aws

Improving application delivery with AWS

Content delivery, network traffic, and edge
security considerations

Table of contents

Ever-changing digital world

In today's ever-changing digital world, your consumers expect web applications to be available 24/7. They also expect those applications to have imperceptible response times, no matter the device they're using or where they're located. To deliver on those expectations, your organization needs a cost-effective way to develop applications with security, performance, and global scale.

Read this eBook to learn how you can build an application that exceeds expectations by moving application delivery onto AWS edge networking services. Whether you are trying to improve latency, protect your applications from attack, or deliver unique, data-intensive customer experiences at scale—discover how you can use customer-vetted AWS services and solutions at an affordable cost.

Continue reading to learn how you can solve your application delivery challenges, with customer examples from Thomson Reuters, InfoSpace, Dow Jones, DishTV, and King.

aws

Application delivery challenges

The early inventors of the internet could never have anticipated the scale of influence their invention would have in nearly every aspect of our lives, nor the critical role it would play in society. By the year 2025, humanity is on track to reach 463 exabytes of new data created every day.[1] And yet, the original design scheme of the internet has not changed with its growth: It remains a best-effort patchwork of thousands of interconnected networks, each with its own vulnerabilities and potential breaking points.

When a user loads a website, data from that website's server has to travel across the internet to reach the user's device. If the user is located far from that server, it will take a long time to load a large file, such as a video or high-resolution image. The path data travels between applications and users is unpredictable and nondeterministic, with multiple network hops that leave data exposed to congestion and outages. Unfortunately, there is no SLA for the internet!

Content delivery networks (CDNs) are one option to help you overcome some of the internet's shortcomings. CDNs help improve the user experience of your applications by reducing latency and increasing reliability and performance. A CDN will cache website content on servers geographically closer to your users, reducing the physical distance it needs to travel to reach the end-user device. This allows for the quick transfer of assets needed for loading application content—including HTML pages, JavaScript files, style sheets, images, and videos.

While this traditional CDN solution is useful, today's applications are more complex—which means reliability, scale, and performance expectations demand more from you than static content caching. If you are looking to **create dynamic user experiences, improve application security, develop advanced traffic routing logic, or deliver real-time workloads**, you will need additional services, features, and infrastructure. Building or finding these solutions can be expensive, potentially requiring additional overhead, additional fees, and the onboarding of more vendors. And managing this complex ecosystem along with compatibility across solutions and vendors can be confusing and time-consuming. After all this, problems can still arise if your CDN is using the public network or has a poor network design, leaving your web applications susceptible to availability issues from physical hardware outages and congestion.

[1] Vuleta, B., "How Much Data Is Created Every Day? [27 Staggering Stats]," Seed Scientific, October 2021.

AWS content delivery and edge networking

The AWS global cloud infrastructure is the most secure, extensive, and reliable cloud platform, offering over 200 fully featured services from data centers globally. As part of this infrastructure, AWS created a cloud-first content delivery network Amazon CloudFront, consisting of 410+ globally distributed edge locations connected to the AWS application infrastructure through redundant 100 GbE metro fiber and transcontinental cables. Combined, AWS provides a single provider to help you build, run, and deliver your applications.

Edge locations allow you to securely onboard all application user traffic onto the AWS-owned network, where it is served by the CDN or encrypted and transmitted to your application servers, improving performance and availability. Whether you are trying to deliver websites, livestream videos, create dynamic customer experiences, route traffic, or protect your application servers—AWS edge locations provide an ecosystem of edge networking and security services that can work together to help you.

You can access the dedicated AWS network by using edge networking services Amazon CloudFront, Amazon Route 53, and AWS Global Accelerator. These services sit at AWS global edge locations to deliver web applications to end users with high availability, security, and performance.

aws

Customizable content delivery networking

When users access your web applications, they expect not only a fast and secure connection, but a customized and engaging user experience as well. To help you meet these expectations, AWS built a cloud-native CDN that offers origin fetch acceleration, added security, and serverless edge computing capabilities in addition to static content caching.

Amazon CloudFront is a highly distributed and scalable CDN that delivers static websites, APIs, video, and dynamic content with consistently low latency. CloudFront reduces latency and improves security by encrypting and moving traffic along the AWS network, delivering application data through a worldwide network of data centers called edge locations. Serverless compute features CloudFront Functions and Lambda@Edge let you insert your own code at edge locations to tailor content and manage server requests based on customer profiles and devices.

Amazon CloudFront use cases:

Improve website content delivery – Reach viewers across the globe in milliseconds with built-in data compression, edge compute capabilities, and field-level encryption. Take advantage of the AWS network backbone and feature-rich edge caching to give users a safe, fast, and reliable experience.

Seamlessly stream live and on-demand videos to any device – Start streams quickly, play them with consistency, and deliver high-quality video with AWS Media Services and AWS Elemental integration.

Customize at the edge – Lambda@Edge and CloudFront Functions are serverless compute features available with CloudFront that provide multiple secure options for customizing your content delivery. Using request information, such as location, account information, or device, you can curate your origin requests.

Faster dynamic content delivery and APIs – Deliver dynamic web content with support for edge termination and WebSockets on the purpose-built AWS global network infrastructure.

aws

> **We are seeing tremendous performance benefits by using both Amazon CloudFront and Lambda@Edge as we are bringing content closer to the user. For our customers, every millisecond counts."**
>
> **Kamal Verma**, Senior Principal Engineer, Dow Jones

CloudFront and Lambda@Edge

Dow Jones, a provider of world-leading data, media, and intelligence solutions, needed to protect its applications from web attacks and have more control over security. It chose AWS to help secure and accelerate its applications

Securing your applications at the edge with AWS

To enable you to build secure, geographically dispersed, fault-tolerant web applications, AWS has implemented a world-class network infrastructure that is carefully monitored and managed. In addition to protecting AWS infrastructure, AWS provides the technology to help you reduce your exposure to security attacks by encrypting data, removing network hops, and controlling application access. You can protect your web applications hosted anywhere globally by deploying Amazon CloudFront on top of your applications.

Prioritize security to help maintain application availability and user trust:

Protection against DDoS attacks – AWS Shield Standard defends traffic transmitted through an AWS edge location from Distributed Denial of Service (DDoS) and malicious web attacks at no additional charge to you. For application protection, you can integrate AWS Web Application Firewall (AWS WAF) using your own rules or leverage Managed Rules for AWS WAF, a preconfigured set of rules managed by AWS or AWS Marketplace sellers.

Apply access control – Amazon CloudFront offers a variety of integrated access control features. Signed URLs and Signed Cookies support token-based authentication to restrict access to authenticated viewers. Geo-restriction prevents users in specific geographic locations from accessing content distributed through CloudFront. With the Origin Access Identity (OAI) feature, access can be restricted to an Amazon Simple Storage Service (Amazon S3) bucket, making it accessible only from CloudFront.

SSL/TLS encryptions and HTTPS – Keep your web applications secured with fresh SSL/TLS certificates to ensure a high standard of authentication, encryption, and integrity. AWS Certificate Manager (ACM) can be used to easily create a custom SSL certificate and deploy it to a CloudFront distribution for free. Secure communication by encrypting content, APIs, and applications over HTTPS.

Compliance – CloudFront infrastructure and processes are all compliant with PCI-DSS Level 1, HIPAA, and ISO 9001, ISO/IEC 27001:2013, 27017:2015, 27018:2019, SOC (1, 2, and 3), FedRAMP Moderate, and more to ensure secure delivery for sensitive data.

Reliable network routing and traffic management

With a globally distributed user population, domain name system (DNS) services need to be instantly available to every user, everywhere. Otherwise, latency and bottlenecks will impact the user experience. And as application architectures become more distributed, it is not enough just to do simple lookups. The DNS needs to make intelligent traffic management decisions and dynamic lookups that ensure users are optimally routed to application delivery endpoints. End-user latency and failed connections can be caused by inefficiencies in network traffic routed between devices and your applications as they initiate and maintain communication.

Amazon Route 53 is a highly available and scalable DNS service. It enables you to customize your DNS routing policies to reduce latency, improve application availability, and adhere to compliance. With Amazon Route 53, you can monitor the health and performance of applications and web servers in real-time and set routing policies to automate responses in case of failure, like redirecting traffic to alternative Availability Zones or Regions. Easily manage how your end users are routed to your application's endpoints—whether in a single AWS Region or distributed across the globe. In addition, you can purchase and manage domain names.

aws

Some of the benefits of using Amazon Route 53:

Highly available and reliable – Configure fail-over policies to reroute your users to an alternate location if your primary application endpoint becomes unavailable using Amazon Route 53 Traffic Flow and routing control.

Cost-effective – Pay for only the resources you use, such as the number of queries the service responses for each of your domains, hosted zones for managing domains through the service, and optional features such as traffic policies and health checks, all at low cost and without minimum usage requirements or any upfront fees.

Flexible – Optimize traffic routes based on multiple criteria, such as endpoint health, geographic location, and latency. Design and configure your traffic policies and decide which policies are active at any given time. The Visual Editor will enable you to create and edit traffic policies easily.

Fast – Automatically route your users to the optimal location depending on network conditions. The result is low query latency for end users and low update latency for your DNS record management needs. Amazon Route 53 Traffic Flow further improves your users' experience by running your application in multiple locations across the globe and using traffic policies to ensure your end users are routed to the closest healthy endpoint for your application.

aws
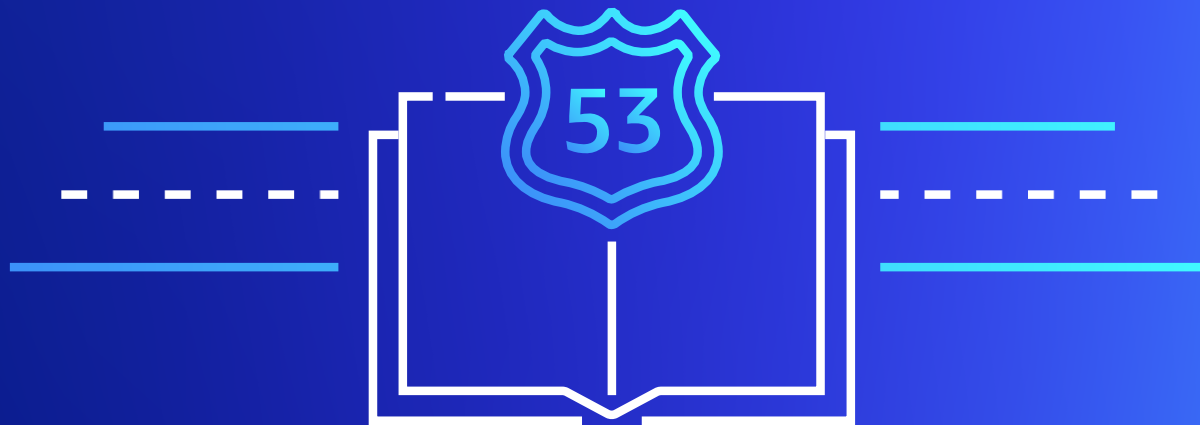
Global traffic acceleration at the network layer

Similar to application content, networking traffic latency is usually driven by the number of networks your user data needs to hop and the bandwidth available along the path. End-user requests can take many networks to reach the application, and the routes to and from the application may differ. Each hop impacts performance and can introduce risks. With applications such as gaming, media, mobile, or financial applications, you need consistent, low latency for great user experiences, but not all of this traffic is cacheable content that can be delivered using a CDN.

AWS Global Accelerator is a service that improves the performance of your users' networking traffic by up to 60 percent using the AWS global network infrastructure. Global Accelerator optimizes the path to your application by routing traffic from the client to the closest AWS edge location and then routing it to the closest regional endpoint over the AWS global network. If you are serving traffic globally from a single region, using AWS Global Accelerator

improves performance, providing a low-cost and easy-to-set-up alternative to replicating applications in multiple AWS Regions for ultra-high availability and performance.

Alternatively, for applications running in multiple AWS Regions, AWS Global Accelerator can help simplify traffic management and improve resilience. As your application architecture grows, so does the complexity, with longer user-facing IP lists and more nuanced traffic routing logic. AWS Global Accelerator solves for this by providing you with two static IPs that are anycast from our globally distributed edge locations, giving you a single entry point to your application, regardless of how many AWS Regions it's deployed in. This allows you to add or remove origins, Availability Zones, or Regions without reducing your application availability.

Some of the benefits of using AWS Global Accelerator:

Accelerate latency-sensitive applications – Internet congestion creates opportunities for network latency that can result in delayed connections and lost data. AWS Global Accelerator combines advanced networking features with the dedicated AWS global network to improve your application network performance for both TCP and UDP workloads. The result is higher throughput, lower packet loss, and reduced jitter.

Improve resiliency and availability – You can run your applications in a single AWS Region across multiple Availability Zones or across multiple AWS Regions. If Global Accelerator detects a failure of your application endpoint, it can automatically trigger traffic re-routing in seconds, sending users to the next available and closest endpoint in another AWS Availability Zone or Region.

Simplify global traffic management – As your application grows, the number of endpoints and IP addresses you need to manage increases. While updating the endpoints, you can run the risk of lowering availability of your application due to firewalls and hardcoded devices. AWS Global Accelerator simplifies global traffic management by providing two static IP addresses that need to be configured by users only once. Behind these IP addresses, you can add or remove AWS origins, enabling endpoint failover, scaling, or testing, without any user-side changes.

Protect your applications – Exposing your application to public internet traffic creates an opportunity for malicious attacks. Use AWS Global Accelerator to decrease the risk of attack by masking your application behind two static entry points. These entry points are protected by default from DDoS attacks with AWS Shield.

Instead of redesigning their applications for multi-region deployment, South Korean gaming company and developer **JoyCity** implemented AWS Global Accelerator. The result? All game traffic improved by over 43 percent, along with one-eighth the timeouts. **Read the story.**

Customer case studies with AWS content delivery and edge networking use cases

Thomson Reuters: Delivering fast, secure news content around the globe

As a global news organization, security, global scalability, latency, speed of delivery, availability, and resiliency are critical factors in Reuters' operation. With Amazon S3 (for cloud storage) and CloudFront providing built-in data compression, edge compute capabilities, and field-level encryption, Reuters can get the news to viewers across the globe in milliseconds.

> **Building new projects on the AWS platform has completely changed how we deliver trusted news content to our customers. Its rich feature set, security, and overall cost of management and delivery have helped us cut costs and latency, not just for ourselves but for our customers too. This leaves us with more time to focus on key business objectives and deliver news faster. Speed for a news agency is critical."**
>
> **Romeo Radanyi,** Solutions Architect at Thomson Reuters

Read the full story

InfoSpace: Accelerating dynamic content delivery and APIs

A major provider of metasearch and monetization solutions, InfoSpace needed to locate its infrastructure close to its partners to improve search response times. To optimize dynamic web content delivery, it migrated to the AWS Global Network infrastructure supporting edge termination and WebSockets. With AWS, search response times for international traffic improved by 20 percent and domestic traffic improved by 10 percent.

> **With careful planning, and by working closely with AWS solutions architects, InfoSpace was able to complete a full data center migration, including its Microsoft Windows stack, within 6 months while supporting an over 30 percent increase in traffic. Using AWS, InfoSpace is able to create a global infrastructure to support its international clients, making our approach to solving problems simpler and quicker."**
>
> **Wayson Vannatta,** Sr. Director of IT & Operations, InfoSpace

Read the full story

aws

DishTV: Streaming live and on-demand video

As the second-largest global provider of satellite TV with some 25 million active subscribers and more than 400 channels in India, DishTV wanted to create "sticky" customer experiences to improve engagement and reduce churn, as well as acquire new subscribers for its over-the-top (OTT) streaming platform. To achieve this, it leveraged AWS Media Services and AWS Elemental integration to enable it to start streams quickly, play them with consistency, and deliver high-quality video to any device. Since going live, its OTT service Watcho has attracted an average of 7,000 new users daily.

> **The OTT platform allows us to reach a larger base of consumers, cater to them in a digital format, explore avenues to expand our revenue stream, and reduce churn while increasing the stickiness of our overall brand. Being on the AWS Cloud helps us create, develop, and deploy software at a much faster pace compared to what we have been doing on premises and achieve larger scale in terms of volumes of customers as well as content distribution."**
>
> **Akash Tyagi,** Head of Product Management, OTT (Watcho), DishTV

Read the full story

King: Distributing patches and updates

King, the leading interactive entertainment company behind popular game franchises such as *Candy Crush, Farm Heroes,* and *Pet Rescue,* chose Amazon CloudFront as its content delivery vehicle. Key factors in choosing AWS included the ability to scale automatically to deliver software, game patches, and IoT over-the-air (OTA) updates with high transfer rates. King's developers and operators use the CloudFront API to manage game content as an integrated part of their application release cycle. This self-service model allows them to get fresh, consistent content out to their users as quickly as possible. CloudFront currently delivers hundreds of terabytes of content for King every day, with spikes to half a petabyte or more when the company launches a new game or large-scale marketing program.

Read the full story

Accelerate innovation and improve user experiences with AWS

AWS edge networking enables you to connect and transmit data to your application users all over the world. By moving traffic onto the purpose-built AWS network through 410+ global edge locations, your business can deliver a secure and reliable low-latency user experience at virtually any scale in new markets and existing ones.

Take control of your application traffic with a portfolio of feature-rich and integrated edge networking services, all available within the same **AWS console** or **CLI**. When you start building, get the support you need with generous **free tiers**, extensive **documentation**, pre-built **CloudFormation templates**, and **developer guides**.

Learn more about AWS Global Infrastructure ›

Learn more about AWS Edge Networking and Content Delivery ›