

# Decision Analysis Report - Cloud Data Store

Project Name: Tresearch

Application Type: Web Application

Trial By Fire

Ian Ho-Sing-Loy

Jessie Lazo

Matthew Chen (Team Lead)

Pammy Poor

Ryan Harrison

Viet Nguyen

Instructor: Vatanak Vong

Submission Date

2/14/2022

California State University, Long Beach

College of Engineering

CECS491B Section 06 Spring 2022

## Purpose

The following is the Decision Analysis and Resolution to analyze which platform provides the best technology that would be most applicable towards the development of the product. The cloud's data store will be used to store user and system admin accounts to user's node content so we will prioritize storage and certain operations. The other storage we will be using is computer storage. Our exact operations and workloads for the data store will be CRUD for accounts and nodes, CRD for logs, and the computer storage will be for compressing and offloading archived log files (bulk operations).

## Technologies

### Computer Storage Technologies

1. Microsoft Azure Blob Storage
  - a. Massively scalable and secure object storage for cloud-native workloads, archives, data lakes, high performance computing, and machine learning
2. Amazon AWS S3
  - a. Amazon Simple Storage Service (S3) is an object storage service offering industry-leading scalability, data availability, security, and performance.
3. GCP Cloud Storage
  - a. Google cloud Storage is a RESTful online file storage web service for storing and accessing data on Google Cloud Platform infrastructure

### Database Store Technologies

1. Microsoft Azure SQL Database
  - a. Part of the Azure SQL family, Azure SQL Database is an intelligent, scalable, relational database service built for the cloud. Optimize performance and durability with automated, AI-driven features that are always up to date. Focus on building new applications without worrying about storage size or resource management with serverless compute and Hyperscale storage options that automatically scale resources on demand.
2. Amazon Relational Database Service
  - a. Amazon Relational Database Service (Amazon RDS) makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while automating time-consuming administration tasks, such as

hardware provisioning, database setup, patching, and backups. It frees you to focus on your applications so you can give them the fast performance, high availability, security, and compatibility they need.

3. Google Cloud SQL
  - a. Fully managed relational database service for MySQL, PostgreSQL, and SQL Server. Run the same relational databases you know with their rich extension collections, configuration flags and developer ecosystem, but without the hassle of self management.

## Preface

In this DAR we will compare Azure, Amazon, and Google Cloud's data store and computer storage. We will be prioritizing storage limits for the datastore primarily, since that is where we will store most of the application information.

## Metrics

### Data Store Metrics

- Duration to use credits (1)
  - The window of time we are given to use our free credits under every cloud provider, because we get free access to the paid features as long as we still have the free credits provided by the cloud.
- SQL Database Engine Storage Limits (0.9)
  - How much storage we are allowed to use under the duration as a free tier user provided by the cloud provider.
- Read/Write cost (0.8)
  - The cost of read, write, and several other operations to the computer storage on the cloud.
- Persistent Computer Storage (0.7)
  - Where persistent data on the VM is stored when the VM

restarts.

- Data Access operation (0.5)
  - What and how many types of operations every cloud provider has in order to access the data in the data store.
- Concurrent Throughput (0.5)
  - How many operations of a certain type (Write/Upload) can be done in one second (for bulk operations).
- Computer Storage Limits (0.3)
  - How much data can be stored on the Cloud provider as a free tier user.
- Computer Storage Duration Limits (0.2)
  - The cost of accessing computer storage that is accessed infrequently but stored for a long period of time.

## DAR

<u><b>Cloud Data Store</b></u>			
Duration to use credits (Months) <b>(1)</b>	12 months <b>(1)</b>	12 months <b>(1)</b>	3 months <b>(.1)</b>
SQL Database Engine Storage limits <b>(0.9)</b>	<b>Free:</b> Storage: 250 gb <b>(1)</b>	<b>Free:</b> Storage: 20 gb + 20 gb backup <b>(.25)</b>	<b>Free (Using provided credits):</b> <b><u>Storage: \$0.170 per GB/month + \$0.080 per GB/month backup</u></b> <b>Min: Using \$0.25 credits (0.17 for storage and 0.08 for backup) =</b> <b>1 gb storage + 1 gb backup</b> <b>Max: Using \$300 credits (150 for storage and 150 for backup) =</b> <b>882 gb storage + 1875 gb backup (0.1)</b>

Read/Write per transaction <b>(DTU) (0.8)</b>	<b>Free:</b> 10 DTUs <b>(1)</b>	S3 Standard PUT,COPY, POST, LIST per 1,000 requests: \$0.005 GET,SELECT, and other requests per 1,000 requests: \$0.0004 <b>(0.6)</b>	GCP Standard Class A operations per 10,000 operations(Insert, update) : \$0.05 Class B operations per 10,000 operations(Get) : \$0.004 <b>(.6)</b>
Persistent Computer Storage <b>(0.7)</b>	Azure Blob <b>(1)</b>	Amazon Instance Store <b>(1)</b>	Compute Engine <b>(1)</b>
Data Access Operation <b>(.5)</b>	Put Blob Get Blob Get Blob Properties Set Blob Properties Get Blob Tags Set Blob Tags Find Blobs By Tags Get Blob Metadata Set Blob Metadata Lease Blob Snapshot Blob Copy Blob Copy Blob From URL Abort Copy Blob Delete Blob Undelete Blob Set Blob Tier Blob Batch <b>(1)</b>	List the buckets in your account Create a bucket List the items in a bucket Upload a file to a bucket Download a bucket item Copy a bucket item to another bucket Delete a bucket item Delete all the items in a bucket Restore a bucket item Delete a bucket List the users with administrator privileges <b>(.75)</b>	Write requests: Uploading, updating, and deleting objects Read requests: Listing objects, reading object data, and reading object metadata <b>(.25)</b>
Concurrent Throughput <b>(.5)</b>	<b>Max per second:</b> 20,000 requests <b>(1)</b>	<b>Max per second:</b> <b>3,500</b> <b>PUT/COPY/POST/DELETE</b> <b>5,500 GET/HEAD</b> <b>(.8)</b>	<b>Max per second:</b> 1,000 WRITE/UPLOAD/UPDATE/DELETE 5,000 READ/LIST/READ OBJ DATA/READ OBJ METADATA <b>(.4)</b>
Computer Storage (Files, Image) <b>(.3)</b>	<b>Free:</b> <b>Stored data: 5gb</b> <b>(1)</b>	<b>Free:</b> <b>Stored data: 5gb</b> <b>(1)</b>	<b>Free:</b> Stored data: 1gb <b>(.2)</b>

Computer Storage Duration limits <b>(LTR)</b> <b>(0.2)</b>	GRS, RA-GRS, GZRS, or RA-GZRS \$0.02/GB <b>(.5)</b>	S3 Standard Infrequent Access (long lived but infrequently accessed data that needs millisecond access) \$0.0125/GB per month <b>(.75)</b>	GCP Archive Storage (Lowest cost, good for data that can be stored for at least 365 days) Cost: \$0.0012/GB per month <b>(.9)</b>
<b>Total:</b>	<b>4.8</b>	<b>3.63</b>	<b>1.935</b>

## Recommendation

Based on the analysis between Azure, Amazon, and GCP data stores, we see that Azure comes out on top. One of the most important aspects to our application was the database size, as well as the cost of reading and writing to said database. We prioritize those because our application will be reading and writing very often, and Amazon and GCP were very lacking in said features and required additional costs. In addition to the data store features, we would like to have computer storage on the cloud to store archived logs. We would like to also have our data persist through VM restarts, and every cloud provider provides a solution, so the scores are even there. The other less important metrics pertain to computer storage, so though the feature might be better than Azures', both providers get overshadowed by Azure's flexibility in both feature and cost.