CLOUDFLARE

DOCUMENTO TÉCNICO

Resiliencia de la red y los servicios de Cloudflare

Contenido

Descripción general

Internet se basa en sistemas imperfectos que están diseñados para priorizar la disponibilidad sobre todo lo demás. Los protocolos como TCP y BGP se basan en los principios de los sistemas distribuidos, esperar fallos y tenerlos en cuenta, y la Internet se diseñó en consecuencia. Sin embargo, aún existen puntos de fallo que pueden afectar de forma negativa. Los proveedores de red deben poder planificar los fallos, detectarlos cuando se produzcan y mitigar el impacto que experimentan los usuarios.

Debido a la naturaleza en tiempo real de muchas aplicaciones en Internet, las redes no solo deben cubrir la mayor parte posible de Internet, sino que también deben responder y mitigar el impacto en tiempo real. Los tiempos de inactividad resueltos en minutos no son aceptables: los clientes esperan una solución en cuestión de segundos.

Cloudflare ofrece servicios de infraestructura de red esenciales para ayudar a las organizaciones a protegerse y comunicarse a través de Internet. Hemos desarrollado nuestra red y los servicios que ofrece para mantener el más alto nivel de excelencia operativa. Cloudflare procesa un promedio de más de 84 millones de solicitudes HTTP y 61 millones de consultas DNS por segundo, lo que brinda servicio a millones de propiedades de Internet y usuarios.

¿Cómo ofrece Cloudflare un servicio confiable a gran escala dadas las características impredecibles de Internet?

La respuesta proviene de la arquitectura de la red de Cloudflare, que opera con capacidades de resiliencia diseñadas para operar de manera independiente y resistir el espectro de interrupciones. Nuestras capacidades de procesamiento, redes y almacenamiento, junto con nuestros procesos operativos, están diseñadas para que Cloudflare sea tan confiable como el tono de marcado de la red de servicio telefónico tradicional. Cloudflare ofrece el "tono de nube" metafórico para los servicios de red y seguridad de los que dependen los clientes, independientemente de las condiciones de Internet en un momento dado.

Este documento técnico analiza los desafíos de operar en tiempo real en Internet, y cómo Cloudflare está en una posición única para resolver esos desafíos.

Vivir en un mundo imperfecto

Internet es un lugar imperfecto y, sin embargo, las organizaciones la necesitan para desarrollar su negocio y gestionar entidades que vinculan usuarios, datos y dispositivos distribuidos con aplicaciones en la nube. Cualquier interrupción del servicio tiene graves consecuencias. A lo largo de los años, Cloudflare ha informado sobre una serie de interrupciones importantes del servicio de Internet en todo el mundo, que van desde accidentes hasta ataques DDoS, desastres naturales y más.

Internet es una red diversa que se ha construido como un grupo débilmente constituido de miles de redes de diversos tamaños y capacidades que participan. Estas redes tienen diferentes niveles de servicio: algunas funcionan lo mejor que pueden, como un apoyo, y otras funcionan como parte de sus servicios comerciales. A través de acuerdos mutuos para intercambiar tráfico, que puede o no estar destinado a hosts en su propia red, Internet funciona como un tejido global a través de la participación benévola de intercambios de Internet y proveedores de servicios regionales y locales.

Sin embargo, esta diversidad conlleva un grado de imprevisibilidad. La ruta que toma cualquier paquete se basa en una secuencia de estimaciones y esfuerzos óptimos para su entrega. En ausencia de datos en tiempo real o modelos predictivos sobre las condiciones reales de la red en un momento dado, un paquete suele dar su siguiente salto utilizando rutas predeterminadas o la ruta presuntamente más corta. Ninguna de estas opciones tiene en cuenta el estado del servicio de red en las decisiones de enrutamiento. Esto significa que los paquetes suelen estar sujetos a una serie de condiciones que las debilitan:

- En el nivel más básico, las redes pueden experimentar fallos temporales y caídas de tensión que reducen el rendimiento y provocan la pérdida de paquetes.
- A medida que las redes se saturan, la congestión perjudica el rendimiento y la experiencia del usuario.
- Si bien estos tipos de desaceleraciones pueden producirse en condiciones operativas normales, un número cada vez mayor de interrupciones son provocadas intencionadamente por agentes hostiles, que consumen de forma maliciosa los recursos disponibles.

En Cloudflare, nuestra misión es ayudar a mejorar Internet. Con ese fin, creamos la infraestructura para que Internet sea más rápida, confiable y segura. Esto es posible gracias a los servicios desde y a través de la red de Cloudflare, una de las redes más grandes del mundo que opera en nuestros propios servidores físicos (sin virtualización), una red troncal privada y una presencia global masiva a la que se puede acceder, en promedio, en 50 ms del 95 % de la población mundial.[1]

Los clientes generalmente quieren entender cómo se hace esto. ¿Cómo diseña Cloudflare servicios a tal escala, sin sacrificar la seguridad y el rendimiento, y con la confiabilidad que necesitan las organizaciones? La respuesta proviene de una serie de objetivos de North Star, que establecen cómo construimos nuestra infraestructura para que siga siendo eficaz a pesar de las fuerzas disruptivas, tanto externas como internas.

Cómo diseña Cloudflare la resiliencia

Muchas organizaciones se centran en mejorar la disponibilidad intentando mejorar su capacidad de reacción o actuar más rápido ante los fallos. Esto requiere invertir y probar constantemente sus capacidades y procesos de conmutación por error. Si bien estos son objetivos válidos, Cloudflare piensa de manera diferente, ya que nuestros equipos de ingeniería de resiliencia dedican un enorme esfuerzo a reducir los escenarios en los que se requiere una respuesta de recuperación ante desastres.

La ingeniería de resiliencia de Cloudflare comienza con una premisa simple: **¿cómo desarrollar una infraestructura crítica que siga siendo operativa suponiendo que se produzcan fallos?**

Cuando inevitablemente ocurren fallos, los servicios resilientes de Cloudflare detectan y aíslan los fallos para que no afecten la disponibilidad del servicio. Las fallas se resuelven fuera de banda desde nuestra prestación de servicios. Nos esforzamos por ser agnósticos ante fallos en todo el conjunto de nuestros servicios.

Para comprender los principios de la resiliencia, es útil definir primero los conceptos clave detrás de las rutas de tráfico con Cloudflare y los objetivos de diseño para los sistemas clave. En el nivel más abstracto, la arquitectura de Cloudflare se puede dividir en dos segmentos: **el plano de control y el plano de datos.** Cada uno tiene una resiliencia y una postura ante desastres únicas.

Resiliencia del plano de control

El plano de control proporciona la interfaz de gestión que establece la fuente de confianza para la configuración de los servicios de red y seguridad en el entorno del cliente. El plano de control en sí no procesa el tráfico (que es la función del plano de datos). Indica al plano de datos qué políticas aplicar y gestiona las configuraciones en diferentes centros de datos.

Los servicios del plano de control de Cloudflare se implementan generalmente en una topografía centralizada, más tradicional, en tres centros de datos relacionados lógicamente pero independientes en una región principal (p. ej., los Estados Unidos). Estos tres centros de datos se replican con una capacidad equivalente en una región secundaria (p. ej., la UE). Los servicios del plano de control están diseñados para ser resistentes y mantener una prestación de servicios constante en caso de fallo de cualquier centro de datos en la ubicación principal. La pérdida de centros de datos adicionales en la ubicación principal desencadenaría una conmutación por error a los centros de datos de la otra región (p. ej., en Europa frente a los Estados Unidos).

De acuerdo con el enfoque de Cloudflare sobre la resiliencia antes de la recuperación, seguimos invirtiendo para fortalecer nuestra postura.

Por ejemplo:

- Utilizamos cada vez más las dos zonas del plano de control en una configuración activa-activa, lo que aumenta simultáneamente nuestra capacidad y la capacidad de respuesta, así como nuestra tolerancia al fallo. En consecuencia, podemos soportar más tipos de fallos sin interrumpir el servicio ni necesitar la conmutación por error.
- También estamos incrementando el detalle de la forma en la que trasladamos los servicios entre los distintos sitios del plano de control, lo que nos permite responder con mayor precisión a los problemas de infraestructura local.

Ingeniería de caos

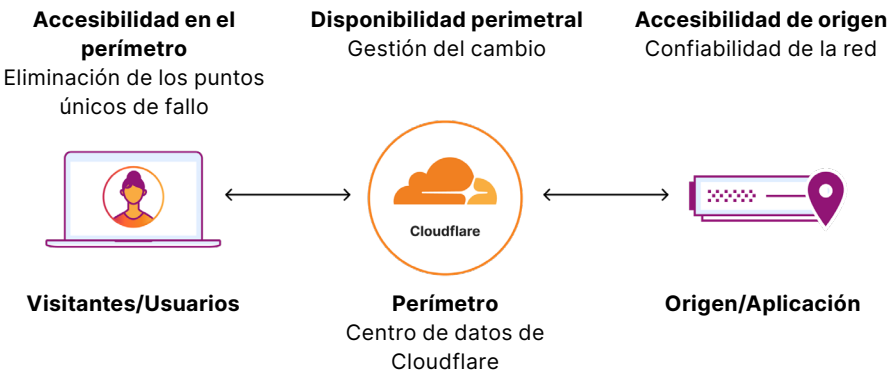
La resiliencia no es una actividad que se configura y se olvida. Incluso los planes de resiliencia más sólidos pueden resultar ineficaces debido a la "desviación" del sistema, la acumulación lenta, muchas veces imperceptible, de cambios que pueden degradar los comportamientos previstos e introducir modos de fallo imprevistos. Para abordar este riesgo de manera proactiva, Cloudflare utiliza la ingeniería de caos con regularidad y analiza sistemáticamente posibles vulnerabilidades relacionadas con la deriva.

Resiliencia del plano de datos

El plano de datos procesa el tráfico de los clientes de Cloudflare de acuerdo con las políticas establecidas desde el plano de control. Aunque los servicios del plano de datos toman la dirección del plano de control, no dependen del plano de control para funcionar. Todas las políticas se mantienen a través de Quicksilver, nuestro almacén de clave-valor distribuido a nivel global, para garantizar que los servicios sigan operativos con una configuración con buena reputación en caso de que se produzca una interrupción de la comunicación con el plano de control.

Anycast desempeña un papel importante en la redundancia de los centros de datos. Los centros de datos de Cloudflare, ubicados en más de 330 ciudades, son autónomos a nivel local y, sin embargo, son intercambiables entre sí a través de Anycast y BGP. Esto se debe a que cada centro de datos puede procesar localmente cualquier servicio, sin ser codependiente de los servicios de otro centro de datos. Con Anycast, cada centro de datos es efectivamente redundante con los demás. Como todos los centros de datos participan en Anycast, no es necesario indicar al cliente que cambie a un centro de datos alternativo en otra dirección IP.

Ya sea un consumidor que visita un sitio web protegido por Cloudflare, un empleado que accede a aplicaciones conectadas a Internet o una oficina que se conecta a su WAN, todos estos escenarios utilizan BGP para encontrar el centro de datos Anycast de Cloudflare más cercano. Si el centro de datos de su elección no estuviera disponible, el BGP se resolvería automáticamente en el siguiente mejor centro de datos de Cloudflare.



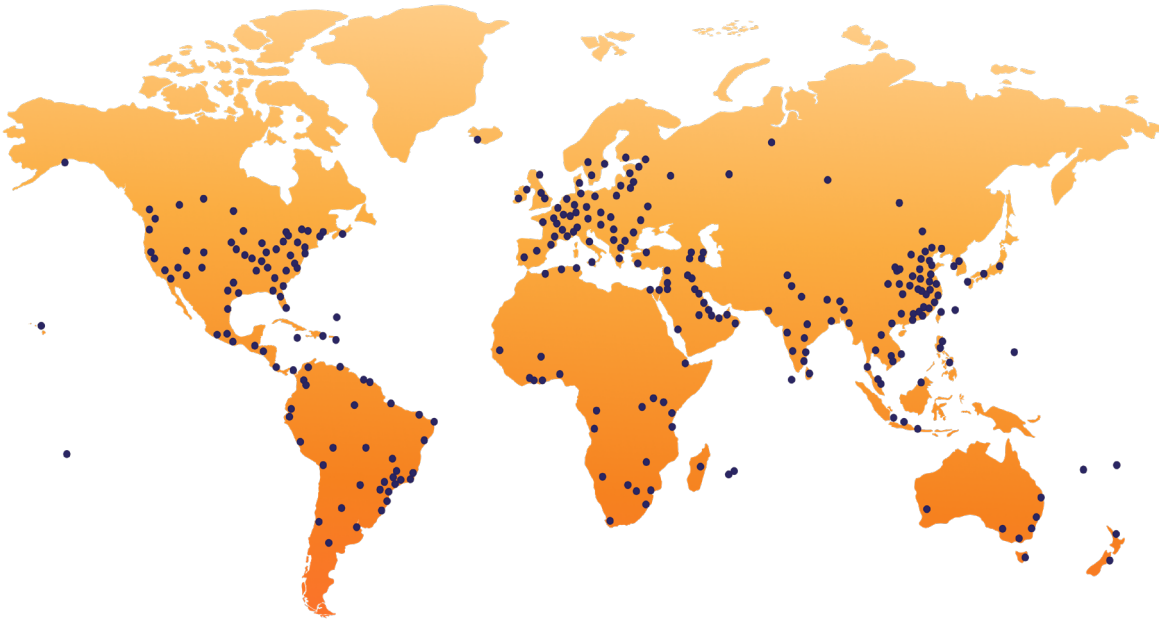
Cloudflare aborda la resiliencia del plano de datos mediante la solución a tres problemas diferentes:

- Accesibilidad en el perímetro: garantizar que el tráfico de los usuarios finales pueda llegar a los centros de datos y eliminar los puntos únicos de fallo en el ingreso del tráfico.
- Disponibilidad perimetral: mantener la calidad del código y el tiempo de actividad del software mediante una gestión de cambios rigurosa.
- Accesibilidad de origen: adaptar el enrutamiento a las aplicaciones del cliente para garantizar que no haya pérdidas en las rutas de salida.

Accesibilidad en el perímetro

La capacidad de alcance en el perímetro es la habilidad de los usuarios finales para llegar a la red de Cloudflare. Podría decirse que es la pieza más importante del ámbito del problema de la resiliencia. Si los proveedores de servicios de Internet (ISP) o los centros de datos dejan de funcionar, la accesibilidad del perímetro se reduce o degrada, lo que impide o disminuye el acceso de los usuarios a Internet. Cloudflare aborda los problemas de accesibilidad en el perímetro de cuatro maneras clave:

1. Red Anycast



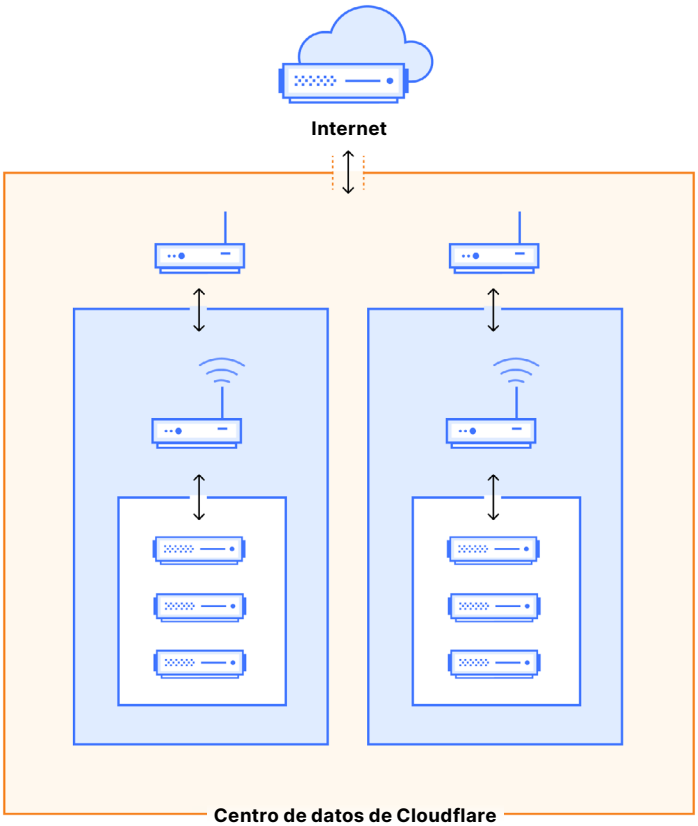
Nuestra arquitectura de red, que depende en gran medida de la tecnología Anycast, incorpora un mejor rendimiento de red con resiliencia. Anycast, la superpotencia de ingeniería de Cloudflare, significa que el espacio IP se anuncia en todas partes: si alguno de nuestros puntos de presencia (PoP) está fuera de línea, el tráfico simplemente se enrutará a otras ubicaciones en lugar de interrumpirse. El hecho de que Cloudflare esté presente en tantas ciudades de todo el mundo e interconectado con redes locales significa que procesamos el tráfico de los clientes lo más cerca posible del usuario. Por ejemplo, incluso si se desconecta el tránsito de un proveedor de servicios de Internet o si un centro de datos se queda sin energía, el tráfico no se ve afectado.

2. Todos los servicios de Cloudflare se ejecutan en todas las ubicaciones.

Además de Anycast, los centros de datos de Cloudflare están diseñados para procesar el tráfico de forma local, sin ser codependientes de las cadenas de proxy con otros recursos de procesamiento. Ejecutamos casi todos los servicios en todas las máquinas. Esto significa que un centro de datos se puede desconectar sin afectar a los clientes. Las máquinas dentro del centro de datos son reemplazables entre sí porque hay cientos que ejecutan el mismo servicio y pueden intervenir en caso de que una falle.

3. Puntos de presencia en múltiples centros de datos de colocación

El diseño y las ubicaciones de los centros de datos de Cloudflare reflejan las necesidades de nuestros clientes. Una red Anycast nos permite agregar y eliminar centros de datos a voluntad, pero debemos monitorear constantemente el rendimiento del cliente. Hemos adaptado las topologías de nuestros centros de datos para permitir que las secciones de capacidad informática (colocaciones) fallen de forma independiente entre sí. Estos centros de datos (con múltiples colocaciones) se denominan puntos de presencia de múltiples colocaciones o ubicaciones MCP.



Estas ubicaciones dividen la conectividad orientada a Internet de la conectividad informática interna para permitir que los centros de colocaciones se desconecten individualmente. Esto significa que incluso si hay un problema con un centro de colocaciones, todo el punto de presencia de una región puede permanecer en línea, lo que proporciona un mayor tiempo de actividad y rendimiento a un cliente. Este tipo de centro de datos también elimina los puntos únicos de falla al tener dispositivos redundantes en la capa orientada a Internet: si un enrutador orientado a Internet (perimetral) falla, el otro enrutador perimetral puede tomar el tráfico y garantizar que la ubicación siga funcionando.

Este modelo operativo ayuda a las ubicaciones MCP a evitar movilizar el tráfico a menos que sea absolutamente necesario y aumenta aún más el tiempo de actividad de los clientes de Cloudflare.

4. Cloudflare Traffic Manager

Los centros de datos MCP trabajan juntos para formar la red de Cloudflare. Esta red aprovecha Anycast para garantizar el servicio de tráfico del cliente. Anycast se ha mejorado con la gestión determinista del tráfico para garantizar que las solicitudes de los clientes se atiendan donde podamos ofrecer el mejor rendimiento posible. Este sistema de gestión de tráfico, Traffic Manager, funciona al sondear de forma continua la red de Cloudflare y alejar automáticamente el tráfico de los centros de datos que experimentan una sobrecarga de la CPU. Esto evita la congestión en los centros de datos de alto tráfico; en su lugar, el tráfico se enruta de forma inteligente a otro centro de datos que puede gestionarlo.

Disponibilidad perimetral

La disponibilidad perimetral se refiere a la capacidad de Cloudflare para procesar el tráfico una vez que llega a nuestra red. Cuando los cambios en las herramientas de red o el software provocan cambios no deseados, la disponibilidad puede disminuir y afectar la experiencia del usuario. Para evitar incidentes derivados de los cambios de código, Cloudflare ha invertido mucho en controles de implementación, como ser:

1. Embudo de implementación

Al implementar un software, garantizar la calidad del código se inicia con monitoreo y limitación de la capacidad de los desarrolladores y clientes para introducir cambios en el ecosistema. Cloudflare limita la cantidad de formas en que se pueden introducir cambios en nuestra infraestructura para que podamos monitorear de cerca cada cambio y asegurarnos de que pasen una serie de pruebas antes de implementarse en producción.

2. Gestión de cambios en el radio de impacto

Otra forma en que Cloudflare respalda la disponibilidad perimetral durante la implementación es mediante la limitación de las implementaciones a los centros de datos de prueba, o a los grupos de prueba, antes de implementarlas de manera generalizada. Nos referimos a esto como la gestión del radio de impacto.

Cuando se implementan los cambios en la red, pueden estar disponibles a nivel global en cuestión de minutos. Al contener los cambios en un entorno de implementación y desplegar más cambios en cascada, podemos monitorear los efectos del cambio en busca de consecuencias previstas o imprevistas, antes de afectar áreas geográficas o poblaciones de usuarios más grandes.

Tenemos dos formas de limitar el impacto de los cambios de código:

- Limitar el número de ubicaciones que reciben cambios.
- Limitar el número de usuarios que reciben cambios.

Al limitar la cantidad de ubicaciones y máquinas que reciben cambios, podemos asegurarnos de que estamos realizando pruebas A/B adecuadas del código dentro de una sola ubicación para evaluar el estado antes de continuar. Al limitar el número de usuarios que reciben los cambios, podemos probar las correcciones en grupos de usuarios más reducidos primero.

3. Implementación mediada por el estado

La implementación mediada por el estado es un sistema que evalúa mediante programación la idoneidad de una versión en función de métricas preestablecidas que dan una señal de "aprobación" o "no aprobación" en función del impacto potencial. Esta serie de controles automatizados no solo puede evitar que se produzca una liberación dañina, sino que también puede revertirla al detectar un impacto.

Cada producto y servicio implementado a través de Cloudflare debe tener un objetivo de nivel de servicio (SLO), que incluye tanto una métrica que representa el estado del producto como un umbral por debajo del cual un producto se consideraría en mal estado.

Los SLO tienen tasas de consumo o límites de fallo aceptables. Cualquier servicio mediado por el estado brindará los SLO a un sistema automatizado como parte de la fusión de un cambio que se implementará. En cada ámbito de implementación establecido (planes gratuitos, un subconjunto de máquinas en Ashburn, etc.), el sistema automatizado:

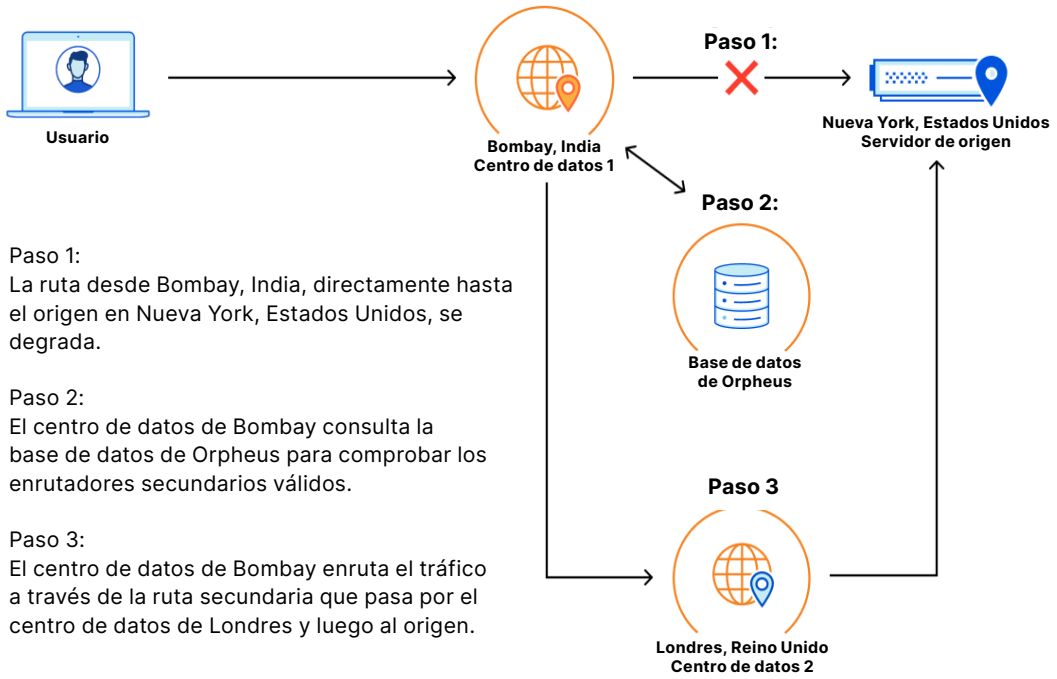
- En primer lugar, supervisara el SLO del servicio durante un periodo de tiempo determinado para asegurarte de que el estado no está por debajo del límite.
- El sistema avanzara automáticamente en la implementación a una etapa más grande, si el SLO se mantiene dentro de los rangos aceptables durante el período de almacenamiento establecido.
- Sin embargo, si se infringe el SLO, el sistema detiene automáticamente la implementación y retrocede para mitigar el impacto de forma automática.

Estos pasos garantizan que el estado de los clientes no se vea afectado por las implementaciones de código y que la duración sea lo más breve posible.

Accesibilidad de origen

La accesibilidad de origen se refiere a la capacidad de Cloudflare para llegar a los destinos, ya sea el origen de un cliente, un sitio SaaS o la Internet pública a través de Cloudflare Gateway. El enrutamiento de las solicitudes a donde deben estar es crucial para que los usuarios accedan a la red de Cloudflare. Por ejemplo, Argo Smart Routing, la herramienta de Cloudflare que optimiza el rendimiento (es decir, el tiempo hasta el primer byte), sondea constantemente la red de Cloudflare para encontrar la ruta más rápida a los orígenes.

Orpheus, la contraparte de Argo, tiene una filosofía similar pero una función diferente. Orpheus existe para establecer conexiones fiables a los servidores de origen. Orpheus analiza específicamente las métricas que afectan la capacidad de Cloudflare para llegar al origen (a diferencia de la ruta más rápida al origen) y encontrará rutas que minimicen la pérdida de paquetes sin afectar el rendimiento de estado estable. Esto significa que cuando surgen problemas, el tráfico se enruta automáticamente alrededor de los errores detectados.



Antes de que Cloudflare lanzara Orpheus en 2021, pudimos enrutar con éxito a los servidores de origen el 99,9 % de las veces. Después de implementar Orpheus, nuestra capacidad de enrutamiento a los orígenes aumentó al 99,99 %. El próximo año, ampliaremos Orpheus para proteger más tipos de tráfico, actuar en más escenarios de fallo y trabajar más rápido para reducir el tiempo que los usuarios se ven afectados.

Compromiso con la transparencia operativa

Incluso las redes más resilientes e innovadoras experimentarán interrupciones. Cuando surgen incidentes y los clientes se ven afectados, Cloudflare sigue una respuesta de comunicación de incidentes que incluye una investigación exhaustiva, un informe interno y un informe externo del incidente y, si es necesario, ofrece actualizaciones del estado durante todo el periodo del impacto.

En ciertos casos en los que los incidentes tengan cierto impacto, o conlleven una innovación, se publicarán análisis post mortem en el Blog de Cloudflare.

Conclusión

El esfuerzo de ingeniería detrás de la red de Cloudflare no es un trabajo fácil, pero es un trabajo que hacemos con orgullo para nuestros clientes. La recompensa es crear una plataforma de red que beneficie a nuestros clientes y a la comunidad de Internet en general.

En última instancia, al priorizar la resiliencia no solo como una preocupación técnica sino como una filosofía operativa central, estamos haciendo algo más que reforzar las defensas: estamos construyendo activamente una empresa inherentemente preparada para el futuro. El enfoque proactivo de pruebas y adaptaciones continuas significa que podemos evolucionar con elegancia junto con las demandas en constante cambio tanto de nuestros clientes como del panorama dinámico de Internet.

Entendemos que muchos de los conceptos que se representan en este documento se centran en conceptos de red a los que muchas empresas no están expuestas, ya que implican la arquitectura interna del funcionamiento de un entorno de nube global de clase operadora.

Para obtener más información sobre la ingeniería de resiliencia de Cloudflare, comunícate con tu representante de Cloudflare.

+55 (11) 3230 4523 | enterprise@cloudflare.com | www.cloudflare.com/es-la/

REV:BDES-8251.2025OCT28