**EX1**

**Abstract:**

Due to the improvements in network communications, the Content Distribution Network (CDN) is becoming more popular and its market has been predicted to grow over 30 billion USD in the coming years. As a result, more and more bandwidths are being utilized. This report explored CDN, how it operates, the distinctions between the traditional and Cloud-based CDN, as well as examples of major Cloud-CDN providers in the marketplace. Furthermore, the Lab1 exercises based on AWS CloudFront was also reported.

Qs1: How do CDNs work? What is the difference between a traditional CDN and a Cloud-based CDN? Provide an in-depth discussion with examples and figures.

Answer:

According to Behrouz et al. (2021), Content Delivery Networks (CDNs) can be defined as a group of interconnected network components that work together to distribute content to users in an efficient and transparent manner. In CDN, contents are copied across numerous mirrored Web servers which both homogeneous and heterogeneous environments support. CDNs deliver services that ehnance performance by maximizing on the network bandwidth, improving accessibility and maintaining the correctness of the replicated contents. Generally, the function of a CDN includes but not limited to: Request redirection and content delivery services, content outsourcing and distribution services, content negotiation services, and management services.

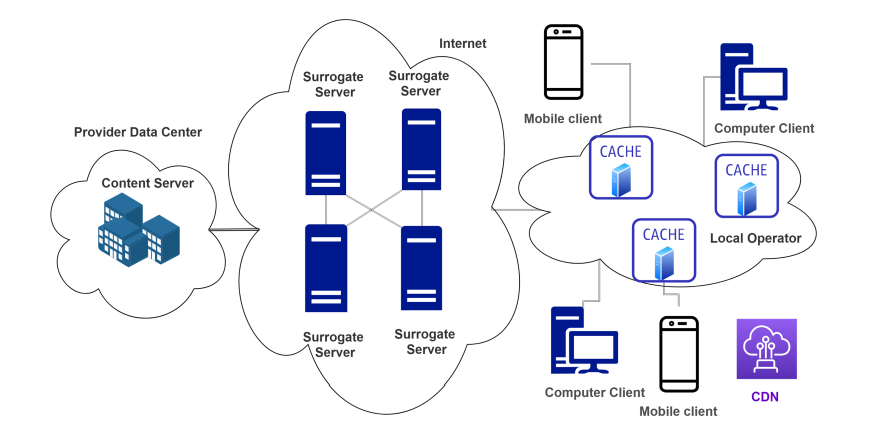
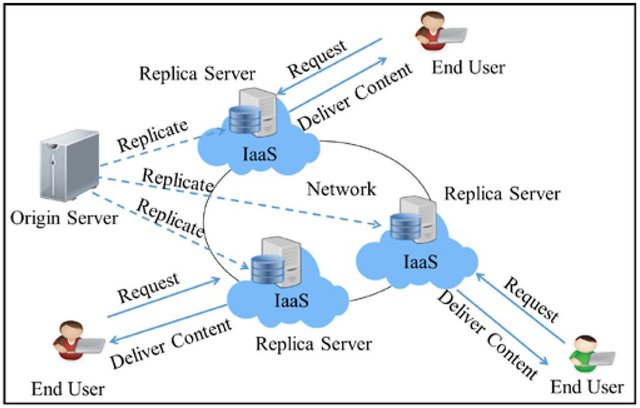


Figure 1: A typical CDN architecture [ref]

A hierarchy of servers aid in the distribution of material in the CDN architecture depicted in Figure 1. Several interconnected surrogate servers that are dispersed around the Internet mirror the content servers. In order to decrease access times and increase service quality, frequently used content items are cached from the closest surrogate server into the neighbourhood edge nodes. Edge caches and surrogate servers can be built and operated by several parties. When a user requests content, the request is forwarded to the closest cache, which is then in charge of providing the requested content. If the content is not already stored in the cache, an attempt will be made to retrieve it from other caches at the same level of the hierarchy. If a miss recurs, the query will be passed up the hierarchy, where the same process will be applied. Moreso, the CDN's provision of content caching lowers traffic on the Internet and within the local operator network.

**Cloud-based CDNs vs Traditional CDNs**

End-users, content providers, and CDN providers are the three primary business actors in traditional CDN (Prathan et al, 2008). The individual who consumes the content (such as videos) from the content provider is known as the end-user. The organisation that either owns the content or has gained the rights to sell it is the content provider, such as YouTube. The organisation that possesses replica servers at advantageous locations and provides content distribution services to the content providers is the CDN provider (for instance, Akamai). The traditional CDN business architecture usually has a connection between a content provider and a CDN provider, in which the content provider pays the CDN provider for hosting its content/material. In addition, a traditional CDN business model additionally involves entities like Internet Service Providers in addition to the business actors mentioned above (ISP). In order to build replica servers, CDN providers may enter into business agreements with ISPs to rent their infrastructure (such as servers and datacenters).



Traditional CDN architecture (Jagruti et al.)

On the other hand, the cloud-based CDNs (usually referred to as CCDNs) offer a customizable platform that enables content providers to intelligently match and arrange material on one or more cloud storage servers and compute resources in accordance with their choices for coverage, spending, and Quality of Services (Wang, et al., 2015). Unlike the traditional CDN, CCDN offers a less expensive way to host and deploy multi-tiered applications that can scale in response to user demands.

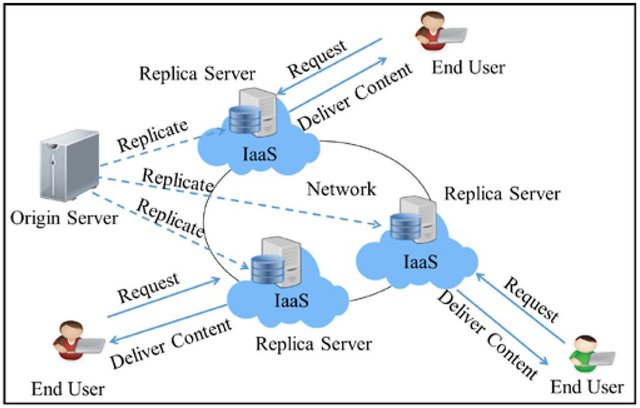


Figure 3: Cloud-based CDN architecture (Jagruti et al.)

Qs2: List major CDN providers in the CDN marketplace.

i. Google Cloud CDN

ii. Cloudflare CDN

iii. CDN77

iv. CacheFly

v. Amazon CloudFront

vi. Azure CDN

vii. Fastly

viii. Verizon Media Platform

ix. Akamai

x. StackPath

Compare two CDN providers based on at least five characteristics. For example, the number of edge locations.

|  | **Amazon CloudFront** | **Cloudflare CDN** |
| --- | --- | --- |
| **Cloud compatibility** | It is integrated with the developer tools and existing AWS services. | With a specific cloud firewall for content security, it operates on any cloud infrastructure. |
| **Edge locations** | Its 225+ points of presence (PoPs) are connected by the AWS backbone, a private network made of 100 GbE metro fibre cables that spans the Mediterranean, Red Sea, South China Sea, and the Atlantic, Pacific, and Indian Oceans. | It supports about 25 million online properties and has data centres in 200 cities spread across more than 100 countries. |
| **Security** | In addition to AWS Shield Standard, it integrates with AWS Shield Advanced and AWS Web Application Firewall (WAF). | One can benefit from a comprehensive set of security tools that cover user, network, and application levels, such as web application firewalls and threat intelligence. |
| **Advanced Analytics** | Amazon CloudFront monitoring provides rich analytics, and Amazon CloudWatch can provide near-real-time alerts. | By using website performance dashboards, it provides information on website cache.With this, data can be filtered using criteria such as hostnames, cache URLs that were missed, etc. |
| **Low latency** | According to tests by AWS team, Amazon CloudFront provides performance with extremely low latency, which is about 10% faster than the industry average for latency across CDNs. | It facilitates extremely quick content delivery and clever content request routing along the quickest route using the Argo Smart Routing feature. |

**Part 2**

**Exercise 1:** **In this exercise, you will learn how to host a static website from a**

**Cloud storage and create a CDN distribution to serve static files such as**

**images. Follow the steps below:**

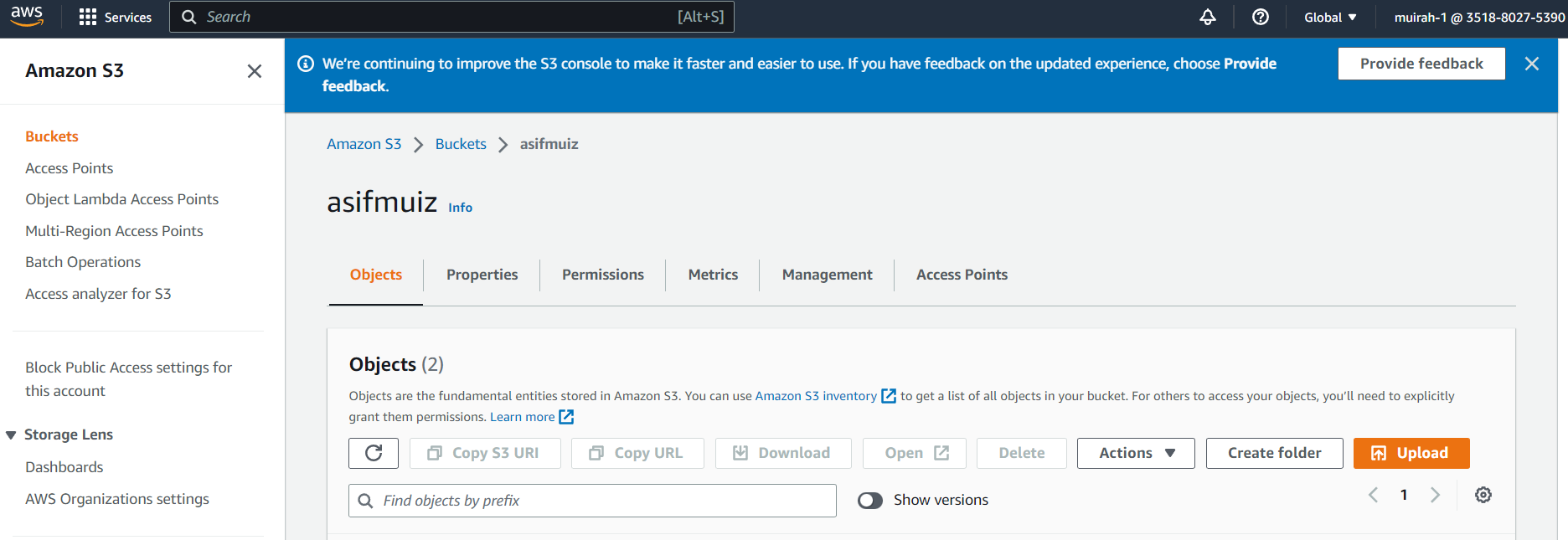
**1. Create a simple website with some static files embedded such as CSS and**

**Images.**

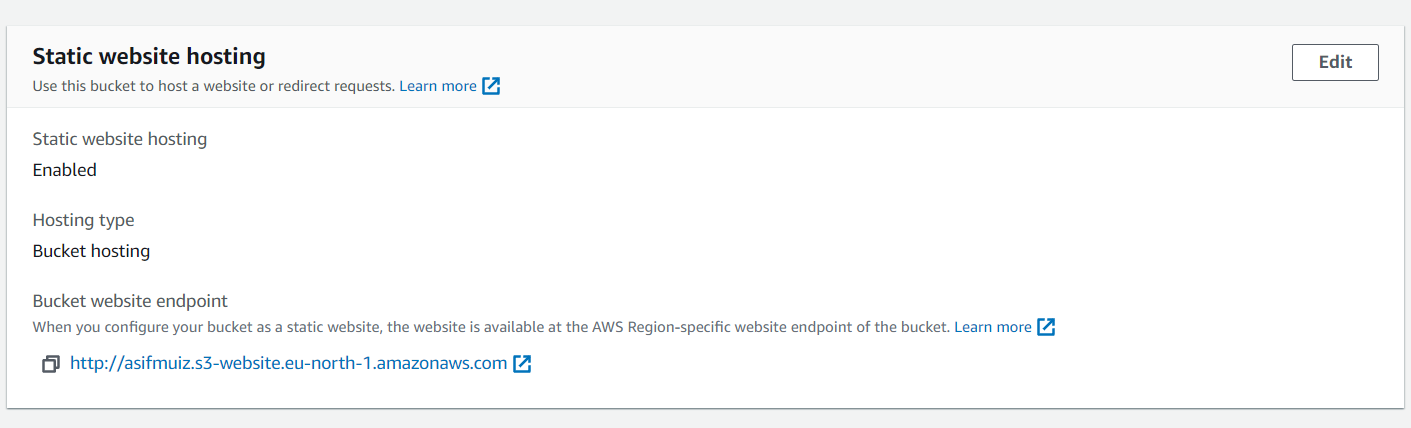
* > A simple website was developed:



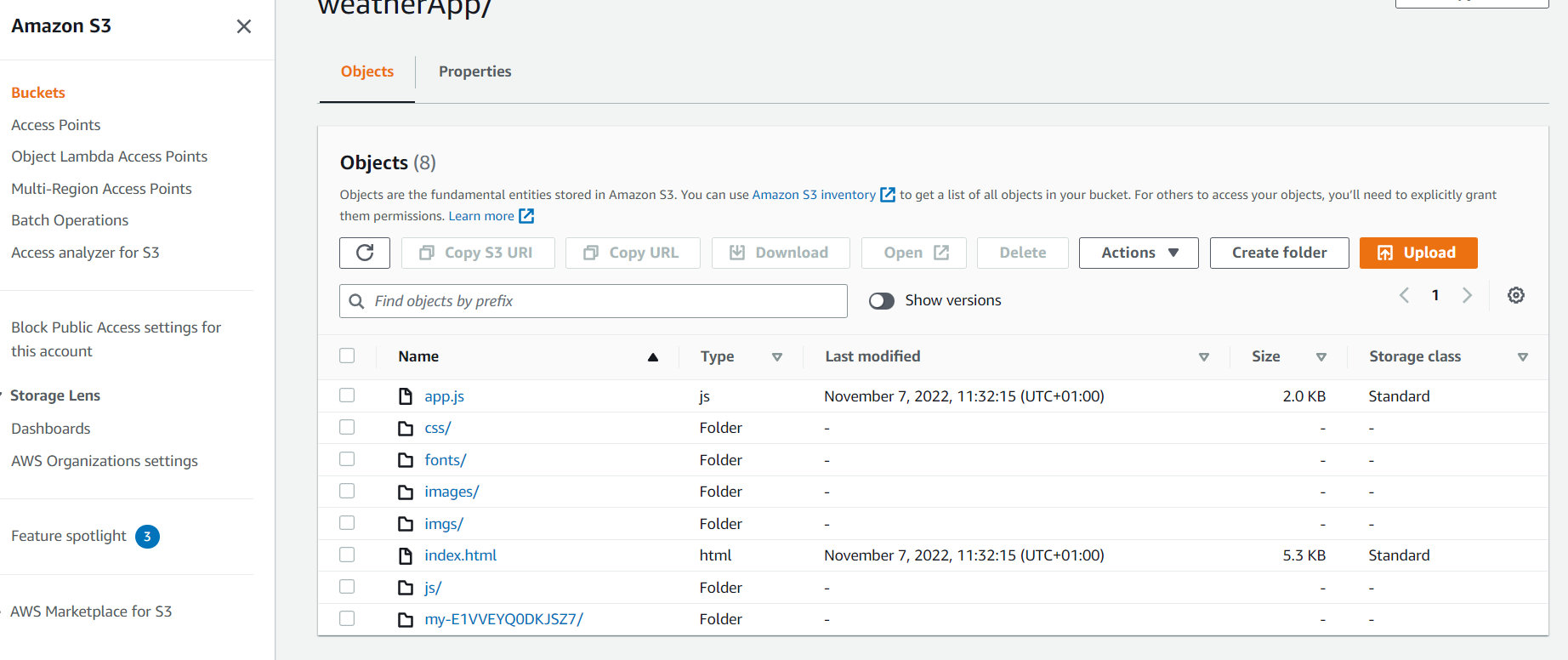
**2. From an Amazon S3 console, create a new bucket.**



**3. Enable your bucket for static hosting.**

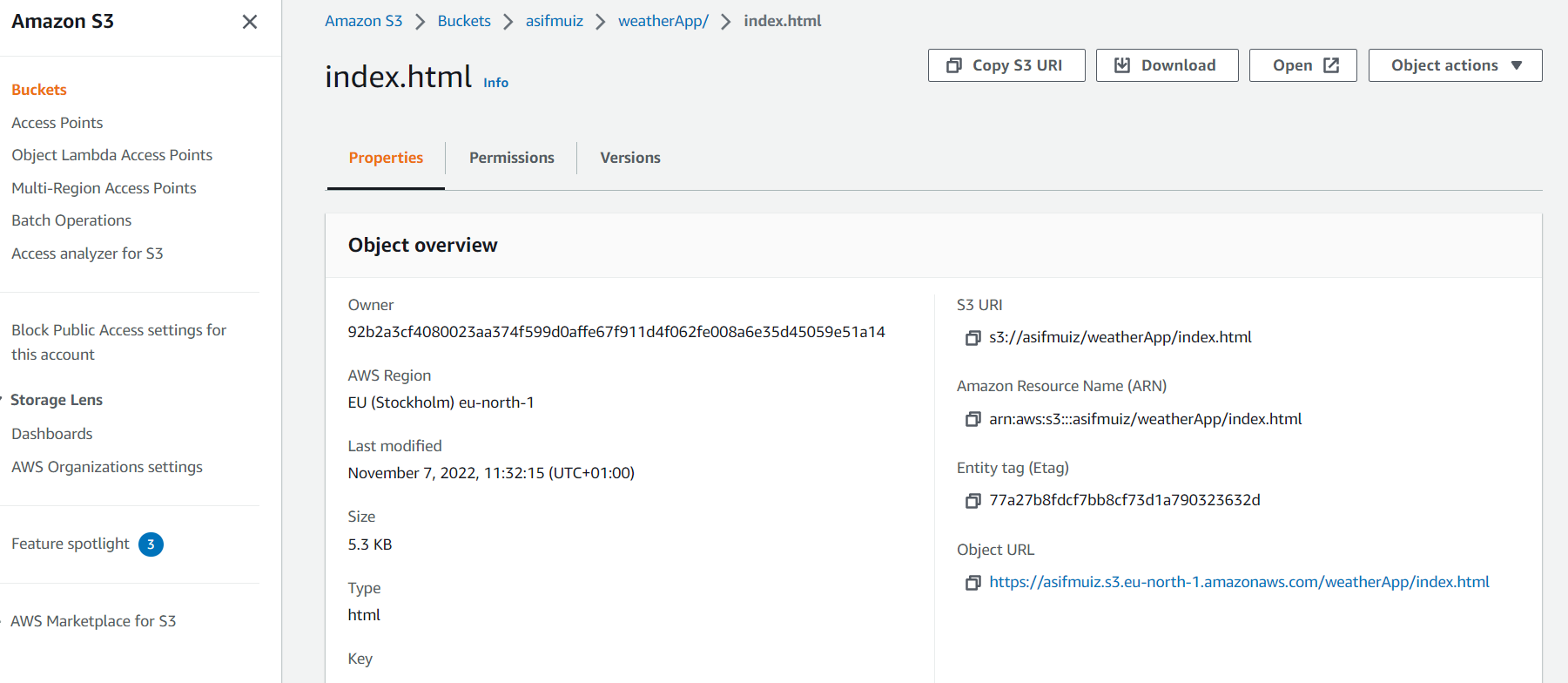


**4. Upload the website to the bucket.**



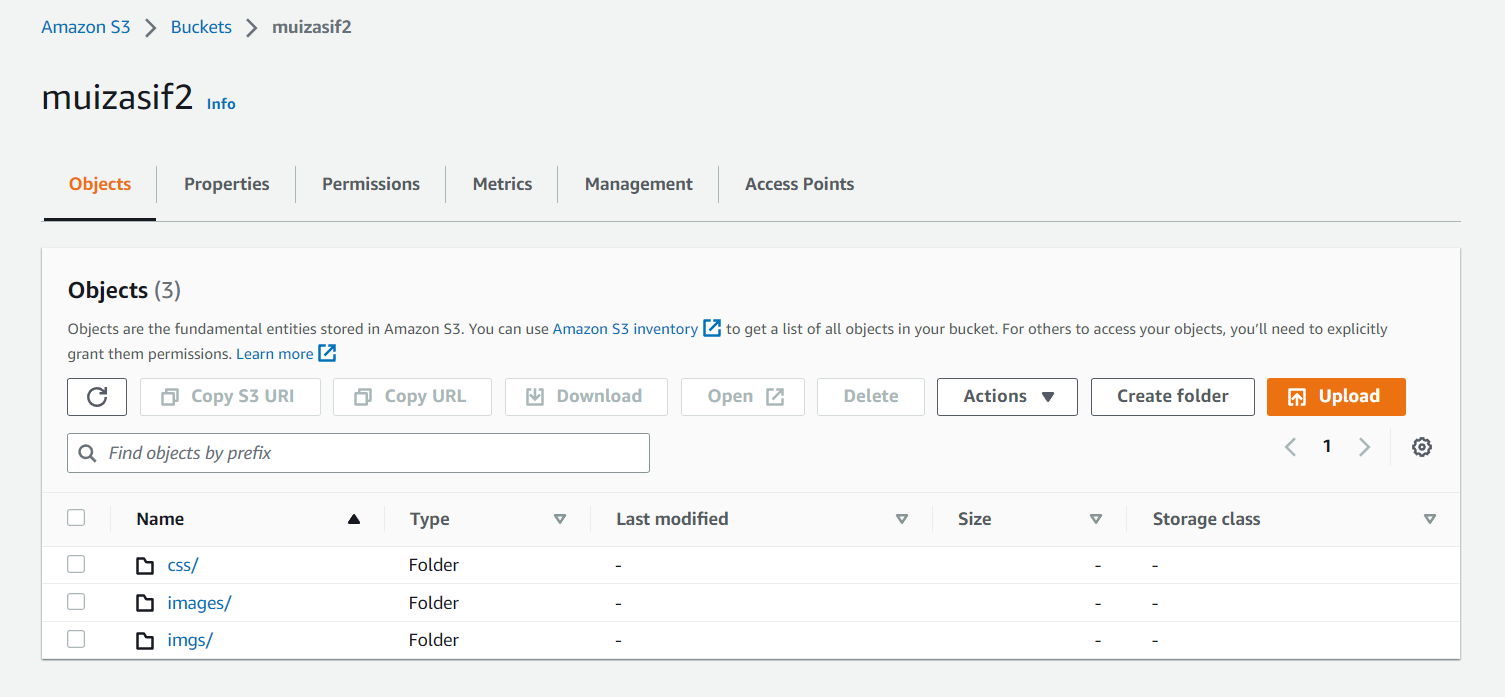
**5. From a Web browser, access the Web page via the Amazon S3 Website**

**endpoint for your bucket.**



**6. Now create another bucket and copy the static files in the webpage (CSS,**

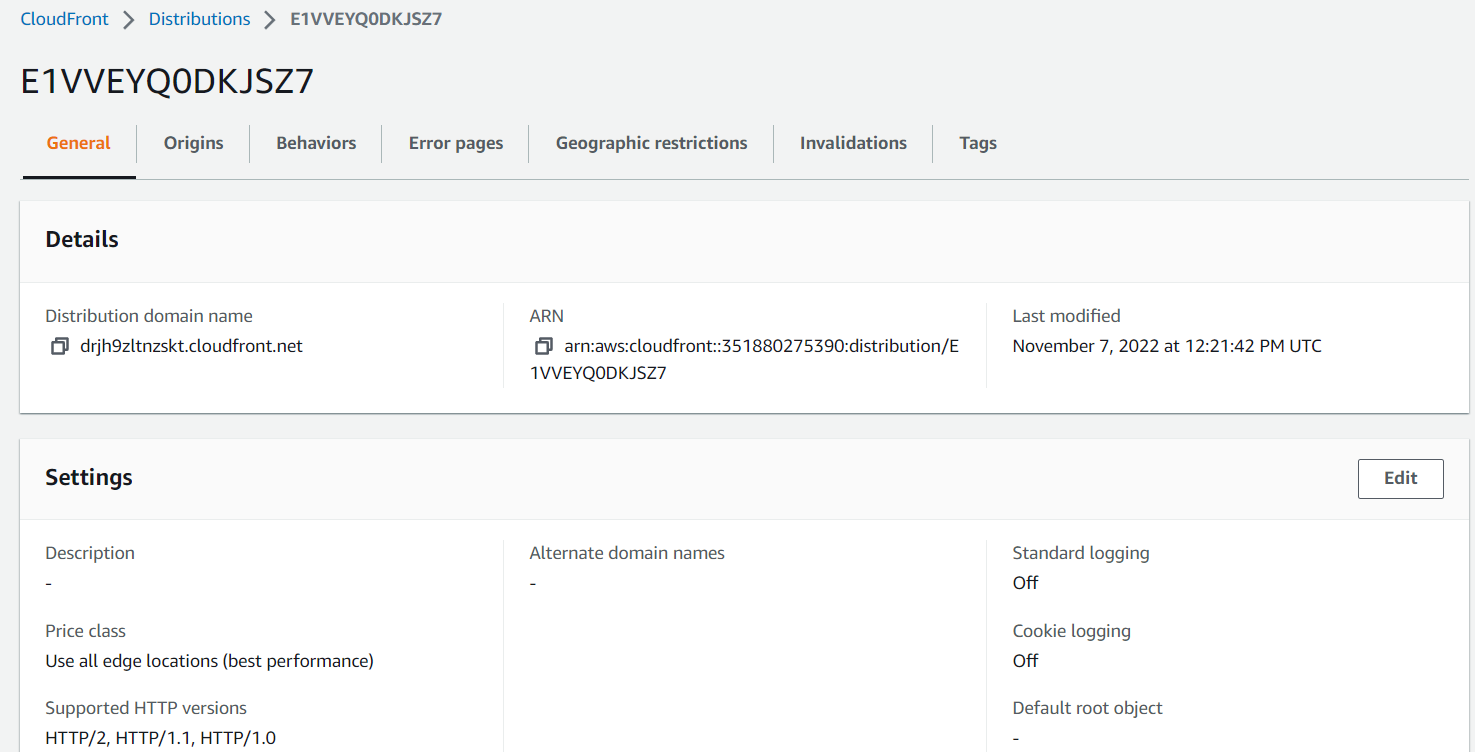
**image, etc.) to the bucket.**



**7. From the Amazon CloudFront console, create a new download distribution**

**and specify the name of the second bucket in which you copied the static**

**Files.**



**8. When the distribution becomes available, note the distribution URL.**

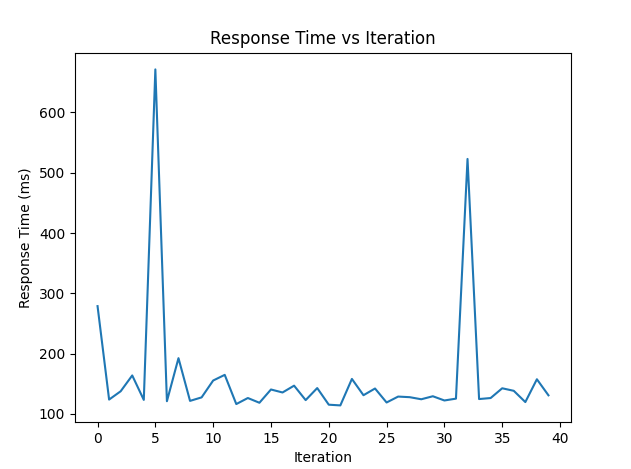
**9. Now edit the URLs of the static files in the Web page in the first bucket**

**and append the CloudFront distribution URL (e.g., files/img.jpg to <mycloudfrontdistURL>/files/img.jpg).**

**10. Open the Web page again in the Web browser and note if the response**

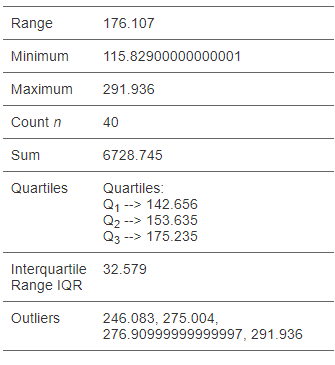
**time improves over several runs.**

* After running the website 40 times we get the following response times:



As we can see from the figure above, the reload time was mostly below 200 ms. After performing a statistical analysis on the response times we observe the following:



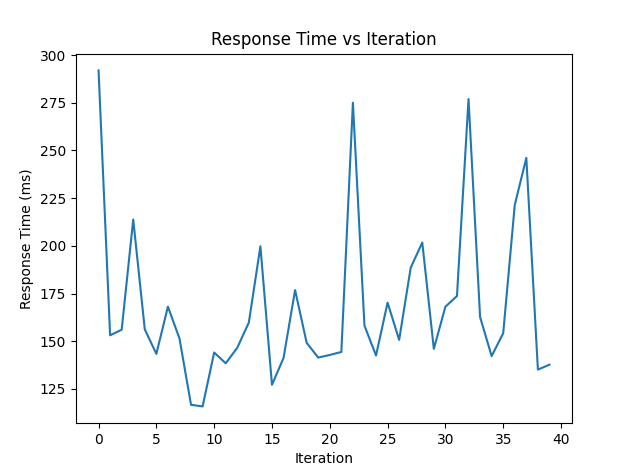


**Exercise 2: Test your CDN distribution using a number of methods and explain the results.**

For this task we made use of different sized static files as an input and observed the following:

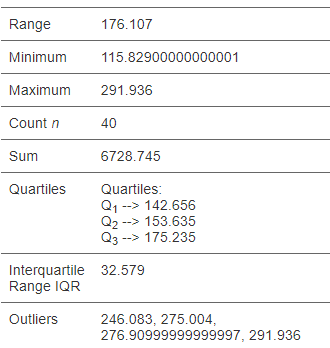
For 1Mb file:

The response times:



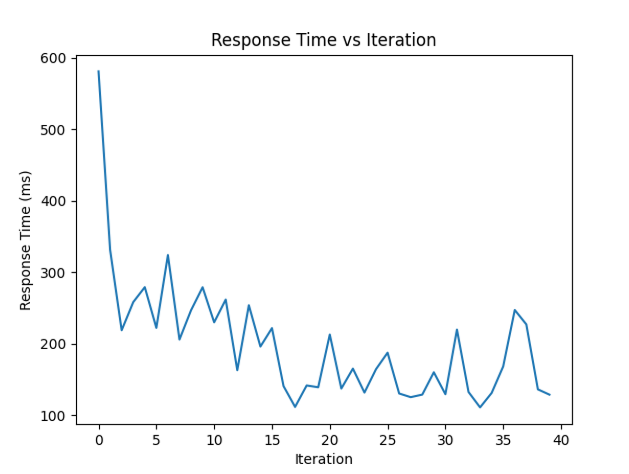
After doing a statistical analysis on the response data we get:





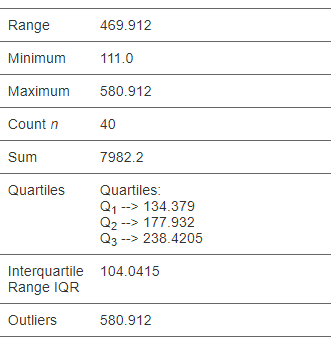
For 10Mb files:

The response times:



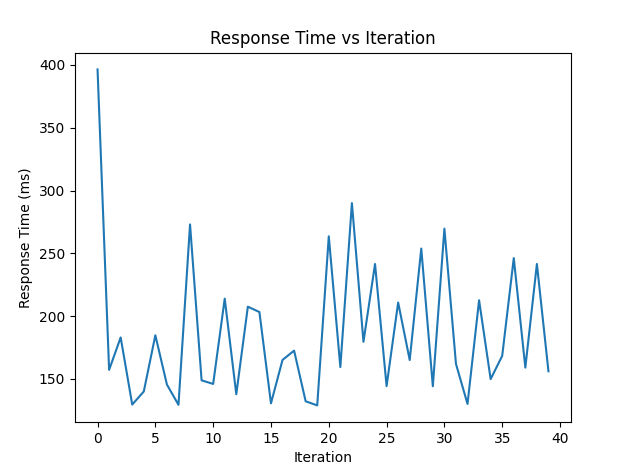
After doing a statistical analysis on the response data we get:



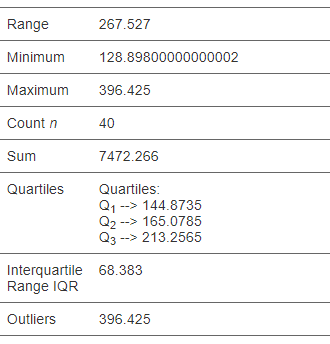


For 20Mb files:

The response times:

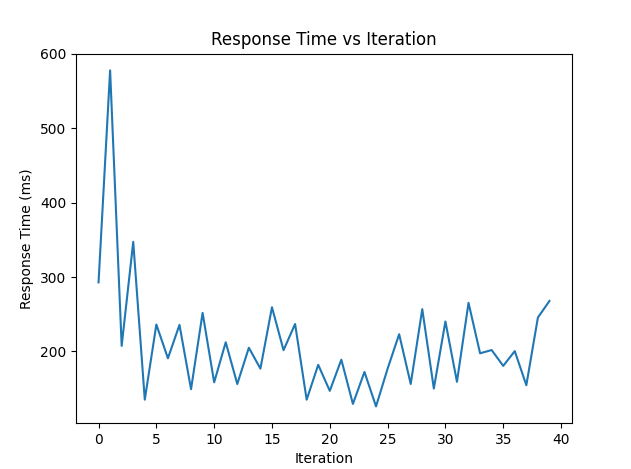


After doing a statistical analysis on the response data we get:

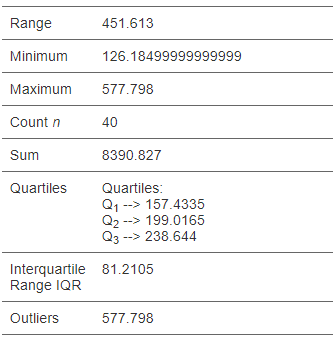
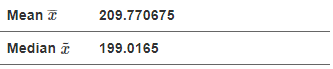


For 50Mb files:

The response times:



After doing a statistical analysis on the response data we get:



**References**

Behrouz Zolfaghari, Gautam Srivastava, Swapnoneel Roy, Hamid R. Nemati, Fatemeh Afghah, Takeshi Koshiba, Abolfazl Razi, Khodakhast Bibak, Pinaki Mitra, and Brijesh Kumar Rai. 2020. Content Delivery Networks: State of the Art, Trends, and Future Roadmap. ACM Comput. Surv. 53, 2, Article 34 (March 2021), 34 pages. https://doi.org/10.1145/3380613

(16) (PDF) A Survey on Replica Server Placement Algorithms for Content Delivery Networks. Available from: https://www.researchgate.net/publication/309738174\_A\_Survey\_on\_Replica\_Server\_Placement\_Algorithms\_for\_Content\_Delivery\_Networks [accessed Nov 17 2022].

M. Pathan, R. Buyya, and A. Vakali, “Content delivery networks: State of the art, insights, and imperatives,” in Content Delivery Networks. Heidelberg, Germany: Springer, 2008, pp. 3–32.

Wang, M., Jayaraman, P. P., Ranjan, R., Mitra, K., Zhang, M., Li, E., ... & Georgeakopoulos, D. (2015). An overview of cloud based content delivery networks: research dimensions and state-of-the-art. *Transactions on large-scale data-and knowledge-centered systems XX*, 131-158.