CLOUDFLARE

WHITEPAPER

Cloudflare network and service resilience

Content

Overview

The Internet is built on imperfect systems that are designed to prioritize uptime over everything else. Protocols like TCP and BGP build upon the principles of distributed systems — expect failures and account for them — and the Internet was designed accordingly. However, points of failure still exist and can cause impact. Network providers need to be able to plan for failures, detect them when they happen, and mitigate the impact users experience.

Due to the real-time nature of many applications on the Internet, networks need to not only cover as much of the Internet as possible, but they also have to respond and mitigate impact in real time. Downtimes on the order of minutes is not acceptable: customers expect remediation in mere seconds.

Cloudflare delivers critical network infrastructure services to support how organizations secure and communicate over the Internet. We have developed our network and the services it provides to maintain the highest level of operating excellence. Cloudflare processes over 84 million HTTP requests and 61 million DNS queries per second on average, providing service to millions of Internet properties and users.

How does Cloudflare provide reliable service at this scale given the Internet's unpredictable characteristics?

The answer comes from the architecture of the Cloudflare network, which operates with resilience capabilities designed to operate independently and withstand the spectrum of disruption. Our compute, networking, and storage capabilities, alongside our operating processes, are designed to make Cloudflare as dependable as the plain old telephone service network's dial tone. Cloudflare delivers the metaphorical "cloud tone" for the networking and security services that customers rely on, despite the conditions of the Internet at any given time.

This whitepaper dives into the challenges of operating in real time on the Internet, and how Cloudflare is uniquely positioned to solve those challenges.

Living in an imperfect world

The Internet is an imperfect place, and yet organizations need it to build their business and run organizations that link distributed users, data, and devices to applications in the cloud. Any service disruption has serious implications. Over the years, Cloudflare has reported on a number of major Internet service disruptions around the world, ranging from accidents to DDoS attacks to natural disasters, and more.

The Internet is a diverse network that has been built as a loosely coupled collective of thousands of participating networks of various sizes and capacities. These networks have different service levels, with some operating as best as they can as an extension of goodwill, and others operating as part of their commercial services. Through mutual agreements to exchange traffic, which may or may not be destined for hosts on their own network, the Internet functions as a global fabric through the benevolent participation of regional and local Internet exchanges and service providers.

However, with such diversity comes a degree of unpredictability. The path that any given packet takes relies on a sequence of best guesses and best efforts for delivery. Absent real-time data or predictive modeling on the actual network conditions at a given time, a packet typically takes its next hop either using default routes or the presumptive shortest path. Neither of these options take the state of the network service in its routing decisions. This means that the packets are often subject to a number of debilitating conditions:

- At the most basic level, networks can experience temporary glitches and brownouts that reduce throughput and cause packet loss.
- As networks become saturated, congestion hurts performance and the user experience.
- While those types of slowdowns can happen under normal operating conditions, a growing number of disruptions are intentionally caused by hostile actors, who maliciously consume the available resources.

At Cloudflare, our mission is to do our part to build a better Internet. Towards that end, we build infrastructure to make the Internet faster, more reliable, and more secure. This is made possible by services from and through the Cloudflare network, one of the largest networks in the world that operates on our own bare metal servers (without virtualization), a private backbone, and a massive global footprint that is reachable, on average, within 50 ms of 95% of the world's population.[1]

Customers often want to understand how this is done. How does Cloudflare engineer services at such a scale, with no trade-offs between security and performance, and with the dependability that organizations need? The answer comes from a set of North Star goals, which establish how we build our infrastructure to remain performant despite disruptive forces, both external and internal.

How Cloudflare architects for resilience

Many organizations focus on improving availability by trying to get better or faster at reacting to failures. This requires constantly investing in and testing their failover capabilities and processes. While these are valid goals, Cloudflare thinks differently — our resilience engineering teams spend enormous effort to shrink the scenarios where a disaster recovery response is required.

Cloudflare resilience engineering starts with a simple premise: **How would you build critical infrastructure that remains operational assuming that failures will happen?**

When failures inevitably happen, Cloudflare's resilient services detect and isolate the failures so that they do not impact service availability. Failures get resolved out-of-band from our service delivery. We strive to be failure agnostic across our entire fleet.

In order to understand the principles for resilience, it's useful to first define key concepts behind traffic paths with Cloudflare and the design goals for key systems. At the most abstract level, Cloudflare's architecture can be separated into two segments: **the control plane and the data plane.** Each has a unique resilience and disaster posture.

Control Plane Resilience

The control plane provides the management interface that establishes the source of truth for the configuration of networking and security services in the customer's environment. The control plane itself does not process traffic (which is the role of the data plane). It tells the data plane which policies to enforce, and manages configurations across different data centers.

Cloudflare's control plane services are generally deployed in a more traditional, centralized topography across three logically related but independent data centers in a primary region (e.g., US). These three data centers are replicated with equivalent capacity into a secondary region (e.g., EU). The control plane services are designed to be resilient and maintain consistent service delivery in the event of a failure of any single data center in the primary location. The loss of additional data centers in the primary location would trigger a failover to the other region's data centers (e.g., in Europe versus the US).

In keeping with Cloudflare's focus on resilience before recovery, we continue to invest in deepening our resilience posture.

For example:

- We are increasingly using the two control plane regions in an active-active configuration, which simultaneously increases our capacity / responsiveness, as well as our failure tolerance. Consequently, we can withstand more types of failures without service disruption or need to failover.
- We are also increasing the granularity of how we move services between the various control plane sites, allowing us to respond with more precision to local infrastructure issues.

Chaos Testing

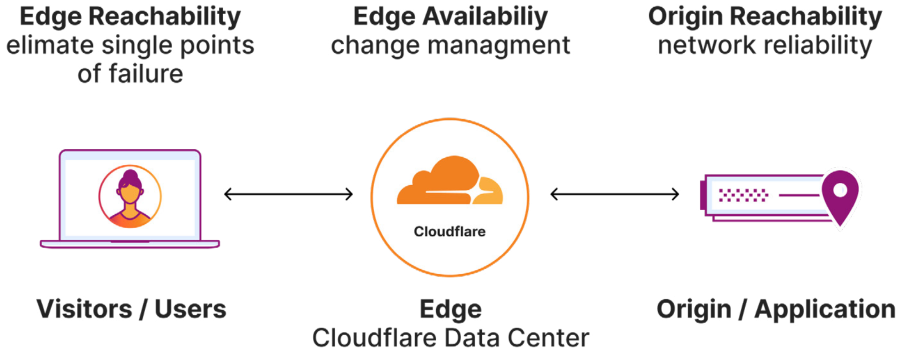
Resilience isn't a set-it-and-forget-it activity. Even the most robust resilience plans can prove to be ineffective because of system "drift" — the slow, often imperceptible accumulation of changes that can degrade intended behaviors and introduce unforeseen failure modes. To proactively address this risk, Cloudflare chaos tests regularly, systematically probing for potential drift-related vulnerabilities.

Data Plane Resilience

The data plane processes Cloudflare customers' traffic in accordance to the policies established from the control plane. Although data plane services take direction from the control plane, it does not depend upon the control plane to operate. All policies are maintained through Quicksilver, our globally distributed key-value store, to ensure that services remain operational with a known good configuration in the event of any communication disruption with the control plane.

Anycast plays an important role in data center redundancy. The Cloudflare data centers, located in over 330 cities, are locally autonomous and yet interchangeable with one another through Anycast and BGP. That's because each data center can locally process any services, without being codependent on services at another data center. With Anycast, every data center is effectively redundant with the others. Because every data center participates in Anycast, there is no need to instruct the client to switch to an alternate data center at another IP address.

Whether it's a consumer visiting a Cloudflare-protected website, an employee accessing Internet-connected apps, or an office location connecting to their WAN, all of these scenarios use BGP to find the closest Cloudflare Anycast data center. Should their data center of choice become unavailable, BGP would automatically resolve to the next best Cloudflare data center.



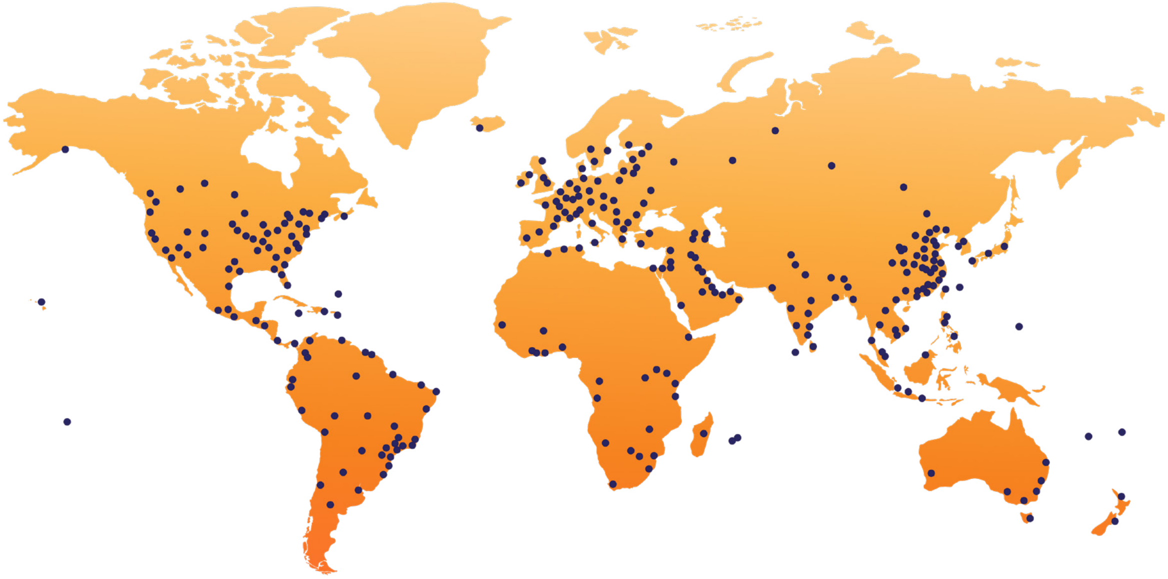
Cloudflare tackles data plane resiliency by solving for three different problems:

- Edge reachability: Ensuring end-user traffic can reach data centers and eliminate single points of failure on traffic ingress
- Edge availability: Maintaining code quality and software uptime through rigorous change management
- Origin reachability: Adaptive routing to customer applications to ensure that there is no loss on the outbound paths

Edge Reachability

Edge reachability is the ability for end users to reach Cloudflare's network. It is arguably the most important piece of the resiliency problem space. If Internet service providers (ISPs) or data centers go down, edge reachability is reduced or degraded, which prevents or slows down users from getting where they need to be on the Internet. Cloudflare addresses edge reachability problems in four key ways:

1. Anycast network



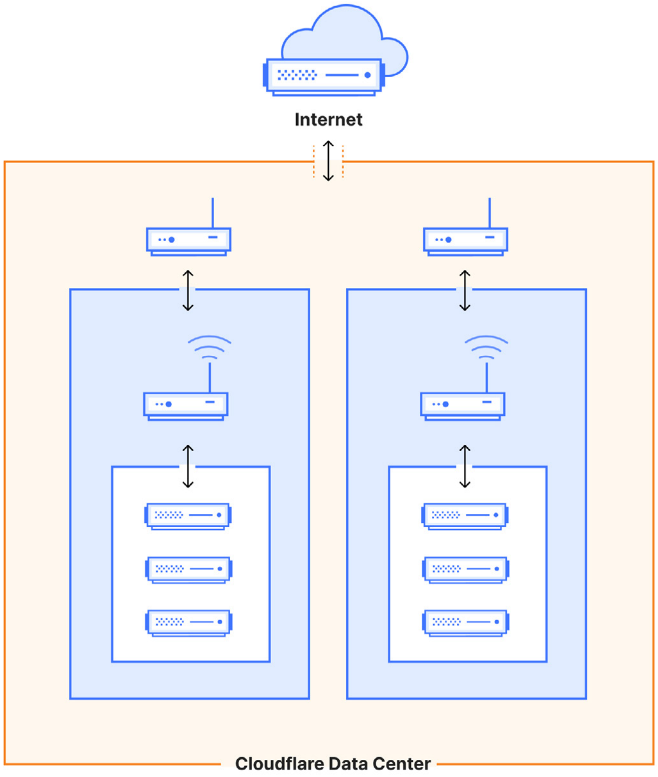
Better network performance with resiliency is built into our network architecture, which relies heavily on Anycast technology. Anycast — Cloudflare's engineering superpower — means that IP space is advertised everywhere: if any of our points of presence (PoPs) are offline, traffic would simply route to other locations rather than dropping. The fact that Cloudflare is present in so many cities worldwide and peered with local networks means that we process customer traffic as close to the user as possible. For instance, even if an ISP transit is disconnected, or a data center loses power, traffic remains unimpacted.

2. Every Cloudflare service runs in every location

In addition to Anycast, Cloudflare data centers are designed in a manner to process traffic locally, without being codependent on proxy chains to other computes. We run almost every service on every machine. This means that a data center can be taken offline without customer impact. Machines within the data center are replaceable with one another because there are hundreds running the same service that are capable of stepping in, in the event that one fails.

3. Multi-colo PoPs

Cloudflare's data center design and locations reflect the needs of our customers. An Anycast network allows for us to add / remove data centers at will, but we must constantly monitor customer performance. We've adapted our data center topologies to allow sections of compute capacity (colos) to fail independently of each other. These data centers (with multiple colos) are called multi-colo points of presence, or MCP locations.



These locations bifurcate Internet-facing connectivity from internal compute connectivity to allow colos to be taken offline individually. This means that even if there is an issue with a colo, the entire PoP in a region can stay online, providing increased uptime and performance to a customer. This data center type also removes single points of failure by having redundant devices at the Internet-facing layer: if one Internet-facing (edge) router fails, the other edge router can take the traffic and ensure that the location remains operational.

This operating model helps MCP locations avoid moving traffic unless it is absolutely necessary, and further increases Cloudflare customers' uptime.

4. Cloudflare Traffic Manager

MCP data centers work together to form the broader Cloudflare network. This network leverages Anycast to help ensure customer traffic will get served. Anycast is enhanced with deterministic traffic management to ensure that customer requests will get served where we can serve them, with the best performance possible. This traffic management system, Traffic Manager, works by continuously probing Cloudflare's network, and automatically moving traffic away from data centers that experience CPU overload. This prevents congestion at high-traffic data centers; instead, the traffic intelligently routes to another data center that can handle it.

Edge Availability

Edge availability refers to Cloudflare's ability to process traffic once it hits our network. When changes to network tools or software result in unintended changes, availability can decline and impact the user experience. In order to prevent incidents resulting from code change from happening, Cloudflare has invested heavily in the following deployment controls:

1. Deployment funnel

When deploying software, ensuring code quality starts with keeping track of and limiting the ability for developers and customers to introduce change into the ecosystem. Cloudflare limits the number of ways anyone can introduce change into our infrastructure so that we can closely monitor every change, and ensure they pass a battery of tests before being deployed to production.

2. Blast radius change management

Another way Cloudflare supports edge availability during deployment is by limiting deployments to test data centers, or test groups, before rolling them out widely. We refer to this as blast radius management.

When changes to the network are rolled out, they can go live globally within minutes. By containing changes to a deployment environment and rolling out further changes in a waterfall, we can monitor the effects of the change for intended or unintended consequences — before affecting larger geographies or user populations.

We have two ways to limit impact from code changes:

- Limit the number of locations that receive changes; and
- Limit the number of users that receive changes

By limiting the number of locations and machines that receive changes, we can ensure that we are properly A/B testing code within a single location to evaluate health before proceeding. By limiting the number of users that receive changes, we can battle test fixes to smaller user populations first.

3. Health mediated deployment

Health-mediated deployment is a system that programmatically evaluates the suitability of a release based on preset metrics that give a “go” or “no go” signal based on potential impact. This series of automated checks can not only prevent a harmful release from going out, but it can roll back a release upon sensing impact.

Every product and service deployed via Cloudflare must have a service level objective (SLO), which contains both a metric representing the health of the product and also a target below which a product would be considered unhealthy.

SLOs have burn rates, or acceptable thresholds of failure. Any health-mediated service will provide SLOs to an automated system as part of merging a change to be deployed. At every set scope of deployment (free plans, a subset of machines in Ashburn, etc.), the automated system will:

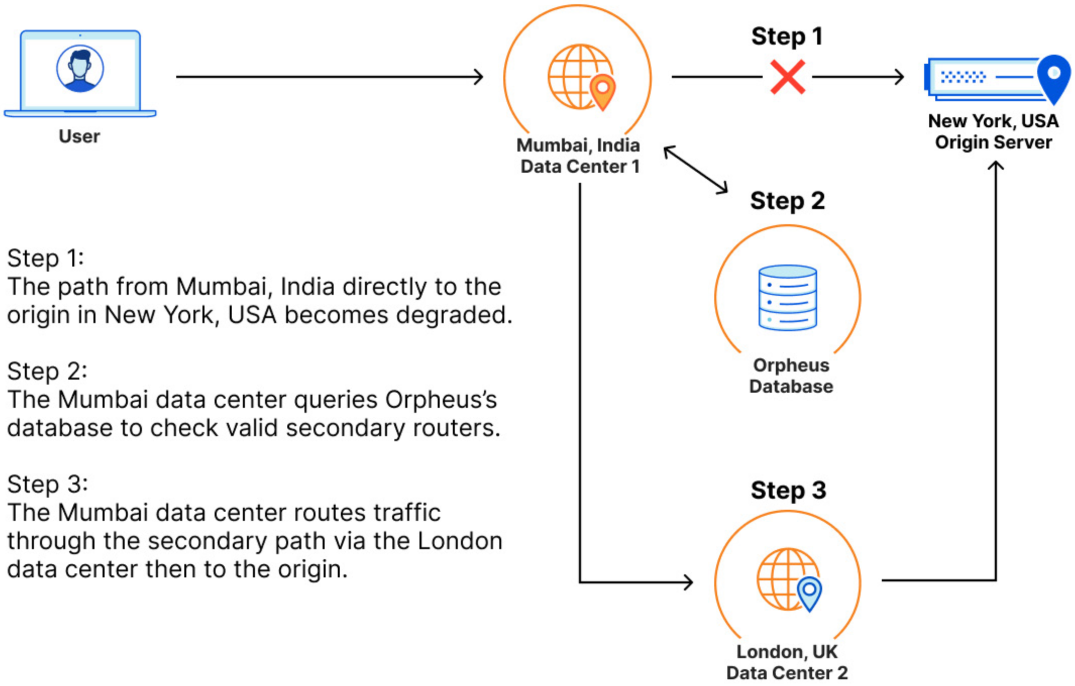
- First, monitor the service’s SLO for a set period of time to ensure that the health is not falling below the threshold.
- If the SLO stays within acceptable ranges for the duration of the set soaking period, the system will automatically progress the deployment to a larger stage.
- However, if the SLO is breached, the system automatically halts the deployment and rolls back so the impact is mitigated automatically.

These steps ensure that customer health will not be impacted by code deployments and that the duration is as short as possible.

Origin Reachability

Origin reachability refers to Cloudflare's ability to reach destinations, whether that is a customer origin, a SaaS site, or the public Internet through Cloudflare Gateway, routing requests to where they need to be is crucial for users accessing Cloudflare's network. For instance, Argo Smart Routing, Cloudflare's tool that optimizes for performance (namely time to first byte), constantly probes Cloudflare's network to find the fastest path to origins. Argo constantly probes Cloudflare's network, finding the fastest path to origins.

Orpheus, the counterpart to Argo, has a similar philosophy but a different function. Orpheus exists to establish reliable connections to origin servers. Orpheus specifically looks at metrics that impact Cloudflare's ability to reach the origin (as opposed to the fastest path to the origin), and will find paths that minimize packet loss without impacting steady state performance. This means that when problems arise, traffic is automatically routed around detected errors.



Step 1:
The path from Mumbai, India directly to the origin in New York, USA becomes degraded.

Step 2:
The Mumbai data center queries Orpheus's database to check valid secondary routers.

Step 3:
The Mumbai data center routes traffic through the secondary path via the London data center then to the origin.

Before Cloudflare released Orpheus in 2021, we were able to successfully route to origins 99.9% of the time. After implementing Orpheus, our ability to route to origins increased to 99.99%. In the coming year, we will be expanding Orpheus to protect more types of traffic, action on more failure scenarios, and work faster to reduce the time that any user is impacted.

Commitment to operational transparency

Even the most resilient and innovative networks will experience outages. When incidents arise and customers are impacted, Cloudflare follows an incident communication response that includes a thorough investigation, an internal incident report, an external incident report, and, if necessary, status updates throughout the window of impact.

In certain cases where incidents result in such impact — or innovation — post-mortems will be published on the Cloudflare Blog.

Conclusion

The engineering effort behind Cloudflare's network is not easy work, but it's work that we proudly do for our customers. The reward is building a network platform that benefits our customers and the broader Internet community as a whole.

Ultimately, by prioritizing resilience not just as a technical concern but as a core operational philosophy, we're doing more than just shoring up defenses: we're actively building a company that's inherently future-ready. The proactive approach of continuously testing and adapting means we can gracefully evolve alongside the ever-changing demands of both our customers and the dynamic landscape of the Internet itself.

We understand that many of the concepts that are represented within this document center around networking concepts that many enterprises are not exposed to, as they involve the inner architecture of operating a global carrier-class cloud environment.

If you'd like an in-depth briefing to learn more about Cloudflare resilience engineering, contact your Cloudflare representative.

This document is for informational purposes only and is the property of Cloudflare. This document does not create any commitments or assurances from Cloudflare or its affiliates to you. You are responsible for making your own independent assessment of the information in this document. The information in this document is subject to change and does not purport to be all inclusive or to contain all the information that you may need. The responsibilities and liabilities of Cloudflare to its customers are controlled by separate agreements, and this document is not part of, nor does it modify, any agreement between Cloudflare and its customers. Cloudflare services are provided "as is" without warranties, representations, or conditions of any kind, whether express or implied.

© 2025 Cloudflare, Inc. All rights reserved. CLOUDFLARE® and the Cloudflare logo are trademarks of Cloudflare. All other company and product names and logos may be trademarks of the respective companies with which they are associated.