

Proposal for an Intelligent Distribution Model of DNS queries according to Time Zones

Jeisson S. Sánchez M.
Jeisson@niclabs.cl

Sandra Céspedes Umaña
scspedes@ing.uchile.cl

Abstract

There are multiple factors that affect optimal response times for web browsing and one of them is the resolution of DNS requests. Currently, at the DNS level, the latency is increasing due to servers location, malicious traffic under-scaling, among others. In this work we propose a novel DNS traffic distribution according to daylight time in specific zones. We demonstrate that DNS-related traffic congestion increases during the day light in each country. Based on these observations, we propose the traffic distribution depending on the time-zone difference to balance the load at the DNS servers and to generate new routes toward the DNS servers with less work-load at these times. The distribution is based on the analysis of a real DNS traffic dataset during a one-day period and the responses of different queries in the main server located in Santiago de Chile.

1 Introducción

Comercialmente desde el punto de vista de proveedores de servicio de Internet (ISP por sus siglas en inglés) el rendimiento de una red se traduce en millones de dólares, de manera positiva o negativa de acuerdo a los servicios recibidos por el usuario final. La principal causa de resultados negativos es determinada por la latencia, que se ha incrementado aproximadamente 30ms en tiempos de respuesta desde el 2013 hasta el 2017, principalmente en servicios Web tanto en redes de Internet fijo como móvil [Incode17]. Por ejemplo, Amazon estima que los ingresos disminuyen en un 1% por cada 100 ms de latencia, mientras que para Google

se estima un descenso del 0,74% en las búsquedas web cuando la latencia aumenta alrededor de 400 ms [Karzand17] y el tráfico disminuye en un 4% por cada 100 ms de latencia [Incode17]. Adicionalmente, el reporte en [Incode17] concreta que esas pérdidas alcanzan hasta los 580 billones de dólares en pérdidas por transacciones electrónicas en redes de Internet fijo para el año 2017, aproximadamente 70 billones más que en 2016.

En consecuencia la prestación y control de servicios de Internet orientados con alta disponibilidad, menor retraso en transmisión, tasa de errores baja, tiempos de respuesta bajos, uso eficiente de ancho de banda, entre otros aspectos son temas de interés para la investigación. Actualmente la mayoría de las investigaciones sobre la calidad de servicio (QoS por sus siglas en inglés) en Internet están enfocadas al desarrollo de nuevos estándares, protocolos y métricas. A través del tiempo se han desarrollado propuestas como la Arquitectura de Servicios Integrados de Internet (IISA por sus siglas en inglés), acuerdos de nivel de servicio (SLA por sus siglas en inglés), gestión avanzada de tráfico (ATM por sus siglas en inglés) y servicios diferenciados (DiffServ por sus siglas en inglés) para garantizar QoS en redes a gran escala. Sin embargo, muchos ISP pasan por alto la importancia de la latencia relacionada a tráfico DNS, debido a que afecta directamente la percepción QoS por parte de los usuarios finales [Akamai17]. En [Asap11] analizan la resolución de solicitudes DNS y cómo se relaciona con los tiempos de respuesta óptimos para la navegación web.

Posteriormente, en [Bozkurt17] realizan una investigación para identificar las causas de la inflación de la latencia en Internet (relación entre el tiempo de recuperación y $latencia-c^1$), entre las que se destaca la resolución DNS como factor de latencia en general. Sobre 1.9 millones de conexiones analizadas, la resolución DNS es uno de los factores más relevantes al igual que el factor de handshake-TLS, con una mediana de 6.3

In: Proceedings of the IV School of Systems and Networks (SSN 2018), Valdivia, Chile, October 29-31, 2018. Published at <http://ceur-ws.org>

¹tiempo que tardaría la luz en viajar de ida y vuelta a lo largo de la ruta más corta entre los mismos puntos finales

y 10.2 veces respectivamente de 36.5 veces el tiempo inflación de latencia total. Investigaciones como estas revelan que el problema no está situado en el ancho de banda del canal como muchos usuarios suponen.

En este trabajo se presenta el estudio de un conjunto de datos reales para analizar el tráfico de resolución y solicitudes DNS. Posteriormente se realizan mapas de calor de acuerdo a la ubicación geográfica de las solicitudes o respuestas DNS. Basados en los resultados de mapas de calor se propone una distribución para apoyar el balance de carga en servidores DNS y generar nuevas métricas de rendimiento de red en función de la hora del día y las solicitudes DNS. Esta propuesta permitiría generar una solución a la congestión de solicitudes DNS en las zonas de más alto tráfico de datos Web de acuerdo al horario de dichas solicitudes DNS. En la RFC 7808 "Time Zone Data Distribution Service" [RFC 7808], se define el protocolo de servicio de distribución de datos de zona horaria, consultas de zona horaria, localización de zona horaria y otros conceptos que son empleados en este trabajo.

Previos desarrollos de protocolos y algoritmos han generado soluciones genéricas pero la evolución de tecnologías como la virtualización en la nube permitirían tener soluciones adaptativas y prácticas como la que se propone en este trabajo. Para efectos de este trabajo se propone replicar servidores DNS de manera local en un nivel de dominio bajo, es decir replicación de DNS caché y previas solicitudes durante un período establecido. Para ello es necesario geolocalizar direcciones IP con alta precisión, lo cual ha sido propuesto en trabajos previos para detección anycast y trabajos relacionados [Cicalese15],[Cicalese16] y [Cicalese18].

2 Propuesta de Distribución Inteligente de acuerdo a zona horaria

El conjunto de datos analizado corresponde a registros de un servidor DNS autoritativo ubicado en Santiago de Chile de un día completo tomados entre los días 2 a 3 del mes de octubre de 2017. Los protocolos contenidos en el conjunto de datos son TCP y DNS, en este trabajo se procesan solo DNS. Las solicitudes DNS corresponden en su mayoría a servicios Web con un 89% del total, el restante corresponde una gran parte a solicitudes Mail Exchange (MX por sus siglas en inglés). Durante el análisis de los registros se pudo observar que la cantidad de respuestas/solicitudes DNS en servidores alojados en diferentes países aumentaba durante horas específicas del día. Esto pudo ser determinado mediante la visualización de diferentes mapas de calor de las respuestas DNS, como se observa en la figura 2. Países como Brasil, Canadá, México, Argentina, entre otros, tienen solicitudes variables de

acuerdo a la zona horaria en la que se encuentran. En Fig.2 (izquierda) se determina que países como Argentina o México no tienen solicitudes en un horario de 04:00 am según Tiempo Universal Coordinado UTC/GMT-4 (UTC por sus siglas en inglés). Mientras que en la Fig.2 (central) se observa como aumenta la cantidad de solicitudes para Brasil, Canadá e inclusive para los países mencionados anteriormente en el horario de 02:00 pm UTC/GMT-4. Finalmente en la Fig.2 (derecha) se observan cambios en un horario de 06:30 pm UTC/GMT-4.

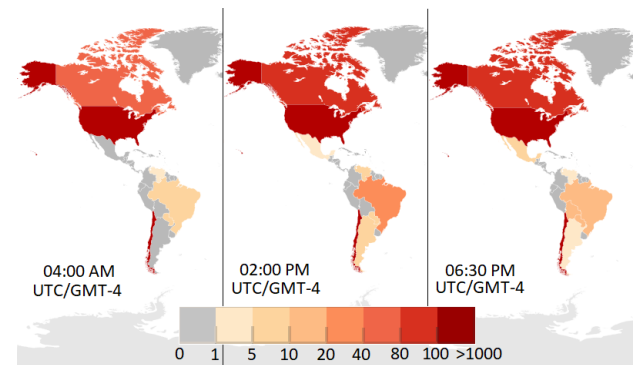


Figura 1. Mapas de calor en diferentes horarios de la cantidad de respuestas DNS en diferentes países de América a un servidor ubicado en Santiago de Chile

De acuerdo a los mapas de calor generados se realiza un análisis de los países con más solicitudes DNS y que obtienen respuesta del servidor DNS en Santiago de Chile. Para geolocalizar los servidores se utilizaron los módulos más conocidos como geoip de geolite2 contenidos en el desarrollo de la industria MaxMind [Maxmind]. Geolite2 es un grupo de bases de datos que relacionan direcciones IP con la ubicación (ciudad o país) del servidor. Python se elige como lenguaje de programación para implementar la geolocalización y la visualización por la cantidad de módulos de visualización de datos que son compatibles y por el desarrollo práctico de algoritmos para realizar extracción de características de grandes cantidades de datos.

Adicional a la distribución por zonas DNS mediante nubes anycast de los servidores instanciados geográficamente a nivel mundial, esta propuesta incorpora un balance de carga para generar un enrutamiento con diferentes saltos entre nodos intermedios o la distribución de las solicitudes realizadas hacia o desde el servidor local en Santiago de Chile. Un ejemplo son las solicitudes provenientes desde Brasil en las que se puede generar una carga significativa en el servidor aparte de las solicitudes internas (Chile) y externas como EU o Canadá. Por este motivo resulta útil dirigir el tráfico a servidores instanciados en una

zona horaria con una carga menor de solicitudes en el mismo horario y que no afecte servicios donde se instancia la replica. Por ejemplo, para el rango de horas entre las 02:00 pm y 04:30 pm UTC/GMT-4, del total de solicitudes en ese día desde Brasil el servidor DNS alojado en Santiago de Chile procesó el 45.47%, es decir casi la mitad en un horario específico. Mientras tanto, en el mismo horario el servidor DNS procesa el 13% de solicitudes del total en Estados Unidos que equivalen a alrededor de 200.000 mil solicitudes.

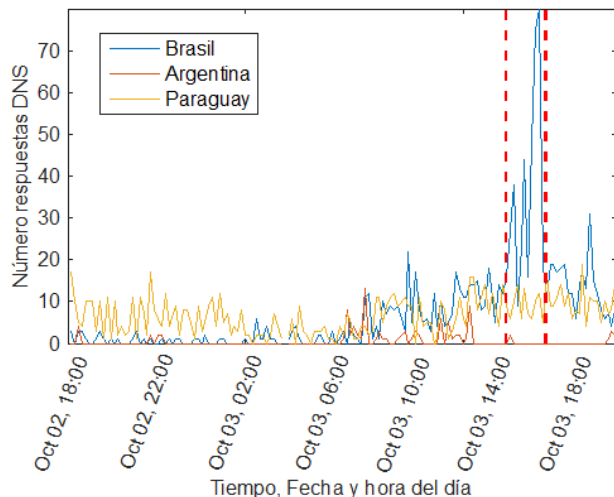


Figura 2. Número de respuestas DNS desde el servidor local en Santiago hacia solicitudes desde Brasil, Paraguay y Argentina durante un día

Aunque actualmente existen servidores con grandes capacidades de procesamiento, día a día la implementación de nuevas tecnologías y aplicaciones basadas en servicios IP en su mayoría servicios Web hacen que se limiten los tiempos de respuesta DNS óptimos. Por este motivo la propuesta de distribución DNS basada en los tiempos de horario de mayor congestión según el país de procedencia busca disminuir los RTT a solicitudes DNS. La distribución horaria propuesta aborda dos estrategias: distribuir carga del servidor DNS localizado en Santiago de Chile en un porcentaje que se basa en reglas y umbrales establecidos por los administradores de red y la segunda es la replicación de servicios por un tiempo determinado en otros servidores, lo cual genera un uso óptimo de servicios por el tiempo de mayor cantidad de solicitudes DNS, lo cual ahorra costos de recursos económicos/computacionales en el uso de servidores virtuales e inclusive locales si es el caso. Para la última estrategia no solo se considera enviar las solicitudes hacia un servidor localizado en otro espacio geográfico, sino que se va a considerar que la carga de solicitudes DNS en ese horario no va a afectar el rendimiento de

los servicios locales, debido a que no coincide el rango de horas con más alto tráfico DNS del país donde se replica el servidor.

3 Trabajos Futuros

Principalmente esta propuesta hace énfasis en una solución generalizada a servidores DNS en otras ubicaciones geográficas, lo expuesto en este trabajo es solo un ejemplo de cómo se comporta el tráfico DNS para el caso de Santiago de Chile porque es de donde se adquirieron los datos de un servidor DNS real. Basados en el análisis previo se plantea desarrollar un algoritmo de distribución de acuerdo a los tiempos durante el día con más tráfico de solicitudes provenientes de diferentes países. Para ello se plantea realizar un análisis estadístico de los datos seleccionados aleatoriamente. Un promedio de las horas con más flujo de tráfico DNS permite establecer el rango de horas en el modelo propuesto. El balance de carga hacia un servidor real en otro país permite disminuir tiempos de respuesta por parte de los servidores de respaldo y servidor local analizado en Chile. Por otro lado estas pruebas se piensan realizar en servidores de distintos países para coordinar los tiempos de funcionamiento del servidor de respaldo. Para replicar el servidor DNS, múltiples trabajos [Castro09], [Narayan14] han planteado la emulación o geo-replicación de servidores con las que se obtienen mayor auto-escalabilidad y disponibilidad (incluso virtual). Adicionalmente se plantea hacer un diagnóstico de las rutas de las solicitudes para determinar si la distribución es más óptima en tiempos de respuesta de acuerdo a la distancia entre el servidor y el usuario/servidor que realiza la solicitud. Finalmente se realizará una comparación con los tiempos de respuesta del servidor actual en Santiago de Chile.

References

- [Incode17] IncodeConsulting .*The cost of network latency* www.incodeconsulting.com/wp-content/uploads/2017/01/inCode-Cost-of-Latency-White-Paper-060114.pdf, Enero 2017.
- [Karzand17] M. Karzand, D. Leith, J. Cloud & M. Médard. *Design of FEC for low delay in 5G*. IEEE Journal on Selected Areas in Communications (pp. 1783-1793), Vol. 35, Issue 8. August 2017.
- [Asap11] W. Zhou, Q. Li, M. Caesar & P. Godfrey. *ASAP: A low-latency transport layer*. Proceedings of the Seventh Conference on emerging Networking EXperiments and Technologies, 2011.

- [Bozkurt17] I. N. Bozkurt, A. Aguirre, B. Chandrasekaran, P. B. Godfrey, G. Laughlin, B. Maggs & A. Singla. *Why is the Internet so slow?!*. International Conference on Passive and Active Network Measurement, Springer (pp. 173-187). March 2017.
- [Castro09] Sebastian Enrique Castro Ávila. *Una emulación para anycast*. Tesis para optar al grado de Magíster en Ciencias mención Computación, Agosto 2009.
- [RFC 7808] M. Douglass (Spherical Cow Group) & C. Daboo (Apple) "*Time Zone Data Distribution Service*", RFC 7808, Marzo 2016.
- [Cicalese15] D.Cicalese, J. Auge, D. Jounblatt, D. Rossi, M. Olivier & T. Friedman. *Lightweight Anycast Enumeration and Geolocation*. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs), 2015.
- [Cicalese18] D.Cicalese & D. Rossi, *A longitudinal study of IP Anycast*. ACM SIGCOMM Computer Communication Review, Enero 2018.
- [Cicalese16] D.Cicalese, J. Auge, D. Jounblatt, D. Rossi, M. Olivier & T. Friedman. *Latency-Based Anycast Geolocation: Algorithms, Software, and Data Sets*. IEEE Journal on Selected Areas in Communications, Volume: 34, Issue: 6, Junio 2016.
- [Akamai17] D. McDonald, Akamai Technologies. Why you should care about DNS latency <https://blogs.akamai.com/2017/06/why-you-should-care-about-dns-latency.html>
- [Maxmind] Maxmind. Paquete de desarrolladores GeoLite2, Geolite2 Redistribution <https://www.maxmind.com/es/geolite2-developer-package>
- [Narayan14] I. Narayanan, A. Kansal, A. Sivasubramaniam, B. Urgaonkar, S. Govindan. Towards a Leaner Geo-distributed Cloud Infrastructure In HotCloud Junio 2014.