

City University of New York (CUNY)

CUNY Academic Works

Open Educational Resources

Hunter College

2020

CSCI 49378: Lecture 9: Cloud Storage and Databases I

Bonan Liu
CUNY Hunter College

NYC Tech-in-Residence Corps

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/hc_oers/20

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

Cloud Storage and Databases I

Bonan Liu

Tech-In-Residence Member, Hunter College, CUNY



Disclaimer

The content of this presentation is being provided for educational and informational purposes only. The views, thoughts, and opinions expressed in this presentation belong solely to the author, and not necessarily to the author's employer.

The content of this presentation is not endorsed by the author's employer.

- Blob Storage: Google Cloud Storage
- SQL Basics
- Google SQL
- Google Spanner
- Google Datastore
- Assignment 4

- Distributed File System (Blob Store)
- Distributed Key-Value Store
- Distributed Relational Databases

We are going to learn their counterparts in cloud platform.

Google Cloud Storage provides world-wide storage service for unstructured data. Applications or end users could read and write data to GCS as normal file systems.

- Basic Concepts
 - Bucket
 - Global namespace
 - Object
 - Geo-Redundancy
 - Object Immutability
- Storage Classes
 - Standard
 - Nearline
 - Coldline

- Locations
 - Single Region
 - Dual Regions
 - Multi-Regions
- Domain-named Buckets
 - Use for static websites
- Interaction with GCS
 - gsutil

Demo 1: Copy/Paste with gsutil

Demo 2: Static website with GCS

Practice: Use gsutil to interact with GCS. Create a small HTML file and upload it to GCS.

SQL stands for Structured Query Language. It is the most common way to interact with relational databases.

- Basic concepts in relational databases
 - Database
 - Tables
 - Rows
 - (Other features omitted...)
- SQL has many different variants and implementations
- We will focus on CRUD operations

- SELECT
 - SELECT field1,field2
FROM tablename
WHERE ... GROUP BY ... HAVING ... SORT BY ...
- UPDATE
 - UPDATE *tablename*
SET *field1=val1*
WHERE ...

- CREATE
 - CREATE DATABASE *dbname*
 - CREATE TABLE *tablename* (*field1 type1, field2 type2*)
- INSERT
 - INSERT INTO *table* (*field1,field2*) VALUES (*val1, val2*)
- DELETE
 - DROP TABLE *tablename*
 - DELETE FROM *tablename* WHERE ...

Demo: Use SQL to query data.

Connect to:

`db.teaching.liubonan.com`

Google Cloud SQL is a managed relational database service which provides MySQL, PostgreSQL and SQL Server backends.

- Basic Concepts:
 - Instance
 - Regional Location
- Two-Level access control
- Difference between two generations
- Replication
 - Read Replica
 - External Replica

Demo: Setup Primary/Replica model in Google Cloud SQL.

Practice: Create your own Cloud SQL instance. Set up primary/replica model. Write to your primary db and read from replica db.

Spanner is Google's scalable, multi-version, globally distributed, and synchronously-replicated database. It is the first system to distribute data at global scale and support externally-consistent distributed transactions.[1]

[1] *Spanner: Google's Globally-Distributed Database.*

<https://static.googleusercontent.com/media/research.google.com/en//archive/spanner-osdi2012.pdf>

- Basic concepts
 - Instance
 - Node
- Five-Nine availability
- Global data replication
- Decide between region or multi-region
 - Availability
 - Latency
 - Cost
 - Replication

Google Cloud Datastore is a **managed** NoSQL database which provides basic key-value data storage service.

- Automatic replication
- Automatic sharding
- SQL-like query

- Basic Concepts
 - Kind
 - Entity
 - Property
- Strong consistency
- Various limitation (replaced by Firestore)
 - Request Size
 - QPS
 - Number of Nested Entities
 - Index

Demo: Tiny URL service with data storage.

Practice: Interact with Google Cloud Datastore via UI.

Spanner: Google's Globally-Distributed Database.

<https://static.googleusercontent.com/media/research.google.com/en//archive/spanner-osdi2012.pdf>

f4: Facebook's Warm BLOB Storage System.

<https://www.usenix.org/system/files/conference/osdi14/osdi14-paper-muralidhar.pdf>