IDC Spotlight, sponsored by Gcore
June 2024

Accelerating Innovation with CDN-Based Edge Computing

Arnal Dayaratna
Research Vice President,
Software Development, IDC

Ghassan Abdo
Research Vice President,
WW Telecom Virtualization & CDN, IDC

Introduction

Content delivery networks (CDNs) have traditionally been optimized to deliver media and web content to end users within strict quality-of-service (QoS) metrics governing factors such as latency, real-time delivery, fidelity, and scale. The architectural underpinnings of CDNs with massively distributed caching servers, supported by content acceleration algorithms, helped CDN providers to deliver on these QoS expectations even during peak traffic occurrences.

Customer expectations are evolving as enterprises accelerate the journey toward digital transformation. According to a recent IDC survey, enhancing customer experience is cited as one of the primary reasons to embark on digital transformation initiatives. Customers expect personalized content, interactivity as they consume content, and rich media experiences, all delivered within the same QoS metrics. To respond to these evolving demands, CDN providers cannot rely on the current architecture, which is highly tuned for one-way delivery of content.

CDN-based edge computing is emerging as the key enabler of personalized, interactive, and rich media use cases. With physical proximity to the end user, the CDN edge is more adept at dealing with use cases that require low latency and provide scale due to its distributed architecture.

This paper will explore the architectural underpinnings of the CDN edge to deliver on innovative use cases. It will highlight the key benefits of embracing this new architecture and elaborate trends supporting the move to smart edge computing solutions.

AT A GLANCE

Key Stats

IDC forecasts worldwide spending on CDN edge services to grow at a CAGR of 34% for the 2022–2027 period.

What's Important?

The conjunction of CDN-based edge with serverless computing and Wasm enables accelerated time to market and cost savings for distributed applications.

Key Takeaways

CDN edge with serverless computing and WebAssembly runtime enables high-performance, personalized, and interactive applications.

Definitions

Edge Computing

Edge computing is a computing paradigm that brings compute resources and data closer to the devices and digital solutions that consume them by means of a distributed computing architecture. The distributed computing architecture specific to edge computing decentralizes compute resources and data with respect to a centralized hub such as cloud computing infrastructure. This decentralization of compute and data resources means that edge devices can run workloads without requiring a roundtrip journey to the central hub but can instead leverage local instances of compute and data to process compute requests more quickly.

Serverless Computing

Serverless computing refers to a computing paradigm that empowers developers to develop and execute applications without provisioning, configuring, or managing infrastructure. This ability to develop and deploy applications without managing infrastructure enables developers to focus on the design of applications while simultaneously empowering a third-party provider to provision and optimize infrastructure for specific digital solutions. Serverless computing has traditionally been associated with functions-as-a-service such as AWS Lambda; however, in recent years, the more widely accepted definition of serverless involves the foregrounding of fully managed computing for application development and deployment.

WebAssembly

WebAssembly (Wasm) is an open standard that allows programming languages such as C/C++ and Rust to be compiled into a binary format that can be executed efficiently in web-based applications. Wasm enables code to be run at near-native speeds in web browsers, server-side applications, edge computing infrastructures, Internet of Things (IoT) devices, and standalone applications. The platform-independent quality of Wasm means that it can run on any operating system and hardware that supports Wasm runtime, thereby reducing the need for developers to maintain a multitude of codebases for different platforms.

Decentralizing Compute Resources to Increase Speed and Availability

The conjunction of CDN edge computing with serverless computing and WebAssembly runtime enables the delivery of high-performance applications as well as personalization at scale. While a CDN-based edge computing infrastructure brings compute resources and data closer to the devices and digital solutions that consume them, serverless computing provides a fully managed deployment that optimizes application performance. Meanwhile, WebAssembly delivers a performant runtime for code execution that features near-native execution speeds, enabled by efficient compilation and optimization for the application's underlying infrastructure.

The distributed computing architecture specific to CDN-based edge computing decentralizes compute resources and data in contrast to a cloud computing infrastructure that features a central hub. This decentralization of compute and data resources means that edge devices can run workloads without requiring a roundtrip journey to the central hub but can instead leverage local instances of compute and data to process compute requests more quickly. The use of locally stored data reduces network bandwidth and latency and enables the rapid processing of real-time streaming data. This conjunction of local compute and data resources enables edge computing architectures to deliver a multitude of benefits that include increased transactional speed, lower latency, high availability, reduced operational costs, increased resiliency, and improved security.

Meanwhile, serverless computing provides a fully managed deployment environment that optimizes application performance by taking responsibility for the selection of the infrastructure, compute, storage, and networking resources required by each application. Because the selection of resources

is dynamically managed for each application, serverless computing avoids overprovisioning or underprovisioning applications, thereby optimizing application performance and minimizing infrastructure-related costs. Moreover, the use of WebAssembly runtime in a serverless computing environment reduces cold start time in ways that amplify the low latency of the CDN-based edge computing infrastructure. As a result, the synergies between CDN-based edge computing, serverless, and WebAssembly runtime enables applications to deliver rapid performance and personalization at scale in ways that recognize the heterogeneous needs of end users and the form factors by which they access digital solutions.

Unlocking Business and Technical Benefits with CDN-Based Edge Computing

Business Benefits

- **Improved Personalization:** The ability to expediently track and leverage details of a user's device type, preferences, geolocation, and visit history enables the enhancement of application personalization. This ability to deliver personalization at scale is enabled by a platform that can expediently leverage CDN-based edge computing to help optimize applications that reflect the specificity of end users and their engagement with digital solutions.
- **Improved Application Responsiveness:** As a low-level assembly-like language, Wasm runs with near-native performance that delivers faster code execution times as compared to interpreted languages such as JavaScript. The faster code execution speed of Wasm is complemented by the bandwidth efficiency of edge computing infrastructures, which reduces the amount of data that needs to be transmitted to centralized server infrastructures. The reduced bandwidth congestion specific to edge computing leads to improved application performance, enhanced responsiveness, and a better user experience.
- **Improved Application Performance:** Edge computing improves application performance by processing data closer to the devices and endpoints that consume them. The ability of edge computing to minimize lag times associated with data processing and the execution of computational instructions enhances user experience while also enabling faster utilization of data-driven insights.
- **Lowering of Capital and Operational Expenditure:** One of the benefits of serverless computing is its reduction of capital expenditure for application development and deployment. Because customers use a utility pricing model to pay for resources that they consume, they are absolved of the need to invest in capital assets such as servers and other IT infrastructure hardware. Moreover, the ability of a CDN-based edge computing infrastructure to leverage a decentralized network infrastructure minimizes round trips to a centralized server in ways that reduce costs related to application delivery and hosting.

Benefits to Developers and Engineers

- **Simplified Operational Overhead:** Serverless computing reduces the operational complexity associated with managing the hosting, deployment, and monitoring of digital solutions, while simultaneously simplifying life-cycle management. By transitioning the operational responsibility of hosting and managing applications to a third-party provider, organizations can focus their developer responsibilities on the design and development of digital solutions. This ability to outsource the operational management of digital solutions empowers independent software vendors (ISVs) and enterprises to focus their resources and budgets on the development of digital solutions for use cases that reflect their core competencies and business specialization.
- **Improved Ability to Manage Application Security:** The decentralization of data specific to edge computing renders it challenging for hackers or bad actors to obtain comprehensive access to data that is of interest to them. Meanwhile, the stateless quality of serverless computing diminishes the risk of server-based attacks because the lack of a persistent server environment reduces potential attack vectors. Moreover, Wasm fortifies security by executing code in sandboxed environments that prevent malicious code from affecting the system it runs on, thereby safeguarding both the application and its data from exploits that could compromise the host.
- **Enhanced Application Portability:** The platform-independent quality of Wasm empowers developers to leverage it to run applications not only in web browsers, but also in server-side applications, edge computing infrastructures, IoT devices, or standalone applications. This ability to run Wasm applications across a wide range of form factors and operating systems means that developers can use it to run code on any environment that supports Wasm, thereby reducing the need to maintain a multitude of codebases across different platforms. Importantly, Wasm is well suited for edge computing infrastructures because of its lightweight quality and efficient execution, which render it well suited for devices that are resource-constrained with respect to computational power and memory.

The Quest to Improve Customer Experience Drives the Growth of CDN-Based Edge Computing

Edge computing uptake by enterprises is increasing due to several business and technology factors. IDC expects worldwide spending on CDN-based edge computing services to grow at a compound annual growth rate (CAGR) of 34% over the 2022–2027 period, with spending in the Europe, Middle East, and Africa (EMEA) region forecast to increase at a CAGR of 61% across the same timeframe. The global CDN-based edge computing services market is expected to reach a value of $3.8 billion in 2027. Several macro factors are driving the increased demand for edge computing services, including:

- **The desire to modernize legacy apps that demand low latency.** The goal here is to improve applications efficiency, increase agility, and realize cost savings.

- **The quest for use cases that enhance customer experience.** Application performance is key to business success for digital enterprises. Deploying these use cases at the edge will improve latency and provide scale with distribution.
- **The need to comply with data sovereignty and localization requirements.** Edge infrastructure is suited to helping enterprises adhere to regulations and data protection laws, ensure privacy, reduce piracy, and achieve risk avoidance.
- **The need to guard against widespread security attacks such as DDoS or malware.** Widespread security attacks are harder to mitigate in a centralized cloud infrastructure. The edge can be the first line of defense against these attacks and prevents further penetration within the enterprise perimeter.

Countering these growth factors are concerns around edge security, performance in contrast to cloud computing, budgetary limitations, and IT skill factors. CDN providers can help mitigate these concerns.

The findings of a recent IDC survey indicated that media applications are among the top use cases for edge computing. Media applications align well with ongoing investments by CDN providers to leverage their points of presence (POPs) as edge computing nodes. In addition to geographic expansion, CDN providers enhance the computing and storage capabilities of these POPs.

The media and entertainment vertical can benefit from increased adoption of edge computing. Some of the emerging use cases include:

- Providing personalized media and web applications. These may include targeted advertising and web content that aligns with user personal needs.
- Delivering better access control, including fast authentication and waiting room functionality.
- Optimizing A/B testing with simplified control and a smooth user experience.
- Enhancing the customer media experience with interactivity. An example includes real-time betting while watching a sports event.
- Adopting new use cases such as AI for inference and improving AR/VR with fast authentication.

While most observers of edge computing envision a multitude of edge use cases, the success of these applications will depend on a sound business case. These use cases must provide measurable and material ROI and/or provide a significant improvement in customer experience.

Gcore Vendor Profile

Gcore, previously known as G-core Labs, is a company that specializes in AI, public cloud services, and content delivery networks. Gcore was founded in 2014 and is headquartered in Luxembourg. Gcore operates a global network with over 180 POPs across six continents, offering a variety of services such as edge computing, hosting, security solutions, AI infrastructure, and DDoS mitigation.

The company intensively invests in hardware and network growth, making it a future-proof asset for tomorrow's apps and emerging workloads. Besides the notable network scale, Gcore demonstrates competency in establishing long-term peering partnerships to enhance infrastructure connectivity. Today, the company has more than 14,000 connections with ISPs, top-tier IXPs, and other networks across the globe.

Gcore has developed several innovative products, including FastEdge, a serverless edge computing solution designed for low latency and high performance. This service leverages Gcore's extensive CDN to distribute custom code across its global edge nodes, supporting workloads like server-side rendering and advanced image editing without the need for traditional server maintenance.

FastEdge is Gcore's serverless edge computing solution designed to leverage its extensive CDN infrastructure. Here are some key features and aspects of FastEdge:

- **Global Distribution:** FastEdge comprises over 180 edge locations or nodes worldwide. This ensures low latency at an average of 30 milliseconds, and high performance for applications deployed on FastEdge.
- **Serverless Architecture:** It enables developers to deploy applications quickly and easily, without worrying about the underlying infrastructure.
- **WebAssembly Runtime:** FastEdge uses the WebAssembly runtime environment, delivering faster application start times compared to container-based solutions.
- **Cloud-Native Development:** The solution is designed to support cloud-native development, catering to developers who need to build and deploy applications globally.
- **General Availability:** As of early 2024, FastEdge is generally available for JavaScript and Rust developers. Gcore plans to expand support to other programming languages in the future.
- **Use Cases:** FastEdge is suitable for a variety of workloads, including server-side rendering, advanced dynamic image editing, personalization, A/B testing, and quick authentication. This provides developers with a valuable tool to deploy applications that require high performance and low latency.

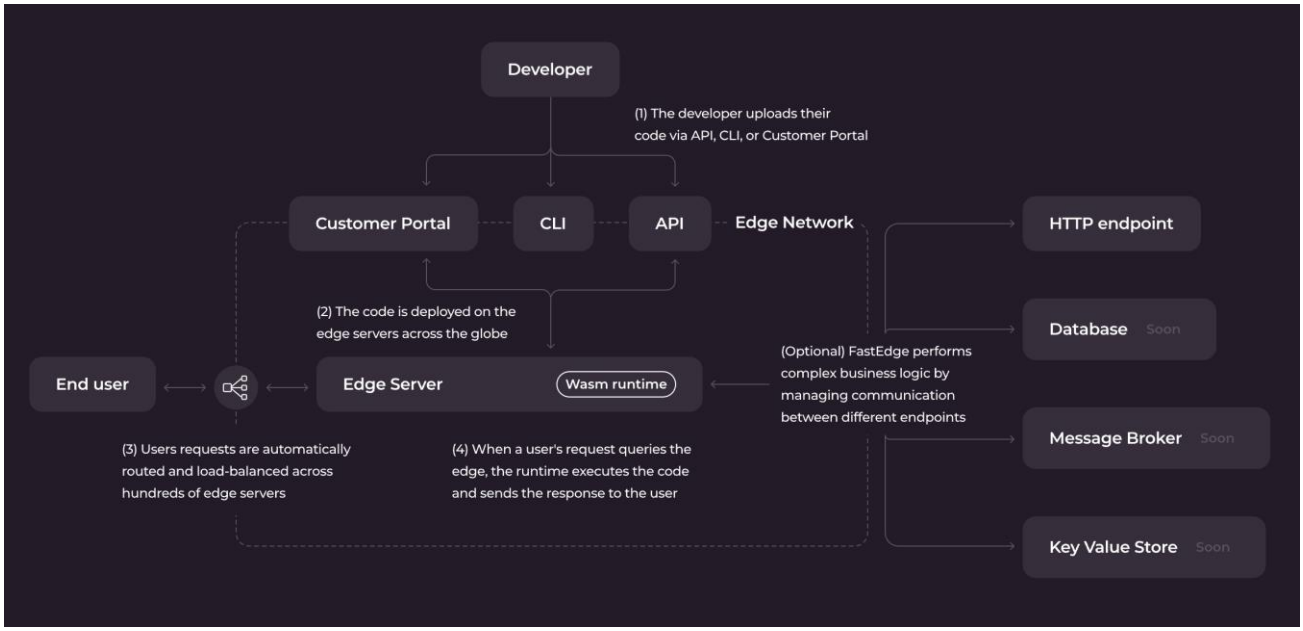
FastEdge provides the following business benefits:

- **Low-Latency User Experience:** With an average start-up time of hundreds of microseconds for the execution environment, FastEdge’s low latency ensures faster content delivery and application responsiveness. This improved performance can lead to better user satisfaction and engagement.
- **Cost Efficiency:** The serverless architecture reduces operational costs and allows businesses to allocate resources more efficiently. The pay-as-you-go model can result in cost savings, especially for businesses with variable demands.
- **Scalability:** With FastEdge, businesses can easily scale their applications up or down based on traffic without worrying about infrastructure limitations. This scalability is beneficial during peak traffic events.

- **Reduced Time to Market:** FastEdge provides the ability for developers to deploy applications quickly without the complexities of traditional infrastructure management.

FIGURE 1

Gcore FastEdge Developer and User Flows



Source: IDC, 2024

Challenges

Gcore faces several challenges as it pursues its goal of becoming a leading provider of smart edge services. These include:

- **Global Competition:** Several of Gcore's global competitors such as Akamai, Fastly, Tencent, and Cloudflare are investing in edge computing. They are appealing to the developer ecosystem to accelerate the development of innovative use cases.
- **Scale and Footprint:** With a more limited footprint, Gcore's main appeal will be in markets where FastEdge is deployed. Enterprises that operate across a broader geography may pivot to global CDN providers with a wider footprint.
- **Adoption of WebAssembly:** Adoption of WebAssembly still has the potential to grow significantly. Developers continue to deepen their knowledge of WebAssembly as well as acquire new skills related to its adoption. Notably, developer-related resources, tutorials, and community support for WebAssembly remain embryonic in comparison to well-established frameworks such as JavaScript.

Conclusion

CDN-based edge computing is emerging as the key enabler of personalized, interactive, and rich media use cases. The conjunction of CDN-based edge computing with serverless computing and WebAssembly runtime enables the delivery of high-performance applications as well as personalization at scale. While a CDN-based edge computing infrastructure brings compute resources and data closer to the devices and digital solutions that consume them, serverless computing provides a fully managed deployment that optimizes application performance. Meanwhile, WebAssembly delivers a performant runtime for code execution that features near-native execution speeds that are enabled by efficient compilation and optimization for the application's underlying infrastructure.

Message from the Sponsor

Gcore – An Edge Platform for Modern Applications

Gcore Edge Network brings low-latency services and distributed workloads to online businesses around the world. Running over a global edge infrastructure, Gcore offers more than 180 points of presence that serve as a backbone for the company’s edge-native products: CDN, DNS, WAAP, Video Streaming, and FastEdge. Locating these services at the network edge reduces roundtrip times for user requests, ensuring a smooth and responsive digital experience. This allows Gcore's customers to build and deploy modern applications that rely heavily on dynamic content processing and high personalization, while end users enjoy fast responses. For more information, visit gcore.com/edge-network.

About the IDC Analysts

Arnal Dayaratna

Research Vice President, Software Development, IDC

Dr. Arnal Dayaratna is research vice president for software development at IDC. Arnal focuses on software developer demographics, trends in programming languages and other application development tools, and the intersection of these development environments and the many emerging technologies that are enabling and driving digital transformation. Arnal's research examines how the changing nature of software development relates to broader trends in the technology landscape.

More about Arnal Dayaratna

Ghassan Abdo

Research Vice President, WW Telecom, Virtualization & CDN, IDC

Ghassan Abdo serves as research vice president within IDC's Telecommunications group, where he covers the evolution of the telco cloud ecosystem as well as emerging virtualized enterprise networking services. His primary focus areas include service provider SD-WAN and managed services, emerging NFV-based virtual networking services, and other managed WAN services. In the hosting and cloud segment, Ghassan covers service provider managed hosting services, including hybrid managed private/public cloud services, colocation services, secure cloud connect, and CDN services.

More about Ghassan Abdo

IDC Custom Solutions

IDC Research, Inc.
140 Kendrick Street, Building B, Needham, MA 02494, USA
T +1 508 872 8200

Twitter @IDC | LinkedIn @IDC | www.idc.com