UOW MALAYSIA KDU PENANG UNIVERSITY COLLEGE

PART OF THE UNIVERSITY OF WOLLONGONG AUSTRALIA GLOBAL NETWORK

ASSIGNMENT COVER PAGE

Validated by: UNIVERSITY OF LINCOLN UNITED KINGDOM

Programme	Course Code and Title
Bachelor of Computer Science (Hons) (UCSE) / Bachelor of Computer Science (Hons) in Computer and Network Technology (UCNT)	COS3043/N System Fundamentals
Student's name / student's id	**Lecturer's name**
0204677 Lim Zhe Yuan	Foo Lye Heng

Date issued	Submission Deadline	Indicative Weighting
20/10/2023 (Week 6)	24/11/2023 (Week 11)	30%
Assignment [1] title	Research report writing	

This assessment assesses the following course learning outcomes

# as in Course Guide	UOWM KDU Penang University College Learning Outcome
CLO2	Evaluate software design issues for advanced computer systems such as multiprocessors or distributed systems
CLO3	Discuss current topics in system fundamentals research by reading and analysis of journal papers.

# as in Course Guide	University of Lincoln Learning Outcome
CLO1	Identify the values of object-oriented design and programming
CLO2	Apply object-oriented principles to the implementation of software programs
CLO3	Apply advanced logical and mathematical techniques in the development of software solutions

Student's declaration

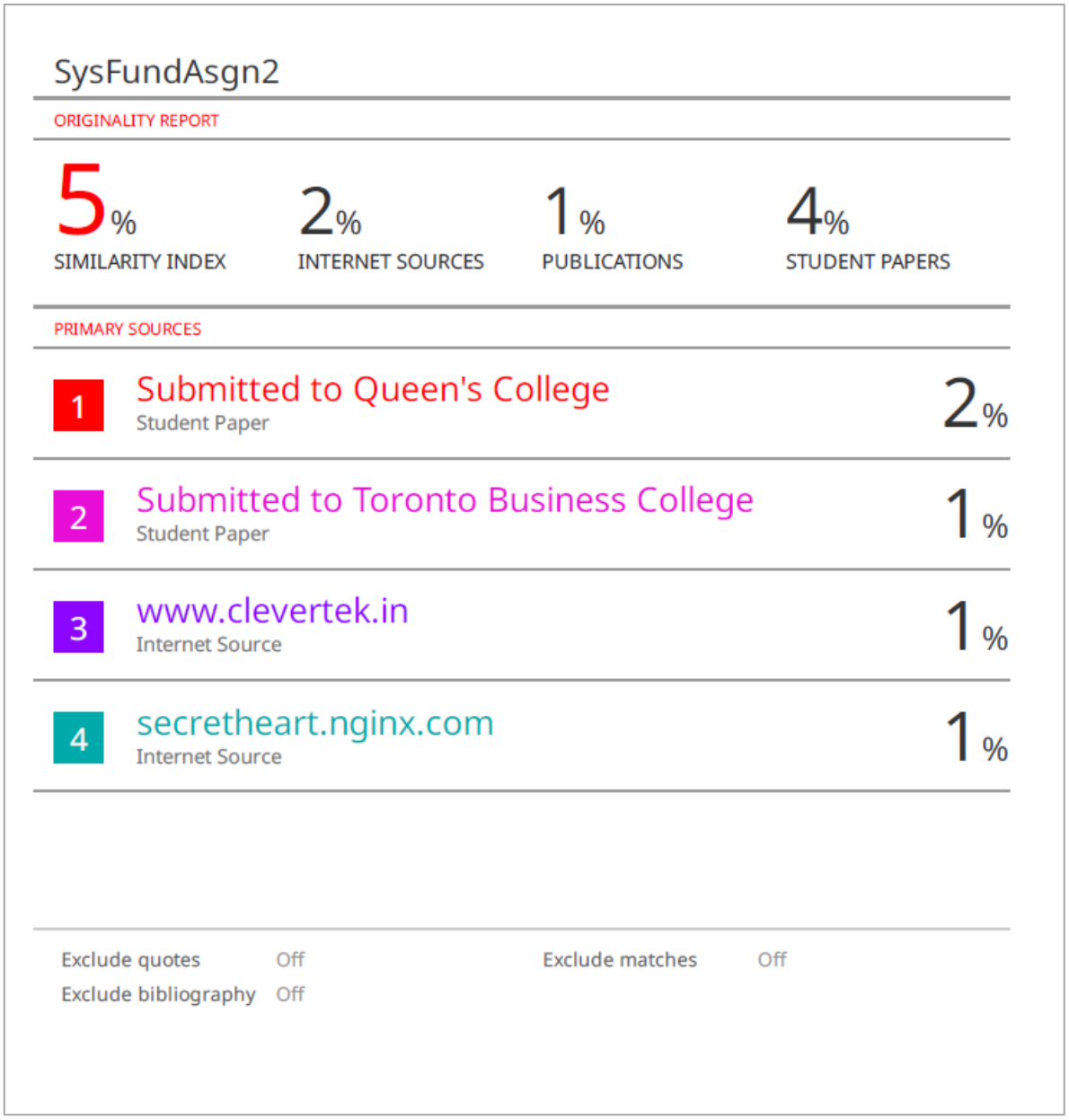
I certify that the work submitted for this assignment is my own and research sources are fully acknowledged.

Student's signature:

Zhe Yuan

Submission Date: 24/11/2023

TurnItIn Similarity Report



SysFundAsgn2

ORIGINALITY REPORT

5%	2%	1%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Queen's College Student Paper	2%
2	Submitted to Toronto Business College Student Paper	1%
3	www.clevertek.in Internet Source	1%
4	secretheart.nginx.com Internet Source	1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

Table of Contents

Main Report

Part 1: Content Delivery Network (CDN)

A content delivery network or content distribution network (CDN) is a cluster of servers that are geographically distributed between end users and host servers around the globe to provide fast, available content delivery (Nanwani, 2021). CDNs enable quicker content delivery by minimizing the travel distance between the content and the user, and this is achieved by utilizing points of presence (PoP), which are edge servers located near to end users (Baeldung, 2023). PoPs replace long distance user requests to the original host server and help cache user requested content from the host servers, so end users can request for content that will be delivered from edge servers instead, which return faster response times (Nanwani, 2021). Reducing travel distance or hops for content data also help minimize packet losses, jitter, latency, and optimize network bandwidths (Yasar, 2023). Among the benefits that CDN offers, increased content availability, website security, faster load times, and hosting cost savings are some of the more prominent ones (Nanwani, 2021; Baeldung, 2023; Yasar, 2023) when CDNs act as proxy. According to Baeldung (2023), there are push CDNs that require web owners to manage and update PoP caches themselves instead of the end user, and pull CDNs that retrieve origin content on a user request to cache it for subsequent requests.

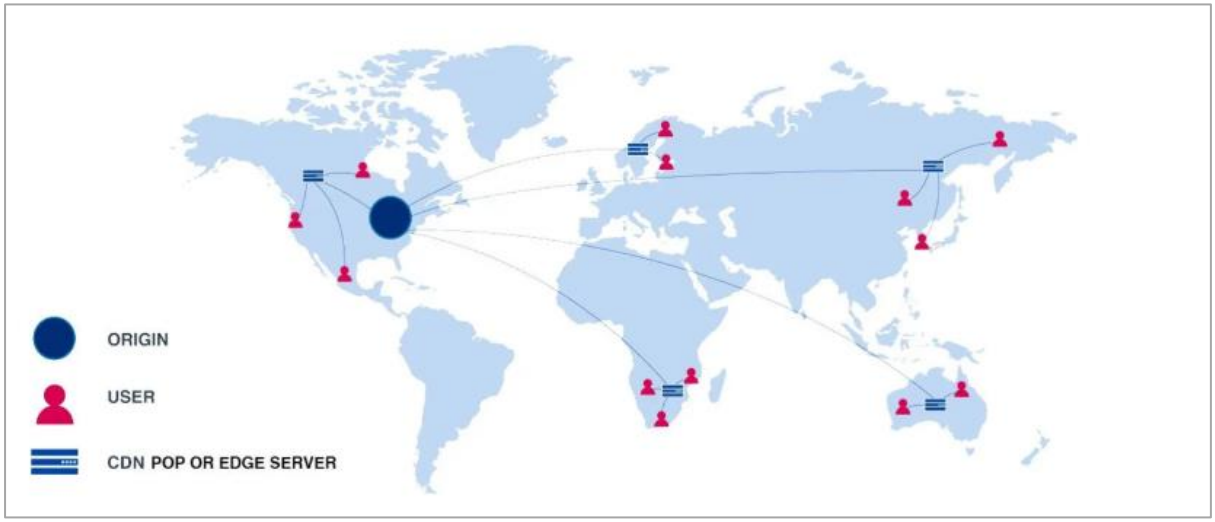


***Figure 1**: Visualization of a CDN (Nanwani, 2021)*

To elaborate on the benefits of CDN, increased content availability can be achieved due to the geographical redundancy of PoPs and their caching mechanism. These characteristics allow content to remain accessible even if a fatal failure has occurred in an edge server or the original host. Security is also ensured by using CDN because they are capable of handling user traffic and distributing flood of requests to multiple edges to prevent server overloads, which help reduce the probability of a DDOS attack. Web contents are able to load faster as well due to rapid responses that traveled a shorter distance between the client and edge server cache than the distance from client to origin host, which improves usability and user experience of an application. Lastly, it was said that hosting costs can be reduced because less data bandwidth were used by the CDN to transfer the content to end users. Hosting providers normally charge based on the amount of data transferred to and from the origin server, however clients only interact with the edge server to retrieve hosted content from the cache instead and rarely make new requests to the origin server when using CDNs (Nanwani, 2021; Baeldung, 2023; Yasar, 2023).

An example of a CDN provider is Cloudflare. Cloudflare (2023a) describes itself as a "large network of servers that improve the security, performance, and reliability" of Internet entities by serving as a reverse proxy for client web traffic. With Cloudflare, specific, interchangeable Cloudflare Anycast IPs will be assigned to origin server domains dynamically and DNS lookups for applications resolve into these IPs instead of the origin server's IP directly, which allows requests that were sent to the proxied hostnames to go to Cloudflare first before being forwarded to the origin host. Consequently, packet transmissions through Cloudflare's CDN makes it possible for users to configure additional rules and optimizations to further improve data flow (Cloudflare, 2023a). This is evident as Cloudflare does not only offers their CDN to customers, but also provide cloud solutions that are inclusive of vast business scenarios, such as application, developer, authentication, and network services that users can setup and use for their Cloudflare applications (Cloudflare, n.d.-b).

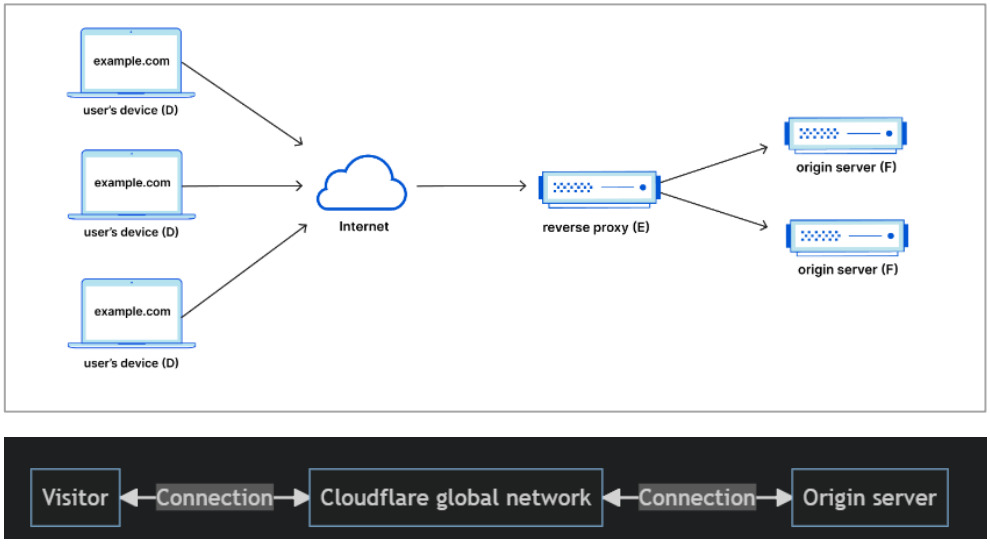


***Figure 2**: Cloudflare network architecture (Cloudflare, 2023a; Cloudflare, n.d.-c)*

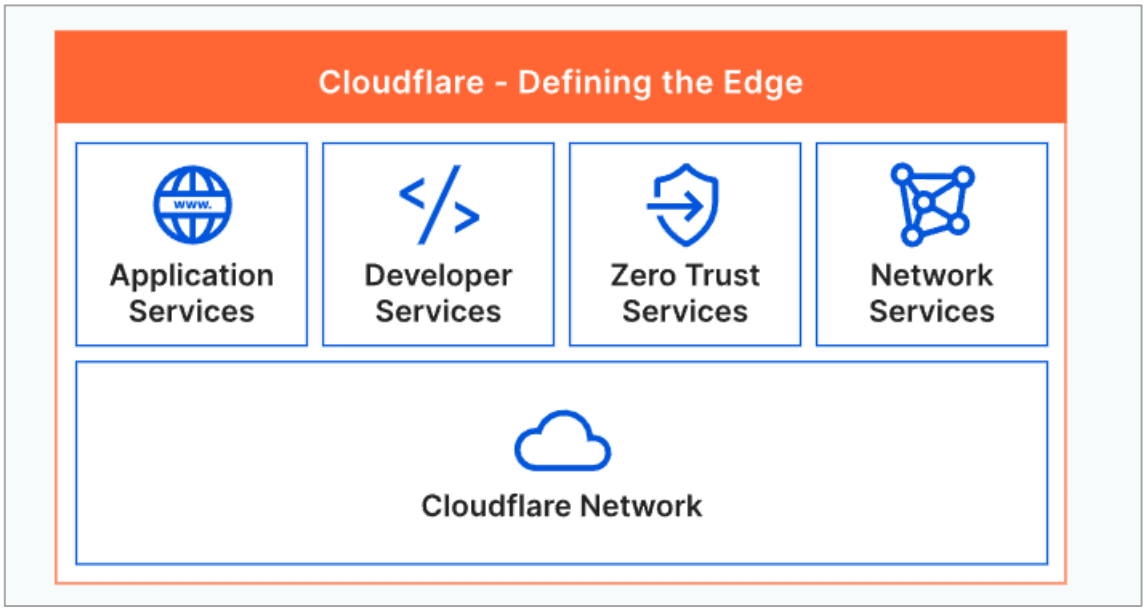


***Figure 3**: Services offered by Cloudflare (Cloudflare, n.d.-b)*

Part 2: Load Balancer

Load balancers are devices that deal with load balancing, which is the process of distributing incoming network traffic efficiently across a group of servers known as a server farm or server pool (NGINX, n.d.-a). According to a NGINX (n.d.-a) article, modern high-traffic applications generally deploy more servers to scale up cost-effectively, hence why load balancers are needed to serve millions of concurrent requests in a rapid and reliable manner by utilizing these added servers. Load balancers act as traffic managers by routing client requests across the available server pool using algorithms that were designed to maximize speed and capacity utilization, while preventing servers from being overwhelmed with unbalanced loads. They are responsible for the dynamic request distributions that scale server usages accordingly when traffic fluctuations are detected. They also provide a failover mechanism that redirects pending requests from downed servers to other server backups (Yasar & Irei, 2023; NGINX, n.d.-a). Load balancers perform health checks to servers to determine their availability. They track the actual number of server instances in the server pool that are available to use for load balancing and identify the server that is best suited for a client request in real time (ProgressKemp, n.d.; Yasar & Irei, 2023).

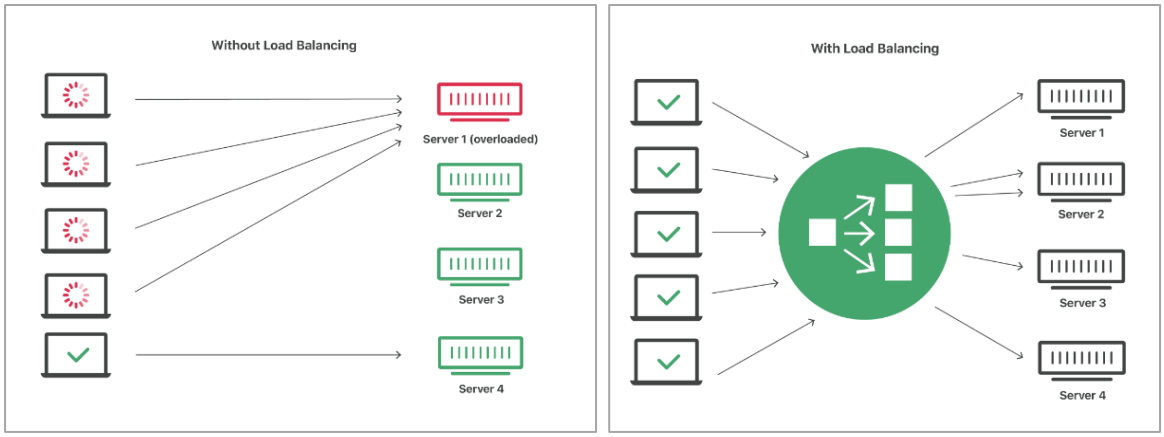


***Figure 4**: Visualization of load balancer (Cloudflare, n.d.-d)*

Load balancers can come as a hardware or software appliance, or a combination of both. A hardware load balancer is a hardware device with specialized and proprietary built-in software designed to handle massive amounts of application traffic, while a software load balancer runs on virtual machines (VMs), cloud environments, or white box servers, most likely as an application delivery controller (ADC) function that allows the user to configure it virtually. In terms of algorithms, load balancers make use of dynamic and static load balancing algorithms to determine server reroutes. Dynamic load balancing algorithms consider the current state of each server to decide on a distribution strategy, while static load balancing algorithms do not put server states into consideration and reroute request traffic based on a predetermined strategy (Yasar & Irei, 2023).

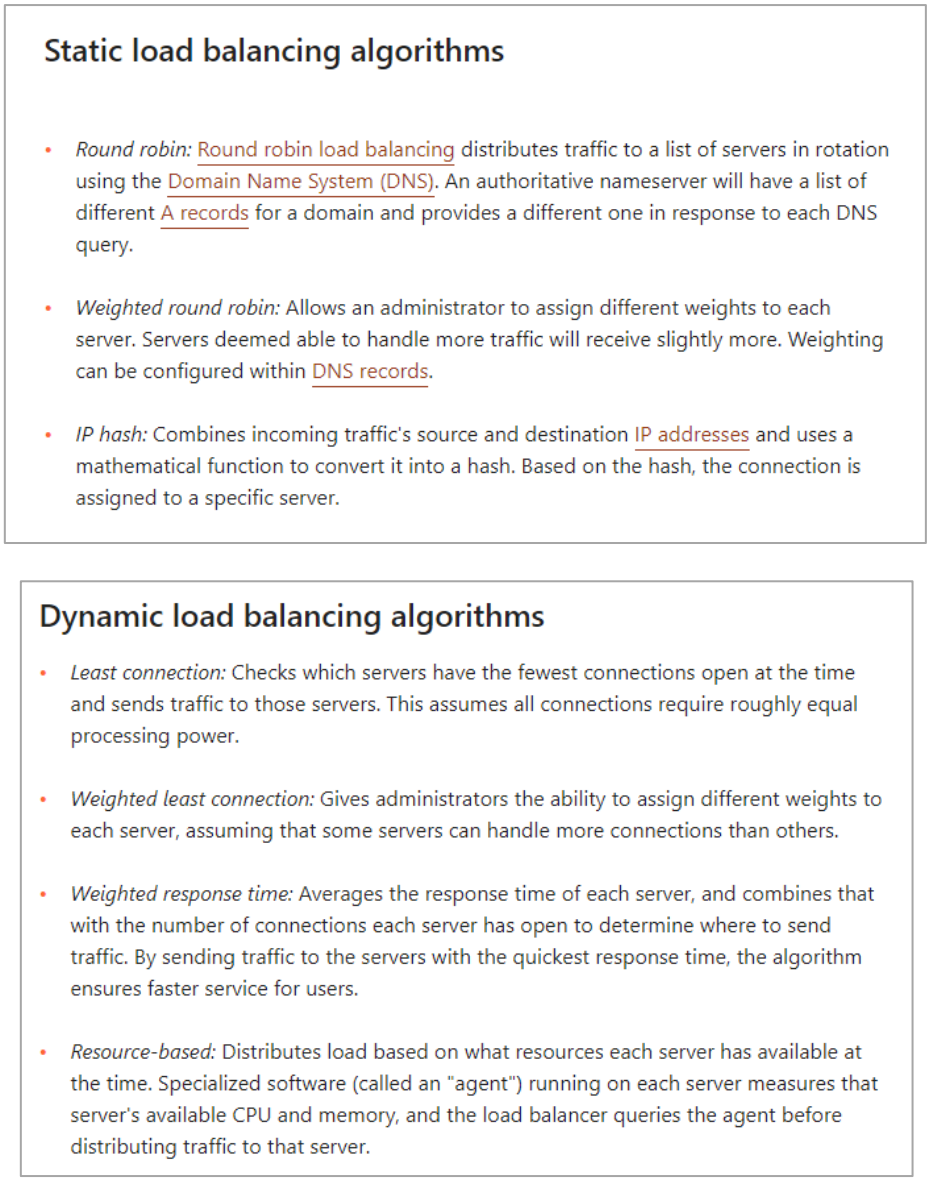


***Figure 5**: Dynamic and static load balancer algorithms (Cloudflare, n.d.-e)*

An example of a load balancer is NGINX Plus. Built on top of its predecessor NGINX Open Source, NGINX Plus is an extended, proprietary solution with advanced functionality and compatibility, integrating load balancer, content cache, web server, security controls, and application monitoring into one feasible software package (NGINX, n.d.-b). In terms of load balancing capabilities, NGINX Plus functions exceptionally well with cross-core synchronization, multiple session persistence methods to refine load balancing decisions, health checks to identify failed servers, and slow-start to reintroduce them (NGINX, n.d.-c). It enables global server load balancing and load balances not only HTTP, but also TCP, UDP and gRPC traffic (NGINX, n.d.-d). When load balancing HTTPs traffic, NGINX Plus terminates each HTTPs connection and processes each request individually. SSL/TLS encryption is stripped to relieve application of the computational load of the encryption (NGINX, n.d.-e), and requests can be inspected, manipulated, queued with rate limits, and selected by a load-balancing policy. For TCP and UDP applications like MySQL, DNS and RADIUS, NGINX Plus terminates these types of connections and creates new connections from itself to the backend (NGINX, n.d.-d).

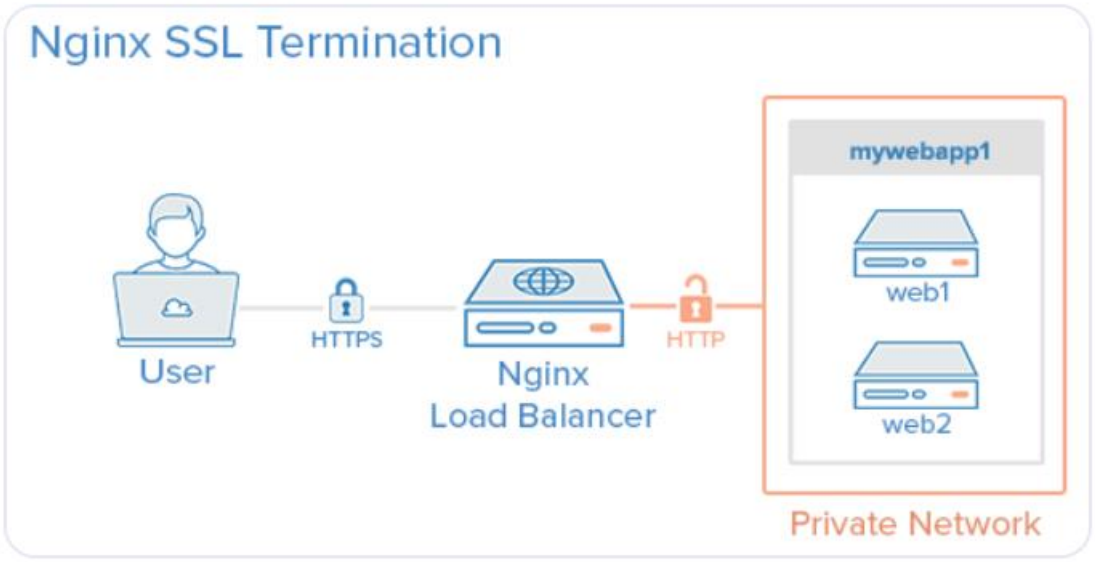


***Figure 6**: Visualization of NGINX SSL termination to relieve server computational stress (Jesin, 2014).*

Part 3: Comparison and Challenges of CDN and Load Balancer

It is important to differentiate between load balancers and CDNs as they share similar characteristics and benefits. Firstly, both technologies can act as a reverse proxy, which sits between the client and affiliated origins, thus also enabling middleware to be configured for preprocessing client requests before handing them over to server processes (ProgressKemp, n.d.). Next, server redundancy, availability, security, and efficiency are recurring benefits that were mentioned for both technologies as well (NGINX, n.d.-a, Yasar & Irei, 2023). This is because both aim to refrain clients from imposing stress onto origin hosts individually, which bogs down their performance. Ultimately, the fundamental factor that distinguishes both technologies is the difference in purpose. The main purpose of using CDNs is to distribute content across a wide geographic area, while the purpose of load balancers is to manage request traffic across a server pool, where servers are within close geographic proximity (Watts, 2022). Load balancers can be an added solution to CDNs to perform load balancing in the network of PoPs to prevent them from overworking too (Cloudflare, n.d.-c).

Contemporarily, CDNs and load balancers are still facing numerous obstacles that prevent them from being entirely reliable. The following tables discuss more about these challenges of implementing CDNs and load balancers.

Challenge	Elaboration	Proposed Solution
Caching	Caching introduces inconsistencies to CDNs which escalates into other risks due to the possibility of outdated content, cache poisoning, and cache invalidation (Front-end Development, 2023).	Caching strategies, such as setting appropriate cache headers, using cache busting techniques, updating caches on intervals, and testing the cache behavior, should be determined (Front-end Development, 2023).
Latency	Factors such as network conditions, node number and locations, content types and sizes, and client device used all affect the latency of a content delivery (Front-end Development, 2023).	Constant monitoring of CDN performances, use tools like *ping*, *traceroute*, and *webpagetest*, and apply web performance best practices, such as minifying, compressing, and bundling content (Front-end Development, 2023).
Bandwidth	Streaming data from CDNs on demand, usually videos, may pose difficulties in sustaining transmission speeds due to different mobile, fixed, and wireless device bandwidths, which affect quality of experiences (QoE) (Thibeault, 2020; Elkotob & Andersson, 2012).	Perform chunking of video content packets to turn them into smaller sizes that are more compatible with most transmission channels.

Interoperability	Different CDN networks may adopt varying infrastructure and distribution frameworks, causing a lack of interoperability (Front-end Development, 2023; Thibeault, 2020).	Ensure that CDN contains external support or internally supports the required features and standards, use fallbacks and polyfills, and test integrated network environments manually (Front-end Development, 2023).
Cost	CDN providers charge customers based on their CDN usages, such as traffic volume, regions, and features used (Front-end Development, 2023).	Understand pricing model and the billing details of the CDN provider, estimate total traffic and usage patterns, and compare different CDN options and plans to reconsider CDN spendings based on actual usage (Front-end Development, 2023).

Table 1: *CDN challenges*

Challenge	Elaboration	Solution
Latency	The distance between distributed nodes, communication delays, network delays affect process speeds due to the scattered nature of geographically distributed nodes (Zenarmor, n.d.).	Close the distance of connected nodes until the best latency can be achieved from using them. Else, a node selection algorithm that reduce delays between nodes need to be employed within the load balancers.
Scalability	It is difficult to determine on a dynamic scale range to accommodate changing consumer demands, compute requirements, storage needs, and system capabilities (Zenarmor, n.d.).	Study peak time usages and consumptions to identify the minimum and maximum boundary of scale needed to accommodate most use cases.
Interoperability	Varying customer needs are normally implemented on heterogeneous nodes in real world business scenarios, and the lack of algorithms for resolving diverse context causes load balancers to lack cross-platform compatibility (Zenarmor, n.d.).	Employ middleware between load balancer and other nodes in order to interpret and convert communication into a compatible form.
Migration	Certain instances, like VMs, live migrate to another location if the original host is overcrowded, and network delays occur when load balancing large VM images over low bandwidths (Zenarmor, n.d.). It also increases service level violations, task rejection ratio, and power consumption (Afzal & Kavitha, 2019).	Use filtering and reduction techniques between load balancers to handle large workloads.

Algorithm complexity	Determining an algorithm can be hard as its complexity is a contributing factor of overall performance. Each algorithm contains performance tradeoffs in terms of simplicity, migration time, fault tolerance, and reaction time (Zenarmor, n.d.).	Compare algorithm tradeoffs with personal use cases to determine best fit, shifting the criteria for choosing an algorithm from complexity to maximizing compensation.

Table 2: *Load balancer challenges*

To summarize, CDNs and load balancers share similar challenges in terms of latency, bandwidth, and interoperability, which are common challenges of distributed systems where nodes communicate with one another over distances (GeeksforGeeks, 2023).

References

Afzal, S., & Kavitha, G. (2019) Load balancing in cloud computing – A hierarchical taxonomical classification. *Journal of Cloud Computing*, *8*(22). Available at: https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-019-0146-7 [accessed 26 October 2023].

Baeldung (2023) *What Is a Content Delivery Network (CDN)?* Available at: https://www.baeldung.com/cs/cdn [accessed 23 October 2023].

Cloudflare (2023a) *How Cloudflare works*. Available at: https://developers.cloudflare.com/fundamentals/concepts/how-cloudflare-works/ [accessed 23 October 2023].

Cloudflare (n.d.-b) *What is Cloudflare?* Available at: https://www.cloudflare.com/what-is-cloudflare/ [accessed 23 October 2023].

Cloudflare (n.d.-c) *What is a reverse proxy? | Proxy servers explained*. Available at: https://www.cloudflare.com/learning/cdn/glossary/reverse-proxy/#:~:text=A%20reverse%20proxy%20is%20a,security%2C%20performance%2C%20and%20reliability. [accessed 23 October 2023].

Cloudflare (n.d.-d) *What is load balancing? | How load balancers work*. Available at: https://www.cloudflare.com/learning/performance/what-is-load-balancing/ [accessed 24 October 2023].

Cloudflare (n.d.-e) *Types of load balancing algorithms*. Available at: https://www.cloudflare.com/learning/performance/types-of-load-balancing-algorithms/ [accessed 24 October 2023].

Elkotob, M., & Andersson, K. (2012) Challenges and opportunities in content distribution networks: A case study. *Globecom 2012 Workshops: The 4th IEEE International Workshop on Mobility Management in the Networks of the Future World*, 1021-1026. Available at: https://www.researchgate.net/publication/261242108_Challenges_and_opportunities_in_content_distribution_networks_A_case_study [accessed 25 October 2023].

Front-end Development (2023) *How can you overcome the challenges of using a CDN?* LinkedIn. Available at: https://www.linkedin.com/advice/0/how-can-you-overcome-challenges-using-cdn#:~:text=CDNs%20can%20improve%20web%20performance,latency%2C%20compatibility%2C%20and%20cost [accessed 25 October 2023].

GeeksforGeeks (2023) *Advantages and Disadvantages Of Distributed Systems*. Available at: https://www.geeksforgeeks.org/advantages-and-disadvantages-of-distributed-systems/ [accessed 25 October 2023].

Jesin, A. (2014) *How To Set Up Nginx Load Balancing with SSL Termination*. DigitalOcean. Available at: https://www.digitalocean.com/community/tutorials/how-to-set-up-nginx-load-balancing-with-ssl-termination [accessed 24 October 2023].

Nanwani, R. (2021) *What is a Content Delivery Network (CDN) - A Beginner's Guide*. ImageKit. Available at: https://imagekit.io/blog/what-is-content-delivery-network-cdn-guide/ [accessed 23 October 2023].

NGINX (n.d.-a) *What Is Load Balancing?* Available at: https://www.nginx.com/resources/glossary/load-balancing/ [accessed 24 October 2023].

NGINX (n.d.-b) *NGINX Plus*. Available at: https://www.nginx.com/resources/datasheets/nginx-plus-datasheet/ [accessed 24 October 2023].

NGINX (n.d.-c) *Frequently Asked Questions*. Available at: https://www.nginx.com/resources/faq/ [accessed 24 October 2023].

NGINX (n.d.-d) *F5 NGINX Plus: Load Balancing*. Available at: https://www.nginx.com/products/nginx/load-balancing/#:~:text=TCP%20and%20UDP%20Load%20Balancing&text=For%20TCP%20applications%2C%20NGINX%20Plus,new%20connections%20to%20the%20backend.&text=The%20configuration%20is%20similar%20to,request%20to%20an%20upstream%20group. [accessed 24 October 2023].

NGINX (n.d.-e) *NGINX SSL Termination*. Available at: https://docs.nginx.com/nginx/admin-guide/security-controls/terminating-ssl-http/ [accessed 24 October 2023].

ProgressKemp (n.d.) *What is load balancing?* Available at: https://kemptechnologies.com/what-is-load-balancing [accessed 24 October 2023].

Thibeault, J. (2020) *Key CDN Challenges*. Streaming Media. Available at: https://www.streamingmedia.com/Articles/ReadArticle.aspx?ArticleID=144171 [accessed 25 October 2023].

Watts, S. (2022) *Content Delivery Networks (CDNs) vs. Load Balancers: What's The Difference?* Splunk. Available at: https://www.splunk.com/en_us/blog/learn/cdns-vs-load-balancers.html#:~:text=At%20the%20end%20of%20the,geographic%20proximity%20to%20each%20other. [accessed 25 October 2023].

Yasar, K. (2023) *CDN (content delivery network)*. TechTarget. Available at: https://www.techtarget.com/searchnetworking/definition/CDN-content-delivery-network#:~:text=The%20following%20is%20an%20example,response%20to%20the%20website%20visitor. [accessed 23 October 2023].

Yasar, K., & Irei, A. (2023) *load balancing*. TechTarget. Available at: https://www.techtarget.com/searchnetworking/definition/load-balancing [accessed 24 October 2023].

Zenarmor (n.d.) *Load Balancing. Benefits, Challenges, Types, Algorithms, and Tools*. Available at: https://www.zenarmor.com/docs/network-security-tutorials/what-is-load-balancing#what-are-the-challenges-when-implementing-load-balancing [accessed 26 October 2023].

COS3043/N System Fundamentals Marking Rubric
ASSIGNMENT [2] Individual Report Writing (Weighted Marks: 30%)

REPORT COMPONENT (100%)

LEARNING OUTCOME	MARKING CRITERIA	SCALE					
		Failed (0% to 49%)	**3rd class (50% to 59%)**	**2nd lower (60% to 69%)**	**2nd upper (70% to 79%)**	**1st class (80% to 100%)**	**YOUR MARKS/COMMENTS**
CLO 2: [Evaluate software design issues for advanced computer systems such as multiprocessors or distributed systems] **CLO 3: [Discuss current topics in system fundamentals research by reading and analysis of journal papers.]**	**1. The depth of descriptive writing in regard to CDN (30%)**	No descriptive writing approach (fact reporting, vague impressions) with lack of detailed development. Ideas are vague with little evidence of critical thinking related to CDN.	Superficial descriptive writing approach (fact reporting, vague impressions) without reflection or introspection. Main points lack detailed development. Ideas are vague with little evidence of critical thinking related to CDN.	Elaborated descriptive writing approach and impressions without reflection. Writing indicates thinking and reasoning. Main points are present with limited detail and development. Some critical thinking is present related to CDN.	Movement beyond reporting or descriptive writing to reflecting (i.e., attempting to understand, question, or analyse the event). Main points are well developed with quality supporting detail and quantity related to CDN.	Exploration and critique of assumptions, values, beliefs, and/or biases, and the consequences of action (present and future). Main points are well developed with high quality and quantity support related to CDN.	
	2. The depth of descriptive writing in regard to Load Balancer (30%)	No descriptive writing approach (fact reporting, vague impressions) with lack of detailed development. Ideas are vague with little evidence of critical thinking related to Load Balancer.	Superficial descriptive writing approach (fact reporting, vague impressions) without reflection or introspection. Main points lack detailed development. Ideas are vague with little evidence of critical thinking related to Load Balancer.	Elaborated descriptive writing approach and impressions without reflection. Writing indicates thinking and reasoning. Main points are present with limited detail and development. Some critical thinking is present related to Load Balancer.	Movement beyond reporting or descriptive writing to reflecting (i.e., attempting to understand, question, or analyse the event). Main points well developed with quality supporting detail and quantity related to Load Balancer.	Exploration and critique of assumptions, values, beliefs, and/or biases, and the consequences of action (present and future). Main points well develop with high quality and quantity support related to Load Balancer.	
	3. The depth of analysis writing on Challenges of CDN and Load Balancer (30%)	No analysis or meaning-making related to Challenges of CDN and Load Balancer.	No analysis or meaning-making. Shows some thinking and reasoning but most ideas are underdeveloped and unoriginal related to Challenges of CDN and Load Balancer.	Little or unclear analysis or meaning-making. Analysis indicates thinking and reasoning applied with an original thought on a few ideas related to Challenges of CDN and Load Balancer.	Some analysis and meaning-making. Critical thinking is weaved into points. Analysis indicates original thinking and develops ideas with sufficient and firm evidence related to Challenges of CDN and Load Balancer.	Comprehensive analysis and meaning-making. Reveals high degree of critical thinking. Analysis indicates synthesis of ideas, in-depth analysis and evidence of original thought and support for the topic related to Challenges of CDN and Load Balancer.	
	4. References, Sources &Citation (10%)	No references. Some sources are not accurately documented. Diagrams and illustrations are not accurate OR do not add to the reader's understanding of the topic. Missing or no citation and major flaws in the format.	References are insufficient and outdated. All sources (information and graphics) are accurately documented, but many are not in the desired format. Some diagrams and illustrations are not accurate OR do not add to the reader's understanding of the topic A very minimal amount of cited works, with an incorrect format.	Reference is sufficient but not recent. All sources (information and graphics) are documented, but an adequate amount is not in the desired format. Diagrams and illustrations are neat and accurate and sometimes add to the reader's understanding of the topic. An adequate amount of cited works, both text and visual, are done in the correct format. Inconsistencies evident.	Some references are appropriate and recent. All sources (information and graphics) are accurately documented, but a few are not in the desired format. Diagrams and illustrations are accurate and add to the reader's understanding of the topic. All, both text and visual, are done with minimal errors in the format.	All references are appropriate and recent. All sources (information and graphics) are accurately documented in the desired format. Diagrams and illustrations are neat and accurate, and they add to the reader's understanding of the topic. All cited works, both text and visual, are done in the correct format with no errors.	
						Total (100%)	