

SERVERLESS APPLICATION: SECURE S3 FILE UPLOAD

¹Anushree Misra, ²Saloni Maheshwari, ³Sukriti Macker, ⁴Ruchika Doda

^{1, 2, 3} Student, ⁴ Assistant Professor

^{1, 4} Department of Electronics and Communication Engineering

² Department of Information Technology

³ Department of Computer Science

Mahavir Swami Institute of Technology, Sonipat, India

Abstract: *Cloud computing services are growing in the IT market. These services cost-effectively provide on-demand IT resources via the Internet. It facilitates the individual to pay-as-they-go rather than buying the entire service. Amazon Web Services (AWS) offer such cloud computing services. The authors of the paper used the AWS platform to create the Serverless Application to read and write a file. With various advantages offered by cloud computing services, such as cost-effectiveness and reliability, there is a significant concern that needs to be addressed, i.e., data security. The authors achieved the objective of the security of information and data by making the object private and the S3 bucket public. Thus, making the cloud hybrid and the application secure. The users can read and write their files securely and effectively in a specific period via this application.*

2. Platform as a Service (PaaS) - As the name suggests, it offers a platform to clients which consists of an operating system, database, web server. It authorizes them to deploy and manage business applications without the necessity of building and maintaining the infrastructure. The client does not control the cloud infrastructure consisting of networks, servers, operating systems, and storage.
3. Infrastructure as a Service (IaaS) - It is a form of cloud computing service that delivers basic building blocks of cloud IT. It is a dynamic and flexible self-service that accesses and monitors infrastructure such as servers, networking, processing, storage, and other resources. Cloud servers are offered to the organization through an API and ease the deployment of IT Cloud infrastructure.

1. INTRODUCTION

Cloud computing is the on-demand model for the availability of convenient computing resources such as data storage, computing power, networks, and applications. It is generally used to describe data centers available to many users over the internet. Following is a brief about cloud computing:

Service Models:-

1. Software as a Service (SaaS) - It is a way of delivering software applications that allow data to be accessed from any device with an internet connection and web browser. In the software as a Service model, the consumer has access to use the provider's applications running on cloud infrastructure without installing the software on the user's system. Cloud providers supervise the platforms that run the applications.

Deployment Models

1. Public Clouds- It is a form of service where the cloud service provider makes resources available to the public via the internet. Public clouds include SaaS, PaaS, and IaaS services. In this model, computing resources are owned by a provider and shared among multiple tenants.
2. Private Clouds- It is a term that is solely dedicated to end-users. Understanding private cloud demands understanding the working of virtualization. The need for virtualization technology is to combine resources sourced from physical hardware into shared pools. This cloud is accessible to only one organization.
3. Community Cloud- It is a multi-tenant platform that allows systems and services to be accessible by different organizations to work on a shared platform. It aims to enable multiple

customers to work on joint projects and applications belonging to a community.

4. Hybrid Cloud- It is an environment that combines a private cloud with one or more public cloud infrastructures. It provides greater flexibility and scalability to business organizations.

Cloud computing provides a large space to store data with good quality service, thereby becoming popular among companies. The servers are easily maintained and ensure that updates are more compatible with the devices. It provides self-service and on-demand service models. Amazon Web Services is the broadly adopted on-demand cloud platform, providing over 175 fully-featured products and services, including cloud storage, networking, database, deployment, and the Internet of things from data centers worldwide. It includes hardware central processing units (CPUs) and graphics processing units (GPUs) for processing. Amazon Web Services' offerings are accessed over HTTP, using the REST architectural style and JSON for newer APIs.

Amazon Web Services offers services from various centers present across availability zones in regions across the world. According to their affordability, a business company will choose one or multiple availability zones; the reasons can be compliance and proximity to end-user

2. RELATED WORKS

A research paper called "Research on Cloud Data Storage Technology and Its Architecture Implementation by Liu, Kun & Dong, Long-jiang" introduces the concept of cloud computing, cloud storage, and architecture of cloud storage. It emphasizes more on cloud storage, how it can be accessed for business and data storage. It talked about enterprises' cloud storage technology, i.e., GFS, consisting of a single master, multiple chunk servers, and various clients. Each of these is a commodity Linux. Next is HDFS (Hadoop Distributed File System), like GFS of Google, as it runs on large amounts of commodity hardware. This research paper also talks about cloud storage architecture based on Hadoop, including EyeOS and system implementation. Cloud storage is the

crux part of cloud computing and brings rich applications to the Internet.

In another paper called "Windows Azure Storage: a highly available cloud storage service with strong consistency", the authors discuss the Windows Azure Storage (WAS). WAS is a system that provides users to store unlimited data for any amount of time. They can access it anytime and pay only for the data they use and/or store. This paper discusses the architecture of WAS, global namespace, and data model. It also throws insights into load balancing, resource provisioning, and replication systems. WAS data is stored both locally and geographically for disaster and recovery management. This paper touches upon the consistency of WAS and on how it is accessible to its users. While the paper called "Unsecured Cloud Object Storage: A Primer On Prevalence & Mitigation Techniques" focuses on the leading four cloud technologies: Amazon Web Services (AWS), Google Cloud, IBM Cloud, and Microsoft's Azure. It looks upon how and why data is leaked and why these cloud-based companies are a popular target for attacks. A tool was developed and tested to look for public objects in the cloud after looking at these. It was found that even though the default settings were private, the admin was able to change the settings. A tester hired by any company could use a wordlist that contains the bucket names that the company may use (company name, company product names, etc.) to see if the buckets are public or private. The tester can then download the files to check whether it should be changed for private or public access.

3. PLATFORMS COVERED

Software - It uses Amazon Web Services that includes: S3, API Gateway, Lambda.

Hardware - It requires a desktop or laptop hardware running recent versions of Windows, 4 gigabytes of memory (RAM), A broadband Internet connection with upload speeds of at least 400 kilobits (50 kilobytes) per sec.

4. LANGUAGES AND SERVICES USED

HTML: HTML stands for Hypertext Markup Language, used for creating web pages. It describes how the structure of the web page would look.

Python: Python is a programming language. It is designed with the aim of readability and has some similarities to the English language. Python can be used for mathematics, system scripting, server-side development, and software development.

JSON: JSON stands for JavaScript Object Notation. JSON is a lightweight format for transporting and storing data.

CSS: CSS stands for Cascading Style Sheets. It describes how the HTML elements would appear on the screen. It handles the layout of multiple web pages together, which saves much work.

S3 Bucket: Amazon Simple Storage Service (Amazon S3) is an object storage service that provides industry-leading scalability, data availability, security, and performance. This means customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases, such as websites, mobile applications, backup and restore, archive, enterprise applications, IoT devices, and big data analytics.

API Gateway: Amazon API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale. API Gateway handles all the tasks involved in accepting and processing hundreds of thousands of concurrent API calls, including traffic management, CORS support authorization and access control, throttling, monitoring, and API version management.

AWS Lambda: AWS Lambda can be used to execute code in response to triggers such as changes in data, shifts in system state, or actions by users. AWS Lambda lets us run code without provisioning or managing servers. We can pay only for compute time we consume. Compared to AWS EC2, the purpose of Lambda is to simplify building smaller, on-demand applications that are responsive to events and new information. Unlike Amazon

EC2, which is priced by the hour but metered by the second, AWS Lambda is metered in increments of 100 milliseconds.

5. HARDWARE SPECIFICATIONS

The only hardware used will be a PC. It will act as a medium for the user to view the web service and post data and save it. While almost any computer that meets the minimum hardware requirements can run AWS that alone does not make a desktop computer a dedicated server. The essential need for web service to work is a stable internet connection.

AWS requires only a basic knowledge of operating systems like Windows OS, Linux, etc. A basic understanding of visualization is also needed. Networking is an essential skill as all operations on cloud platforms involve it. Also, we need to understand the difference between the Public and Private cloud. Last but not least, they must have basic command over coding and python language.

6. METHODOLOGY

This application is deployed on AWS, which tackles the problems of storage, computing, and servers very quickly. These days cloud security is a huge issue. To combat this, we have designed a web application for small file uploads that are visible for a short interval. This system uses AWS architecture involving S3, Lambda, API to read and write the user's data. We created an S3 Bucket and enabled CORS in bucket permission to edit CORS Rule using JSON for the first step. Then, we uploaded an HTML document to the bucket and made this object (ONLY) public. After this, an IAM Role is created, and permissions are attached to it (AWSLambdaExecute). Going into the project's nitty-gritty, a lambda function is then created, and an IAM role is attached to it; API Gateway with POST method acts as a link between them. Lambda function is then attached as a proxy with enabled CORS support. For the final step, the API is deployed, and through a global link generated using S3, we can see our final project.

The workflow of the following is given below:

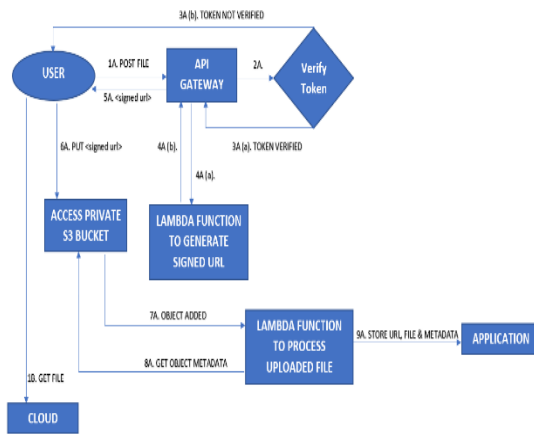


Fig 1.1 Workflow diagram

A. POST THE FILE

When the user uses the POST request to WRITE its file in the APPLICATION, the POST request is sent to the REST API GATEWAY to get verified. The verification is done through the TOKEN method. If the user is authorized then the token will get verified else the token would get rejected, and the user will not be able to write the file in the application. After the token gets approved the REST API Gateway grants the permission to access the Lambda Function. The Lambda Function generated a signed URL. This URL is then sent to the user. By using this signed URL, the user gets the access to the private S3 bucket. After this, the object, i.e the file is processed to be uploaded to the application with the help of the Lambda Function. Also, the object's metadata is stored in the S3 bucket. Then, the URL, file, and metadata are stored in the application.

B. GET THE FILE

When the user wants to READ its file from the APPLICATION, it uses the GET request to fetch it. After the GET request is sent, the file is fetched from the Cloud storage and displayed for the user to access it.

The administrator can also maximize their throughput by analyzing the metrics provided by AWS and make necessary conclusions about the time limit, storage size, and more and make necessary changes that will support their application even more. If any problem arises, it is

catered using AWS services as per demand, which decreases admin duties.

7. RESULT AND CONCLUSION

After completing all the codes and designing the project at the back end, the website and the project would appear as follows in the front end. Below are screen caps depicting how the project will look. After following all the necessary steps and implementing the coding, this illustrates the completed project.

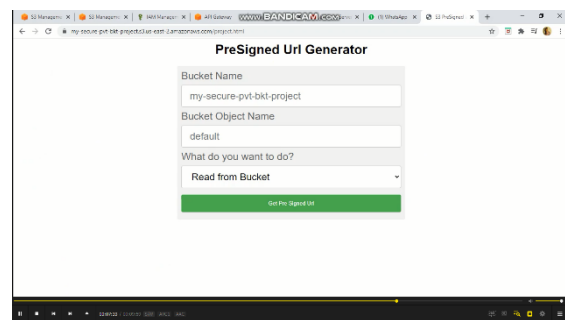


Fig 1.2: Final preview of Project (Read)

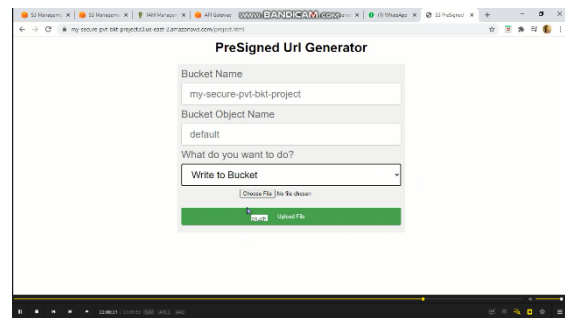


Fig 1.3 =: Final preview of Project (Write)

The project aims to develop an application that allows the user to upload their file securely using Amazon Web Services. We used S3 to upload the object. AWS Lambda is used to execute code to trigger such changes in data or actions by users. API Gateway handles the tasks, including CORS support, authorization, and access control and monitoring.

Overall this system makes users close to their data, i.e., users can see, read, and write their data anytime from anywhere using a global link.

Here we are using a combination of python, JSON, HTML, and CSS used by various AWS

services, making this feasible. A user who wants to use this system's services must have their object connected to the gateway file, which will allow the user to generate a unique global code and hence access the application.

8. FUTURE SCOPE

After the first stage of the project, the followings points are expected:

1. It is expected to be more secure.
2. Read and Write an uploaded file for a fixed time.
3. Highly scalable.

The application can be easily accessed from anywhere through the internet using a public link. People can view their files using the link provided in the application and write them from anywhere worldwide.

Many new features are expected to be added to it to improve its performance, and the ease of accessibility is also likely to improve as cloud computing is a growing technology.

REFERENCES

- [1]. <https://azure.microsoft.com/en-in/overview/what-is-cloud-computing/>
- [2]. <https://aws.amazon.com/what-is-cloud-computing/>
- [3]. https://en.m.wikipedia.org/wiki/Cloud_computing#:~:text=Cloud%20computing%20is%20the%20on,many%20users%20over%20the%20Internet.
- [4]. <https://aws.amazon.com/>
- [5]. https://en.m.wikipedia.org/wiki/Amazon_Web_Services
- [6]. <https://www.google.com/amp/s/searchaws.techtarget.com/definition/Amazon-Web-Services%3famp=1>
- [7]. <https://azure.microsoft.com/en-in/overview/what-is-a-public-cloud/#:~:text=The%20public%20cloud%20is%20defined,storage%2C%20or%20bandwidth%20they%20consume,storage%2C%20or%20bandwidth%20they%20consume.>
- [8]. <https://azure.microsoft.com/en-in/overview/what-is-a-private-cloud/>
- [9]. Ramirez Aguado, Miguel. 2018. "Unsecure cloud storage: prevalence & mitigation." *Projecte Final de Màster Oficial, UPC, Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona, Departament de Teoria del Senyal i Comunicacions* Retrieved September 8, 2020 (<http://hdl.handle.net/2117/127695>).
- [10]. Liu, Kun & Dong, Long-jiang. (2012). Research on Cloud Data Storage Technology and Its Architecture Implementation. *Procedia Engineering*. 29. 133–137. 10.1016/j.proeng.2011.12.682.
- [11] Brad Calder, Ju Wang, Aaron Ogus, Niranjan Nilakantan, Arild Skjolsvold, Sam McKelvie, Yikang Xu, Shashwat Srivastav, Jiesheng Wu, Huseyin Simitci, Jaidev Haridas, Chakravarthy Uddaraju, Hemal Khatri, Andrew Edwards, Vaman Bedekar, Shane Mainali, Rafay Abbasi, Arpit Agarwal, Mian Fahim ul Haq, Muhammad Ikram ul Haq, Deepali Bhardwaj, Sowmya Dayanand, Anitha Adusumilli, Marvin McNett, Sriram Sankaran, Kavitha Manivannan, and Leonidas Rigas. 2011. Windows Azure Storage: a highly available cloud storage service with strong consistency. In *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles* (*SOSP '11*). Association for Computing Machinery, New York, NY, USA, 143–157. DOI:<https://doi.org/10.1145/2043556.2043571>