los siguientes conceptos en mente:

- Uso de software de código abierto
- Habilidad para leer eficientemente desde el disco y escribir en los socket de red
- Envío de alto rendimiento en HTTP
- Habilidad de reunir información a través de BGP (Border Gateway Protocol)

Así, en lo que respecta a software, los OCAs tienen la siguiente composición:

Componente	Descripción	Comentarios
Sistema Operativo	FreeBSD	Colaboración conjunta para mantenerlo actualizado y libre de errores. Escogido por su robustez, comunidad programadora y costo al ser de código abierto.
Web Server	NGINX	Escogido por su escalabilidad y desempeño. El streaming de audio y video a cada cliente es transmitido via HTTP.
Routing Intelligence Proxy	BIRD internet Routing Daemon	BIRD es utilizado para permitir la transferencia de topologías de red desde las redes de los ISPs al sistema de control de Netflix que dirige a los clientes a la fuente del contenido.
IP Support	IPv4/v6	
Other		Otros softwares para administración de contenidos, salud y otras estadísticas de interés.

¿Cómo escoger la ruta óptima entre cada cliente y un OCA? Esta pregunta surge debido a que los OCA necesitan mantener constantemente actualizadas sus tablas de ruteo debido a la diversidad y volatilidad de dispositivos en los que Netflix desea ser ejecutable: Televisores, consolas, dispositivos móviles. Además, necesitan coordinarse con la nube AWS y mantener su posición relativa a otros OCA. **Éste es el interés de nuestra investigación.**

La respuesta a esto es BIRD, un demonio de ruteo dinámico en internet ejecutado en cada OCA.

2. BIRD

Es un *routing daemon* de libre acceso que realiza la tarea dinámica del ruteo en la Internet, determina las tablas de ruteo mediante una diversidad de algoritmos opcionales, las envía al kernel del SO donde se hace efectivo el envío de paquetes. Su nacimiento fue inspirado en superar las limitaciones prácticas de sus competidores. BIRD posee una arquitectura modular, la que permite realizar una configuración con relativa facilidad de sus algoritmos de ruteo, dar soporte a una amplia variedad de sistemas operativos y permitir implementar con facilidad nuevos algoritmos.

BIRD da soporte a la mayoría de los mecanismos de ruteo que existen en la actualidad, entre los más utilizados destacamos:

- IPv4 e Ipv6
- Múltiples tablas de ruteo
- Border Gateway Protocol
- Open Shortest Path First (OSPF)

2.1 ¿Cómo funciona BIRD?

BIRD tiene soporte nativo para múltiples protocolos los cuales se pueden implementar en paralelo en cada terminal y múltiples tablas de ruteo. Los filtros de ruteo son programables por scripts ejecutables de código transparente e intuitivo, en vez de las usuales listas de acceso. Los archivos de configuración son claros y estructurados, además de permitir implementarlos sobre la marcha mediante reconfiguraciones que minimizan el impacto en la red.

Una aplicación básica es usarlo en conjunto con Linux, en donde el kernel del SO actúa en el plano de datos y BIRD en el plano de control ejecutándose en segundo plano comunicándose con los routers de la periferia. Esta implementación resulta suficiente cuando el throughput de paquetes no es tan alto, para grandes cantidades soluciones de hardware se vuelven más efectivas.

Un uso particular de BIRD es usarlo como un router servidor del *Border Gateway Protocol* en puntos de intercambio de Internet, tal como lo hace Netflix en sus servidores intermedios.

2.2 Simulación

Usando **GNS3** se emuló una red de 5 routers (Figura 3.1.) que implementan BIRD. El protocolo elegido para cada nodo de la red fue Open Shortest Path First (OSPF), el cual asignándole un costo a cada enlace determina el camino de menor costo usando el algoritmo Dijkstra. Para esta tarea se comunica enviando mensajes *Hello* (Figura 3.3.) a sus vecinos a una tasa determinada por el administrador, mediante este mecanismo el nodo se puede enterar de la situación de sus nodos adyacentes y en caso de una modificación (como no recibir un mensaje *Hello* en un tiempo determinado) actualizar su tabla de ruteo para su posterior comunicación al kernel.

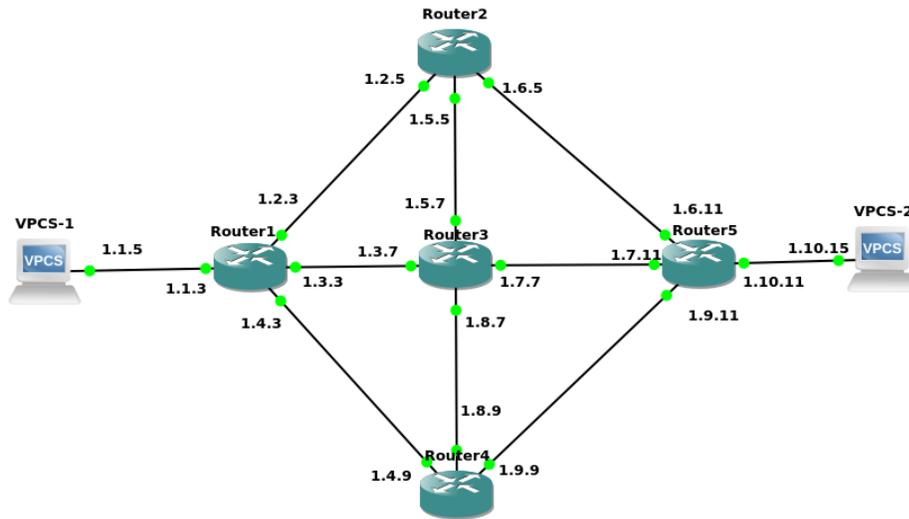


Figura 3.1. Topología de red simulada en **GNS3**

Source	Destination	Protocol	Length	Info
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet
223.1.3.7	224.0.0.5	OSPF	98	Hello Packet
223.1.3.3	224.0.0.5	OSPF	98	Hello Packet

Figura 3.2. Mensajes OSPF de reconocimiento (Antes del corte)

Cada router se configuró con el código mostrado en la Figura 3.2. en donde podemos observar los parámetros que determinan el modo de operar del algoritmo. En este caso se le asignó a cada enlace un costo de 5. Acá podemos observar que la sintaxis del archivo de la configuración es simple y clara, para cargar esta configuración al socket basta ejecutar en el terminal del SO mediante línea de comandos "*bird -c bird_ospf*".

```

protocol ospf {
  import all;
  export filter {
    ospf_metric1 = 1000;
    if source = RTS_STATIC then accept; else reject;
  };

  area 0 {
    interface "eth0", "eth1", "eth2", "eth3" {
      cost 5;
      type pointpoint;
      hello 5; retransmit 2; wait 10; dead 20;
      authentication cryptographic; password "XXXX";
    };

    interface "*" {
      cost 1000;
      stub;
    };
  };
};
- bird_ospf [ReadOnly] 32/32 100%

```

Figura 3.3. bird_OSPF: Configuración BIRD de Router 3

A continuación, para observar el comportamiento dinámico de BIRD en acción, se cortó intencionadamente el enlace entre Router 1 y Router 4, luego observamos la reacción del Router 1 observando sus tablas de ruteo antes y después de este evento (Figura 3.4.) y su secuencia de comunicación (Figura 3.5.) para actualizar su información y la tabla de ruteo posterior a la convergencia del algoritmo Dijkstra.

gns3@box:~\$ netstat -rn		netstat -rn	
Kernel IP routing table		Kernel IP routing table	
Destination	Gateway	Destination	Gateway
127.0.0.1	0.0.0.0	127.0.0.1	0.0.0.0
223.1.1.0	0.0.0.0	223.1.1.0	0.0.0.0
223.1.2.0	0.0.0.0	223.1.2.0	0.0.0.0
223.1.3.0	0.0.0.0	223.1.3.0	0.0.0.0
223.1.4.0	0.0.0.0	223.1.4.0	0.0.0.0
223.1.5.0	223.1.3.7	223.1.5.0	223.1.3.7
223.1.6.0	223.1.2.5	223.1.6.0	223.1.2.5
223.1.7.0	223.1.3.7	223.1.7.0	223.1.3.7
223.1.8.0	223.1.3.7	223.1.8.0	223.1.4.9
223.1.9.0	223.1.2.5	223.1.9.0	223.1.4.9
223.1.10.0	223.1.2.5	223.1.10.0	223.1.2.5

Figura 3.4. Tabla de ruteo antes (Izquierda) y después (Derecha).

223.1.3.3	224.0.0.5	OSPF	98 Hello Packet
223.1.3.3	223.1.3.7	OSPF	174 LS Update
223.1.3.7	223.1.3.3	OSPF	174 LS Update
223.1.3.7	224.0.0.5	OSPF	98 Hello Packet
223.1.3.3	224.0.0.5	OSPF	98 Hello Packet
223.1.3.7	223.1.3.3	OSPF	150 LS Update
223.1.3.7	223.1.3.3	OSPF	150 LS Update
223.1.3.3	223.1.3.7	OSPF	94 LS Acknowledge
223.1.3.3	223.1.3.7	OSPF	94 LS Acknowledge
00:9a:58:49...	00:9a:58:fd:9a:...	ARP	60 Who has 223.1.3.7? Tell 223.1.3.3
00:9a:58:fd...	00:9a:58:49:5a:...	ARP	60 223.1.3.7 is at 00:9a:58:fd:9a:00
223.1.3.7	224.0.0.5	OSPF	98 Hello Packet

Figura 3.5. Secuencia de intercambio ante evento.

Conclusiones

En definitiva, Netflix ha sido una pionera en la implementación y administración de redes para el streaming de datos multimedia. La complejidad de estos procesos se extiende desde la diversidad de formatos de compresión de audio y video, la interactividad del usuario, la disponibilidad del servicio en diversas plataformas estáticas y móviles hasta la administración del streaming.

Por esto, decidimos enfocarnos en este último punto, y en específico en la forma que los servidores encargados del streaming organizan y actualizan sus tablas de ruteo para escoger un cliente óptimo. Así, nuestra investigación se reduce a exponer y demostrar en funcionamiento a BIRD, el demonio ruteador dinámico de internet instalado en cada uno de estos servidores.

Así, descubrimos que BIRD es de las alternativas más utilizadas en el mercado por ser de código abierto y tener una arquitectura que le permite ser adaptable y escalable a cualquier sistema UNIX, como FreeBSD que utiliza Netflix.

Para mostrar al programa en funcionamiento, utilizamos la plataforma GNS3, dado que posee un elemento Router con BIRD instalado. Haciendo unas mínimas modificaciones a lo implementado por Netflix (protocolo OSPF en vez de BGP, router en vez de terminal FreeBSD funcionando como tal) y sin pérdida de generalidad, mostramos cómo BIRD actualiza las tablas de ruteo en un principio y luego de eliminar un enlace entre routers, lo que modificó la topología. Mediante el uso de Wireshark, pudimos observar los paquetes enviados entre 2 nodos para notificar y actualizar dichas tablas.

Mucho del funcionamiento de la red de distribución de contenidos de Netflix queda fuera del alcance de este documento, pero esperamos que haya sido suficiente para mostrar el tipo de soluciones comerciales utilizadas en la industria, y motivar al lector a expandir su espectro de soluciones a problemas de redes, que cada vez se vuelven una parte fundamental del funcionamiento normal de la sociedad.

Contacto:

carlos.antinopai@alumnos.usm.cl

Jose.rojel@alumnos.usm.cl

Referencias:

- <https://media.netflix.com/es/company-blog/how-netflix-works-with-isps-around-the-globe-to-deliver-a-great-viewing-experience>
- https://openconnect.netflix.com/es_mx/software/
- <https://openconnect.netflix.com/Open-Connect-Overview.pdf>
- http://bird.network.cz/?get_doc&v=16&f=bird-1.html#ss1.1
- <https://people.netfilter.org/pablo/netdevo.1/papers/BIRD-Internet-Routing-Daemon.pdf>