



WHITE PAPER

Scaling Content Delivery for a Digital World:
Innovations in CDN Architecture

Introduction

Since their inception in the late 1990s, Content Delivery Networks (CDNs) have emerged as a fundamental component of the digital world, serving an essential function in efficiently distributing content across the internet. Originally designed to mitigate bandwidth challenges and lessen the burden on network infrastructure that accompanied the content surge of the early 2000s, CDNs maintain their critical importance in contemporary network applications nearly three decades later.

“A content delivery network (CDN) is a group of geographically distributed servers that speed up the delivery of web content by bringing it closer to where users are.”

Akamai®¹

First commercialised by Akamai® in 1998, CDNs have evolved significantly beyond their initial purpose as basic caching mechanisms. Incorporating advanced technologies such as the Internet of Things (IoT), edge computing, and Artificial Intelligence (AI) has dramatically enhanced their functionality, bringing data and content closer to users through rapidly advancing technologies and devices. Reductions in latency facilitate real-time capabilities such as live streaming, online gaming, and digital broadcasting, ushering in a new age and changes to how we consume and distribute content.²

Today, streaming video content constitutes 53.7% of overall internet traffic, an increase of nearly 5% over the previous year.¹⁴ However, it is essential to recognise the parallel between advancements in CDN applications and the adoption of new technologies. Innovations such as 5G, wearable technology, Extended Reality (XR), and Machine Learning (ML) have all flourished as CDNs have grown, today serving as tools to monopolise the opportunities they present.

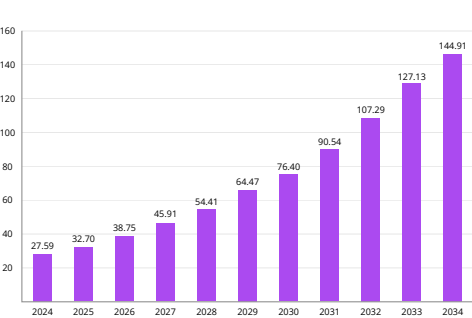


Figure 1: CDN Market Size 2024-2034 (USD Billions)²

Shifts in consumption habits, particularly those spurred by the COVID-19 pandemic, have significantly influenced the CDN market. In 2024, it was reported that the market would account for \$27.59 billion and exceed \$140 billion by 2034, reflecting a Compound Annual Growth Rate (CAGR) of 18.04% over the next decade, as shown in Figure 1.¹⁴ This remarkable growth can be attributed to factors such as the rapid expansion of internet access, the rising prevalence of e-commerce, the increasing demand for Over-The-Top (OTT) content delivery through mobile networks, and the escalating need for cloud-based services. As a result, CDNs have evolved into a transformative technology, prompting major providers to enter the industry themselves.¹⁴

Drawing upon our extensive experience in the media and broadcast industry, we'll examine the dynamic relationship between user habits and the advancement of CDNs, how industry trends are fuelling innovation, and how our recent collaborations in the field have fuelled go-to-market strategies.

Content Delivery Network Architecture

At the core of CDN technology lies the capability to effectively manage and distribute traffic across a network, thereby minimising the distance between the end user and the requested data. Traditionally, such content was hosted on an origin server, making it vulnerable to the routine pressures of network traffic. However, a CDN addresses this issue by storing cached versions of the server's content at geographically dispersed Points of Presence (PoPs).

The efficiency of CDNs can be measured using Round-Trip Time (RTT), which gauges how quickly a request can traverse the network, receive a server response, and deliver the data back to the user - typically within milliseconds. CDNs optimise this process by strategically caching content based on usage patterns and analytical insights. Still, factors such as PoP distribution, data centre capacity, and overall network load significantly influence the measurement of a CDN's success.⁵

CDN architecture is adept at managing substantial traffic volumes through caching, load balancing, and various routing components. This capacity aids in reducing latency, minimising bandwidth, and scaling infrastructure to adapt to varying volumes, ultimately enhancing the end-user experience.

While network topologies may differ among providers, scope, and application, most CDNs encompass the following (see Figure 2):

- **Origin Server**

The central repository for the original version of the content serves as the primary source from which CDN edge servers cache data for distribution.

- **Point of Presence (PoP)**

Generally housed within data centres, a PoP contains a localised network of edge servers strategically positioned geographically to improve RTT and delivery speeds, thus decreasing the distance between the end user and the stored data.

- **Edge Servers**

These servers, located at the CDN's periphery and closest to end users, can be found within PoPs, storing cached content or retrieving it from the origin server as necessary.

- **Cache Servers**

Cache servers typically possess multiple drives and substantial RAM and are responsible for storing frequently requested content to alleviate bandwidth bottlenecks and accelerate response times.

- **Load Balancer**

Responsible for distributing traffic across servers; load balancers direct requests to the most suitable server based on current network conditions, including load, response time and latency, ensuring no particular server becomes overloaded.

The precise configuration of CDN architecture plays a crucial role in determining market value. Mid-capacity PoPs can be particularly effective in densely populated areas, providing reliable bandwidth without overwhelming the existing network infrastructure. Whereas transitioning towards larger data centres and consolidating architecture into more comprehensive PoPs can enhance the agility of network management.

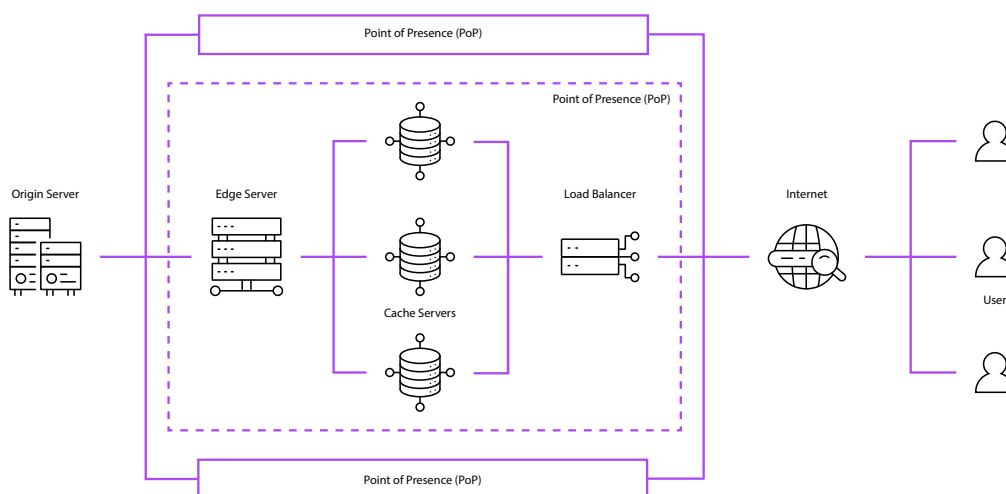


Figure 2: A typical CDN architecture pipeline

CDN design is founded on four principal tenets: performance and speed, scalability, trust and reliability, and responsiveness.⁵ Based on our experience in the media and broadcast industry, we add a fifth - innovation:

- **Performance and Speed**

A CDN's primary objective is to minimise latency between user requests and the resultant data or content (see Figure 3). While we, as hardware partners, can tailor the architecture to specific requirements, external factors such as data centre locations, connection quality, carrier agreements, and the underlying network infrastructure must also be considered.

- **Scalability**

Scalability is vital in accommodating evolving user demands and the subsequent increases in traffic. Our experience with new product introductions has underscored the importance of preparing for volume early in the development process, equipping the CDN with the necessary resources to operate efficiently at scale and optimise operations where needed.

- **Trust and Reliability**

Nearly half of the world's 1,000,000 most popular websites use CDNs for content distribution, making the potential for outages a critical concern.¹¹ This issue is amplified in broadcast production, where requirements for ultra-low latency, minimal buffering, and high-definition outputs are not merely desirable but are regarded as non-negotiable standards.

- **Responsiveness**

The other side of the scalability discussion encompasses the speed at which configurations can be managed and changes implemented throughout the network. Given the global distribution of architecture, even the most minor adjustments must be communicated at a network level, making scale a significant factor in a network's responsiveness.

- **Innovation**

A CDN is a formidable and versatile tool capable of transforming a fleeting idea into a global phenomenon. From real-time content personalisation driven by edge computing to XR experiences on mobile devices, immersive solutions can be delivered to audiences worldwide, manifesting unique market propositions realised through CDNs.

The future of content delivery isn't necessarily how fast or how far it can be sent but the ability to adapt to use cases that can change as quickly as the users who drive them. As consumer expectations become more sophisticated and the data supporting them becomes more intricate, CDNs must embrace innovative architectures and enhance their edge computing capabilities.

Future networks will redefine how users interact with that content - personalised, interactive, and delivered in real-time, regardless of location - making the question no longer *"What can a CDN do?"* but *"What can be done with it?"*

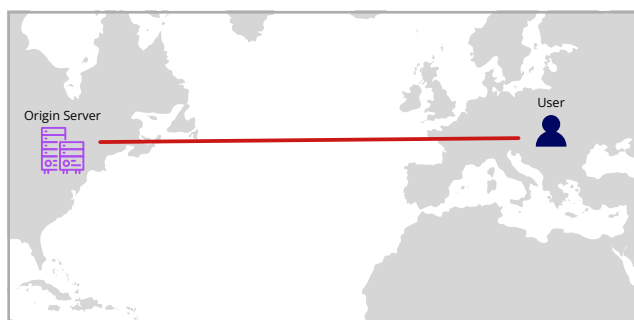


Figure 3: (1) A traditional connection path can put long distances between the origin server and user, causing slow load times, increased latency, and a higher dependency on a single source.

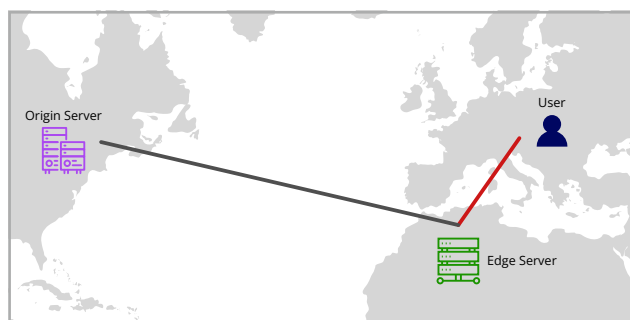


Figure 3: (2) With a CDN, connection paths to locally cached edge servers can be much shorter, therefore reducing the distance data must travel, improving load times and decreasing latency.

History and Trends

Changes in user behaviour regularly parallel the pace of technological evolution, a dynamic that has persisted for generations and well before the creation of CDNs. As the internet age spurred growth across industries, the rise of CDN architecture would forever catalyse usage habits.³

1G - A More Accessible World Wide Web

In the mid-1990s, while at the Massachusetts Institute of Technology, Tim Berners-Lee, the inventor of the World Wide Web, challenged his peers to devise a more effective method of delivering internet content, transitioning it from the static aggregation of web pages characteristic of the preceding decade.² Commercial CDNs were born from here, using existing technologies like server farms to create a faster and more accessible internet.

2G - The Rise of Broadband

The transition from dial-up connectivity to broadband substantially increased the number of internet users, elevating the demand for interactive content. This era witnessed the early proliferation of mobile devices and social media, fundamentally altering distribution practices and dynamic content delivery methods well into the 2000s.

3G - A New Way to Shop

The following decades experienced exponential growth, with the online community expanding to four billion

users by the latter part of the 2010s.⁸ The maturation of CDN security allowed online transactions to become commonplace, thus fostering the rise of e-commerce. The simultaneous increase in cyber threats underscored the need for enhanced security measures, driving advancements in encryption technology.

4G - The Modern CDN

CDNs have played a crucial role in advancing IoT and, subsequently, the proliferation of AI. Initially, a CDN served merely as a rapid and adaptable cloud storage solution; however, the integration of edge computing at the periphery has unlocked real-time processing capabilities that have spurred innovation across industries. Incorporating AI and ML within CDN architecture has led to significant strides in content personalisation, adaptive streaming technologies, and improved data governance. Today, CDNs reach around the world through strategically positioned resources that can put content and data within easy reach of users (see Figure 4).

As the digital landscape continues to evolve, so must the infrastructure that powers it. CDNs are no longer just caching static content; they have become the backbone of seamless and high-demand digital experiences.

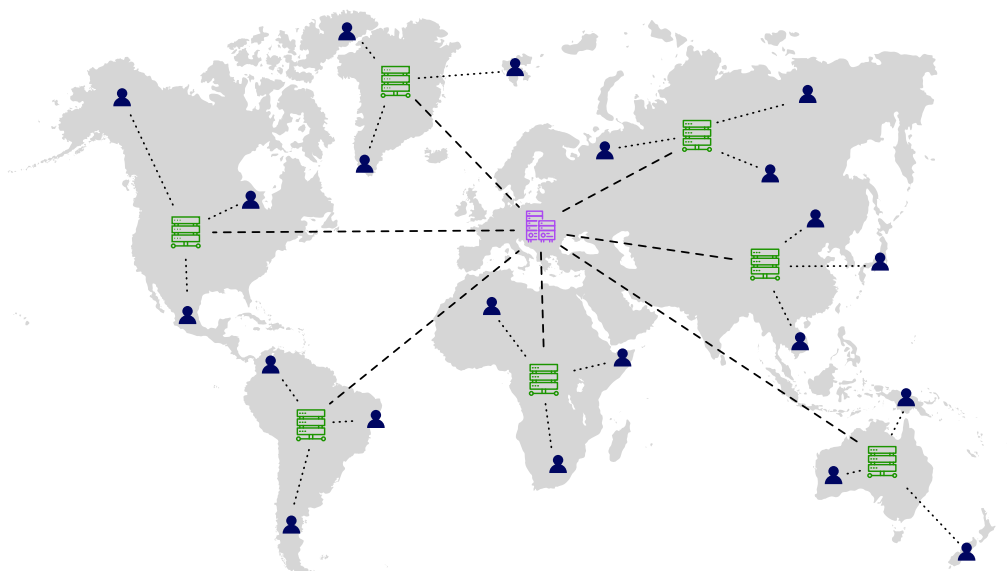


Figure 4: CDN architecture at a global level

Industry Landscape

Exponential growth in the CDN industry has led to significant market acquisitions, particularly within the media delivery segment.¹⁴ This growth, propelled by increasing demands for seamless, low-latency, and interactive content delivery, has resulted in the evolution of CDNs beyond the traditional storage architecture, thereby solidifying their critical role in contemporary technologies.

Drawing upon industry forecasts, insights from our partner network, and our own experience in the media and broadcast industry, we examine the emerging trends shaping the future of CDNs in greater depth.

Edge Processing

Localised data processing combined with the global connectivity offered by CDNs - whether through existing infrastructure or third-party hosting - is improving end-user speed and performance, reducing latency, and lowering bandwidth consumption from both ends of the network, making edge processing valuable for applications like online gaming where connection lag can hinder end-user experience.

We're already seeing this shift towards edge-based solutions on a global scale, with Redge Technologies and MedOne partnering in Israel, Comcast® collaborating with Qwilt® on a new edge distribution in the US, and BT® testing their EE® TV platform on Edgio™ in the UK - indicating a promising market for edge-based CDNs.^{6,9,13}

AI-Optimised Content Distribution

User demands are evolving at unprecedented speeds, making the time- and resource-heavy approach to hardware-driven scalability not necessarily the best solution. The performance enhancements and real-time insights derived from AI and ML could make software the more viable answer. Integrations are becoming more sophisticated, yielding advancements in predictive caching, intelligent load balancing, traffic route optimisation, and real-time bot detection.

The speed of adaptation is further accelerated by the emergence of edge AI, which enhances CDN response times and the agility with which AI language models can be trained. Projections indicate that at least 60% of edge computing deployments will incorporate composite AI by 2029, up from 5% in 2023, underscoring the growing momentum of AI-driven innovation.⁷

Mobile Content Delivery with 5G

5G is up to 100 times faster than 4G, making the difference in milliseconds a game-changer.¹⁰ Improvements in 5G modem technology yield speeds comparable to fibre

networks, unlocking opportunities for mobile applications that are no longer constrained by wires or static connections and do not rely on congested networks.¹² High-capacity use cases such as autonomous vehicles and remote healthcare diagnoses are just a few that put one-millisecond latencies to use. Moreover, uncoupling network speeds from physical infrastructure drives innovation in previously inaccessible industries like precision agriculture and maritime management.

Optimising CDNs for interactive content has led to increased network slicing, which creates dedicated topologies that allocate resources to specific applications or use cases. This shift from a one-size-fits-all approach enhances CDN architecture efficiency, allowing for tailored application performance based on particular speed, bandwidth, and volume requirements, thereby improving cost-effectiveness.

Personalised Streaming Experiences

The combination of AI and 5G technology facilitates the development of ultra-fast, intelligent, and personalised content experiences that prioritise audience resonance over content delivery speed. From real-time adjustments to web pages based on user preferences to hyper-personalised entertainment experiences driven by ML insights, content personalisation is projected to enhance user engagement, foster trust, and boost revenue.

Insights from the realm of sports indicate a substantial shift towards personalised experiences, with 37% of Generation Z viewers expressing a desire for real-time game analytics and 34% seeking preferred camera angles.⁴ Our expertise in this domain highlights the growing trend towards experiences beyond conventional broadcasting, equipping software innovators with the necessary tools to engage a technologically progressive audience.

Latency and scalability continue to pose significant challenges, much like in previous generations, albeit within the context of rapidly evolving user behaviours. Whereas a five-second delay in traditional broadcasting was more accepted in the 2000s, today, a lag of milliseconds can be the difference between a positive viewing experience and a negative one. Similarly, before the rise of digital streaming, terrestrial networks had limited bandwidth to accommodate broadcast viewers, restricting access to content. Today, viewer figures can soar into the tens of millions, imposing considerable strain on service providers.

HIPER Global's 1U Video-on-Demand Server

With the CDN industry progressing towards increasingly personalised and immersive viewing experiences, HIPER Global is committed to continued innovation as a reliable CDN technology partner. Our successful implementations in virtual production, real-time graphics rendering, and video wall technology empower our clients to maintain a competitive edge, fulfilling current demands while seamlessly scaling for future opportunities.

As one of the UK's leading broadcast and video-on-demand providers responsible for designing, implementing, and maintaining on-premise architecture across three continents, our client was tasked with reducing cost and footprint without compromising thermal efficiency. With ambitious targets for future video-on-demand products, it became clear that long-term performance and platform stability would be the key to success in the short term as a driver for new go-to-market strategies and long into the future as a tool for customer growth.

With new geographies and markets on the horizon, an energetic focus on future innovation, and a clear vision for ecosystem growth, our collaboration resulted in the bespoke design of a new long-life, storage-dense, and high-performance 1U Video-on-Demand server that could scale with "Copy Exact" production.



Figure 5: HIPER Global's 1U Video-on-Demand Server

Equipped with an AMD EPYC™ CPU, fast DDR5 5600MHz memory, and up to 248TB of high-speed NVMe® storage manageable via a hot-swappable drive system, this 1U solution reduces footprint without compromising output thanks to an airflow-optimised chassis. Each unit supports:

- Single socket with a 5th gen. AMD EPYC™ processor
- 768GB DDR5 5600MHz ECC memory
- NVIDIA® Mellanox graphics capabilities
- 248TB of high-speed NVMe® storage on read-intensive data centre drives designed for cloud and virtualisation workloads
- High-volume cache supporting up to 256MB per CPU
- Dual 860W redundant power supplies with 80+ titanium level certification

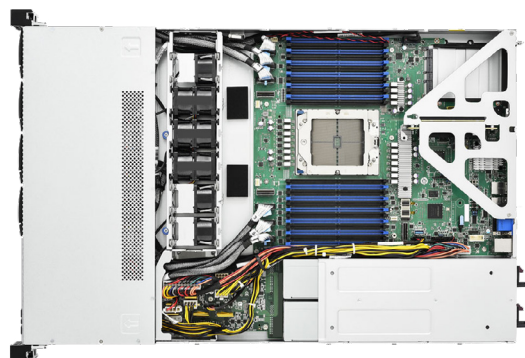


Figure 6: Internal view of the chassis used for HIPER Global's 1U Video-on-Demand Server

In the context of go-to-market projects, the end product is the final piece of a bigger puzzle. As HIPER Global Engineers meticulously refined the technical specifications, benchmarked performance, and conducted field tests, our Operations and Marketing teams simultaneously addressed the elements necessary to transform server-based technologies into a market-ready product:

- UL, CE, and UKCA certification testing of pre-production models to ensure compliance with safety, health, and environmental protection standards.
- Comprehensive regional shipping and customs forecasting, guaranteeing that all touchpoints remained compliant throughout product lifecycles.
- Development of custom packaging components that align cohesively with the broader ecosystem and the client's established brand identity.

By facilitating both Media Asset Management (MAM) and Production Asset Management (PAM) workflows for efficient content distribution across diverse topologies, we helped our client take a decisive step into the market with a long-life foundation that would complement their IP, augment their video-on-demand portfolio, and fuel their content technology goals. This initiative enables product managers, broadcast engineers, and commercial leadership to make more informed decisions, drive strategic enhancements, and improve operational excellence.

Focus Areas

This 1U Video-on-Demand server joins our client's family of content technologies and plays a pivotal role in their global ecosystem. A new streamlined, thermally efficient, high-performance footprint helps production teams to manage and organise content more efficiently, lowering data centre costs thanks to a 50% reduction in server unit allocation.

Built for Growth

As the core of our client's PoP infrastructure and the powerhouse of their video-on-demand portfolio, this scalable 1U solution is incredibly well-equipped thanks to high storage densities, lightning-fast DDR memory, and a long-life AMD EPYC™ CPU. Our strict focus on airflow optimisation will ensure that thermal efficiency remains stable, providing a strong foundation for scalability at higher volumes.

An End-to-End Solution

Our capacity to streamline the often resource-intensive stages of product development has led to substantial savings and improvements in efficiency early in production. This success illustrates the advantages of nurturing long-term partnerships based on collaboration and trust. Consistency in technical standards, established communication channels, and simplified vendor management have cultivated a streamlined procurement process. This allows our clients to reallocate resources from in-house development toward future innovations and strategic initiatives.

Global Deployment

With PoPs in multiple countries, our client's established relationships with content and broadcast networks are key to their worldwide services. The effectiveness of our global supply chain was particularly crucial during the final stages of deployment, as the complexities of international logistics required specialist expertise capable of transcending borders and overcoming logistical challenges.

Reduced Support Overheads

Transitioning from older and inefficient form factors has significantly improved temperature regulation, boosting reliability and diminishing support requests. Furthermore, hot-swappable drive systems have reduced post-deployment overheads, enabling our clients' in-house support teams to change drives and expand storage volumes as/when required. Considering the long-term nature of this deployment, HIPER Global's in-house resources remain aligned with internal operations for targeted and system-specific support based on a unit's location, configuration, and network load.

HIPER Global's 1U Video-on-Demand Server is more than a technology upgrade - it is a strategic enabler for our client's growing content-on-demand infrastructure. Our expertise in system optimisation and product development helps our client focus on their core mission: to better connect the world with content, innovation, and technology as one of the world's leading media and entertainment broadcasters.

Conclusion

Content Delivery Networks are no longer just facilitators of faster content delivery; they are foundational to the future of digital experiences, transforming how content is consumed. From edge AI computing to 5G-enabled content distribution, CDNs have shaped how businesses, broadcasters, and consumers interact with digital media.

Our 1U Video-on-Demand server is a prime example of HIPER Global's commitment to innovation, addressing short-term requirements and long-term goals with scalable, storage-dense, and high-performance technologies that will continue to deliver for years to come. By integrating cutting-edge technology with a deep understanding of our clients' on-demand workflows, we help our clients to stay ahead of the curve, ensuring seamless, low-latency, and high-quality content delivery at any scale.

As the industry moves towards more personalised, interactive, and intelligent digital experiences, the role of CDNs will continue to change as trends and user habits fuel the exploration of new technologies and applications. The future of content delivery is not just about speed - it's about adaptability, efficiency, and innovation. HIPER Global is committed to driving that future forward, enabling our partners to confidently navigate an ever-changing digital landscape.

To learn more about HIPER Global, our CDN technologies, and how we can help you build the future faster, visit HIPER Global at www.hiper-global.com.

References

- ¹ Akamai. "What is a CDN (Content Delivery Network)?" Akamai. Published 2020. <https://www.akamai.com/glossary/what-is-a-cdn>
- ² Baptiste JL. "Content delivery networks explained." Medium. Published 2023. <https://medium.com/globant/content-delivery-networks-explained-5a1feaa224c8>
- ³ Carmel S. "The history of content delivery networks (CDN)." GlobalDots. Published 2012. <https://www.globaldots.com/resources/blog/the-history-of-content-delivery-networks-cdn/>
- ⁴ Giorgio P. "2023 sports fan insights: The beginning of the immersive sports era." Deloitte. Published 2023. <https://www2.deloitte.com/us/en/insights/industry/media-and-entertainment/immersive-sports-fandom.html>
- ⁵ Imperva Learning Center. "The essential CDN guide." Imperva. Published 2024. <https://www.imperva.com/learn/performance/cdn-guide/>
- ⁶ Jackson M. "EE UK to trial new content delivery network tech with TV services." ISPreview. Published 2024. <https://www.ispreview.co.uk/index.php/2024/08/ee-uk-to-trial-new-content-delivery-network-tech-with-tv-services.html>
- ⁷ Lawrence D. "Edge AI success demands technology curated for the edge." VMware. Published 2024. <https://blogs.vmware.com/sase/2024/10/31/edge-ai-success-demands-technology-curated-for-the-edge/>
- ⁸ Ritchie H, Mathieu E, Roser M, Ortiz-Ospina E. "Internet." Our World in Data. Published 2023. <https://ourworldindata.org/internet>
- ⁹ SatNews.com. "Video CDN services come to Israel with partnership of MedOne and Redge Technologies." SatNews.com. Published 2024. <https://news.satnews.com/2024/10/09/video-cdn-services-come-to-israel-with-partnership-of-medone-and-redge-technologies/>
- ¹⁰ Thales Group. "5G vs 4G: what's the difference?" Thales Group. Published 2020. <https://www.thalesgroup.com/en/worldwide-digital-identity-and-security/mobile/magazine/5g-vs-4g-whats-difference>
- ¹¹ Viggiano J, Girinathan J, Moening A. "CDN." The Web Almanac by HTTP Archive. Published 2024. <https://almanac.httparchive.org/en/2024/cdn>
- ¹² Vodafone Press Office. "Vodafone, Qualcomm and Ericsson complete 5G mmWave trials in the UK." Vodafone. Published 2024. <https://www.vodafone.co.uk/newscentre/press-release/5g-mmwave-trials-uk-newbury-stadium/>
- ¹³ Yadav N. "Comcast deploys Qwilts Open Edge platform for distributed CDN." Data Center Dynamics. Published 2024. <https://www.datacenterdynamics.com/en/news/comcast-deploys-qwilts-open-edge-platform-for-distributed-cdn/>
- ¹⁴ Zoting S. "Content delivery network (CDN) market size, share, and trends 2025 to 2034." Precedence Research. Published 2025. <https://www.precedenceresearch.com/content-delivery-network-market>



HIPER Global UK
Unit 5 Millars Business Centre
Fishponds Close
Wokingham
RG41 2TZ
Tel: +44 (0)118 929 4990
Email: uk.info@hiper-global.com

HIPER Global (Global Headquarters)
10 Hamelacha Street, Afek Business Park
Rosh Haayin
4809118
Israel
Tel: +972 3 910 9105
Email: info@hiper-global.com

HIPER Global EDGE
10 Hamelacha Street, Afek Business Park
Rosh Haayin
4809118
Israel
Tel: +972 3 910 9105
Email: info@hiper-global.com

HIPER Global US (Marlborough)
Building 2
24 Saint Martin Drive
Marlborough
MA 01752
Tel: +1 978 486 0300
Email: us.info@hiper-global.com

HIPER Global ENTERPRISE
20 Haluzei Hata'asiya Street
Haifa
2629420
Israel
Tel: +972 3 910 9105
Email: info@hiper-global.com

HIPER Global US (US Headquarters)
Suite 1001
1616 Osgood Street
North Andover
MA 01845
Tel: +1 978 486 0300
Email: us.info@hiper-global.com

HIPER Global US (Atlanta)
Suite 225
1225 Old Alpharetta Road
Alpharetta
GA 30005
Tel: +1 978 486 0300
Email: us.info@hiper-global.com