HSBC Pune A - Notes - 1 & 2 July 2020

Start time: 9:00am Pune Time

Link to the Slide Deck:

https://tinyurl.com/ybk3qhmt

Link to this document (Classroom Notes):

https://tinyurl.com/y8psetf9

Link to the course Evaluation

https://bit.ly/3di3tU4

Link to Google's Cloud Architect Certification Site: https://cloud.google.com/certification/cloud-architect

Break until 4:40pm

On Exam

50 Questions - either Multiple Choice (One answer out of 4 or 5 options) OR Multiple-select (select 2 of the following that meet the requirements)

On the EXAM - you will see 1, 2 or ALL 3 Case Studies - with 4 or 5 questions on each

There are THREE Case Studies - assume you will get ALL THREE on the exam

Mountkirk Games

Dress4Win

Terramearth

Common Patterns

Batch Transforming and Load Data into Data Warehouse (ETL - Extract, Transform, Load)

GCS (Cloud Storage) -> Dataflow -> BigQuery

Streaming Data into Data Warehouse

Pub/Sub -> Dataflow -> BigQuery

OnPremise Hadoop to GCP

3 Steps:

1. OnPrem Hadoop -> DataProc (Compute - hadoop, Storage - HDFS, Maybe using Hive)
2. HDFS -> GCS, and update hdfs:// references in code to gs:// reference in code
    a. NOW our DataProc Cluster is JUST COMPUTE - so we can create/delete it as needed
3. Now we can do NEW development using Dataflow, and land data into BigQuery (instead of a HiveDB)
    a. GCS (we moved our HDFS) -> Dataflow -> BigQuery

Application Deployment Pattern

Run code - but I don't want to manage my web servers (Python, Java, Php, etc) (not containerized)

AppEngine Standard

These products go REALLY WELL with AppEngine: GCS (blob/static content), Firestore/Datastore for Structured Data Storage

AppEngine + GCS + Firestore - Autoscaling infrastructure that can scale to infinity and beyond

Application Code -> Artifacts

Application -> Docker Container (Continous Integration)

Source Repositories (git repos in your project)->Cloud Build->Google Container Registry (GCR)

Source Repos->Cloud Build (CI/CD on GCP)->GCR (docker container)

Normal GIT->JENKINS->Artifact (Container, .jar/.war)

GOOGLE LIKES JENKINS (CI Integration) AND SPINNAKER (CD for Containers)

How do I run JENKINS on GOOGLE: VM (GCE) or KUBERNETES (both can be found in the Marketplace)

Deciding on Compute

Code or Containers

- Code: GCE, App Engine Standard or Cloud Functions
  - Full Control over OS and above: GCE
    - (need for hours or permanently - that when GCE makes sense)
  - Don't want to manage OS, Autoscaling, etc - AppEngine Standard
    - (app starts at first request, runs, and turns off after 15 minutes if no activity)
  - Just want to run FUNCTIONS (not applications) - Cloud Function
    - (using compute for ms to seconds)

- Containers: Kubernetes, App Engine Flex or Cloud Run
  - Kubernetes CLUSTER (3+ nodes) - Complex Containers (lots of em)
    - Need for hours or permanently

  App Engine Flex - google will start my container in a VM (one container per VM) always run at least ONE instance (example: a java app running on tomcat might take 10 seconds to start - first access waiting 10 seconds - no way! - so google will run the container in a VM for you - and if needed, they will start MORE - but ALWAYS ONE VM)

  - Stateless Docker Containers - Cloud RUN (more like a Cloud Function)
    - (run the container for ms to seconds)

Data Storage

- BLOB (images, files, pdfs, backups, etc)
  - GCS
  - Host NFS (NFS that can be mounted across the network)
    - (unix equivalent NFS is Unix's version of a shared folder)
    - (uses the NFS protocol - /secondisk/ -> remote file system)
  - Windows File Share (Windows VM, sharing the folder)
    - (uses SMB protocol)

- SQL
  - CloudSQL (mysql, posgresql, or MS SQL 2017)
  - Spanner
  - VM and install the DB
  - BigQuery (Analytic ONLY - NOT Transactional - Data Warehouse)

- NoSQL
  - Pattern: Key-Value (Redis, Memory Store (managed Redis)
  - Pattern: DocumentDB (Mongo, Firestore)
  - Pattern: Single Key, Wide-column (HBase, BigTable)

How we Deploy

Single Region, Single Zone - development, things that don't need High Availability
Single Region, Two Zone - Primary + Failover (High Availability SQL)
Single Region, ALL Zones (Regional Bucket, 3-Zone Kubernetes Cluster) - High Availability
Multiple Regions, Multiple Zones (Active/Active, Active/Failover) - Disaster Recovery

Bucket Multi-Region - are deployed to AT LEAST 2 REGIONS in that part of the world (US, EU, Asia)

API - Global Services

Google Compute Products

Compute Engine - a VM runs in a Single Zone
Uses Persistent disks for OS and additional, pay for what we allocate
Example: n1-standard-1 - 1 CPU/standard ration of memory, n1-standard-4 - 4 CPUs
Adding more CPUs give more network capacity on GCP (2gbps/per CPU)
Increasing Disk size gives better performance, more IOPS
Persistent Disk (spinning) or Persistent SSD (ssd-based, faster, but more $$)
Local SSD is NOT a Persistent Disk, CANNOT hold the Operating system. Erased upon shutdown
But Local SSD are VERY fast

Could have Regional (dual-zone) Disk - so we could bring it up in ANOTHER zone

Kubernetes - Cluster is Single or Multiple Zones in a REGION (but NOT multi-regional)

AppEngine Standard - Regional

AppEngine Flex - Regional (google is running at least ONE VM - so technically it's zoneal, but they can bring up in another zone should a ZONAL issue arise)

Cloud Functions - Regional

Cloud Run - Regional

Data Storage

Cloud SQL - Single Region, Single Zone or Single-Region Dual-Zone (High Availability
Can run a READ-Replica in ANOTHER Region - but failover would NOT be automatic
RPO = Recovery Point Objective - how much data can I lose
RTO =- Recovery Time Objective - how much time do I have to fail over
We would have to fail over

Spanner - ANSI SQL - Single Region, Three zone (high Availability or Multi-Region, Multiple Zone (Disaster Recovery - loss of entire region - still operational - no downtime)

ANSI SQL Transactional Database - Highly Relational, Strongly Consistent and Transactional
CloudSQL and Spanner

NoSQL Document DB (Firestore/Datastore) - is NOT built for Highly Relation - but it is Strongly Consistant and can be Transactional - it doesn't know ANSI SQL - it has it's own query language

If they say ANSI SQL, SQL or Highly Relations - you MUST stay with SQL

Is BIGQUERY a Fully SQL Database? It's an ANALYTIC Database NOT a TRANSACTIONAL Database
Understands ANSI SQL
But it is NOT a transactional database (OLTP) - it's a ANALYTIC Database
You don't want to be UPDATING rows in BigQuery. You want to be adding rows to BigQuery

- 40 web application servers providing **micro-services** based APIs and **static content**

    Tomcat - Java - containering

    Nginx - what is nginx??? Web Server/Load Balancer/Cache/etc

Internet Users -> NGINX (front-end webserver/load balancer/API GW)→ Tomcat/Running Java (microservices) -> Data storage

Nginx - Load Balancer, Static Content, CDN, Cloud Armor - API GW Compute, Cloud EndPoints (really just an nginx API gateway)

Tomcat/JAva Containers - simple: appengine flex, complex: kubernetes
Stateless containers - cloud run

20 Apache Hadoop/Spark servers:

- Data analysis
- Real-time trending calculations
- Eight core CPUs
- 128 GB of RAM
- 4x 5 TB HDD (RAID 1)

Apache - step 1 - Lift and Shift into DataProc w/enough Disk to handle the HDFS
Step 2 - move the HDFS into GCS and update hdfs:// references to gs://
Now we can create/delete the cluster as needed
Step 3 - write NEW jobs as Dataflow jobs - using GCS as the source (but old jobs still run on dataproc)

Google Builds GFS, MapReduce Infrastructure, BigTable

GFS Paper -> Yahoo Reads Paper -> Yahoo Creates Hadoop Distributed File System (HDFS)
MapReduce Paper -> Yahoo Reads Paper -> Yahoo Create Hadoop (parallel processing workload)
Bigtable Paper -> Yahoo Read Paper -> Yahoo Creates Hadoop Database (Hbase)

Three RabbitMQ servers for messaging, social notifications, and events:

Lift and Shift:

Replatforming: Move RabbitMQ to Pub/Sub

Pub/Sub is a messaging infrastructure for streaming

Pub/Sub -> Dataflow -> BigQuery or BigTable or write results to GCS

Dataflow is a Python/JAva-based PROCESSING ENGINE

Misc Servers:

Jenkins (Continuous Integration) -> Lift & Shift -> VM or Kubernetes

Replatform: Source Repos -> Cloud Build -> Container Registry (artifact Registry if it NOT a container)

Monitoring - Stackdriver (now called Operations) Logging and Monitoring

Stackdriver logging by default captures ONLY THINGS THE HYPERVISOR Does

Does NOT capture activity inside the VM

Stackdriver monitoring only captures what the Hypervisor can see (CPU, Network bandwith, Disk IO Operations.thruput - it does NOT know about application activity

bastion hosts

security scanners

Cloud Security Scanner - ORGANIZATIONAL LEVEL - Scan projects for assets, vulnerabilities, mis-configurations, firewall rules, etc

Web Security Scanner - Application Scanner

Forseti Security Open source security tools for GCP

Other Product we need AWARENESS / What they DO knowledge for the exam

Some other products to know about (just one or two lines about what the product does)

Cloud Armor - Firewall / Web Application Firewall (layer 5-7) for Cloud Load Balancers

Identity Aware Proxy (IAP) - forces a LOGIN before you can get thru to a web application running on Google

VPC Service Controls - Firewall for API Calls (things that don't have IP address - that we want to restrict access to - based on where the request is coming from (not WHO, but WHERE)

Example: HSBC wants to limit bucket access such that you MUST be inside the HSBC network

Cloud Build - Continuous Integration & Continuous Deployment

Data Loss Prevention API (DLP) - Text and Images - find and redact PII
(other APIs as well - Vision (extract text and information from image), Speech API (extract text from audio)

Data Fusion - code-less platform for writing ETL jobs that run on DataProc underneath

Data Prep - code-less platform for data wrangling for writing ETL jobs that run on Dataflow underneath - in partnership with Trifacta (you write recipes for transforming your data using their syntax)

Web Security Scanner - scan web applications for XSS, SQL Injection, other application testing

Security Command Center (Organization - saw grant's bitcoin mining)

Forseti - open source security tools for GCP - document/manage policy security

IOT Core - Internet of Things Core - handled the Device integration (two-directions - to/from the device) for obtain data and then tuning the devices. Uses Pub/Sub as the communication pathway for a large portion

If you want individual DEVICE information - that data needs to be in something like BigTable
If you want AGGREGATION, ANALYTICS about ALL devices - use BigQuery

Dress 4 Win

Dress4Win is moving their development and test environments

They are also building a disaster recovery site, because their current infrastructure is at a single location.

Day 1: Static Website serviced from a bucket (we're sorry - our website is down)
(HTTP Only) DNS-> CNAME (alias) that points to `c.storage.googleapis.com`

```
bucket name = website
Set permissions to allUsers
```

(HTTP or HTTPS) DNS-> IP of Load Balancer (SSL, Rules) -> Static Content (Bucket)
Set permissions to allUsers
Begin to add more functionality as time goes by

When onpremise disaster strikes - we change DNS record to point to the Load Balancer

They are not sure which components of their architecture they can migrate as is and which components they need to change before migrating them.

The Dress4Win application is served out of a single data center location. All servers run Ubuntu LTS v16.04.
Solve for their Ubuntu LTS 16.04 - Easy to Lift and Shift to GCE VMs

Databases:

- MySQL. One server for user data, inventory, static data,
  - MySQL 5.7
  - 8 core CPUs
  - 128 GB of RAM
  - 2x 5 TB HDD (RAID 1)

Lift & Shift this: (2 options)
CloudSQL (move to CloudSQL is a Lift and Shift - export->import into cloudSQL, update their connections to the new IP/Name

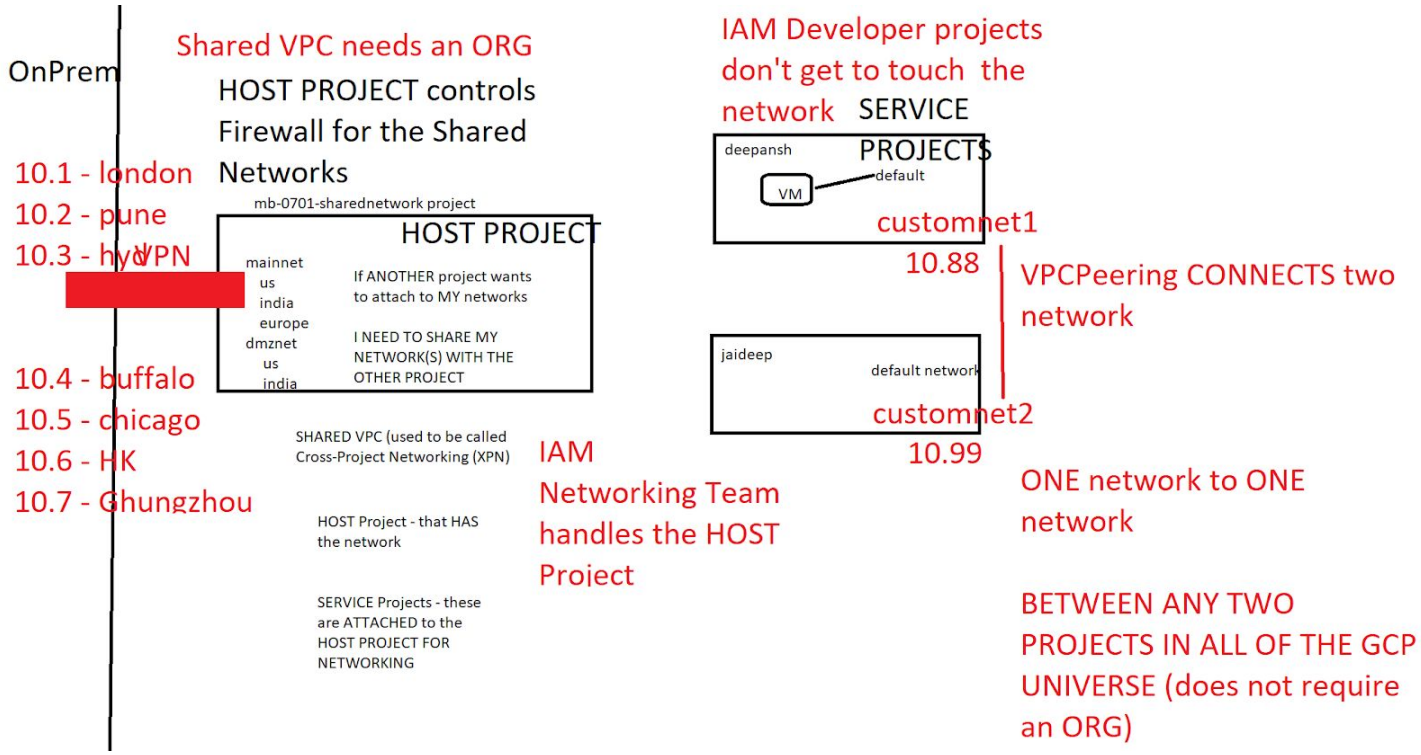
VM and load MySQL 5.7 into a VM
CloudSQL - High availability with 1 Region, 2 Zones (CloudSQL HA is)
To change the server config (more/less CPU/RAM - require a restart)

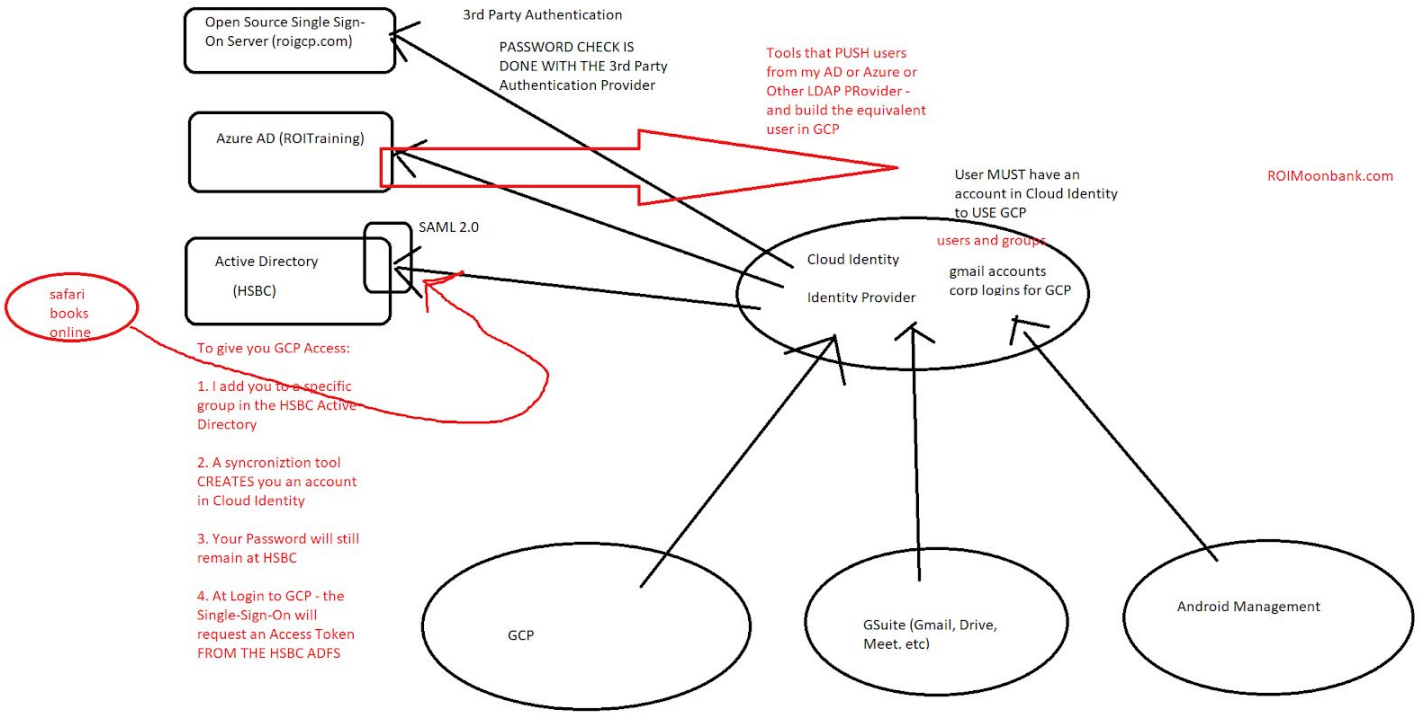
ANSI SQL is a language standard - they all understand insert, update, delete, select, drop table, create table
But MySQL->PostgreSQL - still has a LOT of programming changes

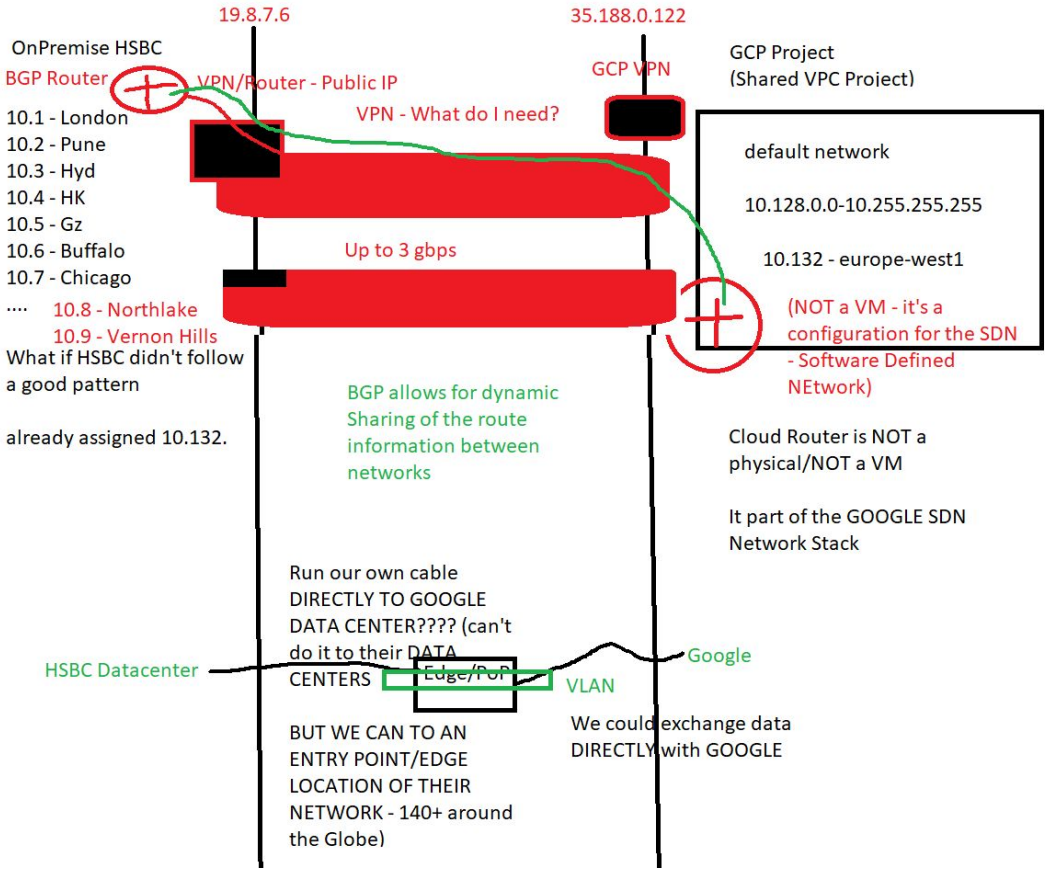
MySQL->Microsoft SQL - has a LOT of programming changes
MySQL->Spanner - has a LOT of programming changes

Replatform - keeping ANSI SQL structures/queries
Move to a DIFFERENT DB engine - Spanner
Why? Multi-Regional, No Maintenance Window, horizontal scalability (add/remove nodes without downtime)
We need permission/funding/code-changes OK/more cloud native/etc

Replatform - cloud native, TRANSACTIONAL + STRONGLY CONSISTENT (but they don't ask for RELATIONAL) - this is permission to use Firestore/Datastore


OnPrem
Shared VPC needs an ORG
HOST PROJECT controls Firewall for the Shared Networks
10.1 - london
10.2 - pune
10.3 - hyd
VPN
10.4 - buffalo
10.5 - chicago
10.6 - HK
10.7 - Ghungzhou
mb-0701-sharednetwork project
HOST PROJECT
mainnet
us
india
europe
dmznet
us
india
If ANOTHER project wants to attach to MY networks
I NEED TO SHARE MY NETWORK(S) WITH THE OTHER PROJECT
SHARED VPC (used to be called Cross-Project Networking (XPN)
HOST Project - that HAS the network
SERVICE Projects - these are ATTACHED to the HOST PROJECT FOR NETWORKING
IAM Networking Team handles the HOST Project
IAM Developer projects don't get to touch the network
SERVICE PROJECTS
deepansh
VM
default
custom-net1
10.88
jaideep
default network
custom-net2
10.99
VPCPeering CONNECTS two network
ONE network to ONE network
BETWEEN ANY TWO PROJECTS IN ALL OF THE GCP UNIVERSE (does not require an ORG)


Open Source Single Sign-On Server (roigcp.com)
3rd Party Authentication
PASSWORD CHECK IS DONE WITH THE 3rd Party Authentication Provider
Tools that PUSH users from my AD or Azure or Other LDAP PRovider - and build the equivalent user in GCP
Azure AD (ROITraining)
User MUST have an account in Cloud Identity to USE GCP
ROIMoonbank.com
SAML 2.0
Active Directory
(HSBC)
Cloud Identity
Identity Provider
users and groups
gmail accounts
corp logins for GCP
safari books online
To give you GCP Access:
1. I add you to a specific group in the HSBC Active Directory
2. A syncroniztion tool CREATES you an account in Cloud Identity
3. Your Password will still remain at HSBC
4. At Login to GCP - the Single-Sign-On will request an Access Token FROM THE HSBC ADFS
GCP
GSuite (Gmail, Drive, Meet, etc)
Android Management



Service Project

Lift and Shift is cheaper and faster than re-platforming/changing the infrastructure - also less risky

BUT - it doesn't necessarily take advantage of the cloud/cloud-native

---

Compute on Google - which is it:

Compute Engine - Infrastructure as a Service
Kubernetes - Containers as a Service (CaaS)
AppEngine Standard - Platform as a Service
Cloud Functions - somewhere between PaaS and SaaS

Gmail - Software as a Service
Drive is a Software as a Service
A Receipt Tracking PRogram running on the cloud is Software as a Service

Cloud Identity - IDaaS - Identity as a Service

Bucket has
    AVAILABILITY (your ability to access the file in the bucket)
    DURABILITY (even if hard drive(s) fail - your data is still there)

A REGIONAL BUCKET HAS THE SAME DURABILITY AS A MULTIPLE REGIONAL BUCKET
(99.999999999 durability)
BUT - a Multi-Region bucket has a higher AVAILABILITY SLA

Visualize the Directory Structure for GCS

Filename	datestamp	size	billing class
Moon.jpg	7/2/2020 1102	56K	standard
Duncan.jpg	7/1/2020 0853	100K	nearline

I want to store medical X-rays for a hospital into GCS (I have an app that read/writes them)

Make sure I can retrieve it (is 50 extra ms going to matter - NO)
Location: Do I need to PROTECT it from disaster (or is it derivative, and I could re-create)
Storage Class: How many times will I read it?

X-Ray is probably seen 4 times after it's been taken
(tech check it, the radiologist look at it, the Dr looks at it, and the dr shows the patient)

Standard - 2.6c/GB to store (month) - no per read cost - would be egress cost
Nearling - 1c/GB to store (month) - 1c/GB to read

Maybe after two weeks - we use an API - change the storage class from standard to nearline
Standard 2.6c/GB ($26TB) - no minimum time in standard, no per read GB charge
Near 1c/GB ($10TB) to store, 1c/GB to retrieve - minimum of 1 month storage charge
Coldline is 0.7c/GB ($7TB)to store, 2c/GB to retrieve - 3 month minimum
Archive is 0.4c/GB ($4TB) to store, 5c/GB to retrieve - 1 year minimum

Moonbank wants to store customer PDFs (1MB in size each is the target size)

1 - 1mb
10 - 10mb
100 - 100mb
1000 - 1g
1000000 1TB (per million)

Goal is 10 million customers - 10TB a month of PDF statements

Multi or Single Region - Multi-Regional

Standard or Nearline

Initially - store Multi-Regional + Nearline (1c/GB storage, 1c/GB retrieval)
Why? Most customers don't look at them or only look at the ONCE
$10 TB/$100 for 10TB (month worth of statement)
Budget another $100 for retrieval
A Month TOTAL LIFECYCLE COST for 10 million customers - FOR ONE MONTH HELD FOR 7 years ~
$100 for Jan statement held in Feb
$100 for Jaunary statements held in MArch
$8400 total storage cost for 7 years of a months statement

BEFORE CHOOSING - MAKE THEM CONVINCE YOU YOU NEED THESE

LocalSSDs - swap/temporary/non-persisent, but VERY FAST Disks on compute engine
(only are attached when the VM is running, and erase when the VM is stopped)
Spanner - minimum cost $1000month/ meant for TB to PB
BigTable - minimum cost $1500month/ meant TB to PB
Regional Buckets - don't survive the ASTERIOD NOT good for backups
Regional is the preferred bucket for a DATA ENGINEER to minimize latency (bucket in central, compute in central is better than MR bucket (might be west and east) and compute in Central (adds latency)

SERVERLESS PRODUCTS LIKE GCS and FIRESTORE - you just call an API - pay for your usage!
Using a database where you are paying per read/write operation is GREAT for smaller / burst/un-predictable use cases

Google defines HIGH AVAILABILITY is
Single Region, Multiple Zones (or better)

Disaster Recovery is
Multiple Regions

A Disaster at Google is NOT loss of a Zone, It’s LOSS of a REGION

MountKirk Games

Requirements for game backend platform - we KNOW they want COMPUTE ENGINE + NoSQL DB

Dynamically scale up or down based on game activity

Compute Engine scales with Managed Instance Groups + Load Balancing

Connect to a transactional database service to manage user profiles and game state

They don't say RELATIONAL, and the story says they don't want to use MySQL/SQL anymore - they want a NoSQL - Firestore/Datastore

Store game activity in a timeseries database service for future analysis

BigQuery (for aggreation/all player analytics)

BigTable is NOT good for aggregation - it's GREAT for individual and range of row retrieval

Looking at a single-players activities - NOT ALL Player Activity

BigTable is WIDE-COLUMN but ONLY ONE KEY / ONE INDEX - doesn't support multiple indexes - because the data is STORED ON THE DISK BASED ON THE KEY/INDEX

As the system scales, ensure that data is not lost due to processing backlogs

MEans buffer the data somewhere - pub/sub most common way to do this (streaming architecture), Redis or other cacheing database

Run hardened Linux distro

Compute Engine give full OS control - so they could control IPTables, Firewalls, inside the VM

Shielded VMs - hardened, more secure interactions with the Hypervisor, vTPM, UEFI BIOS, attestations (boot 1, end of boot1, boot 2, end of boot 2, etc)

Classic Data Engineering Pipeline (but wait - we're doing the Cloud Architect Exam!)

- Requirements for game **analytics** platform

    1. Dynamically scale up or down based on game activity

       Dataflow can scale up/down (it's an autoscaling data processing engine that understands python or java) if 100 messages in the Q - the one base node probably can process it all. If there are 100,000 message in the Q - dataflow will a bunch more works and scale up

    2. Process incoming data on the fly directly from the game servers

       Pub/Sub->Dataflow->BigQuery (streaming architecture on GCP)

3. Process data that arrives late because of slow mobile networks

   Dataflow (via Apache Beam - does Windows, Deal with Late Arriving Data, Reordering of Data, offset to account for latencies in the system, DEALS WITH LATE ARRIVING DATA)

4. Allow **queries** to access at least **10 TB** of **historical** data

   Analytics NOT Transaction (no MySQL, no Spanner, no CloudSQL)

   Analytics = BigQuery - run Queries (run ANSI SQL Queries even better hint)

   At least 10 TB of Data - that used to rule out CloudSQL (summer 2019 - 10TB was max size of CloudSQL - it's 30 TB)

5. Process files that are regularly uploaded by users' mobile devices

   GCS->Dataflow->BigQuery (batch processing architecture on GCP)

Dataflow understands/work with the Apache Beam object class.

Apache Beam does

Horizontal scaling of code/ Parallelization

Encryption

Google Managed Encryption
Cusotmer Managed Encryption Keys (CMEK) - Google has Key, Customer can set key/rotate key/audit key
Customer Supplied Encryption Keys (CSEK) - CUSTOMER PROVIDES THE KEY
On a VM - CSEK - customer must provide the key for the VM to boot

What should I be doing now?

Go thru the Slide Deck from this course - and create your breakdown document

Try the little quizzes we have in our slide deck (end of each chapter)

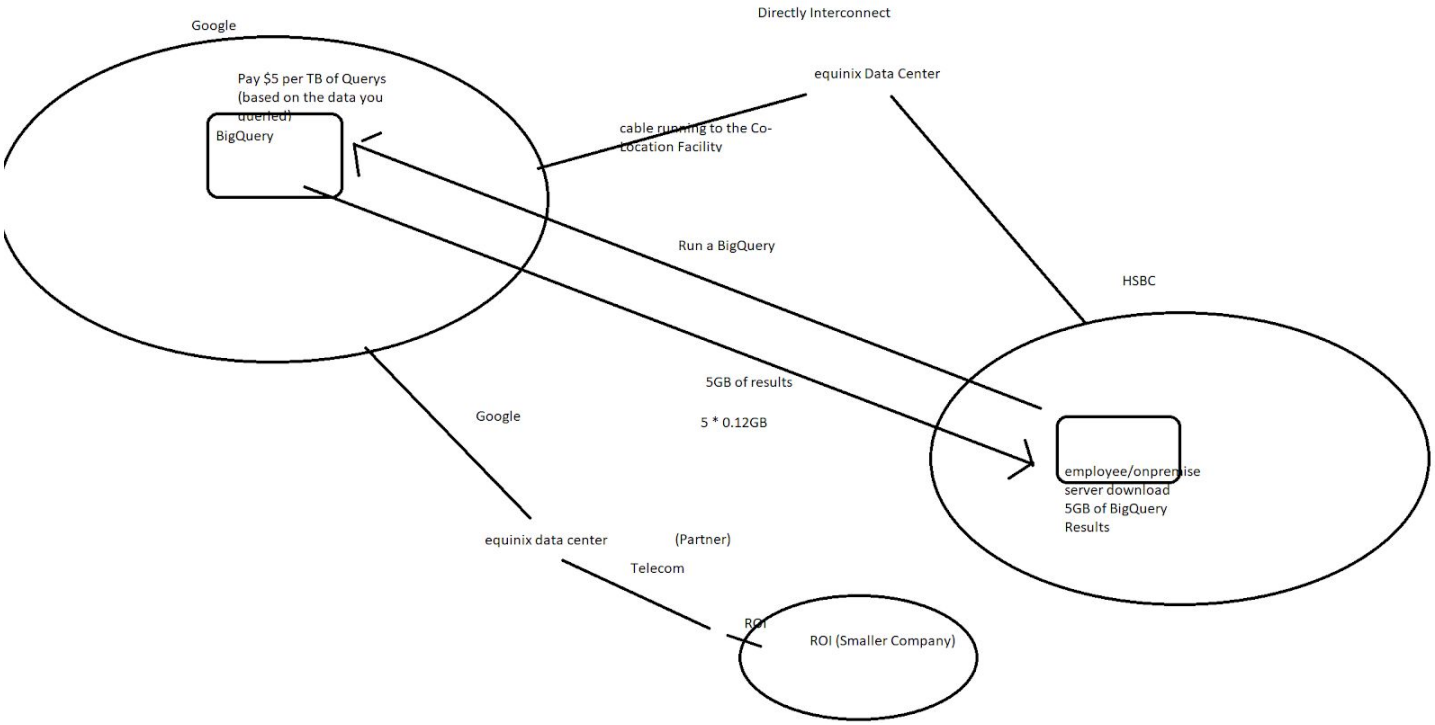
Do the Google Pracitce Exam for Cloud Architect (20 Questions - found on the Google Website)
https://cloud.google.com/certification/practice-exam/cloud-architect

After doing official 20-question practice exam - watch Grant's Practice Exam video (1.5 hours)

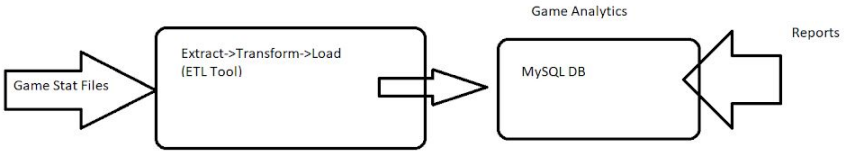
Then - work more on your breakdown document

Maybe try some of the OTHER certification exams:

- Associate Cloud Engineer (20 questions each)
- Data Engineering
- Networking
- Security



MySQL, PostGres, MS SQL, Spanner	Online Transaction (OTLP)	
BigQuery / Data warehouses	Online Analytic (OLAP)	

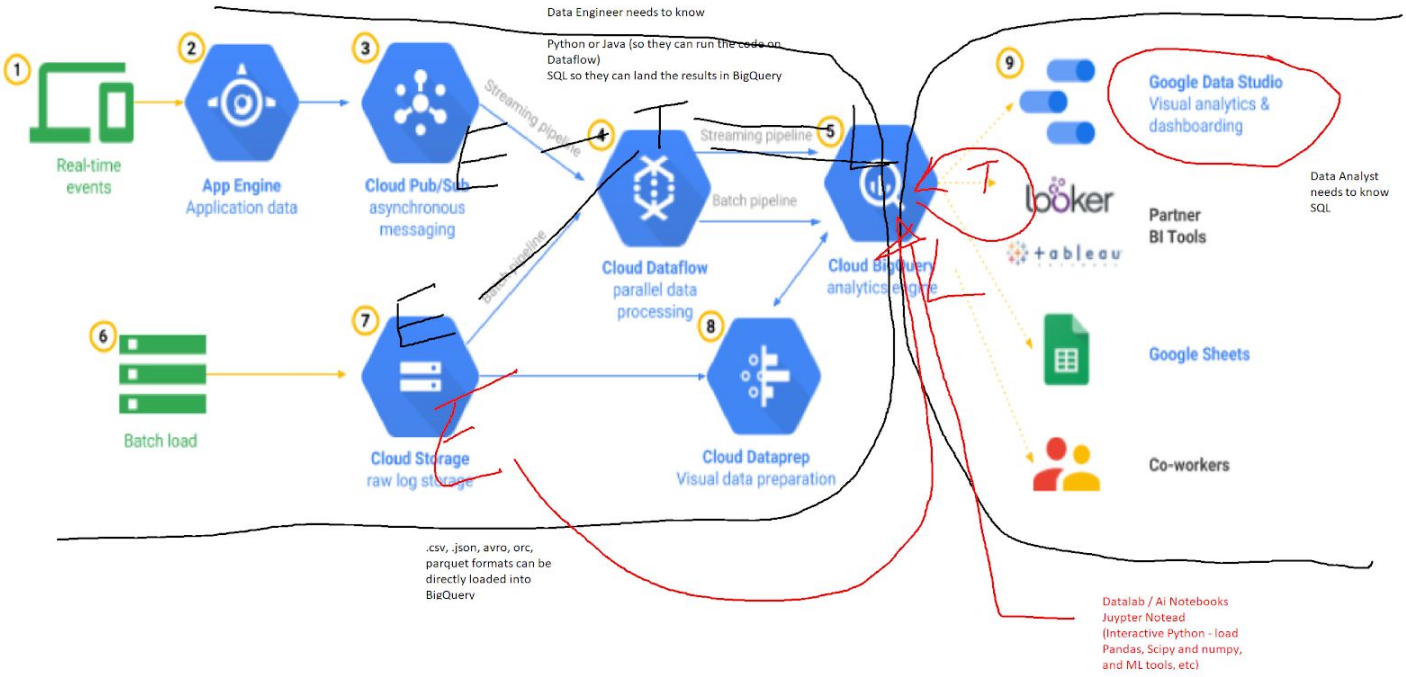


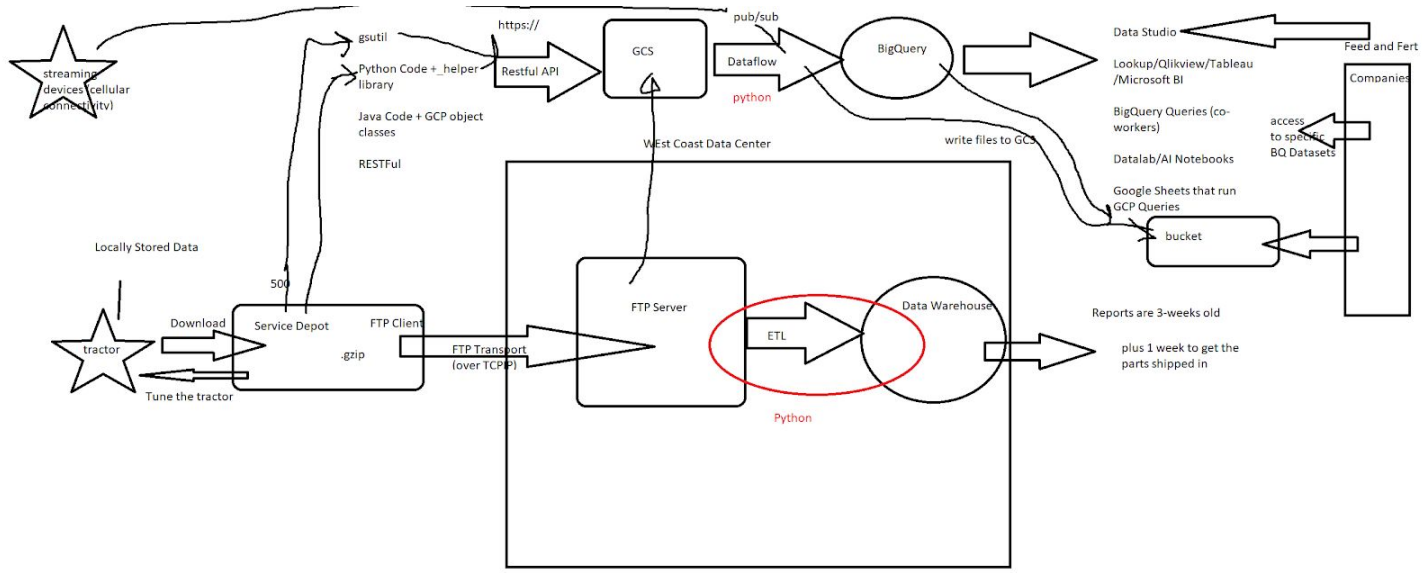
Constraints:

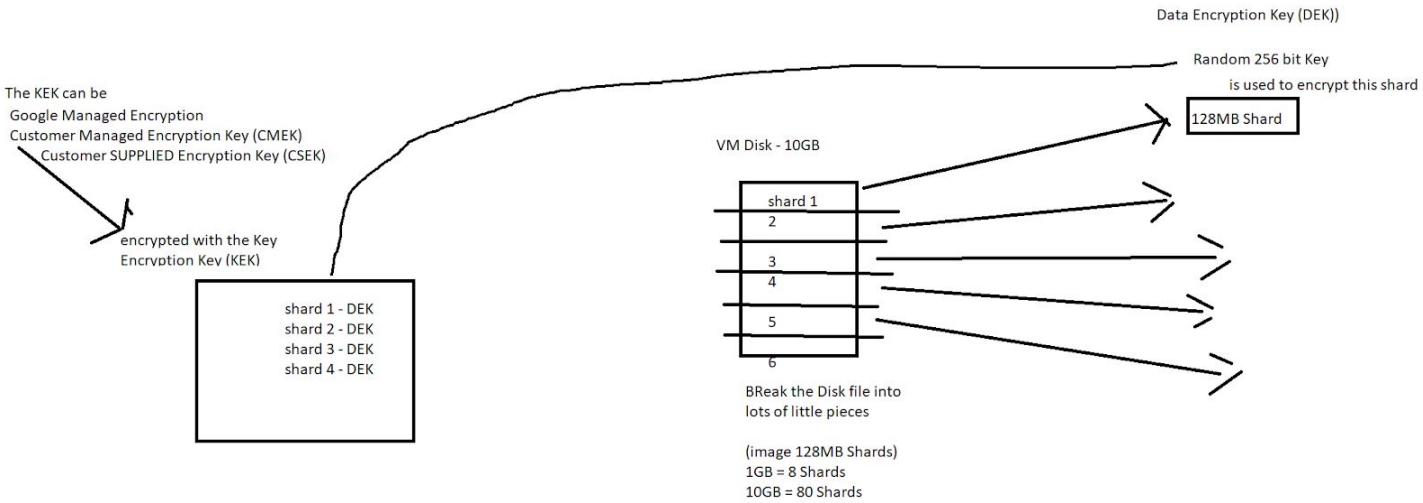
Compute Engine
NoSQL Database (they don't like MySQL anymore)

Run Hardened Linux

TPM _ Trusted Platform Module (place to store secrets and keys - that's seperate from the Hard Drive)

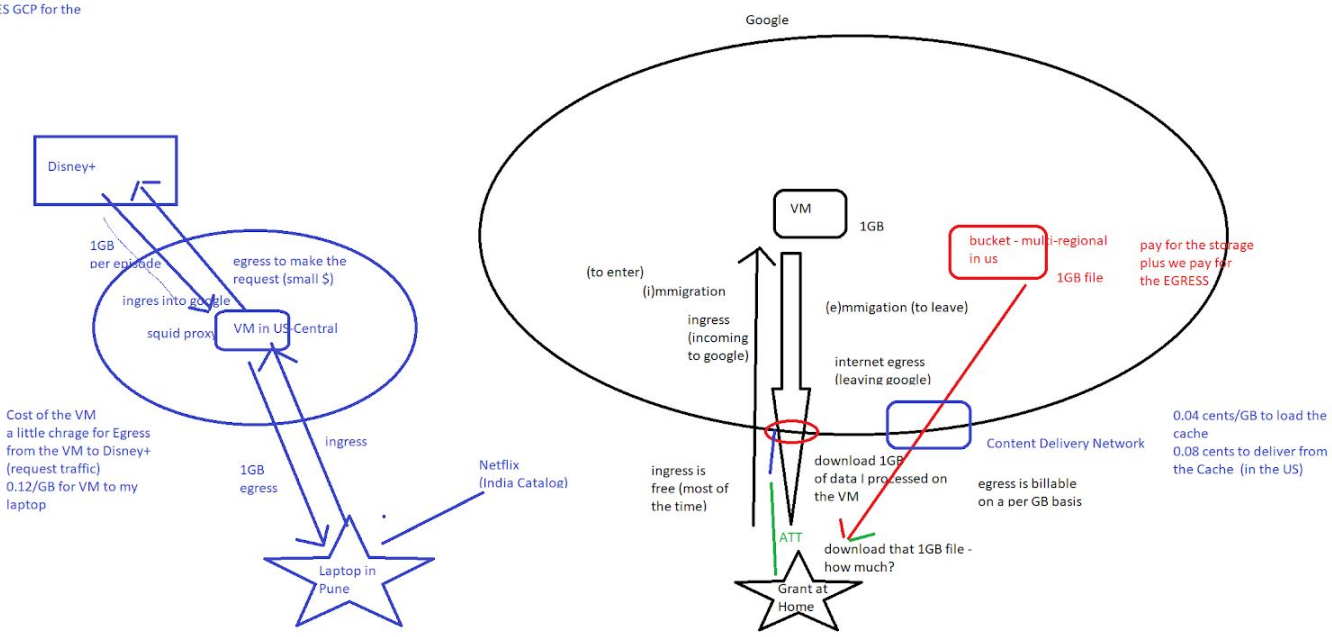

Data Engineer needs to know
Python or Java (so they can run the code on Dataflow)
SQL so they can land the results in BigQuery
Real-time events
App Engine
Application data
Cloud Pub/Sub
asynchronous messaging
Streaming pipeline
Batch pipeline
Cloud Dataflow
parallel data processing
Cloud BigQuery
analytics engine
Batch load
Cloud Storage
raw log storage
Cloud Dataprep
Visual data preparation
Google Data Studio
Visual analytics & dashboarding
Data Analyst needs to know SQL
looker
tableau
Partner BI Tools
Google Sheets
Co-workers
.csv, .json, avro, orc, parquet formats can be directly loaded into BigQuery
Datalab / Ai Notebooks
Juypter Notead
(Interactive Python - load Pandas, Scipy and numpy, and ML tools, etc)


streaming devices (cellular connectivity)
gsutil
Python Code + _helper library
Java Code + GCP object classes
RESTFul
https://
Restful API
GCS
pub/sub
Dataflow
python
BigQuery
Data Studio
Feed and Fert
Companies
Lookup/Qlikview/Tableau /Microsoft BI
BigQuery Queries (co-workers)
Datalab/AI Notebooks
Google Sheets that run GCP Queries
access to specific BQ Datasets
write files to GCS
West Coast Data Center
bucket
Locally Stored Data
500
tractor
Download
Service Depot
FTP Client
.gzip
Tune the tractor
FTP Transport (over TCPIP)
FTP Server
ETL
Data Warehouse
Python
Reports are 3-weeks old
plus 1 week to get the parts shipped in


Data Encryption Key (DEK))
Random 256 bit Key
is used to encrypt this shard
128MB Shard
The KEK can be
Google Managed Encryption
Customer Managed Encryption Key (CMEK)
Customer SUPPLIED Encryption Key (CSEK)
encrypted with the Key Encryption Key (KEK)
shard 1 - DEK
shard 2 - DEK
shard 3 - DEK
shard 4 - DEK
VM Disk - 10GB
shard 1
2
3
4
5
6
BReak the Disk file into lots of little pieces
(image 128MB Shards)
1GB = 8 Shards
10GB = 80 Shards

Internet Egress is data that LEAVES GCP for the Internet

0.12gb
0.11gb
0.08gb



Pub/Sub messages on July 2nd at 1am

CF are 0.40 per million invocations
Plus the CPU/RAM Time

BigTable needs a 3-node cluster
$1500/month

many coilums - it NOT a document store

do I need 30,000 rows per second?

BigTable

Transaction DB

1. Retrieve all rows for Customer Grant from June 2020

Code: Cloud Function
or I could VMs
Container: Cloud Run
or if I've got a Kubernetes Cluster - Kubernetes

CPU Choice

query grant's rows
merge into a PDF template

Bucket - Multiregional
\202006\

create the PDf _ store into Bucket
Create a DB record - pointing to the file in the bucket

$24/month

In DB somewhere

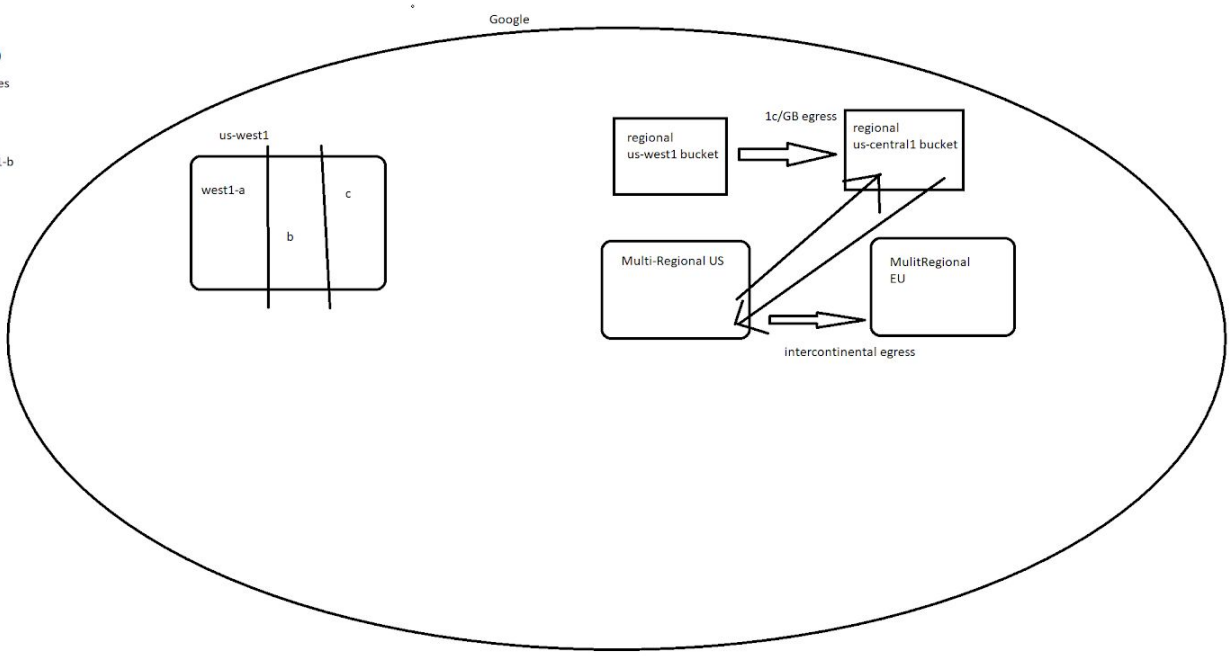
Grant Key | June 2020 | Path in GCS

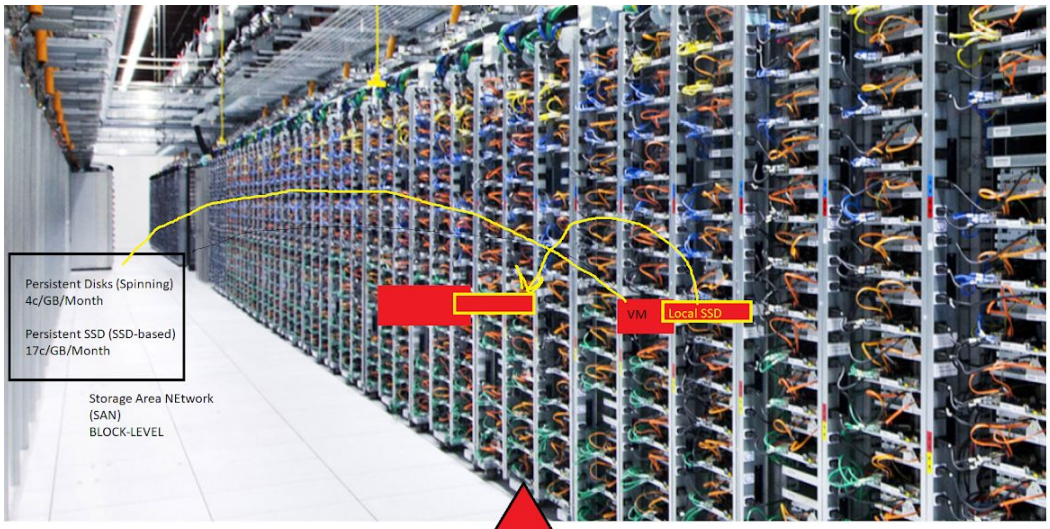
JSON BLOB about statemtents avialable

If I storerd the information about the customer and their statements

0.60 per million reads
1.80 per million writes

18.00 to create all their ROWS
if everyone reads their statements - it's $6





Inbound Traffic gets to the NEW VM

What is a VM

CPU
RAM
Networking

Attached to a Disk

To Live Migrate

1. Create a new w/Same CPU and RAM allocation

2. Copy over what CPU and RAM are doing

3. Detach Disk from old VM and attach to new VM

4. Flip the MAC address from the OLD VM to the NEw VM - so traffic arrives on the NEw VM

LocalSSD is a PHYSICAL DISK IN THE HOST

up to 8 of them per Host

When create a VM - you can ask for 1 to 8 of them

(there are now 16 and 24 available as well)

ephemeral, NON-Persisent, swap/temporary disk

CANNOT PUT YOUR OS ON A LOCAL SSD - they are only avaialble as additional disks