Threats and Security Issues in Cloud Storage and Content Delivery Networks: Analysis

Jean Daniel Kouam Waguia, Alexander Menshchikov
ITMO University
Saint-Petersburg, Russia
ledanielin@gmail.com, menshikov@itmo.ru

Abstract–**The rapid growth of technology tends to improve human's life, to save time and to impose a rhythm to human day to day activity. More than ever, the world is facing tremendous difficulties to acquire, ingest, process, save and visualize information. This is how big companies like AWS Amazon, Microsoft Azure, Google Cloud and Alibaba are coming into the game by providing cloud computing to remediate to such difficulties. To help companies, organizations, corporates and users to store huge amounts of data, Cloud Service Providers provide storage in their data centers and even Content Delivery Network service to keep data to the closest location of users. Among those companies, some are using cloud storage as Content Delivery Network. Beside Cloud Service Providers, other big companies like Akamai or Cloudflare are specialized in Content Delivery Network. Therefore, which kind of threats are facing companies which use cloud storage as Content Delivery Network? Are those threats similar to companies specialized only on Content Delivery Networks? And what can be the impact of Content Delivery Network attacks on cloud storage used as Content Delivery Networks? The aim of this paper is to propose a new threat model of cloud storage used as Cloud Delivery Network and determine the level of risk on this last one.**

I. Introduction

Cloud computing can be defined in many ways. There is no universal definition for it. NIST's (National Institute of Standards and Technology) definition of cloud computing is considered as the de facto definition. NIST defines cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Nowadays, various companies are still using several types of devices to store, process, analyze, visualize and archive their data.

The apparition of cloud services incited many organizations to move their infrastructure in the cloud, thus their data. The flexibility and elasticity of cloud services gives organizations the ability to centralize and store their data into the cloud usually called Cloud Storage (CS) for any kind of operation. On CS are stored sensitive and critical data for archive purpose or for further exploitation. The average business uses 2,145 cloud services and apps, and the most popular app categories are cloud storage, collaboration, Web mail, consumer, and social media. The most popular apps overall are Google Drive, YouTube, Microsoft Office 365 for Business, Facebook, Gmail, SharePoint, Outlook, Twitter, Amazon S3, and LinkedIn [1]. Attackers are targeting popular cloud applications and services to exploit the growing trust in commonly used enterprise platforms. This clearly indicates the need to raise more awareness of consumers and Cloud Service Providers (CSP).

Today more embedded devices joined the Internet to monitor and connect everything (traffic facilities, buildings, environment, and lakes), enlarging the size of the data generation. All those data are stored in CS as data sets that include heterogeneous formats: structured, unstructured and semi-structured data which refers to Big Data [4]. Which is one of the CS use cases. Another important use case of CS is Content Delivery Network (CDN).

CDN usually deploys edge servers or surrogates in different geographical locations that are often across the global Internet backbone, working as a distributed network service with a significant capacity of both computational resources and network bandwidth. For example: on Google Cloud Platform, information received by Internet of Things (IoT) devices are stored in CS, then they can be analyzed, processed and visualized using various tools leading to Big Data. Furthermore, in that same environment, consumers can host their websites and applications which can use a global load balancer technology which gives the ability to use CDN in order to cache content to the closest location of web users. A correlation resulting from it is that, CS used for CDN can also be used for Big Data, which may result in a catastrophic data breach or data loss.

Several contributions have been proposed to address the problem such as industry best practices and standards, security and accountability audits, Service Level Agreement management, security incident management, virtual machine introspection, development of a cloud framework for security audit [11]. For example, Cloud Storage Brokers provide seamless and concurrent access to multiple Cloud Storage Services while abstracting cloud complexities from end-users [3]. However, these approaches have not considered that CS essentially works as CDN [6], which means if an attacker is able to successfully perform an attack in the CS and gains access to the data storage, then he can modify files cached by the CDN and launch other attacks like website defacement or Cache Poisoned Denial of Service which may lead to a huge loss of money. Therefore we propose to perform an analysis of security threats that CDN and CS can face.

We summarize our contributions as follows:

- We show issues and challenges that CS and CDN are facing.
- We formulate a threat model of CS used as CDN which involves CDN and CS.
- We raise the awareness of CSP, CDN providers (CDNP) and consumers of those specific services.

The rest of this paper is structured as follows, a presentation of commonly used architectures in Section II, followed by a presentation of related works in Section III. In Section IV, we perform an analysis of security issues or threats of CS and CDN. A threat model of CS integrated with CDN, in other words CS used as CDN to determine the level of risk in case the threat occurs in Section V will be done and the paper is concluded in Section VI.

II. Presentation of commonly used architectures

Cloud storage is a model of networked online storage where data is stored on multiple virtual servers. It uses virtualization technology to aggregate resources of various servers to create pools of storage that can be easily replicated, provisioned or deprovisioned and resized. It's the fundamental technology which gives the cloud the property of elasticity.

An origin server is a computer that hosts the original version of web files. The purpose of an origin server is to process and respond to incoming internet requests from internet clients.

An edge server or cache server is a computer that can cache content retrieved from an origin server.

Web caching is the activity of storing data for reuse, such as a copy of a web page served by a web server. It is cached or stored the first time a user visits the page and the next time a user requests the same page, a cache will serve the copy, which helps keep the origin server from getting overloaded. Some benefits from web caching are:

- Decreased network costs: content can be cached at various points in the network path between the content consumer and content origin. When the content is cached closer to the consumer, requests will not cause much additional network activity beyond the cache.
- Improved responsiveness: caching enables content to be retrieved faster because an entire network round trip is not necessary. Caches maintained close to the user, like the browser cache, can make this retrieval nearly instantaneous.
- Increased performance on the same hardware: for the server where the content originated, more performance can be squeezed from the same hardware by allowing aggressive caching. The content owner can leverage the powerful servers along the delivery path to take the brunt of certain content loads.
- Availability of content during network interruptions: with certain policies, caching can be used to serve content to end users even when it may be unavailable for short periods of time from the origin servers.

Web caching implements its performance by Cache servers. When a web page has been requested by a user, that one is immediately cached by the server. When this page is once more requested, the page cached is served to the user instead of the original one, because cache servers are usually deployed near users.

CDN act as an overlay network that offers high-performance delivery of common web objects, static data, and rich multimedia content by distributing content load among servers that are close to the users [7].

Commonly, end users are not most of time aware of the fact that they can use CDN either a free or a paid version. Therefore, most web application deployment are not behind a CDN. Which means, each request sent to the application hits directly the origin server and sometimes huge amount of requests may increase server's response time or loading page time. For example if a user is located in Africa and the web server or the application is hosted in America, then it will take more time for that user to see a fully loaded page of the website, than a user in America.

Fig. 1 presents an example of architecture without CDN. In this case, the server's time response for a user in Africa is 3 seconds, meanwhile it is 0.5 second for a user in America and 2 seconds for a user in Europe. It simply means that, the closer is a user to a location, the faster he can retrieve files from the hosting or the CSP. In some cases, server's time response can be even greater when the server is loaded and worse if the server is under a Denial of Service attack (DoS) or a Distributed DoS (DDoS). In the last case, the server is no more responding, because of requests' overflow and that's why CDN is important.

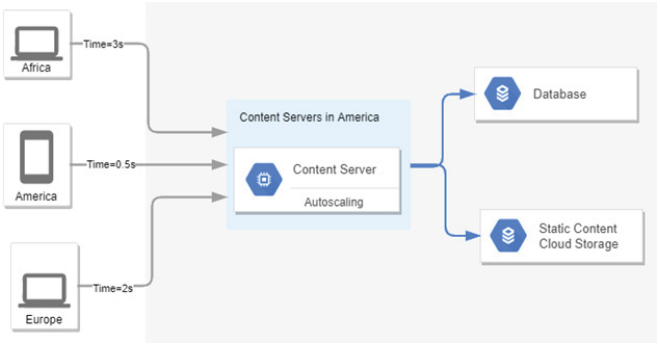

Fig. 1. Architecture a site without Content Delivery Network

Some major advantages of a CDN are: DDoS mitigation attack, very short response time of a page or file already cached and the ability to retrieve some static files of a website even if this one is down. Fig. 2 presents an example of architecture with CDN. In this case, the server's time response of each user is the same, because website's files have been cached by the CDN. In this case, even if there is a DDoS attack from America, users from Africa or Europe will not be affected, because the CDN will absorb via servers located in America.

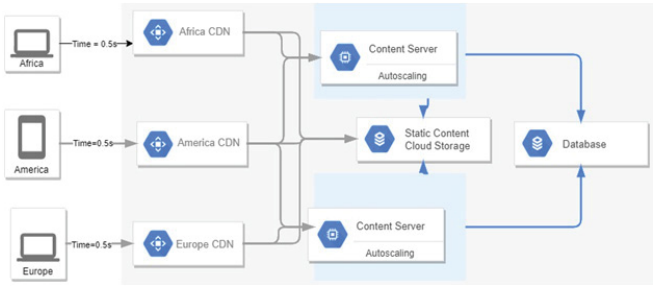

Fig. 2. Architecture of a site with Content Delivery Network

III. Related works

CDN security and CS security involve network and information security. CS a type data storage, which uses for example Storage Area Network (SAN) to host various type of data. A small analysis can be done between CDN and SAN (which can refer to CS or Data Storage in this case):

- In CDN, C goes for Content and S in SAN goes for Storage. Storages are used to keep content, which involves information security.
- D of CDN stands for Delivery and A of SAN stands for Area, which may imply the delivery in a location.
- Lastly, N of CDN and SAN means Network, which implies network security.

In information and network security, few areas have been explored with regard to CDN. The author of reference [8] mentioned relevant research works such as:

- Ownership Claiming: resource claiming as well as copyright and content claiming.
- Intrusion Detection: anomaly detection.
- Penetration Testing: packet eavesdropping and censorship evasion.
- Privacy: multi-CDN an approach to improve privacy.
- Fault-Tolerance: framework to reduce stall times through controlling affecting parameters such as caching space and bandwidth allocation.

In cloud storage, some research works related to security are presented as follow:

- Usage of a cloud storage broker auditor to audit the entire cloud storage broker system for suspicious activities, unauthorized changes in the cloud storage broker cloud accounts and performing detailed risk analytics [3].
- Threat detection and incident response to detect potential CS threats and take some measures to detect an unauthorized user and mitigate attacks [13].
- Development of security framework to protect data in cloud storage Framework (CSSF) to support an integrative approach to understanding and evaluating security in cloud storage. The framework enables understanding of the makeup of security in cloud storage and measures the understanding of security in cloud storage [12].

IV. Threats

NIST Security Glossary defines threat as: any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service.

1) Cloud Storage Threats: The Cloud Security Alliance (CSA) in [14] describes eleven threats to cloud services in general. As cloud storage is a specific cloud service, these threats can be applied to CS. Therefore, the following threats are the top ones reported by CSA for the year 2019:

- Data breaches may occur due to the failure of using multifactor authentication and strong passwords, a lack of ongoing automated rotation of cryptographic keys, passwords and certificates and a lack of scalable Identity and Access Management (IAM) systems.
- Misconfiguration and Inadequate Change Control occurs when computing assets are set up incorrectly, often leaving them vulnerable to malicious activity.
- Lack of Cloud Security Architecture and Strategy can be seen by a poorly implementation of appropriate security architecture to withstand cyberattacks. This aspect is related to Shared Technology Vulnerability and Inadequate Cloud Planning/Design.
- Insufficient Identity, Credential, Access and Key Management include tools and policies that allow organizations to manage, monitor and secure access to valuable resources. An appropriate identity and access management is typically the first line of defense against data breaches and other attacks.
- Account Hijacking is a threat in which malicious attackers gain access to accounts and abuse accounts that are highly privileged or sensitive.
- Insider Threat: Carnegie Mellon Computer Emergency Response Team (CERT) defines an insider threat as, the potential for an individual who has or had authorized access to an organization's assets to use their access, either maliciously or unintentionally, to act in a way that could negatively affect the organization. It can be current or former employees, contractors, or other business partners that may intentionally exceed or misuse access to data storage in order to harm the user of the CS.
- Insecure Interfaces and Application Programming Interface (API): can be a serious threat because they can contain bugs, vulnerabilities and design failures.
- Weak Control Plane is system architect who is not in full control of the data infrastructure's logic, security and verification. In this scenario, he doesn't know the security configuration, how data flows and where architectural blind spots and weak points exist. These limitations could result in data corruption, unavailability, or leakage.
- Metastructure and Applistructure Failures: API calls disclose this information and the protections are incorporated in the metastructure layer for the CSP. The metastructure is considered the CSP/customer line of demarcation.
- Limited Cloud Usage Visibility occurs when an organization does not possess the ability to visualize and

analyze whether cloud service use within the organization is safe or malicious.
- Abuse and Nefarious Use of Cloud Services: malicious actors may leverage cloud computing resources to target users, organizations or other cloud providers or even host malwares on cloud services.

2) Content Delivery Network threats: Surrogates servers in CDN deployment can be used by an attacker as proxies to access some resources restricted, to perform some attacks like reconnaissance or DDoS. Therefore it is important for CDNP to be aware of such attacks and take measures to mitigate them. In the same vein, some threats can be enumerated but not as much as they can be found in CS:

- Outages of a provider or due to a congestion in a network [15].
- Internal threat due to misconfiguration of equipment by technical staff or bad software deployment [15]. For example, an outage occurred on Cloudflare network because, while working on an unrelated issue with a segment of the backbone from Newark to Chicago, the network engineering team updated the configuration on a router in Atlanta to alleviate congestion. That configuration contained an error that caused all traffic across the backbone to be sent to Atlanta. This quickly overwhelmed the Atlanta router and caused Cloudflare network locations connected to the backbone to fail. The affected locations were San Jose, Dallas, Seattle, Los Angeles, Chicago, Washington, Richmond, Newark, Atlanta, London, Amsterdam, Frankfurt, Paris, Stockholm, Moscow, Saint-Petersburg, São Paulo, Curitiba, and Porto Alegre. Other locations continued to operate normally.
- Malicious insider: some serious attacks like Request header overflow or Replay DoS (ReDOS) and Cached Poisoned Denial of Service (CPDoS) can be launched by attackers to simulate downtime of websites which use CDN [10].
- Law enforcement and legal aspect: some publishers may publish some content on their websites which are accepted in their country, but when they use a CDN, that content is published in various geographical locations over the world. If that content is censured in some countries, then the Government of that country may oblige the CDNP to give them all data that they have about the publisher or even to take down publisher's site.
- Dynamic Content Attacks: attackers develop day after days, methods and techniques to circumvent systems' security. One weak point discovered about CDN is that they can't handle dynamic requests [16]. It makes the treatment of dynamic content requests a significant blind spot in CDN services. Since the dynamic content is not stored on CDN servers, all the requests for dynamic content are sent to the origin's servers. Therefore, attackers are taking advantage of this behavior to generate attack traffic that contains random parameters in requests sent to CDN servers. Because those servers can't handle such requests, they immediately redirect those requests to the origin, expecting the origin's server to handle them. But, usually the origin's servers do not have enough resources to handle all those attack requests. Thus, they fail to provide online services to legitimate users, creating a DoS situation. Several CDNP provide the option to limit the number of dynamic requests sent to the origin server under attack. If that functionality is not properly configured, the CDN could even block legitimate users, because it cannot distinguish attackers from legitimate users and the rate limit will result in legitimate users being blocked.
- Secure Sockets Layer (SSL)-based DDoS attacks target the secured online services of the victim [16]. These attacks are easy to launch and difficult to mitigate, making them attackers' favorites. In order to detect and mitigate DDoS SSL attacks, CDN servers must first decrypt the traffic using the customer's SSL keys. If the customer is not willing to provide the SSL keys to its CDN provider, then the SSL attack traffic is redirected to the customer's origin, leaving the customer vulnerable to SSL attacks. SSL attacks that hit the customer's origin can easily take down the secured online service. During DDoS attacks when Web Application Firewall (WAF) technologies are involved, CDN networks also have a significant weakness in terms of the number of SSL connections per second from a scalability capability, and serious latency issues can arise.
- Compromising of third-party hosting facilities: CDN services are often offered only for Hypertext Transfer Protocol (HTTP), Hypertext Transfer Protocol Secure (HTTPS) and Domain Name Server (DNS) applications [16]. Other online services and applications in the customer's data center such as Voice over Internet Protocol (VoIP), mail, File Transfer Protocol (FTP) and proprietary protocols are not served by the CDN and therefore traffic to those applications is not routed through the CDN. In addition, many web-based applications are also not served by CDN. Attackers are taking advantage of this blind spot and launch attacks on applications that are not routed through the CDN, hitting the customer origin with largescale attacks that threaten to saturate the internet pipe of the customer. Once the internet pipe is saturated, all the applications at the customer's origin become unavailable to legitimate users, including the ones that are served by the CDN.
- Direct Internet Protocol (IP) assault: even applications that are served by a CDN can be attacked once attackers launch a direct attack on the IP address of the web servers at the customer origin [16]. These can be network based floods such as User Datagram Protocol (UDP) floods or Internet Control Message Protocol (ICMP) floods that will not be routed through CDN services, and will directly hit the servers of the customer at the origin. Such volumetric network attacks can saturate the internet pipe, resulting in taking down all the applications and the online services of the origin, including the ones that are served by the CDN. More often the misconfiguration of the firewall the data center can leave the applications directly vulnerable to attack.
- Compromising of web apps: CDN protection for web applications threats is limited and exposes the web applications of the customer to data leakage, data thefts and other threats that are common with web applications

[16]. Most CDN-based web application firewall capabilities are minimal, covering only a basic set of predefined signatures and rules. Many of the CDN-based WAF do not learn HTTP parameters, do not create positive security rules and therefore it cannot protect from zero day attacks and known threats.

In addition to the significant blind spots identified earlier, most CDN security services are not responsive enough, resulting in security configurations that take hours to manually deploy and to spread across all its network servers.

To our knowledge, there is no other paper, study, or publication performing a threat analysis of CS used as CDN. However, there are several papers, studies, and publications on security issues and threats to CS and specific CDN attacks and mitigations.

V. Threat Model of CS used as CDN

CS used as CDN offers some advantages and it could be assimilated to CDN Storage, nevertheless the attack surface is wider than the attack surface of a common CDN. Some use cases can be illustrated for this purpose. For example, when the CS can be used as CDN, that storage can be used as origin as well as cache, to serve static content. In another case, the storage can be used as CDN only or cache, when the origin is a dynamic website and only static files are hosted on the storage. In both cases, static content can be accessed via FTP, Secure FTP (SFTP) or a synchronization software. The usage of those protocols or applications expose already the CS or CDN to a brute force attack. This means that some attacks that could not be performed on a common CDN can already be performed on a CS used as CDN. As a result, threats that can occur on CS or CDN can occur on CS used as CDN. The third figure presents a simplified architecture of a CS used as CDN.

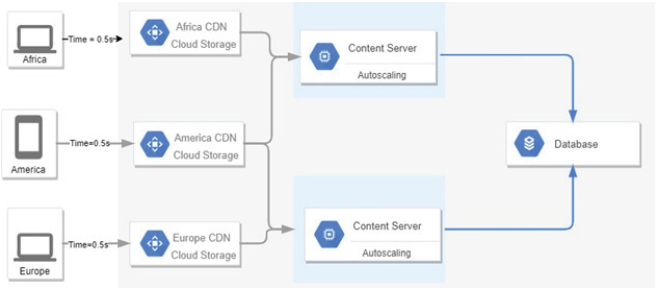

Fig. 3. Architecture of a Cloud Storage used as Content Delivery Network

The novel threat model is illustrated in Table I. It presents the level of risk that a threat of a common CDN could affect a CS used as CDN with the full assumption that all threats of a CS are also surely affecting a CS used as CDN. Each check box represents common CDN threats which affect cloud storage used as CDN and tend to increase risk level or the surface of attack of existing threats of cloud storage when it's not used as CDN.

For example:

- If the CDNP has an outage it may lead to a data unavailability. On the common CDN, cached files would have been lost or unavailable, but original data could still be found on the origin server and cached once more on CDN. In this case the risk level is low. But in the last architecture where CS is used as CDN, due to an outage, those data may have been recovered in the best of case, altered, or lost in a worse case. Thus, such loss can be evaluated to thousands or even millions of dollars, therefore it increases the risk level to high.

TABLE I. Matrix of common Content Delivery Network threats in Cloud Storage used as Content Delivery Network

CDN Threats Risk / Cloud Storage Threats	Outages of a Provider	Internal threat due to misconfiguration of equipment by technical staff	Internal threat due to a bad software deployment	Request header overflow	Law enforcement and Legal aspect	Denial of Service	Risk Level
Data breaches	·	☑	☑	·	☑	·	High
Data loss	☑	☑	☑	·	·	·	High
Insecure API	·	☑	☑	·	·	·	High
Account hijacking	·	☑	☑	☑	·	☑	High
Malicious insider	·	☑	☑	☑	·	☑	Medium
Denial of Service	·	☑	☑	☑	·	☑	High
Insufficiency Due Diligence	☑	☑	☑	·	·	·	High
Shared Technology Vulnerability	☑	☑	☑	·	·	·	High
Hardware failure	☑	☑	☑	·	·	·	Medium
Abuse of Cloud Storage	·	☑	☑	·	·	☑	High
Natural Disaster	☑	·	·	·	·	·	Low
Closure of the Cloud Provider	☑	☑	☑	·	☑	·	Low
Cloud Malware	·	☑	☑	·	·	·	High
Inadequate Cloud Planning/Design	·	☑	☑	☑	·	☑	High

- Internal threat due to misconfiguration of equipment by technical staff is a serious threat that could severely affect a CS used as CDN, because it may create a bottleneck in a whole network that can lead to disruption of services, data loss.
- Internal threat due to a bad software deployment: such threat can paralyze the CS used as CDN, create a disruption and make data unavailable to users.
- Law enforcement and legal aspect: it can affect CS used as CDN because a request from a Government to the CDNP to provide data of client could lead to a data breach and in the worst case to closure of the CDNP or even the company owning those data.
- Request header overflow, DoS: when a DoS attack like CPDoS is launched on the CDN, only cached files are touched and they can be replaced by original files after the attack. But if the same attack is launched on a CS used as CDN it will immediately affect original files and data integrity will be lost.

From the previous illustrations and from the table, CSP, CDNP, companies and users should be aware that there are more challenges to face when the CS is used as CDN. In most cases, if an attack succeeds, then all parties could be severely impacted. The CDNP or CSP could lost reputation, huge amount of money and even can bankrupt in case the biggest client decide to leave because he is not feeling safe after the attack or he may have lost also huge amount of money due to the attack and interruption of services.

As a result from this table, we note that there are 3 major threats from common CDN that could severely affect CS used as CDN: internal threat due to misconfiguration of an equipment, a bad software deployment and DoS.

VI. Conclusion

This paper has presented a novel threat model of a Cloud Storage used as a Content Delivery Network. It has been shown that the same threats that affect common Content Delivery Network can also affect Cloud Storage used as Content Delivery Network as illustrated in Table I by the matrix. Moreover those threats have high risk level. This work intend to raise awareness to providers, in order to take necessary measures to build more secure systems and improve the security level of such architecture. Because the usage of Cloud Storage as a Content Delivery Network may reduce network and traffic costs, but it increases the surface of attack, thus more security issues. This work can be extended to a risk assessment of Cloud Storage used as Content Delivery Network and also a performance comparison between common Content Delivery Network and Cloud Storage used as Content Delivery Network.

References

[1] Dark Reading official site, 44% of Security Threats Start in the Cloud, Web: https://www.darkreading.com/cloud/44--of-security-threats-start-in-the-cloud/d/d-id/1337088.

[2] NIST official site, The NIST Definition of Cloud Computing, Web: https://www.nist.gov/publications/nist-definition-cloud-computing.

[3] Torkura, Kennedy A., et al. "Csbauditor: Proactive security risk analysis for cloud storage broker systems." 2018 IEEE 17th International Symposium on Network Computing and Applications (NCA). IEEE, 2018.

[4] Oussous, Ahmed, et al. "Big Data technologies: A survey." Journal of King Saud University-Computer and Information Sciences 30.4 (2018): 431-448.

[5] Guo, Run, et al. "Abusing CDNs for fun and profit: Security issues in CDNs' origin validation." 2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS). IEEE, 2018.

[6] Google Inc, Serving files from Cloud Storage, Web: https://cloud.google.com/appengine/docs/flexible/python/serving-static-files#serving_files_from.

[7] Wang, Yuedui, et al. "The content delivery network system based on cloud storage." 2011 International Conference on Network Computing and Information Security. Vol. 1. IEEE, 2011.

[8] Zolfaghari, Behrouz, et al. "Content Delivery Networks: State of the Art, Trends, and Future Roadmap." ACM Computing Surveys (CSUR) 53.2 (2020): 1-34.

[9] Müller, Steffen. "Security trade-offs in Cloud storage systems." (2017).

[10] Nguyen, Hoai Viet, Luigi Lo Iacono, and Hannes Federrath. "Your cache has fallen: Cache-poisoned denial-of-service attack." Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019.

[11] Ismail, Umar Mukhtar, and Shareeful Islam. "A unified framework for cloud security transparency and audit." Journal of Information Security and Applications 54 (2020): 102594.

[12] Yahya, Farashazillah. A security framework to protect data in cloud storage. Diss. University of Southampton, 2017.

[13] Torkura, Kennedy A., et al. "SlingShot-Automated Threat Detection and Incident Response in Multi Cloud Storage Systems." 2019 IEEE 18th International Symposium on Network Computing and Applications (NCA). IEEE, 2019.

[14] Cloud Security Alliance official website, Cloud Security Challenges in 2020, Web: https://cloudsecurityalliance.org/blog/2020/02/18/cloud-security-challenges-in-2020/.

[15] Cloudflare official website, Post Mortem, Web: https://blog.cloudflare.com/tag/postmortem/.

[16] Inxy official website, Five main CDN security threats in 2020: forewarned is forearmed, Web: https://inxy.com/knowledge-base/five-main-cdn-security-threats-in-2020-forewarned-is-forearmed/.