**Content Delivery Network Report**

Name: Nathaniel Teskey Student number: 20247672 Date: 15 March 2022

1. **Overview of how Content Distribution Networks operate**

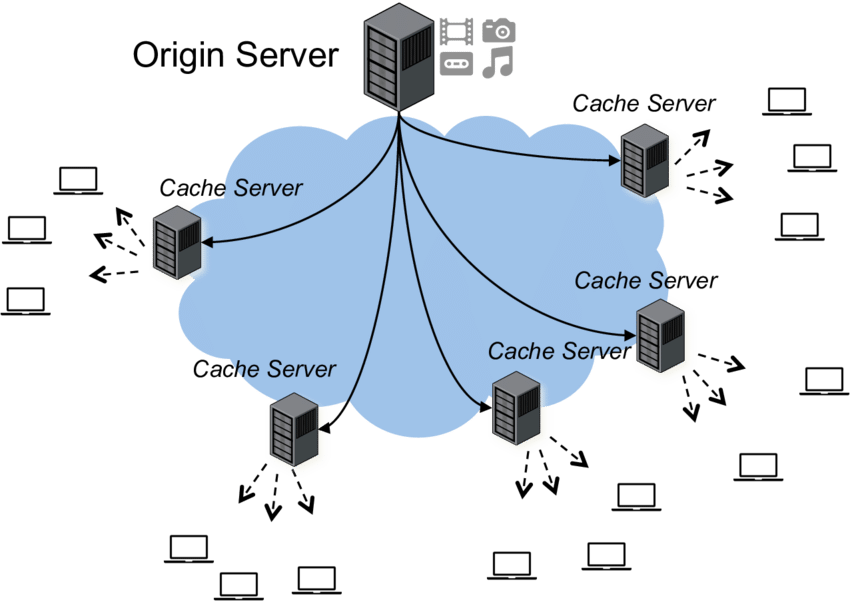
Akamai is one of the world leaders in content delivery network (CDN) services for software and media delivery, the company also offer cloud security solutions. In order to understand what Akamai do, we first need to understand what a content delivery network is. A content delivery network is a group of servers that a distributed around the world that work together to deliver Internet content quickly with low latency. A CDN enables data to be transferred at a very quick rate, this data is required for loading various types of internet content such as HTML pages, stylesheets, JavaScript files and different types of media such as images and videos. Content Delivery Networks are becoming more and more popular, currently the majority of internet traffic is served through content delivery networks. CDN architecture reduces the need to haul data over long distances, therefore reducing network latency. Removing latency from the equation gets more important as more dynamic content such as software as a service and video are delivered to an increasing number of mobile devices, CDNs create a “super motorway” to accelerate this delivery of this content across a long distance.

Another benefit of content delivery networks, apart from them allowing for fast delivery of internet content is that they also assist in protecting websites against common cyber-attacks such as Distributed Denial of Service (DDoS) attacks. To prevent against DDoS attacks the CDN employs techniques such as HTTP load balancing. CDNs also allow for better encryption, authentication and integrity standards due to issuing fresh TLS and SSL certificates. Content availability is improved and if one server goes offline, others can be recruited to deal with the web traffic. Similarly Distributed Denial of Service attacks can be mitigated by the CDN as it has the ability to distribute malicious requests across the network.

CDN provides store cached content in their own network points of presence (PoPs) or in third party data centres. There are lots of these strategically located data centres situated around the world, within each PoP are thousands of caching servers. In principle this works when a user requests content from a website, it then checks if the content is cached on a CDN nearby and if it is it redirects the request to the nearest server to the user and delivers the cached content from its original location at the network edge, all of this happens is milliseconds and it is entirely invisible to the user which allows for a seamless experience. If the CDN does not have the files that were requested by the user, it will load from the origin as required. If the origin is very far away from the user then latency will obviously become an issue. Content distribution networks are especially useful for businesses that attract a large amount of web traffic, such as video streaming platforms like Netflix and e-commerce giants such as Amazon.

CDN Caching is one of the main parts of what makes content distribution networks work. In a nutshell it is the process of keeping a copy of files that were delivered to a user for the first time and reusing these stored copies of the files instead of the originals for all requests after the initial request. Edge servers are where the data is cached in a content distribution network.

The steps for CDN caching are as follows:

1. A user requests files on your web page for the first time.
2. The assets are pulled from the origin server and delivered to the user, but once they are delivered a copy of them is also stored in the Point of Presence edge caching server that is located closest to the end user.
3. When the same user requests the same files the next time the request doesn’t go to the origin server but it searches through the cached files in the PoP server to see if the files you accessed before and that were stored are still available. Obviously if the files are not available the request is sent to the origin server again.

[[1]](#footnote-1)CDN Diagram

CDNs are also beneficial to engineering teams and website owners to manage access to your platform or services. If your business has users or consumers who are spread out across the globe, you may be required to allow access for some regions and deny access to others. Using CDNs you can also collect logs and analyse data about your users. This can be very important if your business attracts a lot of traffic and you need to perform real time analyses of the web traffic.

The main CDN providers globally by customer count are Cloudflare, Amazon Web Services (AWS), Akamai, Net DNA, Fastly, Imperva, Verizon and Microsoft Azure. The top three providers have 89% of customers.[[2]](#footnote-2) Cloudflare and AWS CloudFront have the most customers by count however, Akamai has more enterprise customers. Akamai has been the market leader in CDN since the 1990s. They have been in business 10 years longer than any of their competitors. Akamai own upwards of sixty five percent of the market in revenue, this is due to the fact that it has many large-scale enterprise partners. Akamai may have fewer customers, but they are the only provider that has more enterprise level than mid-market customers. These customers are major users with heavy traffic.

Currently there are 8 CDNs that have one or more PoPs in Ireland. These are all based in Dublin with their PoP location being in Dublin too, however Cloudflare have a PoP placed strategically at the bottom of the country, located in Cork. Some of the CDN providers in Ireland are bunny.net, CDNetworks, CDNvideo, Cloudfront, Cloudflare and Lumen.[[3]](#footnote-3) Ireland is a small country, so it does not need many PoP locations to serve the entire country. I think that the reason they could be all based in Dublin is due to the high population density, many servers are required to handle all of the requests due to the large number of users in the region.

1. **Analysis of “The Akamai Network: A Platform for High-Performance Internet Applications (PDF)” by Erin Nygren, Ramesh Sitaraman, and Jennifer Sun**

The Akamai team acknowledge that the complexity of a globally distributed delivery network brings about a unique set of challenges in architecture, operation and management. Network management and data collection needs to be scalable and fast across thousands of server clusters, many of which are located in unmanned, third-party data centres, and any number of which might be offline or experiencing bad connectivity at any given time. Configuration changes and software updates need to be rolled out across the network in a safe, quick, and consistent manner, without disrupting service. we begin with the assumption that a significant number of failures (whether they be at the machine, rack, cluster, connectivity, network levels) is expected to be occurring at all times in the network. [[4]](#footnote-4)

Akamai have designed their delivery networks with the thinking that failures are normal and happen often and that the network must operate seamlessly despite issues. One of the main design goals that Akamai strive for is a design for reliability. Their goal is to have extremely close to 100% end-to-end availability. They must ensure full redundancy of components and build in multiple levels of fault tolerance. They use protocols such as PAXOS and decentralized leader election to allow for the event of failing system components. [[5]](#footnote-5)

Akamai also strives for scalability. They have more than 60,000 machines across the globe. Scaling means that they need to account for handling more traffic, content and customers. They also need to consider handling increasing large volumes of data as well as building communications, control and mapping systems that are required to support an ever-growing number of distributed machines.[[6]](#footnote-6)

They design the system to be autonomic, reducing the need for human management. If they didn’t do this the human operational expenses would become too high. Hence, the system must be able to respond to faults, handle shifts in load and capacity and self-tune for performance with minimal human intervention. At the time the report was published the Akamai network operations centre employed 60 people to work all year round for 24 hours a day, to keep the system maintained.[[7]](#footnote-7)

They also carry out continual work to ensure the performance of the systems critical paths to improve end user experience, increase performance, cache hit rates and network resource utilization. An example of this is kernel and software optimizations to enable greater capacity and more traffic served with fewer machines.[[8]](#footnote-8)

When a user types a URL into their browser, the domain name of the URL is translated by the mapping system into the IP address of an edge server to serve the content to the end-user. The mapping system bases its answers on large amounts of historical and current data that have been collected and processed. The data is used to choose an edge server that is located nearby to the end user in order to generate minimal latency.[[9]](#footnote-9)

A content delivery network is different to an application delivery network (ADN). This is a key difference that matters to the Akamai network. A CDN can make the task of application performance management (APM) more complex, however they are considered to be better value and a relatively inexpensive way to handle content and scalability for a website. CDNs can sometimes be sloppy and there can be potential for confusion which is not ideal in an AND. Many people generalise the two terms under the term CDN. CDNs work by caching frequently accessed digital content at geographically distributed edge locations. When a client browser requests the cached content, it comes from the nearest edge location. Comparatively, an ADN is a combination of features that provide application availability, security, visibility, and acceleration.  ADNs work with dynamic remote applications, requiring the delivery of real-time data, analytics, and user preferences between the application server and the client. Because every client has different data, each request is fetched from the origin server. Remote applications also require an intelligent traffic monitoring and management solution for distributing network traffic over multiple servers. The Akamai AND helps customers to avoid connectivity problems and allows for high-performance communication between servers which allows for path optimization, packet loss reduction, transport protocol optimizations along with application optimizations which can compress content in order to increase performance.[[10]](#footnote-10)

A peer-to-peer network would not suffice for the Akamai network as it would not ensure guaranteed performance and availability of content. The team do not want to lose control of their content delivery. The client-side model of the Akamai network allows for the platform to seed content to users and is able to fall back once the content has been issued successfully. The key here being that Akamai require full control and knowledge regarding their content at all times, this ensures integrity and guarantees availability. [[11]](#footnote-11)

The Akamai platform is effective in overcoming stress on the origin server. Instead it distributes the load to a variety of PoPs which help to manage the high volumes of traffic and prevent the origin from crashing. The Akamai CDN can also mitigate DDoS attacks as I outlined previously. The way servers are spread out means none are in danger of being overloaded. The platform also increases speed and website performance. It helps businesses to deliver content quickly and effectively by mitigating latency and reducing load times as users can access edge servers that are close to them. If a website is slow to load for a user the business might lose a customer as they will go to a competitor’s site, the platform rectifies this by ensuring minimal latency. Content can be delivered quickly as it is cached on nearby CDN servers. The platform also provides insights on the target audience as it collects a range of data about each user and it features advanced analytical software which aids businesses in marketing and decision making. It can also allow business to reach larger audiences as it is highly scalable and it keeps end users away from the origin website which may not be able to handle high volumes of traffic.

The Akamai platform must be efficient at overcoming the issues listed above in order to keep its position as a market leader as there are many other CDN services available, such as Cloudfront, AWS and Microsoft Azure. If Akamai is not able to carry out these optimisations then customers/ users of the platform will jump ship and go to another competitor instead.

A CDN can be applied to other sectors such as manufacturing, health and gaming. In manufacturing it could be used for pulling files for 3d printing or CNC machining. These files could be of a large size and if the origin is very far away (in a different country or continenant) then the latency could be huge, resulting in crippling download speeds. The CDN could help here by pulling the files or content from a nearby cache-server that is storing the files for ease of access globally. Similarly, it could be used in healthcare for doctors accessing patient information in different parts of the world. It allows for quick response times and ensures that the information can be accessed at all times. If the HSE had been using a CDN to store patient information it would have been much harder for the organisation to be hacked as its files would have been stored on various servers around the world, instead of being stored in one place. Gaming also benefits from a CDN if there is an update required etc. For example as we are in Ireland there would be a lot of latency involved in pulling an update from an origin server in the USA. With a CDN gamers around the world could download updates and service packs for their games much more quickly as it would be pulling the update from a nearby PoP or data centre, therefore reducing latency.

1. **Analysis of how CNN.com and Akamai use DNS to redirect requests for JPEG files so that files are retrieved from an Akamai server close to the end-user.**

First the end user types in [www.cnn.com](http://www.cnn.com) into their search engine. This sends a request to the host, which is CNN. CNN is also the content provider and the origin. Next the CNN website starts loading the website to be displayed on the end-user’s computer. CNN.com has its JPEG images stored on Akamai DNS servers, as Akamai has a range of servers located around the world. CNN know that not everyone is in close proximity to the origin server that the website is hosted on however it can very quickly send the text content from the CNN site anywhere in the world as texts files are very small. However JPEG files are larger and require more bandwidth to be sent and received so they have outsourced this to Akamai, who already have the infrastructure built up around the world to deal with this type of content. Latency becomes a bigger issue when the content being delivered has a larger size. CNN store its JPEG files on various Akamai clusters around the world, in order to be in close proximity to all of its end users. The CNN website sends an instruction to the end-user’s computer to connect to a DNS TLD server, this server then sends an instruction to the end-user which tells it to connect to a large Akamai global DNS server. That server then sends instruction to the end user to connect to an Akamai regional DNS server which then figures out which cluster is nearest to the end-user’s computer. The end-user receives this information from the regional DNS server and sends a GET request to the nearby Akamai cluster to track down the JPEG files required to finish loading the CNN website. This nearby Akamai cluster contains a cache of the JPEG images required by the end user that are for the CNN website. The Akamai nearby cluster then sends the JPEG content to the end user via the lowest latency route as it is much closer than the origin CNN server. The nearby Akamai cluster then sends a message to the origin CNN server (content provider) that it has issued the end user with the JPEG images it requested. The origin then responds to the nearby Akamai cluster and acknowledges that it knows that the end-user has received the JPEG images and the whole CNN.com website has loaded in a matter of milliseconds on the end-user’s computer. This all happens so quickly that the end user is completely unaware of any of it. It allows for a seamless link between the origin CNN.com and the Akamai server that is closest to the end user which provides them with the large JPEG files that the CNN.com origin want to be displayed to the end user on their website ([www.cnn.com](http://www.cnn.com)). If the end user had to wait for the origin CNN server to send the required JPEG files to their computer it could take a long time due to latency and the end user might leave the CNN website and go to a competitors site instead such as Fox News.

1. ResearchGate. (2017). High level understanding of content delivery networks [online].

   Available from: <https://www.researchgate.net/figure/High-level-understanding-of-Content-Delivery-Networks_fig2_321791733>

   [accessed 15 March 2022]. [↑](#footnote-ref-1)
2. Intricately. (2020). CDN Industry: Trends, Size, And Market Share [online].

   Available from: <https://blog.intricately.com/cdn-industry-trends-market-share-customer-size>

   [accessed 15 March 2022]. [↑](#footnote-ref-2)
3. CDN Planet. (2022). Ireland CDN [online].

   Available from: <https://www.cdnplanet.com/geo/ireland-cdn/>

   [accessed 15 March 2022]. [↑](#footnote-ref-3)
4. Nygren E, Sitaraman R, Sun J. (2010). The Akamai network: a platform for high-performance internet applications. pp.5 [↑](#footnote-ref-4)
5. Nygren, Sitaraman, Sun. The Akamai network. pp5 [↑](#footnote-ref-5)
6. Nygren, Sitaraman, Sun. The Akamai network. pp5 [↑](#footnote-ref-6)
7. Nygren, Sitaraman, Sun. The Akamai network. pp5 [↑](#footnote-ref-7)
8. Nygren, Sitaraman, Sun. The Akamai network. pp5 [↑](#footnote-ref-8)
9. Nygren, Sitaraman, Sun. The Akamai network. pp4 [↑](#footnote-ref-9)
10. Nygren, Sitaraman, Sun. The Akamai network. pp8-9 [↑](#footnote-ref-10)
11. Nygren, Sitaraman, Sun. The Akamai network. pp15 [↑](#footnote-ref-11)