Observatorio Latinoamericano del DNS

Primer Informe
Agosto 2016

INTRODUCCIÓN

El sistema de nombres de dominio (DNS) es una capa crítica de la infraestructura de Internet que permite el funcionamiento del resto de capas por encima de ella, incluyendo aplicaciones de cliente final como por ejemplo los sitios web, correos electrónicos, mensajería instantánea, etc. El DNS está por debajo de cada una de estas tecnologías, dando soporte y permitiendo cambios dinámicos en tiempo real para efectos de adaptación a tráfico, evasión de ataques, balanceo de carga, etc.

Siendo un sistema tan crítico es importante preocuparnos de su correcto funcionamiento a nivel país, considerándola una de los fundamentos de la presencia en Internet de una comunidad. Cada día las personas se hacen más dependientes de Internet, y debemos corresponder con mejoras en la robustez y redundancia del DNS,protegiéndonos contra fallas y eventos inesperados contra los que existen técnicas probadas que permiten salvaguardar el funcionamiento del DNS.

Es por esto que se han definido ciertas métricas que permiten medir objetivamente el cumplimiento correcto de estándares de seguridad y robustez en el DNS. Estas medidas, aunque no sean requeridas ni exigidas por algún registro de nombres de dominio, sí son importantes para entregar un servicio de calidad. Tener una visión completa del estado del DNS basado en estas métricas permite tener una fotografía del estado actual del DNS y compararlo a través del tiempo mediante medidas recurrentes. Además de ir estimando las mejoras a través del tiempo, es posible tomar medidas correctivas a largo y mediano plazo que se deberían ir reflejando en mejoras de las mediciones.

El Observatorio Latinoamericano del DNS es un intento de obtener medidas continuas y objetivas de parámetros conocidos del DNS a nivel de los nombres de dominio latinoamericanos, con el fin de obtener informes y reportes que mejoren la gestión de este importante recurso.

LOS DATOS

Para obtener una medición exhaustiva es necesario disponer de la lista de todos los nombres de dominio de la región. En este primer informe preliminar se optó por utilizar una fuente pública alternativa, la "Alexa 1M website list" que entrega un ranking del millón de sitios web más populares de Internet, que fue filtrado para solo dejar los que corresponden a nombres de dominio latinoamericanos. Con ese listado de nombres de dominio se alimentó al Observatorio y se tomaron las medidas que considera este informe.

Este listado de Alexa sirve como prototipo, pero hay que hacer la salvedad que los resultados no son comparables al análisis completo de la región. Los dominios en Alexa comparten ciertas características particulares que no se pueden extrapolar al resto de dominios. Por tratarse de sitios altamente populares, es esperable que cuenten con una infraestrucutra mejor que los dominios más normales, por lo que en realidad representan un "tope superior" en las medidas.

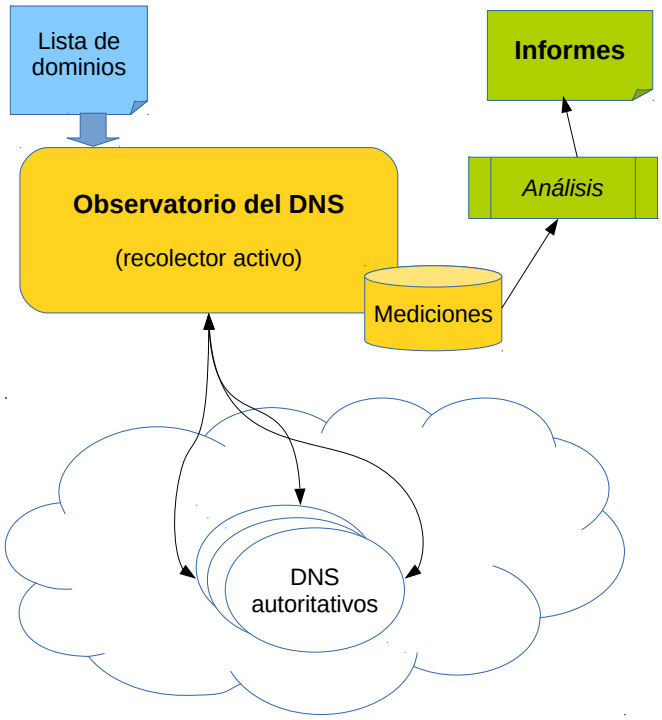
En resumen se quiere dejar en constancia que estos resultados no serán muy parecidos al informe final, pero representan una muestra que como prueba de concepto permite mirar los tipos de métrica y su comportamiento.

Otro punto importante es que los resultados se presentan agrupados por la región completa. No se separa ni se compara por país, porque el objetivo es tener tendencias regionales y no entrar a la realidad de cada ccTLD.

CÓMO FUNCIONA EL OBSERVATORIO

A partir de este listado de nombres de dominio en LAC, se alimenta un recolector de datos que está ubicado en Santiago de Chile, el que en forma periódica (1 vez a la semana) recorre todos los servidores de nombre de cada uno de estos dominios realizando consultas DNS que permiten registrar estas métricas. Además utiliza otras fuentes de información (por ejemplo los números autónomos ASN), que permiten completar otro tipo de análisis.

Con estos datos se arma un repositorio de Big Data que trimestralmente es resumido, graficado y analizado para la elaboración del informe.



Resultados

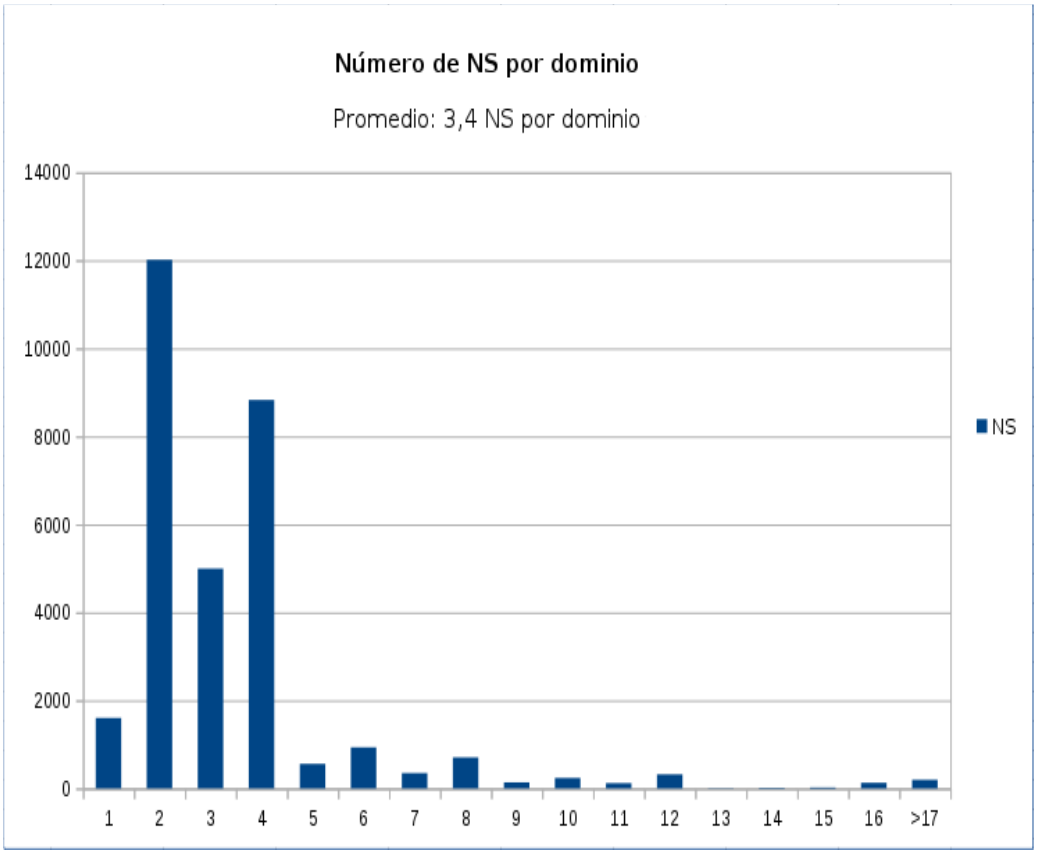
1. Total de dominios

Según estadísticas de LACTLD[1], el organismo que agrupa a los registros de dominio de la región Latinoamérica y el Caribe, existen más de 8 millones de nombres de dominio en esta región. En el listado de Alexa se puede encontrar un poco más de 30 mil nombres de sitios web de la región, lo que representa un 0,38% del total.

Para este análisis, se analizaron 31.872 nombres de dominio de la región de LACTLD[2].

2. Número de NS por dominio

Para entregar redundancia a la correcta y rápida resolución de un nombre de dominio, el DNS permite que se definan "servidores de nombre" (NS) que son los servidores finales encargados de la traducción nombre-número IP. El estándar indica que deben ser al menos dos NS por dominio[3]. Algunos registros obligan por política de contrato que se definan estos dos, pero otros lo dejan a criterio del titular del dominio. De cualquier manera, a esta altura del desarrollo de Internet se considera que dos NS es muy poco, siendo el mínimo tres de ellos[4]. La cantidad adecuada dependerá del uso que se le dará al nombre, de la capacidad de cada NS y de la ubicación de cada uno de ellos.

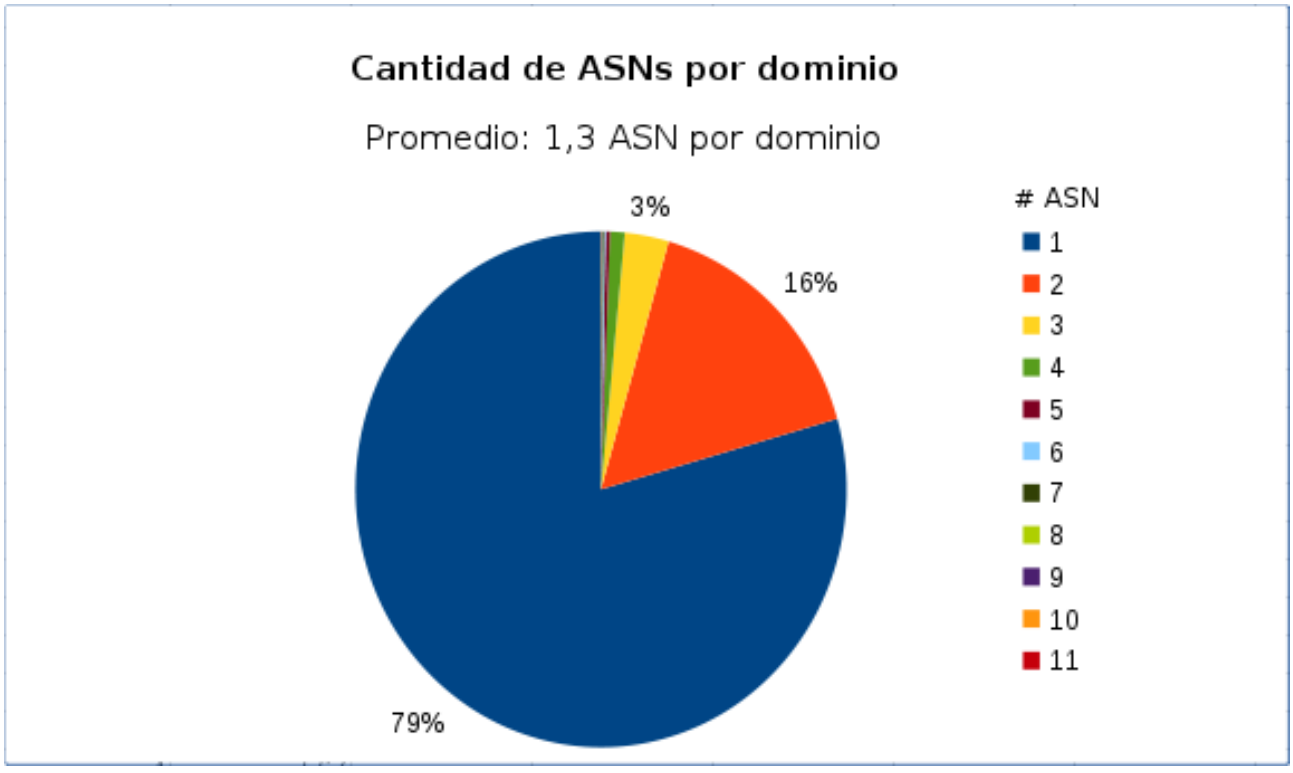


Como se observa en el gráfico, existe una correcta implantación de la buena práctica, al tener un promedio superior a tres NS por dominio. Sin embargo existe una gran cantidad (38%) que sólo tiene 2, lo que aunque cumple la recomendación es demasiado riesgoso.

Es importante advertir que este buen resultado puede estar influenciado con los datos que provienen de dominios en Alexa, por lo que representan los sitios más visitados y por lo mismo los mejor preparados en su infraestructura. Es esperable que un muestreo más completo y diverso de la región pueda disminuir este promedio.

3. Número de ASN por dominio

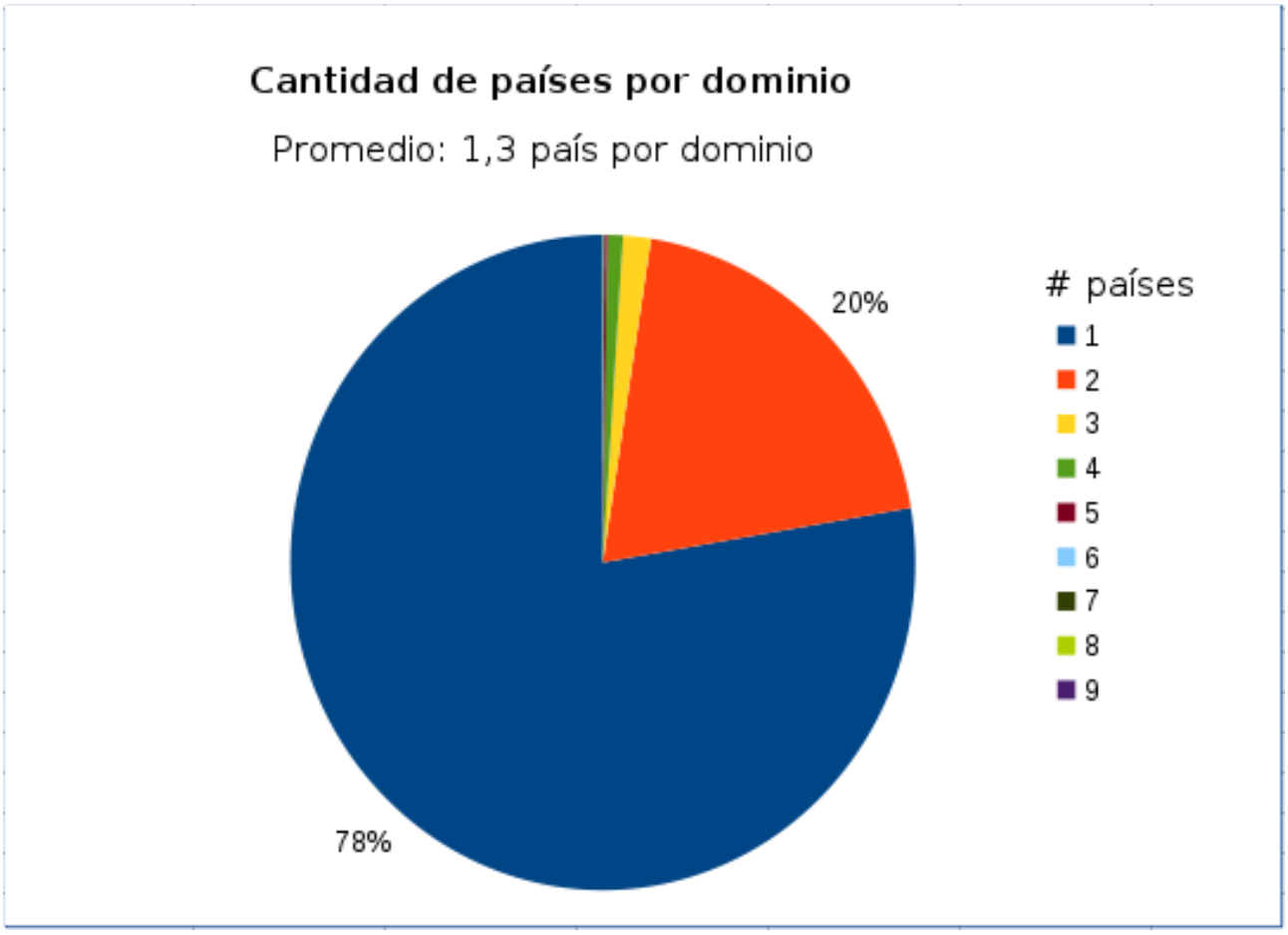
El número de NS por dominio representa una primera aproximación al estado de robustez y redundancia del DNS. Sin embargo una segunda mirada más profunda permite distinguir qué tan distribuidos se encuentran estos NS a nivel de ruteo. Hay ciertos dominios que poseen 3 NS, lo que parecería indicar que son bastante resilientes, pero resulta que los 3 NS están en el mismo datacenter, o uno al lado de otro, incluso que comparten la misma máquina física! Por lo que es una ilusión pensar que esos 3 NS harán seguro el dominio. Ese dominio sigue siendo vulnerable a cortes localizados de energía, enlaces, etc. Es por esto que la métrica de "ASN" permite ser más estricto y contar el número de sistemas autónomos que resuelven al dominio. Los sistemas autónomos representan organizaciones distintas, por lo que tener NS en más de 1 ASN indica que los servidores de nombre efectivamente están en lugares distintos, con ruteos distintos, por lo que en realidad sí entregan diversidad.



Acá vemos que lamentablemente el buen resultado anterior de 3,4 NS por dominio es un tanto engañoso desde el punto de vista de resiliencia. Casi un 80% de los dominios tienen sus NS es 1 solo número autónomo ASN, lo que quiere decir que están en la misma organización y lugar físico. Esto hace su dominio muy vulnerable a cortes por problemas a nivel de una organización. El 16% que tiene sus NS en 2 números autónomos es mucho más robusta, ya que la probabilidad de falla en dos organizaciones es muchísimo más baja.

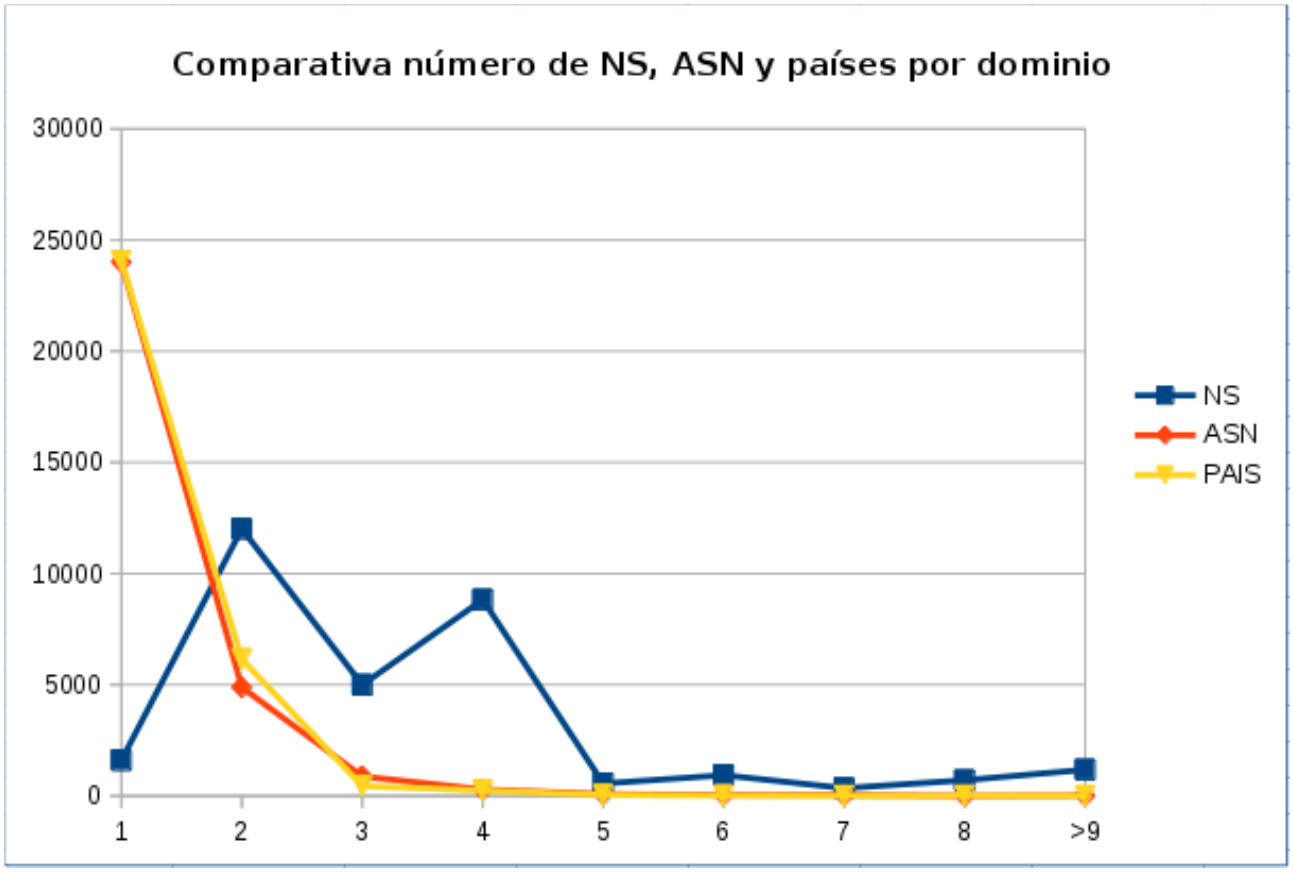
4. Dispersión de servidores por país

Siguiendo con un análisis más profundo de la ubicación de los NS de un dominio, se analiza la ubicación física a nivel de país.



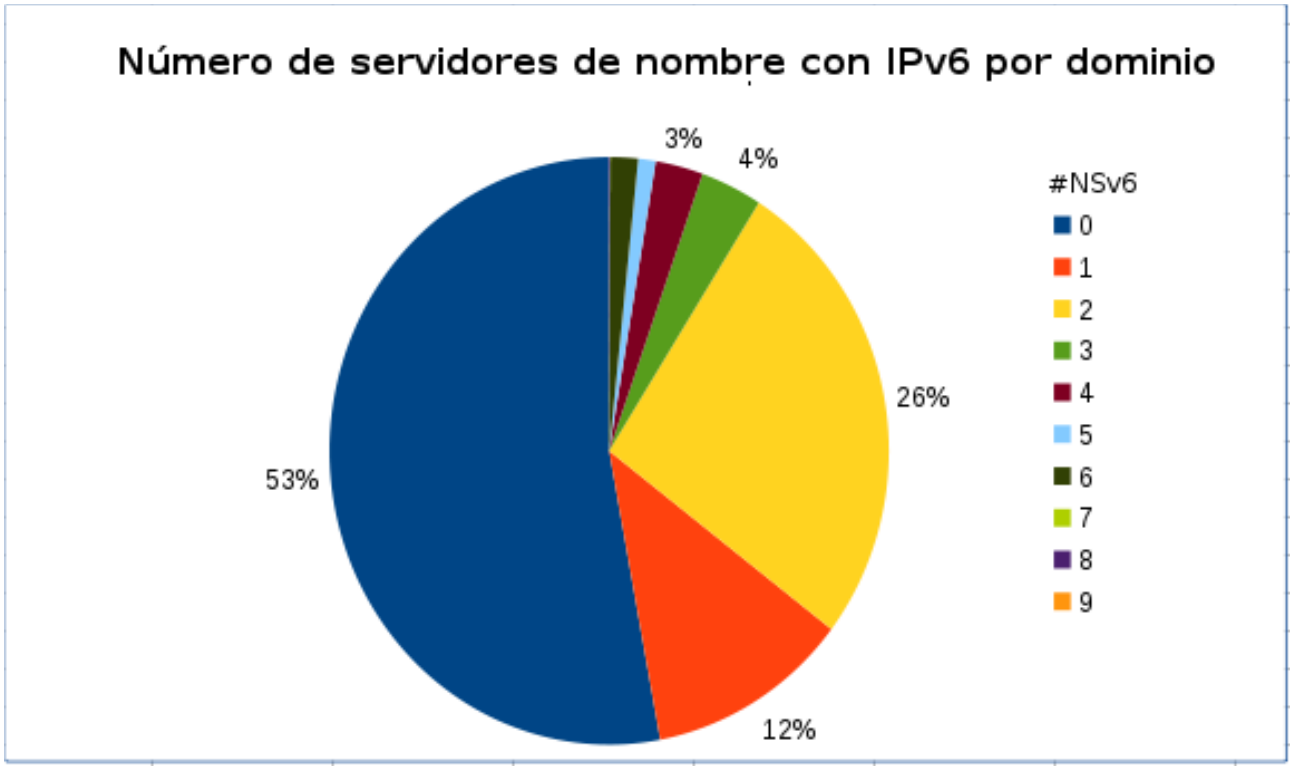
El resultado es bastante similar al del número de ASN, con casi un 80% que tiene todos sus NS en el mismo país, y un 20% que lo tiene en dos países. Nuevamente es recomendable tener los distintos NS en ojalá más de un país, considerando que ya ha ocurrido cortes de Internet a niveles nacionales, que pueden resolverse con este mecanismo.

Para finalizar este análisis de NS por dominio, acá tenemos una gráfica comparativa de las 3 métricas anteriores:



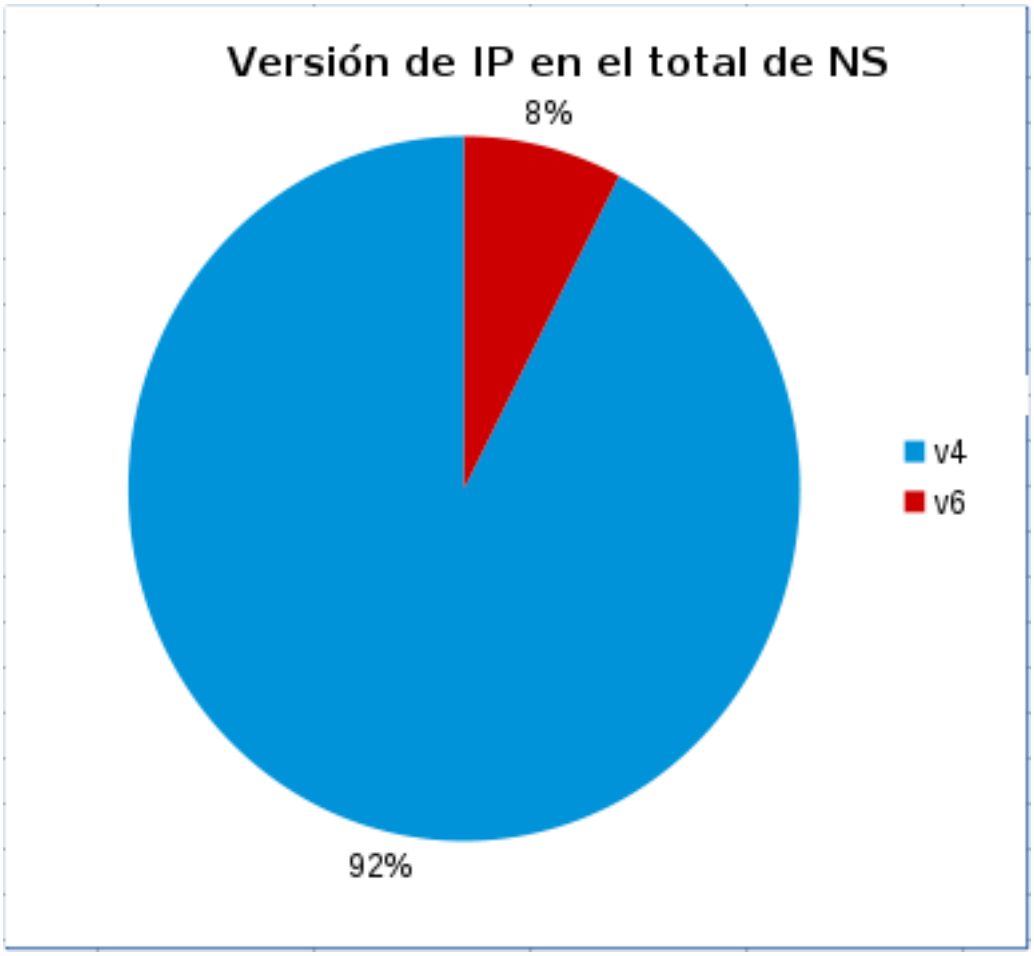
5. Uso de IPv6

El sistema de nombres de dominio, al tratarse de una infraestructura crítica para Internet, debe ser uno de los pioneros en desplegarse en el nuevo protocolo IPv6, para dar el soporte a todas las aplicaciones que dependen de ella. Analizamos para cada nombre de dominio si sus NS tienen direcciones en este nuevo protocolo (registro AAAA y conectividad vía IPv6):



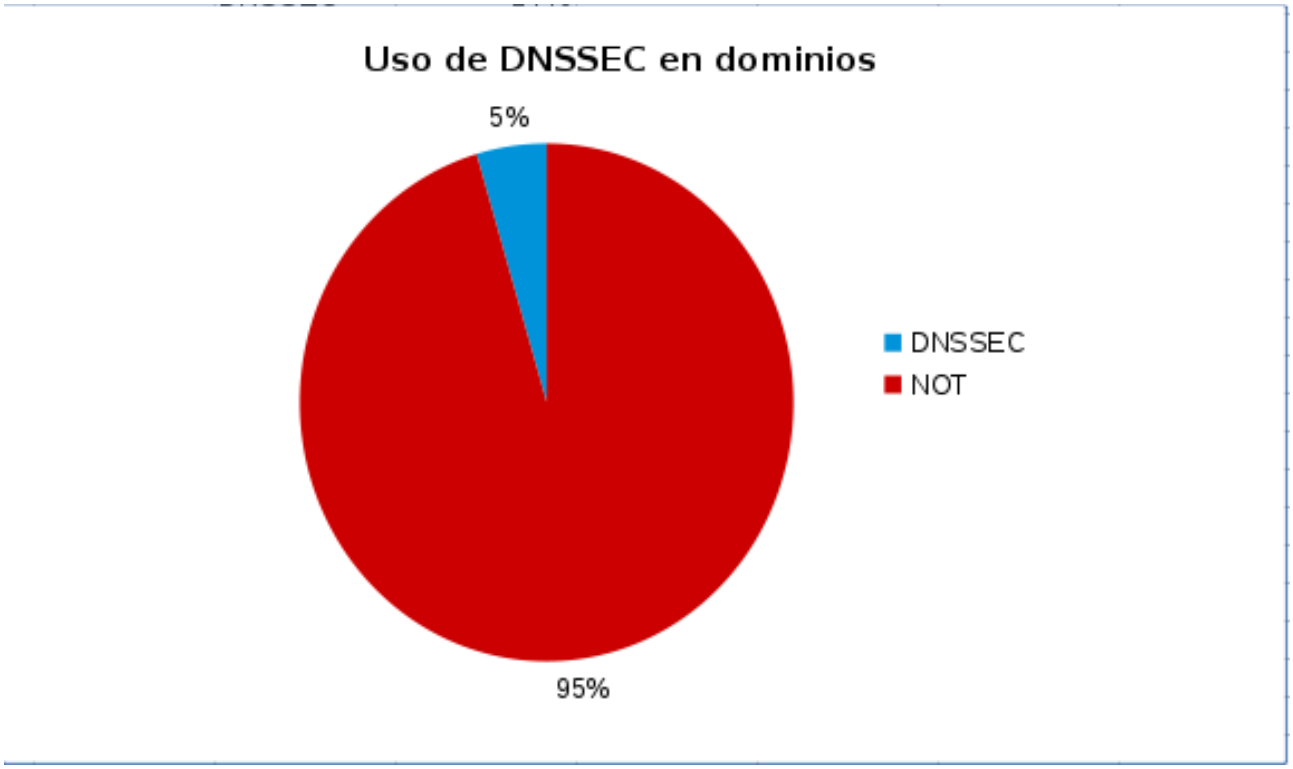
Lamentablemente se ve una poca participación en esta nueva tecnología. Más de la mitad de los nombres de dominio de la región no tiene ningún NS alcanzable por Ipv6, lo que significa que estos nombres de dominio no funcionarán cuando aparezcan redes conectadas vía IPv6 puro. Un 12% tiene al menos uno y 26% tienen dos servidores de nombre en IPv6.

Si miramos a un nivel más global, del total de NS de la región un 92% utiliza IPv4, y un 8% IPv6:



6. Uso de DNSSEC

Las extensiones de seguridad al DNS, llamadas DNSSEC, permiten agregar a esta tecnología las propiedades de autenticidad e integridad de datos. De esta forma es posible estar seguro de que la información que se obtiene no fue modificada en el camino y que viene de la fuente real y autoritativa. Gracias a DNSSEC se pueden mitigar distintos ataques y mejorar la confianza global en Internet.



Sólo un 5% de los nombres de dominio analizados tiene activo DNSSEC.

Conclusiones

Tenemos buenos resultados en las métricas básicas de número de servidores de nombre por dominio, y regulares en el caso de dispersión por sistema autónomo y por diversificación en países. Esto da pie a que sea posible seguir avanzando en mejorar la robustez de nuestro sistema de nombres de dominio en la región, a través de este tipo de estudios. También la comparación a través del tiempo permitirá ir analizando si las políticas de incentivo y educación dan los resultados esperables.

Sin embargo el avance en nuevas tecnologías como IPv6 y DNSSEC está recién comenzando. No es un tema particular de nuestra región. Resultados similares se encuentran a nivel mundial. Acá es importante sumarnos a las campañas de otras regiones y continuar con el despliegue de estas importantes tecnologías.

Este primer reporte es un prototipo y prueba de concepto del tipo de análisis que podemos obtener con los datos reales, y que sirvan de base para mediciones continuas y permanentes.

APÉNDICES

A. Glosario

DNS: Sistema de Nombres de Dominio. Arquitectura que permite la asociación entre nombres de hosts como www.ejemplo.com con una dirección IP como 2001:1398::1 ó 200.7.7.3.

NS: servidor de nombres del sistema DNS. En este documento se refiere a uno de los servidores autoritativos, en control del titular de un nombre de dominio.

Dominio: Subdivisión dentro de las etiquetas de un nombre de host que representa una unidad administrativa única.

IP: número que representa la ubicación dentro de Internet de un servicio. Existen de dos tipos: la versión 6 (IPv6) y la antigua versión 4 (IPv4).

ASN: Número de Sistema Autónomo, es un identificador que representa a una organización dentro de Internet que se encarga del enrutamiento de la información entre distintos servidores.

DNSSEC: Extensiones de seguridad del DNS que permiten entregar autenticidad e integridad en las respuestas.

B. Bibliografía

[1]: “Encuesta Estadística LACTLD 2015”, Segundo trimestre 2015.
[2]: Los miembros de LACTLD en el momento del estudio son los ccTLDs ai, ar, aw, bo, br, bz, cl, co, cr, cu, cw, do, ec, gt, gy, hn, ht, mx, ni, pa, pe, pr, py, sv, uy, y ve.
[3]: RFC1912, “Common DNS Operational and Configuration Errors”
[4]: RFC2181, “Selection and Operation of Secondary DNS Servers”

C. Agradecimientos

El presente estudio nace como un proyecto del Plan Estratégico para la región LAC de ICANN. El grupo de diseño original estaba integrado por Alejandro Acosta (LACNIC), Juan Manuel Rojas (LACRALO), Antonio Alberti y Victor Hugo Fernandes (INATEL), y Hugo Salgado (NIC Chile) como líder. En este grupo se definieron las métricas a analizar y el mecanismo de recolección.

Agradecemos el apoyo de ICANN con la donación de un servidor donde se instaló el recolector activo, para realizar las mediciones y almacenar los datos.

En el soporte para análisis de los datos agradecemos el apoyo de NIC Chile Labs, especialmente a su director Javier Bustos y la ingeniero Maite González.

Por último al apoyo y confianza del equipo de ICANN, especialmente Rodrigo Saucedo y Rodrigo de la Parra, junto a la directora ejecutiva de LACTLD, Carolina Aguerre.

La herramienta de recolección de datos utilizada es DNSdelve, parte del suite DNSwitness, realizada por AFNIC y puesta a disposición de la comunidad.

D. Autor y contacto

Hugo Salgado Hernández
Ingeniero de Proyectos
NIC Chile – Universidad de Chile
Miraflores 222, piso 14, Santiago, CL
hsalgado@nic.cl

Versión 1, 22 de agosto de 2016.