

WHITEPAPER

Ausfallsicherheit des Netzwerks und der Dienste von Cloudflare



Inhalt

3	Übersicht
4	Leben in einer unvollkommenen Welt
5	So sorgt Cloudflare für Ausfallsicherheit
5	Ausfallsicherheit auf Steuerungsebene
6	Ausfallsicherheit auf Datenebene
7	Edge-Erreichbarkeit
9	Edge-Verfügbarkeit
11	Erreichbarkeit des Ursprungsservers
12	Verpflichtung zu betrieblicher Transparenz
12	Fazit

Übersicht

Das Internet beruht auf unvollkommenen Systemen, die darauf ausgelegt sind, der Verfügbarkeit absoluten Vorrang einzuräumen. Protokolle wie [TCP](#) und [BGP](#) folgen den [Prinzipien verteilter Systeme](#): Ausfälle sind zu erwarten und müssen eingeplant werden. Nach dieser Logik ist auch das Internet aufgebaut. Das heißt aber nicht, dass es im Web keine Schwachstellen gibt, die schwerwiegende Konsequenzen haben können. Netzwerkanbieter müssen Ausfälle einplanen, sie erkennen, wenn sie auftreten, und die Auswirkungen für die Nutzer mindern können.

Weil es sich bei bei vielen Webanwendungen um Echtzeitsysteme handelt, müssen Netzwerke nicht nur so viel wie möglich vom Internet abdecken, sondern auch in Echtzeit reagieren und Auswirkungen abmildern. Minutenlange Ausfälle sind nicht hinnehmbar: Die Kunden erwarten, dass das Problem innerhalb von Sekunden gelöst wird.

Cloudflare stellt wichtige Netzwerkinfrastrukturdienste zur Unterstützung der Absicherung und Kommunikation von Unternehmen über das Internet bereit. Unser Netzwerk und die darüber verfügbaren Dienste sind so angelegt, dass sie ein Höchstmaß an Betriebsqualität gewährleisten. Cloudflare verarbeitet im Durchschnitt gut 84 Mio. HTTP-Anfragen und 61 Mio. DNS-Anfragen pro Sekunde für Millionen von Internetauftritten und Nutzern.

Wie gelingt es Cloudflare trotz der Unvorhersehbarkeit des Internets, einen zuverlässigen Service dieser Größenordnung zu bieten?

Die Erklärung findet sich in der Architektur des [Cloudflare-Netzwerks](#). Dieses verfügt über Resilienzfunktionen, die unabhängig arbeiten und sämtlichen, wie auch immer gearteten Störungen standhalten sollen. Unsere Rechen-, Netzwerk- und Speicherkapazitäten sowie unsere Betriebsprozesse sind darauf ausgelegt, Cloudflare so zuverlässig zu machen wie den Wählton eines klassischen Festnetztelefons. Cloudflare liefert den metaphorischen „Cloud-Ton“ für die von den Kunden genutzten Netzwerks- und Sicherheitsdienste – unabhängig von den Bedingungen, die im Internet gerade herrschen.

In diesem Whitepaper gehen wir auf die Herausforderungen des Echtzeit-Betriebs im Internet ein und erläutern, weshalb Cloudflare einzigartig gut aufgestellt ist, um diese zu lösen.

Leben in einer unvollkommenen Welt

Das Internet ist nach wie vor alles andere als perfekt, doch Unternehmen sind darauf angewiesen: nicht nur für den Aufbau ihres Geschäfts, sondern auch zur Verbindung von Nutzern, Daten und Geräten an verschiedenen Orten mit Anwendungen in der Cloud. Jede Unterbrechung hat schwerwiegende Folgen. Im Lauf der Jahre hat Cloudflare [über eine Reihe von größeren Internetstörungen auf der ganzen Welt berichtet](#), deren Ursachen von Unfällen über DDoS-Angriffe bis hin zu Naturkatastrophen reichten.

Das Web ist ein facettenreiches, loses Gefüge aus Tausenden Netzwerken unterschiedlicher Größe und Kapazität. Diese bieten verschiedene Servicegrade. Manche operieren kostenlos und werden so gut wie eben möglich aufrechterhalten, andere sind Teil des gewerblichen Angebots eines Betreibers. Mittels bilateraler Vereinbarungen über den Austausch von Traffic, der nicht zwingend für Hosts im eigenen Netzwerk bestimmt ist, funktioniert das Internet als globales Gefüge auf Grundlage der gutwilligen Beteiligung regionaler und lokaler Internet-Knotenpunkte und Dienstleister.

Allerdings bringt diese Vielfalt auch eine gewisse Unvorhersehbarkeit mit sich. Welchen Weg ein bestimmtes Datenpaket einschlägt, wird durch eine Reihe von Vermutungen bestimmt. Dabei wird die größtmöglichen Anstrengungen unternommen, damit das Paket sein Ziel erreicht. Mangels Echtzeitdaten oder einer vorausschauenden Modellierung der tatsächlichen Netzwerkbedingungen zu einem bestimmten Zeitpunkt folgt ein Paket für seinen nächsten Hop normalerweise Standardrouten oder dem aller Wahrscheinlichkeit nach kürzesten Weg. Der Status des Netzwerkdiensts wird bei diesen [Routing-Entscheidungen](#) nicht berücksichtigt. Somit sind die Pakete während der Zustellung oft mit einer Reihe schier unüberwindbarer Hürden konfrontiert:

- Um mit dem Fundament zu beginnen: Kurzzeitige Spannungsabfälle können den Durchsatz einschränken und Paketverluste verursachen, wodurch vorübergehende Störungen auftreten.
- Mit zunehmender Beanspruchung der Netzwerke leiden Performance und Nutzererfahrung.
- Solche Beeinträchtigungen treten zwar auch unter normalen Betriebsbedingungen auf, doch inzwischen werden Störungen zunehmend absichtlich und böswillig verursacht, um die verfügbaren Ressourcen zu beanspruchen.

Bei Cloudflare sehen wir es als unsere Aufgabe an, zu einem besseren Internet beizutragen. Deshalb bauen wir eine Infrastruktur für ein schnelleres, zuverlässigeres und sichereres Web auf. Möglich wird dies durch Dienste, die entweder direkt vom Cloudflare-Netzwerk oder darüber bereitgestellt werden. Es handelt sich um eines der größten Netzwerke der Welt, das wir über unsere eigenen physischen Server (ohne Virtualisierung) und ein privates Backbone betreiben. Es weist eine massive globale Präsenz auf und ist im Durchschnitt innerhalb von 50 ms von 95 % der Weltbevölkerung erreichbar.¹

Kunden fragen sich oft, wie uns das gelingt. Wie kann Cloudflare Dienste in dieser Größenordnung und mit der von Unternehmen benötigten Zuverlässigkeit entwickeln, ohne Abstriche bei der Sicherheit oder bei der Performance in Kauf nehmen zu müssen? Die Antwort ergibt sich aus einer Reihe von übergeordneten Zielen für den Aufbau unserer Infrastruktur. So wird gewährleistet, dass sie selbst dann ordentlich funktioniert, wenn sie externen und internen Störfaktoren ausgesetzt ist.

So sorgt Cloudflare für Ausfallsicherheit

Viele Unternehmen versuchen, besser oder schneller auf Ausfälle zu reagieren, um die Verfügbarkeit zu erhöhen. Dafür müssen sie kontinuierlich in [Failover](#)-Kapazitäten und -Prozesse investieren und diese testen. Eine solche Strategie hat zwar durchaus ihre Berechtigung, aber Cloudflare verfolgt einen anderen Ansatz: Unsere für Ausfallsicherheit verantwortlichen Technikabteilungen treiben enormen Aufwand, um die Zahl der Szenarien zu verringern, in denen eine Notfallwiederherstellung erforderlich ist.

Bei Cloudflare setzen wir zum Aufbau der Ausfallsicherheit mit einer einfachen Frage an: **Wie müsste kritische Infrastruktur aufgebaut sein, die bei Ausfällen einsatzfähig bleibt?**

Da Ausfälle nun einmal unvermeidlich sind, werden sie von den robusten Cloudflare-Diensten rechtzeitig erkannt und isoliert, sodass die Verfügbarkeit nicht darunter leidet. Fehler werden außerhalb („Out of band“) unserer Servicebereitstellung behoben. Wir streben Ausfallsicherheit für unser gesamtes System an.

Um die Prinzipien für die Gewährleistung der Ausfallsicherheit zu verstehen, ist es sinnvoll, zunächst die Schlüsselkonzepte für die Traffic-Routen innerhalb des Cloudflare-Netzwerks und die Ziele für die Gestaltung der wichtigsten Systeme zu erläutern. Auf abstraktesten Niveau lässt sich die Architektur von Cloudflare in zwei Segmente unterteilen: **die Steuerungsebene** und **die Datenebene**. Diese weisen jeweils eine einzigartige Ausfallsicherheit und Positionierung für den Katastrophenfall auf.

Ausfallsicherheit auf Steuerungsebene

Bei der Steuerungsebene handelt es sich um die Verwaltungsschnittstelle, durch die eine gemeinsame Grundlage für die Konfiguration der Netzwerk- und Sicherheitsdienste in der Umgebung des Kunden geboten wird. Auf der Steuerungsebene selbst wird kein Traffic verarbeitet (das geschieht auf Datenebene). Sie teilt vielmehr der Datenebene mit, welche Richtlinien durchgesetzt werden sollen, und übernimmt die Steuerung der Konfigurationen für verschiedene Rechenzentren.

Auf Steuerungsebene erfolgt die Bereitstellung der Cloudflare-Dienste im Allgemeinen im Rahmen einer herkömmlichen, zentralisierten Topographie. Diese setzt sich aus drei logisch miteinander verbundenen, aber unabhängigen Rechenzentren in einer Hauptregion (z. B. den USA) zusammen. Diese werden mit gleicher Kapazität in einer alternativen Weltregion (z. B. der Europäischen Union) nachgebildet. Die Dienste auf Steuerungsebene sind so konzipiert, dass sie bei einem Ausfall eines einzelnen Rechenzentrums im Hauptgebiet weiterhin zuverlässig funktionieren und eine beständige Verfügbarkeit gewährleistet ist. Beim Ausfall zusätzlicher Rechenzentren in der Hauptweltregion würde ein Failover zu den Rechenzentren der anderen Weltregion erfolgen (also beispielsweise eine Verlagerung von den USA nach Europa).

Da wir bei Cloudflare den Fokus auf Ausfallsicherheit legen, um eine Notfallwiederherstellung nach Möglichkeit zu vermeiden, investieren wir weiter in den Ausbau unserer Resilienz.

Hier ein paar Beispiele:

- Wir nutzen zunehmend die beiden Weltregionen der Steuerungsebene für eine Active/Active-Konfiguration. Dadurch erhöhen sich gleichzeitig unsere Kapazität/Reaktionsfähigkeit und unsere Ausfalltoleranz. So sind wir in der Lage, ohne Unterbrechungen oder Failover mehr Arten von Ausfällen standzuhalten.
- Wir erhöhen auch die Granularität bei der Verlagerung der Dienste zwischen den verschiedenen Standorten der Steuerungsebene. So können wir präziser auf lokale Infrastrukturprobleme zu reagieren.

Chaos-Tests

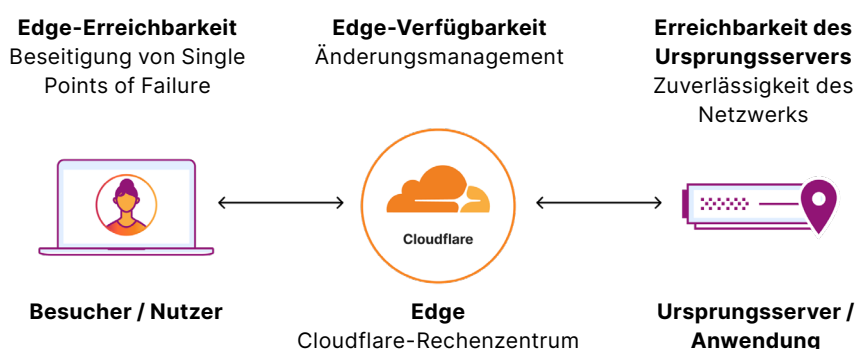
Ausfallsicherheit ist kein Ziel, das sich einfach so einmalig abhaken lässt. Selbst die besten Ausfallsicherheitspläne können sich als nutzlos erweisen, wenn es zu einem Abdriften eines Systems kommt. In solchen Fällen sorgt eine große Zahl von Änderungen über einen langen Zeitraum und oft unmerklich dafür, dass ein System sich nicht so verhält wie erwartet, und unter Umständen unvorhergesehene Fehler auftreten. Cloudflare führt daher zur Vorbeugung regelmäßig Chaos-Tests durch und sucht systematisch nach Schwachstellen, die durch ein solches Abdriften potenziell entstehen können.

Ausfallsicherheit auf Datenebene

Auf Datenebene wird der Traffic der Cloudflare-Kunden in Übereinstimmung mit den auf Steuerungsebene festgelegten Richtlinien verarbeitet. Die Dienste auf Datenebene erhalten zwar Anweisungen von der Steuerungsebene, sind aber für ihren Betrieb nicht von dieser abhängig. Alle Richtlinien werden via [Quicksilver](#) verwaltet, unsere über die ganze Welt verteilte Schlüssel-Werte-Datenbank. Auf diese Weise soll sichergestellt werden, dass die Services im Falle einer Unterbrechung der Kommunikation mit der Steuerungsebene noch auf eine bewährte Konfiguration zurückgreifen können und somit weiterhin einsatzbereit sind.

Anycast spielt bei der Rechenzentrumsredundanz eine wichtige Rolle. Die in mehr als 330 Städten angesiedelten Cloudflare-Rechenzentren sind zwar lokal autonom, dank Anycast und BGP aber trotzdem untereinander austauschbar. Das liegt daran, dass jedes Rechenzentrum die Verarbeitung für beliebige Dienste lokal durchführen kann, ohne von Services in einem anderen Rechenzentrum abhängig zu sein. Durch Anycast kann de facto jedes Rechenzentrum bei einem Ausfall durch jedes andere ersetzt werden. Da jedes Rechenzentrum an Anycast beteiligt ist, muss der Client nicht angewiesen werden, zu einem alternativen Rechenzentrum mit einer anderen IP-Adresse zu wechseln.

Ob es sich um einen Kunden handelt, der eine von Cloudflare geschützte Website besucht, einen Mitarbeiter, der auf mit dem Internet verbundene Anwendungen zugreift, oder ein Bürostandort, der sich mit seinem WAN verbindet – in all diesen Fällen wird mit BGP das nächstgelegene Anycast-Rechenzentrum von Cloudflare ermittelt. Ist das Rechenzentrum der Wahl nicht mehr verfügbar, leitet BGP automatisch zum nächstbesten Cloudflare-Rechenzentrum weiter.



Cloudflare sorgt durch die Lösung von drei Problemen für Ausfallsicherheit auf Datenebene:

- **Edge-Erreichbarkeit:** Es wird sichergestellt, dass der Endnutzer-Datenverkehr die Rechenzentren erreichen kann, und Single Points of Failure beim Traffic-Eingang werden beseitigt
- **Edge-Verfügbarkeit:** Durch rigoroses Änderungsmanagement werden Quellcodequalität und Software-Verfügbarkeit gewährleistet
- **Erreichbarkeit des Ursprungsservers:** Mit adaptivem Routing zu den Kundenanwendungen wird dafür gesorgt, dass auf den Ausgangspfaden keine Verluste entstehen

Edge-Erreichbarkeit

Mit Edge-Erreichbarkeit ist gemeint, dass die Endnutzer das Netzwerk von Cloudflare erreichen können. Sie ist der wohl wichtigste Aspekt bei der Sicherstellung der Ausfallsicherheit. Beim Ausfall von Internet Service Providern (ISP) oder Rechenzentren wird die Edge-Erreichbarkeit eingeschränkt. Die Nutzer gelangen dann entweder gar nicht mehr an den Ort im Internet, den sie erreichen wollen, oder sie brauchen dafür länger. Cloudflare bekämpft Probleme mit der Edge-Erreichbarkeit auf vier Wegen:

1. Anycast-Netzwerk



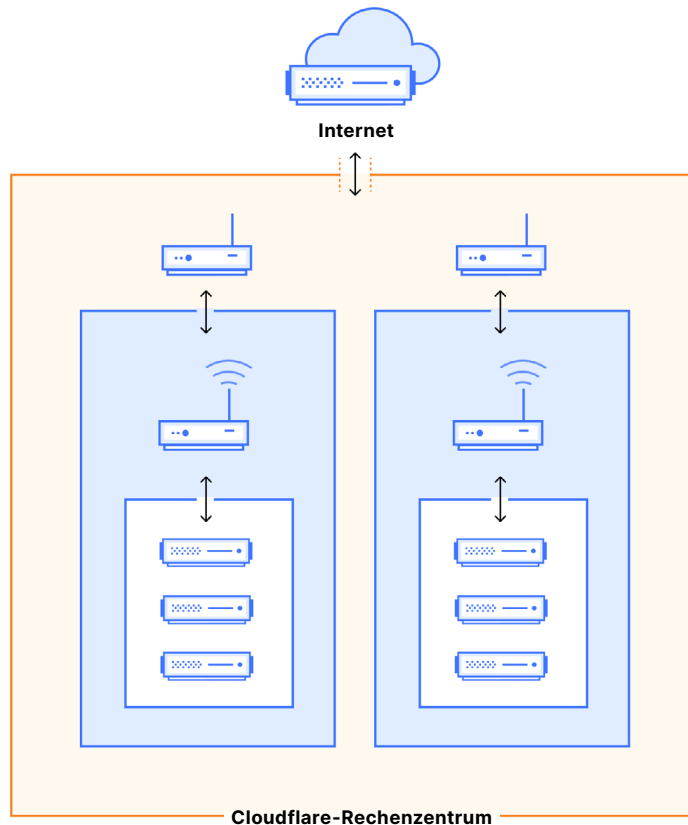
Bei unserer Netzwerkarchitektur, die sich stark auf die Anycast-Technologie stützt, ist eine bessere Performance mit Ausfallsicherheit integriert. Anycast ist eine Art technische Superkraft von Cloudflare. Damit ist gemeint, dass der IP-Adressraum überall bekannt gegeben wird: Falls einer unserer Points of Presence (PoP) offline ist, wird der Traffic somit einfach an andere Standorte weitergeleitet, anstatt verloren zu gehen. Weil Cloudflare in so vielen Städten rund um den Globus präsent ist und [Peering](#) mit lokalen Netzwerken betreibt, können wir den Traffic unserer Kunden in größtmöglicher Nutzernähe verarbeiten. Wenn also beispielsweise die Verbindung zu einem ISP unterbrochen wird oder in einem Rechenzentrum der Strom ausfällt, hat das keine negativen Auswirkungen auf den Datenverkehr.

2. Jeder Cloudflare-Dienst kann an jedem Netzwerkstandort betrieben werden

Abgesehen von der Anycast-Technologie sind die Cloudflare-Rechenzentren so konzipiert, dass sie den Traffic lokal verarbeiten können, ohne von Proxy-Ketten zu anderen Rechenleistungsanbietern abhängig zu sein. Wir betreiben fast jeden Dienst auf jedem Server. Somit kann ein Rechenzentrum ohne Auswirkungen auf die Kunden vom Netz genommen werden. Die Geräte innerhalb des Rechenzentrums sind untereinander ersetzbar, weil auf Hunderten weiteren derselbe Dienst läuft und diese bei einem Ausfall in die Bresche springen können.

3. Multi-Colo-PoP

Die Struktur und die Standorte der Cloudflare-Rechenzentren orientieren sich an den Bedürfnissen unserer Kunden. Ein Anycast-Netzwerk ermöglicht es uns, nach Belieben Rechenzentren hinzuzufügen oder zu entfernen. Allerdings müssen wir die dem Kunden zur Verfügung stehende Performance ständig überwachen. Wir haben die Topologie unserer Rechenzentren so angepasst, dass Abschnitte der Rechenkapazität (Colos) unabhängig voneinander ausfallen können. Diese Rechenzentren (mit mehreren Colos) werden als Multi-Colo-Points of Presence bzw. MCP-Standorte bezeichnet.



An diesen Standorten wird die Internetkonnektivität von der internen Rechenkonnektivität getrennt, damit Colos einzeln offline genommen werden können. So kann bei einem Problem mit einer Colo der gesamte PoP in einer Region online bleiben. Damit wird dem Kunden eine höhere Verfügbarkeit und Performance geboten. Mit dieser Art von Rechenzentrum werden auch Single Points of Failure beseitigt, weil es über redundante Geräte auf internetseitiger Ebene verfügt. Fällt also ein internetseitiger (Edge-)Router aus, kann stattdessen der andere Edge-Router den Traffic entgegennehmen und so dafür sorgen, dass der Standort weiterhin einsatzfähig ist.

Dieses Betriebsmodell hilft MCP-Standorten, die Verlagerung von Traffic zu vermeiden, sofern dies nicht unbedingt erforderlich ist. Dadurch verbessert sich die Verfügbarkeit für die Cloudflare-Kunden nochmals.

4. Traffic-Manager von Cloudflare

Zusammenarbeitende MCP-Rechenzentren bilden das übergeordnete Cloudflare-Netzwerk. Dieses stellt mit Anycast sicher, dass der Kunden-Traffic sein Ziel auch erreicht. Anycast wird durch ein deterministisches Traffic-Management erweitert, damit Kundenanfragen mit der bestmöglichen Performance dort abgefertigt werden, wo wir sie bearbeiten können. Dieses System zur Verwaltung des Datenverkehrs – der Traffic-Manager – prüft das Netzwerk von Cloudflare kontinuierlich und leitet den Datenverkehr automatisch von Rechenzentren weg, bei denen eine CPU-Überlastung verzeichnet wird. Dadurch wird die Überlastung von Rechenzentren mit hohem Trafficaufkommen verhindert, weil der Datenverkehr auf smarte Weise zu einem anderen Rechenzentrum umgeleitet wird, das ihn verarbeiten kann.

Edge-Verfügbarkeit

Edge-Verfügbarkeit bezieht sich auf die Fähigkeit von Cloudflare, Datenverkehr zu verarbeiten, sobald er unser Netzwerk erreicht. Wenn Änderungen an Netzwerktools oder Software unbeabsichtigte Folgen haben, kann das die Verfügbarkeit und die Nutzererfahrung beeinträchtigen. Zur Vermeidung von Zwischenfällen, die von Codeänderungen verursacht werden, hat Cloudflare stark in die folgenden Bereitstellungskontrollen investiert:

1. Bereitstellungstrichter

Beim Bereitstellen von Software beginnt die Sicherstellung der Qualität des Quellcodes mit der Beobachtung und Begrenzung der Möglichkeiten für Entwickler und Kunden, Änderungen am Ökosystem vorzunehmen. Cloudflare begrenzt die Möglichkeiten für Änderungen an unserer Infrastruktur durch jeden beliebigen Kunden oder Nutzer, damit wir jede Anpassung genau überwachen und sicherstellen können, dass sie eine Reihe von Tests besteht, bevor sie im Produktivbetrieb angewandt wird.

2. Änderungsverwaltung nach dem Blast Radius-Prinzip

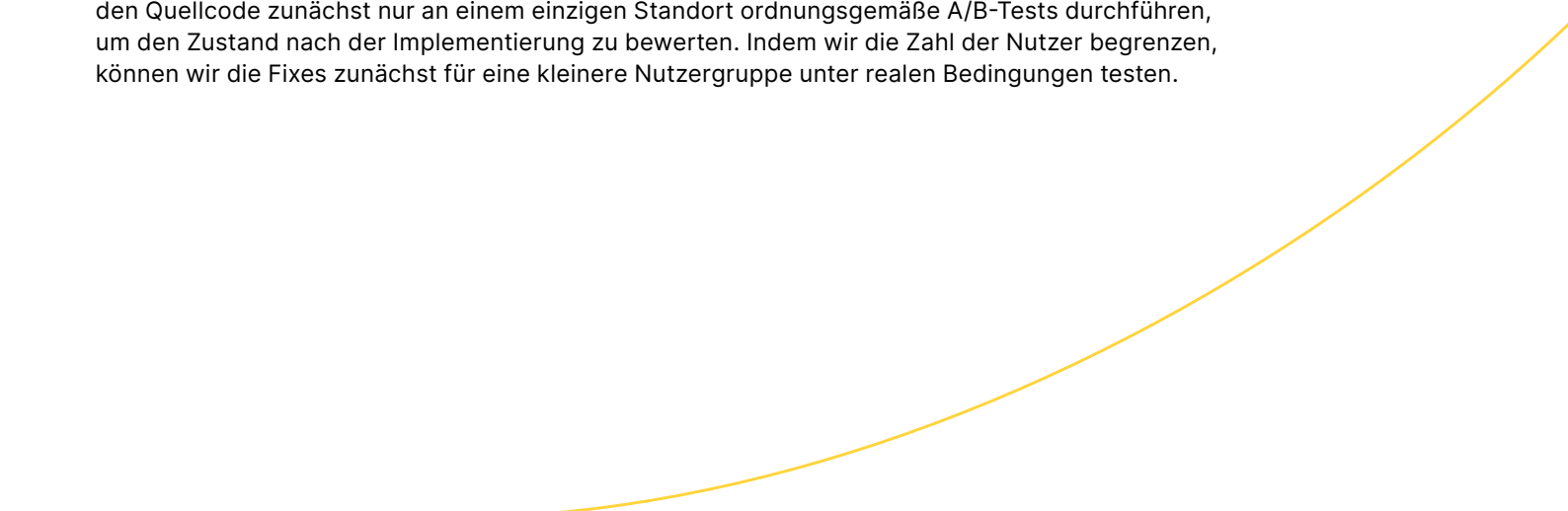
Eine weitere Cloudflare-Maßnahme zur Unterstützung der Edge-Verfügbarkeit während der Bereitstellung besteht darin, vor der breiteren Einführung die Bereitstellung auf Testrechenzentren oder Testgruppen zu beschränken. Man spricht von Änderungsverwaltung nach dem Blast Radius-Prinzip.

Wenn Änderungen am Netzwerk eingeführt werden, können diese innerhalb von Minuten weltweit in Kraft treten. Indem wir Änderungen auf eine Bereitstellungsumgebung beschränken und weitere Anpassungen kaskadenartig ausrollen, können wir beabsichtigte oder unbeabsichtigte Auswirkungen überwachen, bevor größere Regionen oder Nutzergruppen davon betroffen sind.

Es gibt für uns zwei Möglichkeiten, die Auswirkungen von Quellcode-Änderungen zu begrenzen:

- Begrenzung der Zahl der Standorte, bei denen diese Änderungen angewandt werden
- Begrenzung der Zahl der Nutzer, für die diese Änderungen wirksam werden

Durch die Begrenzung der Zahl der Standorte und Server können wir dafür sorgen, dass wir für den Quellcode zunächst nur an einem einzigen Standort ordnungsgemäße A/B-Tests durchführen, um den Zustand nach der Implementierung zu bewerten. Indem wir die Zahl der Nutzer begrenzen, können wir die Fixes zunächst für eine kleinere Nutzergruppe unter realen Bedingungen testen.



3. Zustandsorientierte Bereitstellung

Bei der zustandsorientierten Bereitstellung handelt es sich um ein System, mit dem die Eignung eines Releases auf Grundlage voreingestellter Kennzahlen programmatisch bewertet wird. Diese Metriken signalisieren je nach den potenziellen Auswirkungen entweder ein „Go“ oder ein „No Go“. Mit einer solchen automatischen Prüfungsreihe lässt sich die Veröffentlichung einer schädlichen Version verhindern und eine Version zurücksetzen, wenn unerwünschte Auswirkungen festgestellt werden.

Für alle über Cloudflare bereitgestellten Produkte und Dienste muss ein Service-Level-Ziel (Service Level Objective – SLO) existieren. Dieses umfasst eine Kennzahl, die den Zustand des Produkts abbildet, und eine Schwelle, unterhalb derer ein Produkt als nicht intakt gilt.

In SLO sind akzeptable Ausfallsschwellenwerte festgelegt, sogenannte Burn Rates. Jeder zustandsorientierte Dienst übermittelt im Rahmen des Zusammenführens einer zu implementierenden Änderung SLO an ein automatisiertes System. In jedem festgelegten Umfang (Gratis-Tarifoption, eine Untergruppe von Rechnern in Ashburn usw.) führt das automatisierte System folgende Schritte aus:

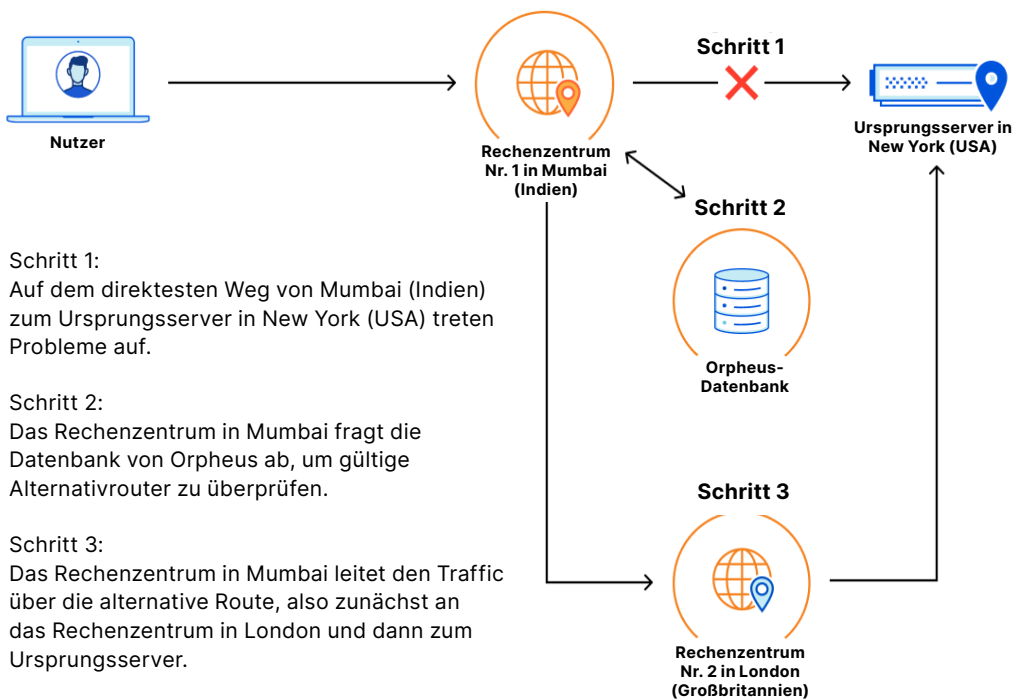
- Zunächst werden die SLO-Vorgaben des Diensts über einen festgelegten Zeitraum überwacht, damit der Schwellenwert nicht unterschritten wird.
- Wenn sich die im SLO angeführten Parameter für den festgelegten Zeitraum innerhalb akzeptabler Bereiche halten, geht das System automatisch zu einer weitreichenderen Implementierung über.
- Werden die Vorgaben des SLO jedoch nicht eingehalten, wird die Bereitstellung automatisch abgebrochen und ein Rollback durchgeführt, wodurch die Auswirkungen automatisch abgemildert werden.

Durch diese Schritte wird sichergestellt, dass Kundensysteme nicht durch die Implementierung von Quellcode beeinträchtigt werden und dieser so kurz wie möglich im Einsatz ist.

Erreichbarkeit des Ursprungsservers

Die Erreichbarkeit des Ursprungsservers bezieht sich auf die Fähigkeit von Cloudflare, Ziele zu erreichen – sei es der Ursprungsserver eines Kunden, eine SaaS-Website oder das öffentliche Internet über Cloudflare Gateway. Dass Anfragen auch das richtige Ziel erreichen, ist für Nutzer, die auf das Netzwerk von Cloudflare zugreifen, entscheidend. Beispielsweise testet [Argo Smart Routing](#), das Cloudflare-Tool zur Optimierung der Performance (also die Time to First Byte), ständig das Netzwerk von Cloudflare, um den schnellsten Weg zum Ursprungsserver zu finden.

[Orpheus](#), das Gegenstück zu Argo, verfolgt einen ähnlichen Ansatz, hat aber eine andere Funktion. Die Lösung dient dazu, zuverlässige Verbindungen zu den Ursprungsservern herzustellen. Konkret untersucht Orpheus Kennzahlen, die Einfluss darauf haben, ob Cloudflare den Ursprungsserver erreichen kann. Es geht hier also nicht um den schnellsten Weg zum Ursprungsserver. Das Tool findet Routen, bei denen der Paketverlust möglichst gering ist und die Performance nicht beeinträchtigt wird. So wird der Traffic automatisch umgeleitet, wenn auf der Route Fehler erkannt werden.



Schritt 1:
Auf dem direktesten Weg von Mumbai (Indien) zum Ursprungsserver in New York (USA) treten Probleme auf.

Schritt 2:
Das Rechenzentrum in Mumbai fragt die Datenbank von Orpheus ab, um gültige Alternativrouter zu überprüfen.

Schritt 3:
Das Rechenzentrum in Mumbai leitet den Traffic über die alternative Route, also zunächst an das Rechenzentrum in London und dann zum Ursprungsserver.

Bevor Cloudflare 2021 Orpheus auf den Markt gebracht hat, ist es uns in 99,9 % der Fälle gelungen, den Traffic zu den Ursprungsservern zu leiten. Nach der Implementierung von Orpheus hat sich dieser Wert auf 99,99 % erhöht. Im kommenden Jahr werden wir Orpheus ausbauen, um noch mehr Arten von Traffic zu schützen, auf mehr Ausfallszenarien zu reagieren und schneller zu arbeiten, damit wir den Zeitraum verkürzen können, in dem Nutzer betroffen sind.

Verpflichtung zu betrieblicher Transparenz

Selbst bei den widerstandsfähigsten und innovativsten Netzwerken kommt es zu Ausfällen. Wenn Kunden von solchen Zwischenfällen betroffen sind, führt Cloudflare gründliche Nachforschungen durch, auf Grundlage derer wir einen internen und einen externen Vorfallbericht erstellen. Falls erforderlich, liefern wir außerdem während der gesamten Zeit, in der Auswirkungen zu spüren sind, [Updates zum Sachstand](#).

Wenn diese Vorfälle entsprechende Auswirkungen oder Innovationen nach sich ziehen, veröffentlichen wir nachträglich im [Cloudflare-Blog](#) eine Analyse dazu.

Fazit

Der Aufbau und die Pflege des Cloudflare-Netzwerks sind mit erheblichem technischen Aufwand verbunden, aber für unsere Kunden ist uns jede Mühe wert. Die Arbeit zahlt sich aus, weil die daraus entstehende Netzwerkplattform unseren Kunden und der gesamten Internet-Community zugute kommt.

Letztendlich betrachten wir Ausfallsicherheit nicht einfach nur als technisches Problem, sondern weisen ihr in unserer Betriebsphilosophie eine zentrale Rolle zu. Deshalb tun wir mehr, als die Abwehr zu stärken: Wir bauen aktiv ein in seinen Grundfesten zukunftsfähiges Unternehmen auf. Da wir vorausschauend kontinuierlich Tests durchführen und Anpassungen vornehmen, können wir uns mühelos an die sich ständig wandelnden Anforderungen unserer Kunden und die dynamische Weiterentwicklung des Internets selbst anpassen.

Uns ist bewusst, dass sich viele der in diesem Dokument vorgestellten Konzepte auf Netzwerktypen konzentrieren, mit denen viele Unternehmen nichts zu tun haben, weil sie für die interne Architektur des Betriebs einer globalen Carrier-Cloud-Umgebung relevant sind.

Wenn Sie an einer ausführlichen Erläuterung interessiert sind und mehr über unsere technische Struktur zur Gewährleistung der Ausfallsicherheit erfahren möchten, [wenden Sie sich einfach an Ihren Ansprechpartner bei Cloudflare](#).

Dieses Dokument dient nur Informationszwecken und ist Eigentum von Cloudflare. Es begründet Ihnen gegenüber keine Verpflichtungen oder Zusicherungen von Cloudflare oder verbundenen Unternehmen. Sie sind dafür verantwortlich, die Informationen in diesem Dokument selbst und unabhängig zu bewerten. Diese können sich ändern. Das Dokument erhebt keinen Anspruch auf Vollständigkeit oder darauf, alle Informationen zu enthalten, die Sie möglicherweise benötigen. Die Pflichten und die Haftung von Cloudflare gegenüber den eigenen Kunden werden durch gesonderte Vereinbarungen geregelt, und dieses Dokument ist weder Teil von Vereinbarungen zwischen Cloudflare und den eigenen Kunden, noch werden solche Vereinbarungen davon berührt. Die Cloudflare-Dienste werden ohne ausdrückliche oder stillschweigende Mängelgewähr, Zusicherungen oder Bedingungen jeglicher Art erbracht.

© 2026 Cloudflare, Inc. Alle Rechte vorbehalten. CLOUDFLARE® und das Cloudflare-Logo sind Marken von Cloudflare. Alle anderen Firmen- und Produktnamen und -logos können Marken der jeweiligen Unternehmen sein, mit denen sie verbunden sind.