



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE ELECTRÓNICA

Proyecto ELO-322

Content Delivery Network: Cloudflare

Carlos Fernández
Javier Ortiz
Phillipa Paredes
Loreto Romero

9 de Septiembre de 2019

Resumen

Las redes de distribución de contenidos (CDN, *content delivery networks* en inglés) proporcionan los medios necesarios a los servidores de Internet para aproximarse a hosts en distintas ubicaciones. A pesar de interactuar con ellos de manera regular, la mayoría de los usuarios solo es levemente consciente de la tecnología que hay detrás y de cómo operan conjuntamente para proveer de seguridad y reducir los tiempos de espera al acceder a páginas web, en especial si dichos servidores operan a una gran distancia.

En el presente estudio, se aborda el funcionamiento de las CDN desde un punto de vista práctico y se establece relación con conceptos del área de redes de computadores. Se comienza por hablar de su origen y los problemas que busca resolver. Luego se explica su funcionamiento y la manera en que está implementado en la realidad. Por último, se hace una mirada general al mercado mundial de proveedores de CDN y se centra la atención en la empresa Cloudflare.

Acorde a lo anterior, se incluye una demostración experimental que muestra el tiempo de carga de un sitio web con el servicio de Cloudflare activado y se compara con ejecutar la misma acción con este desactivado. De la prueba, se obtiene una contundente diferencia entre ambas mediciones, lo cual aporta a la comprobación de la reducción de latencia producto de incorporar dicha plataforma.

Introducción

Tiempo atrás, era menos imprescindible el servicio de redes de distribución de contenido para servidores de sitios Web locales o pequeños. Más bien, su uso se ajustaba a servidores más masivos como Google o Youtube. Actualmente, con el fenómeno de la globalización y el rápido incremento en la cantidad de hosts con acceso a Internet, se hacen notar los efectos de la capacidad limitada de los servidores y de la demora de la información por recorrer grandes distancias.

Se hizo pertinente el mejoramiento de la eficiencia en el reparto de los archivos que componen las páginas Web, tales como imágenes o videos, junto con un acercamiento del contenido hacia puntos locales que redujesen el tiempo de descarga de estos. En respuesta a ello, se produce una expansión de las redes de distribución de contenido hacia los servidores que antes no lo requerían, promoviendo un uso global e implementando nuevas características conforme sigue creciendo el mercado.

La técnica de añadir puntos de presencia de CDN distribuidos globalmente resulta bastante efectiva, pues permite desprestigiar tiempos de espera a los usuarios en más lugares del planeta, así como también a los servidores llegar con rapidez a más clientes, reducir costos operacionales y mejorar así su servicio. Por otro lado, desde el punto de vista del distribuidor de contenidos, el entrar en el mercado tiene un alto costo pues parte de su éxito está directamente relacionado con la cantidad de puntos de presencia que disponga. Sin embargo, mientras más sitios Web apunten a usuarios lejanos, más crece la popularidad de las CDN y decrece el precio de su servicio. Al mismo tiempo, al expandirse el mercado de las CDN y tener un costo menor, más servidores buscarán adquirir su prestación.

Qué son las CDN

CDN, abreviación de *content delivery network*, se refiere a un sistema de servidores distribuidos que trabajan en conjunto con el fin de entregar contenido de manera rápida. Este servicio es utilizado por sitios como *Facebook* o *Netflix*. El utilizar este sistema puede presentar diversas ventajas si el servicio que se entrega cumple con ciertas características. La principal ventaja de las CDN es el aumento de velocidad al poder entregar contenido desde el servidor más cercano geográficamente al usuario. A grandes rasgos su funcionamiento es el siguiente:

1. Usuario realiza requerimiento a un servidor "X".
2. El requerimiento pasa por el servidor CDN POP para verificar si el contenido requerido esta actualmente almacenado en algún servidor CDN cercano.
3. Si se encuentra el contenido, el requerimiento es redirigido al servidor CDN que lo almacena y es este el que hace el envío del contenido.

Si no se hubiese encontrado el contenido del requerimiento, el servidor CDN lo habria descargado para tenerlo disponible en caso de ser necesario en futuro.

Esta ventaja del sistema CDN en cuanto a tiempo en la entrega del contenido solo se presenta si los requerimientos son realizados en un ambiente globalizado, debido a que, en caso de que fuese requerido de manera local, el utilizar este sistema no presentaría un aumentos en la velocidad, incluso, podría hacerlo mas lento debido a que tendría que buscar si el contenido se encuentra almacenado en algún servidor cercano.(1)

Las primeras CDN fueron implementadas por *Akamai Technologies*. Sus CDN son responsables de 15-30 % del tráfico global de Internet.

Cómo se implementa

Como se menciona previamente, los CDN se constituyen como servidores esparcidos físicamente por el mundo a los que se les denomina puntos de presencia (POP, *Points of Presence*). Estos servidores poseen elementos claves que los componen que se definen a continuación:

- **Delivery Nodes:** Estos nodos son los mas cercanos a los servidores del usuario final y es el encargado de repartir el contenido.
- **Storage Nodes:** Son los encargados de almacenar el contenido que entrega el distribuidor de contenido, por lo general se compone de una serie de Caches jerarquizadas.
- **Origin Nodes:** Corresponden a los servidores de originales de donde proviene el contenido a distribuir.
- **Control Nodes:** Estos nodos son encargados de mantener el estado de la red CDN, generando reportes de velocidad y de estados para poder informar al usuario.

El flujo de contenido desde el nodo de origen hasta el usuario final puede implementarse de dos maneras, mediante una estructura de *Pull CDN* o *Push CDN*.

Pull CDN refiere a que son los servidores CDN los encargados de pedir el contenido a los servidores de origen, si es que no se encuentra almacenado en el servidor, cada vez que el usuario final lo requiere, de esta forma en un principio no se percibe la mejora de tener un servicio CDN, pero esto cambia a medida que se almacenen los elementos en el caché. Por otro lado, Push CDN refiere a que es el servidor de origen el encargado de mantener una copia actualizada del contenido en los servidores CDN, por lo siempre habrá una copia vigente en el servidor cuando un usuario final haga un requerimiento.

Otro factor que corresponde a la implementación es la distribución física de los servidores CDN o topología de la red, esto se puede enfrentar mediante un *CDN distribuido* o mediante un modelo de *CDN consolidado*.

El CDN de topología distribuida se implementa mediante una gran cantidad de servidores de capacidad media/baja de manera de disminuir al mínimo las distancias físicas, este enfoque es preferido en áreas de baja conectividad, pero su mantenimiento e implementación representan grandes costos.

El CDN de topología consolidada, en cambio, busca aprovechar al máximo la conectividad e implementa pocos servidores, pero de gran capacidad, de esta forma se logra centralizar el servicio disminuyendo los costos de mantención y aumentando la seguridad de la red. El beneficio que se alcanza con este tipo de topología es en gran medida al aumento en la tecnología de conexiones como lo es la fibra óptica.

Proveedores de CDN y su presencia en Chile

Es relevante mencionar las principales compañías con presencia mundial. A partir del ranking actualizado de CDNPerf, servidor que se dedica a analizar el desempeño de proveedores comerciales de CDN, se tiene los resultados mostrados en la Figura 1 y Figura 2.(2)

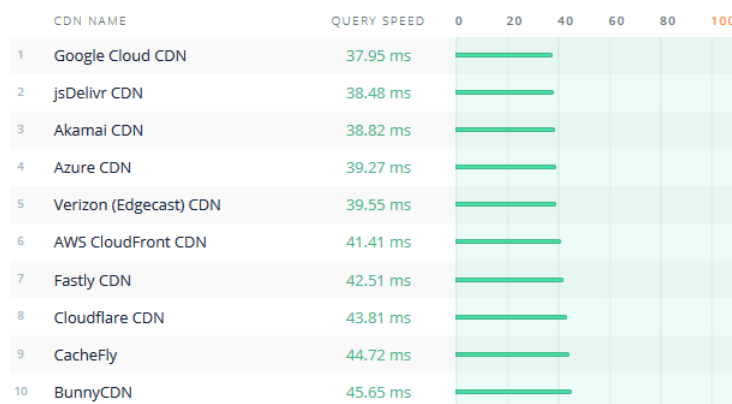


Figura 1: Ítem *Performance*, según tiempo de usuarios para acceder a determinado archivo

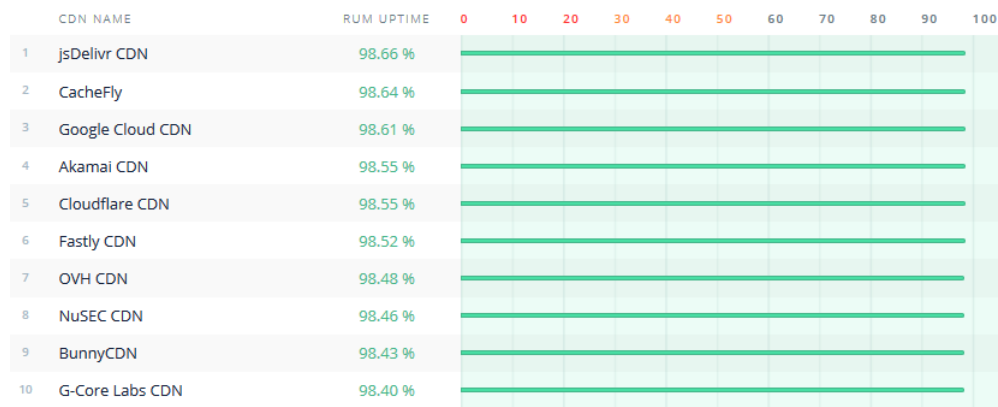


Figura 2: Ítem *Rum Uptime*, en base a calificación promedio obtenida

De las compañías anteriores, se destaca a *Akamai CDN*, líder del mercado, y a *Google Cloud CDN*, por ser pionero en la implementación de nuevas tecnologías para aumentar su eficacia y por estar presente en más de cien instalaciones de interconexión del mundo.

Entre los proveedores gratuitos de CDN mejor señalados se encuentra *Cloudflare*, empresa estadounidense que provee servicios de DNS a unos 6 millones de sitios web y con 20.000 nuevos clientes por día. Además, ofrece alta velocidad y protección a sitios web, y posee una vasta cobertura global.

Según reporta CDN Planet, se cuenta con ocho proveedores de CDN con uno o más puntos de presencia en Chile, todos ubicados en las regiones de Santiago y Valparaíso. Algunos de estos son CDN77, CDNetworks y Cloudflare.

Para empresas con clientes de diversos orígenes resulta provechoso el contar con servicios de CDN dado que este le asistirá con reducir el tiempo de llegada a cada uno de ellos. Por ejemplo, algunos de los usuarios de Cloudflare son la empresa de retail Falabella, el comercializador de software Hubspot y el sitio argentino Taringa.(4) A su vez, en Chile múltiples empresas forman parte de esa comunidad como es el caso de MediaWeb, el cual anunció abiertamente su asociación con Cloudflare mediante un comunicado en su página.

¿Qué es Cloudflare?

En el año 2010 comenzó a funcionar Cloudflare, el cual es un sistema que ofrece servicio de CDN en el que se destaca el uso de un proxy reverso. Este administrador DNS almacena temporalmente una copia de la página web a nivel de contenido estático (CSS, JS, HTML, imágenes, etc.), para luego hacer un despacho directo de forma rápida y segura a sus clientes desde uno de sus 193 centros de datos de todo el mundo. En la Figura 3, se aprecia un mapa mundial que señala todos los puntos de presencia de Cloudflare, incluyendo el primer y único proxy de Chile, ubicado en la ciudad de Valparaíso.(5)

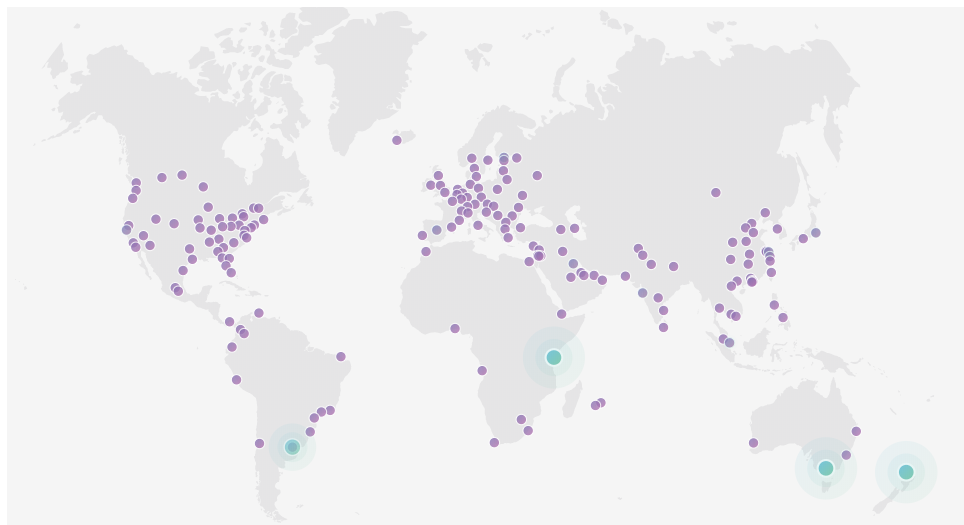


Figura 3: Mapa con puntos de presencia de Cloudflare

Con el tiempo Cloudflare fue evolucionando logrando ampliar sus funciones, obteniendo características positivas, tales como:

- Al usar CDN, la carga de la página web aumentará su velocidad drásticamente.
- Tiene protección contra ataques DDoS (del inglés, *Distributed Denial of Service*).
- Incorpora un modo *anti-hotlinking* que permite proteger imágenes y otros recursos estáticos para evitar que se pueda hacer peticiones a ellos desde sitios desconocidos.
- En el caso de que el proveedor de hosting falle, Cloudflare puede avisar a los usuarios que el sitio está offline temporalmente.
- Tiene un sistema de estadísticas que informa sobre su propio rendimiento, por ejemplo, cuántos kbs o mbs se ahorraron en la transferencia.
- Posee un VPN WARP que permite tener conexiones móviles privadas con una mayor velocidad de navegación, con una menor latencia y de forma gratuita.
- Limita spam y detecta tráfico malicioso.

A pesar de sus múltiples ventajas, no todo es perfecto. Sus clientes son a menudo muy concurridos y muchísimas empresas trabajan en mayor o menor medida con esta plataforma. El problema radica cuando Cloudflare falla y las páginas que dependen de su red deben mostrar el mensaje *502 Gateway Error*. Es decir, aunque el sitio web en cuestión funcione correctamente, su intermediario no ofrecerá el contenido que se le pide debido a un desperfecto en su propia infraestructura. La última caída registrada fue el 2 de julio de 2019, en la cual los visitantes a páginas de Cloudflare recibieron mensajes de error durante 30 minutos.

Resultados de parte práctica

Mostrando el funcionamiento de las CDNs, se utiliza una página web de prueba(www.daily-valpo.site), la cual se inscribió en los servidores de Cloudflare pudiendo acceder a un paquete gratuito del servicio CDN que ofrece la compañía. Para comprobar la disminución del RTT, se hizo *ping* a dicha página utilizando el servicio y luego se repite con el servicio pausado. En la Figura 4, se observa el ping realizado sin utilizar el servicio CDN, donde se puede ver que el RTT medio es de 167[ms] y la dirección IP a la que se dirige es 72.167.191.69, la cual corresponde a los servidores de GoDaddy, justamente donde se encuentra alojada la página de prueba.

```
λ ping daily-valpo.site
Haciendo ping a daily-valpo.site [72.167.191.69] con 32 bytes de datos:
Respuesta desde 72.167.191.69: bytes=32 tiempo=168ms TTL=242
Respuesta desde 72.167.191.69: bytes=32 tiempo=168ms TTL=242
Respuesta desde 72.167.191.69: bytes=32 tiempo=166ms TTL=242
Respuesta desde 72.167.191.69: bytes=32 tiempo=166ms TTL=242

Estadísticas de ping para 72.167.191.69:
  Paquetes: enviados = 4, recibidos = 4, perdidos = 0
  (0% perdidos),
  Tiempos aproximados de ida y vuelta en milisegundos:
  Mínimo = 166ms, Máximo = 168ms, Media = 167ms
```

Figura 4: Ping a pagina web sin servicio CDN

En la Figura 5, se observa ahora el ping realizado a la página cuando se activa el servicio CDN. El RTT medio es de 18[ms] lo que es una disminución notable en comparación al caso en el que no se usa CDN. Como se puede ver, la dirección IP a la que se dirige es 104.31.64.181, la cual esta asociada a uno de los servidores de Cloudflare.

```
λ ping daily-valpo.site
Haciendo ping a daily-valpo.site [104.31.64.181] con 32 bytes de datos:
Respuesta desde 104.31.64.181: bytes=32 tiempo=24ms TTL=54
Respuesta desde 104.31.64.181: bytes=32 tiempo=18ms TTL=54
Respuesta desde 104.31.64.181: bytes=32 tiempo=15ms TTL=54
Respuesta desde 104.31.64.181: bytes=32 tiempo=17ms TTL=54

Estadísticas de ping para 104.31.64.181:
  Paquetes: enviados = 4, recibidos = 4, perdidos = 0
  (0% perdidos),
  Tiempos aproximados de ida y vuelta en milisegundos:
  Mínimo = 15ms, Máximo = 24ms, Media = 18ms
```

Figura 5: Ping a pagina web utilizando servicio CDN

Mediante esta sencilla prueba se corroboro que el servicio estaba activo y uno de los beneficios inmediatos que presenta que es al disminución de los tiempos de respuesta de la página web.

Conclusiones

La latencia que se produce entre el instante en que se solicita una página web y el segundo en que ella finalmente se despliega ante el usuario constituye un aspecto molesto a solucionar. Frente a ello, las CDN cumplen un rol relevante, implementando puntos de acercamiento de la información que disminuyen virtualmente el retraso producto de la distancia física entre el cliente y el servidor.

De la presente investigación, lo más relevante fue notar la relevancia de la función que cumple las redes de distribución de contenido para ambas partes, pues no solo reduce el tiempo de carga del sitio de Internet y otorga protección a ataques cibernéticos sino también disminuye un gasto de recursos importante en el servidor, tales como ancho de banda y disponibilidad. Algunos de estos beneficios se pudieron observar al preparar la demostración práctica. No obstante, el resultado más concluyente fue obtener cifras apreciablemente inferiores para la reducción de la latencia por medio del uso de un proveedor de este servicio, Cloudflare.

A medida que la globalización va en aumento, más necesario se vuelve el uso de CDNs. Para el futuro, se contempla un mundo aún más conectado, con mayor cantidad de usuarios y demanda de información en tiempo real. Por tanto, queda como desafío continuar mejorando las tecnologías de comunicación actuales y añadir robustez ante posibles fallas del servicio.

Referencias

- [1] Content Delivery Network Explained. (s.f). Recuperado Septiembre 9, 2019 de <https://www.globaldots.com/content-delivery-network-explained/>
- [2] CDNPerf. (2019, Septiembre, 4). CDN Performanc Analytics & Comparison. Recuperado Septiembre 4, 2019 de <https://www.cdnperf.com/>
- [3] CDN Planet. (s.f). CDNs in Chile. Recuperado Septiembre 4, 2019 de <https://www.cdnplanet.com/geo/chile-cdn/>
- [4] Cloudflare. (s.f). Casos prácticos de clientes. Recuperado Septiembre 5, 2019 de <https://www.cloudflare.com/es-es/case-studies/>
- [5] Cloudflare. (s.f). The Cloudflare Global Anycast Network. Recuperado Septiembre 1, 2019 de <https://www.cloudflare.com/network/>