

# La visión 50/50 para el tráfico de Internet

Versión 2.0, enero de 2024

## Introducción

La Visión 50/50 define la estrategia de Internet Society para reunir esfuerzos de múltiples partes interesadas y recursos nacionales e internacionales para garantizar que al menos el 50 % de todo el tráfico generado de manera local en economías seleccionadas siga siendo local para 2025. Alcanzar este ambicioso objetivo fortalecerá la calidad de Internet y reducirá los costos de acceso para los individuos.

Un [estudio de 2012](#) estableció una base de referencia en los centros de puntos de intercambio de Internet (IXP) en Kenia y Nigeria, y un [estudio de seguimiento en 2020](#) mostró que los niveles de tráfico local aumentaron del 30 % al 70 %. El esfuerzo ayudó a aumentar la comprensión del impacto del peering en la infraestructura local. Se descubrió que el aumento del tráfico local generó importantes ahorros en costos para las redes participantes y colocó a estos dos países en una posición más sólida para participar en la economía digital. Mientras tanto, Sudáfrica localizó más del 80 % de su tráfico local y ahora disfruta de una Internet estable, resiliente, de alta calidad y asequible.

El objetivo del estudio de visión 50-50 es medir el porcentaje de tráfico de Internet que se ofrece desde un servidor o caché del país. Este documento describe la metodología de Internet Society para lograr el objetivo anterior.

## Definiciones

- Tráfico local: se refiere al tráfico que permanece local y no sale del país.
- Contenido local: contenido de Internet alojado dentro del país (por ejemplo, sitios web de noticias, servicios de gobierno electrónicos, etc.) y
- RIPE Atlas: es un marco de medición de Internet operado por RIPE NCC que consta de sondas (puntos de referencia) y anclas (objetivos). Las sondas se pueden utilizar para ejecutar varias mediciones (latencia, traceroute, DNS, SSL, HTTP) desde el extremo.

- CDN: las Redes de entrega de contenido son responsables de entregar contenido hasta el extremo.
- Caché de contenido: es un equipo de alojamiento de contenido colocado por un proveedor de contenido cerca de los usuarios finales.
- Punto de vista: desde donde se realiza la medición.
- Red perimetral: se refiere a la red de acceso donde se encuentran los eyeballs (usuarios).

## Falta de datos sobre el volumen de tráfico

Medir los niveles de tráfico local y no local en un país no es sencillo. Los proveedores de acceso suelen conocer los servicios más utilizados y cuáles son los mayores consumidores de ancho de banda dentro de su red. Los datos sobre los servicios más utilizados rara vez se comparten públicamente. En algunos casos raros y basándose en un acuerdo previo, los ISP proporcionan datos de tráfico s-flow (flujos de muestra), de los cuales se puede extraer la IP de destino. En ausencia de datos reales, utilizaremos el porcentaje de sitios web ofrecidos de manera local (de forma nativa o a través de una caché CDN) analizando los 1000 sitios web principales de un país.

## Google CruX

Utilizamos los 1000 sitios web principales del [Informe UX de Google Chrome](#) (CrUX). Es un conjunto de datos públicos y una iniciativa de Google que tiene como objetivo proporcionar métricas de rendimiento del mundo real sobre cómo los usuarios reales experimentan la web. Agrega datos de rendimiento anónimos de los usuarios que han optado por compartir sus datos de navegación con Google, principalmente a través del navegador Chrome. CrUX elabora un informe mensual de los sitios web más populares por país.

## Cómo determinamos la localidad del contenido

1. Para cada sitio web en la lista de los 1000 principales, debemos determinar qué servicio CDN se está utilizando. Para ello, inspeccionamos el sitio web para extraer lo siguiente:
  - a. Encabezados del servidor HTTPS: los encabezados como "servidor" o "vía" suelen contener pistas geográficas.
  - b. Registros CNAME: un registro de nombre canónico (CNAME RECORD) es un tipo de registro en servidores de nombres de dominio (servidores DNS) que crea un alias de un dominio a otro. Al observar un registro CNAME, podemos ver el nombre de dominio vinculado a una CDN, antes de dirigirlo al recurso deseado.
  - c. Datos de WHOIS: utilizando la dirección IP del dominio, haremos una búsqueda de WHOIS e inferiremos el uso de las CDN.

- d. Basado en encabezados de servidor: algunos encabezados de respuesta HTTP proporcionarían pistas geográficas.
2. Utilizando puntos de referencia locales dentro del país, realizamos mediciones activas hacia un conjunto seleccionado de objetivos (CDN, aplicaciones, sitios web) para extraer pistas geográficas y otra información de geolocalización. Con la información anterior, podemos marcar con confianza un sitio web como local.
3. Para los sitios web alojados de forma nativa, utilizamos la geolocalización para determinar la ubicación de la dirección IP. La dirección IP se obtiene ejecutando la consulta desde un punto de vista local a través de un solucionador de DNS local.
4. Luego calculamos la proporción de sitios web alojados localmente frente a sitios web alojados de forma remota.

## Ejemplo

A continuación, se detallan los pasos para calcular el porcentaje de sitios web alojados de manera local:

1. Tomamos los 1000 sitios web principales de Google CrUX para el país X.
2. Clasificamos los sitios web por <nombre de CDN> si están alojados en una CDN o "Nativo" si el sitio web está alojado de forma nativa.
3. Utilizamos la geolocalización de la IP para determinar la ubicación de los dominios alojados de forma nativa.

CDN	Número	Ubicación
Cloudflare	454	Local
Akamai	92	Externo
Cloudfront	55	Local
Google	51	Local
Fastly	25	Externo
Facebook	8	Local
Yahoo	6	Externo
Amazon AWS	5	Externo

EdgeCast	4	Externo
Level3	3	Externo
Alojado de forma nativa	250	Local: 100 Externo: 150

Tabla 1: Distribución del uso de CDN para los 1000 sitios web principales en el país X

4. El resultado es la proporción entre sitios web locales y externos.

## Resultado

Planeamos producir un informe detallado sobre cuánto tráfico permanece local según la metodología mencionada para fines de febrero de 2024. También desarrollaremos un widget en nuestra plataforma Pulse para visualizar los datos en el segundo trimestre de 2024. Dividiremos los 1000 sitios web principales en diferentes conjuntos de datos para presentar diferentes perspectivas sobre la localidad del tráfico: (1) los 1000 sitios web principales (2) las 10 aplicaciones principales (Netflix, Disney+, etc.) (3) los 100 sitios web principales que utilizan el ccTLD del país (4) los sitios web gubernamentales (5) los sitios web de noticias locales (6) los principales participantes de CDN.