Product Questions: 334

Version: 25.0

Topic 1, Mountkirk Games Case Study 1

Company Overview

Mountkirk Games makes online, session-based. multiplayer games for the most popular mobile platforms.

Company Background

Mountkirk Games builds all of their games with some server-side integration and has historically used cloud providers to lease physical servers. A few of their games were more popular than expected, and they had problems scaling their application servers, MySQL databases, and analytics tools.

Mountkirk's current model is to write game statistics to files and send them through an ETL tool that loads them into a centralized MySQL database for reporting.

Solution Concept

Mountkirk Games is building a new game, which they expect to be very popular. They plan to deploy the game's backend on Google Compute Engine so they can capture streaming metrics, run intensive analytics and take advantage of its autoscaling server environment and integrate with a managed NoSQL database.

Technical Requirements

Requirements for Game Backend Platform

1. Dynamically scale up or down based on game activity.

2. Connect to a managed NoSQL database service.

3. Run customized Linx distro.

Requirements for Game Analytics Platform

1. Dynamically scale up or down based on game activity.

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

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

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

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

6. Use only fully managed services

CEO Statement

Our last successful game did not scale well with our previous cloud provider, resuming in lower user adoption and affecting the game's reputation. Our investors want more key performance indicators (KPIs) to evaluate the speed and stability of the game, as well as other metrics that provide deeper insight into usage patterns so we can adapt the gams to target users.

CTO Statement

Our current technology stack cannot provide the scale we need, so we want to replace MySQL and move to an environment that provides autoscaling, low latency load balancing, and frees us up from managing physical servers.

CFO Statement

We are not capturing enough user demographic data usage metrics, and other KPIs. As a result, we do not engage the right users. We are not confident that our marketing is targeting the right users, and we are not selling enough premium Blast-Ups inside the games, which dramatically impacts our revenue.

Question: 1

For this question, refer to the Mountkirk Games case study.

Mountkirk Games wants you to design their new testing strategy. How should the test coverage differ from their existing backends on the other platforms?

A. Tests should scale well beyond the prior approaches.

B. Unit tests are no longer required, only end-to-end tests.

C. Tests should be applied after the release is in the production environment.

D. Tests should include directly testing the Google Cloud Platform (GCP) infrastructure.

Answer: A

Explanation:

From Scenario:

A few of their games were more popular than expected, and they had problems scaling their application servers, MySQL databases, and analytics tools.

Requirements for Game Analytics Platform include: Dynamically scale up or down based on game activity

Question: 2

For this question, refer to the Mountkirk Games case study.

Mountkirk Games has deployed their new backend on Google Cloud Platform (GCP). You want to create a thorough testing process for new versions of the backend before they are released to the public. You want the testing environment to scale in an economical way. How should you design the process?

A. Create a scalable environment in GCP for simulating production load.

B. Use the existing infrastructure to test the GCP-based backend at scale.

C. Build stress tests into each component of your application using resources internal to GCP to simulate load.

D. Create a set of static environments in GCP to test different levels of load — for example, high, medium, and low.

Answer: A

Explanation:

From scenario: Requirements for Game Backend Platform

Dynamically scale up or down based on game activity

Connect to a managed NoSQL database service

Run customize Linux distro

Question: 3

For this question, refer to the Mountkirk Games case study.

Mountkirk Games wants to set up a continuous delivery pipeline. Their architecture includes many small services that they want to be able to update and roll back quickly. Mountkirk Games has the following requirements:

- Services are deployed redundantly across multiple regions in the US and Europe.
- Only frontend services are exposed on the public internet.
- They can provide a single frontend IP for their fleet of services.
- Deployment artifacts are immutable.

Which set of products should they use?

A. Google Cloud Storage, Google Cloud Dataflow, Google Compute Engine

B. Google Cloud Storage, Google App Engine, Google Network Load Balancer

C. Google Kubernetes Registry, Google Container Engine, Google HTTP(S) Load Balancer

D. Google Cloud Functions, Google Cloud Pub/Sub, Google Cloud Deployment Manager

Answer: C

Explanation:

Question: 4

For this question, refer to the Mountkirk Games case study.

Mountkirk Games' gaming servers are not automatically scaling properly. Last month, they rolled out a new feature, which suddenly became very popular. A record number of users are trying to use the service, but many of them are getting 503 errors and very slow response times. What should they investigate first?

A. Verify that the database is online.

B. Verify that the project quota hasn't been exceeded.

C. Verify that the new feature code did not introduce any performance bugs.

D. Verify that the load-testing team is not running their tool against production.

Answer: B

Explanation:

503 is service unavailable error. If the database was online everyone would get the 503 error.

https://cloud.google.com/docs/quota#capping_usage

Question: 5

For this question, refer to the Mountkirk Games case study

Mountkirk Games needs to create a repeatable and configurable mechanism for deploying isolated application environments. Developers and testers can access each other's environments and resources, but they cannot access staging or production resources. The staging environment needs access to some services from production.

What should you do to isolate development environments from staging and production?

A. Create a project for development and test and another for staging and production.

B. Create a network for development and test and another for staging and production.

C. Create one subnetwork for development and another for staging and production.

D. Create one project for development, a second for staging and a third for production.

Answer: D

Explanation:

Question: 6

For this question, refer to the Mountkirk Games case study.

Mountkirk Games wants to set up a real-time analytics platform for their new game. The new platform must meet their technical requirements. Which combination of Google technologies will meet all of their requirements?

A. Container Engine, Cloud Pub/Sub, and Cloud SQL

B. Cloud Dataflow, Cloud Storage, Cloud Pub/Sub, and BigQuery

C. Cloud SQL, Cloud Storage, Cloud Pub/Sub, and Cloud Dataflow

D. Cloud Dataproc, Cloud Pub/Sub, Cloud SQL, and Cloud Dataflow

E. Cloud Pub/Sub, Compute Engine, Cloud Storage, and Cloud Dataproc

Answer: B

Explanation:

A real time requires Stream / Messaging so Pub/Sub, Analytics by Big Query.

Ingest millions of streaming events per second from anywhere in the world with Cloud Pub/Sub, powered by Google's unique, high-speed private network. Process the streams with Cloud Dataflow to ensure reliable, exactly-once, low-latency data transformation. Stream the transformed data into BigQuery, the cloud-native data warehousing service, for immediate analysis via SQL or popular visualization tools.

From scenario: They plan to deploy the game's backend on Google Compute Engine so they can capture streaming metrics, run intensive analytics.

Requirements for Game Analytics Platform

Dynamically scale up or down based on game activity

Process incoming data on the fly directly from the game servers

Process data that arrives late because of slow mobile networks

Allow SQL queries to access at least 10 TB of historical data

Process files that are regularly uploaded by users' mobile devices

Use only fully managed services

Reference: https://cloud.google.com/solutions/big-data/stream-analytics/

Topic 2, TerramEarth Case Study

Company Overview

TerramEarth manufactures heavy equipment for the mining and agricultural industries: About 80% of their business is from mining and 20% from agriculture. They currently have over 500 dealers and service centers in 100 countries. Their mission is to build products that make their customers more productive.

Company Background

TerramEarth formed in 1946, when several small, family owned companies combined to retool after World War II. The company cares about their employees and customers and considers them to be extended members of their family.

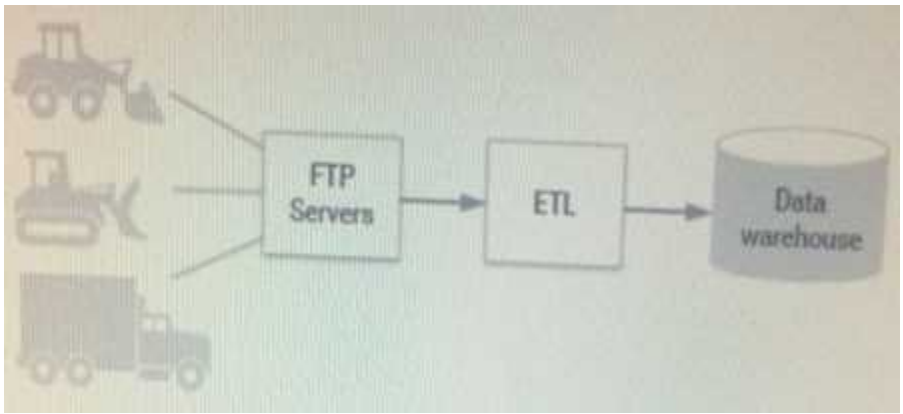
TerramEarth is proud of their ability to innovate on their core products and find new markets as their customers' needs change. For the past 20 years trends in the industry have been largely toward increasing productivity by using larger vehicles with a human operator.

Solution Concept

There are 20 million TerramEarth vehicles in operation that collect 120 fields of data per second. Data is stored locally on the vehicle and can be accessed for analysis when a vehicle is serviced. The data is downloaded via a maintenance port. This same port can be used to adjust operational parameters, allowing the vehicles to be upgraded in the field with new computing modules.

Approximately 200,000 vehicles are connected to a cellular network, allowing TerramEarth to collect data directly. At a rate of 120 fields of data per second, with 22 hours of operation per day. TerramEarth collects a total of about 9 TB/day from these connected vehicles.

Existing Technical Environment



TerramEarth's existing architecture is composed of Linux-based systems that reside in a data center. These systems gzip CSV files from the field and upload via FTP, transform and aggregate them, and place the data in their data warehouse. Because this process takes time, aggregated reports are based on data that is 3 weeks old.

With this data, TerramEarth has been able to preemptively stock replacement parts and reduce unplanned downtime of their vehicles by 60%. However, because the data is stale, some customers are without their vehicles for up to 4 weeks while they wait for replacement parts.

Business Requirements

• Decrease unplanned vehicle downtime to less than 1 week, without increasing the cost of carrying surplus inventory

• Support the dealer network with more data on how their customers use their equipment IP better position new products and services.

• Have the ability to partner with different companies-especially with seed and fertilizer suppliers in the fast-growing agricultural business-to create compelling joint offerings for their customers

CEO Statement

We have been successful in capitalizing on the trend toward larger vehicles to increase the productivity of our customers. Technological change is occurring rapidly and TerramEarth has taken advantage of connected devices technology to provide our customers with better services, such as our intelligent farming equipment. With this technology, we have been able to increase farmers' yields by 25%, by using past trends to adjust how our vehicles operate. These advances have led to the rapid growth of our agricultural product line, which we expect will generate 50% of our revenues by 2020.

CTO Statement

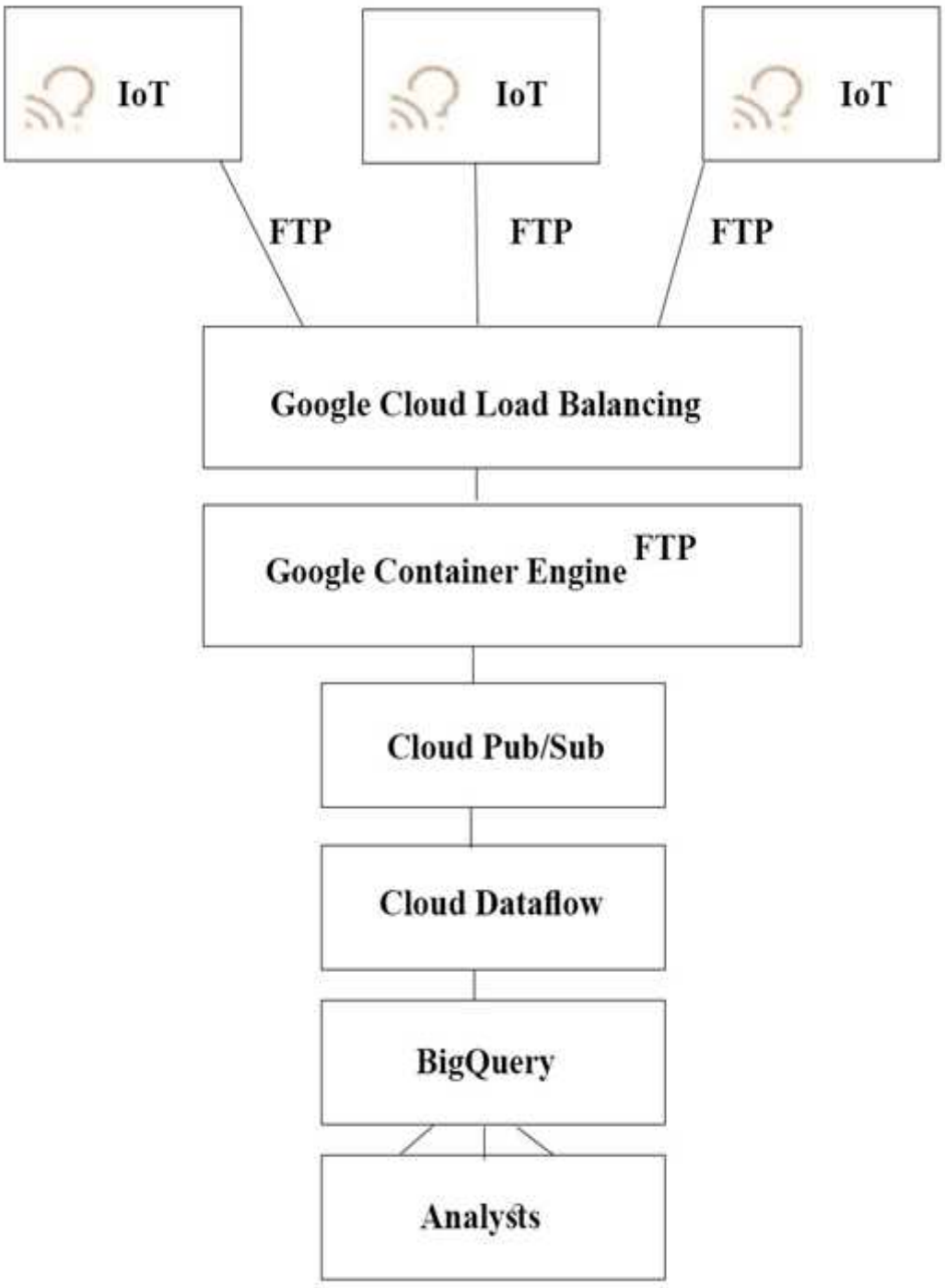
Our competitive advantage has always been in the manufacturing process with our ability to build better vehicles for tower cost than our competitors. However, new products with different approaches are constantly being developed, and I'm concerned that we lack the skills to undergo the next wave of transformations in our industry. Unfortunately, our CEO doesn't take technology obsolescence seriously and he considers the many new companies in our industry to be niche players. My goals are to build our skills while addressing immediate market needs through incremental innovations.

Question: 7

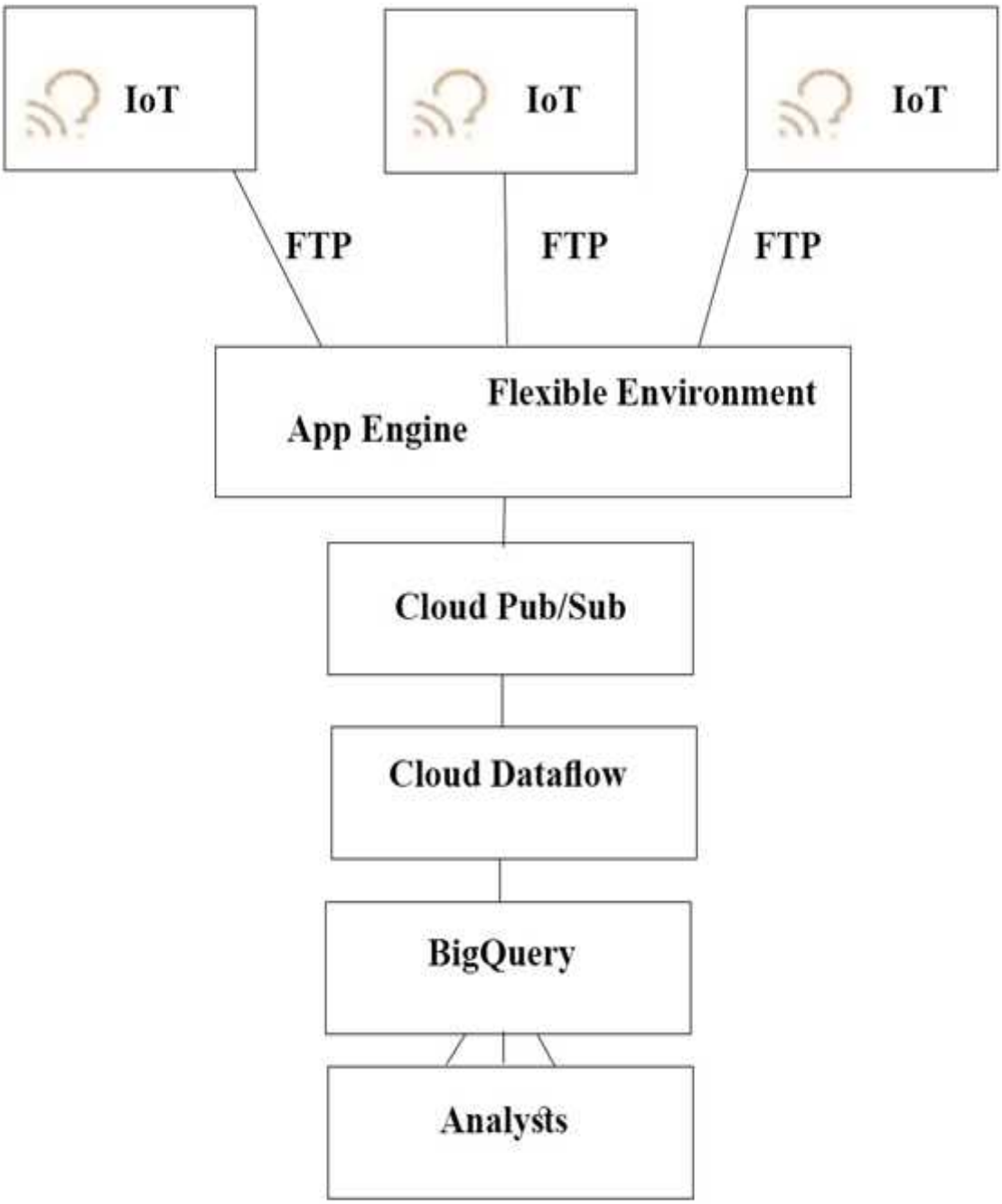
For this question, refer to the TerramEarth case study.

TerramEarth's CTO wants to use the raw data from connected vehicles to help identify approximately when a vehicle in the development team to focus their failure. You want to allow analysts to centrally query the vehicle data. Which architecture should you recommend?

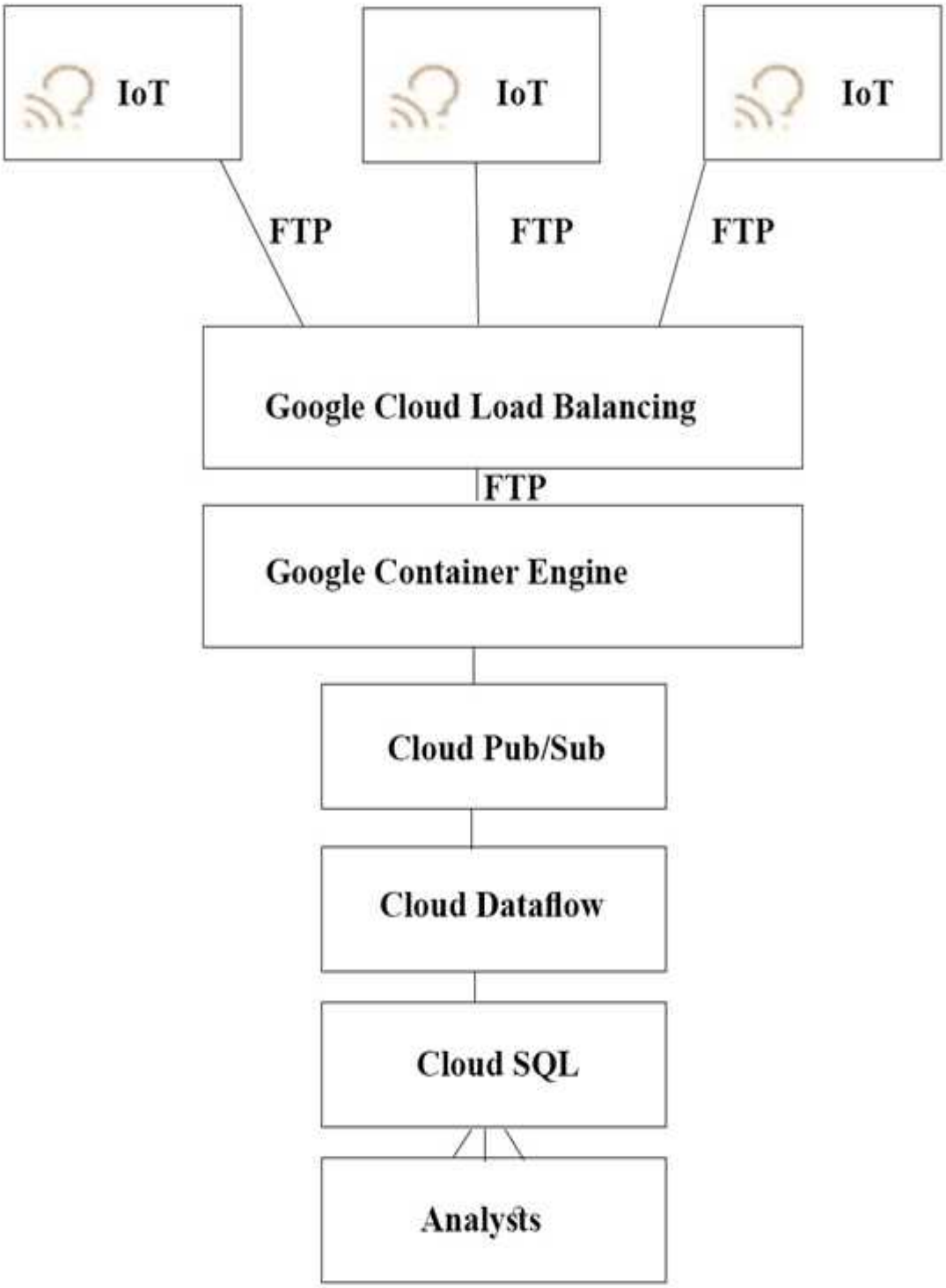
A)



B)


IoT
IoT
IoT
FTP
FTP
FTP
Flexible Environment
App Engine
Cloud Pub/Sub
Cloud Dataflow
BigQuery
Analysts

C)



D)

A. Option A

B. Option B

C. Option C

D. Option D

Answer: A

Explanation:

https://cloud.google.com/solutions/iot/

https://cloud.google.com/solutions/designing-connected-vehicle-platform

https://cloud.google.com/solutions/designing-connected-vehicle-platform#data_ingestion

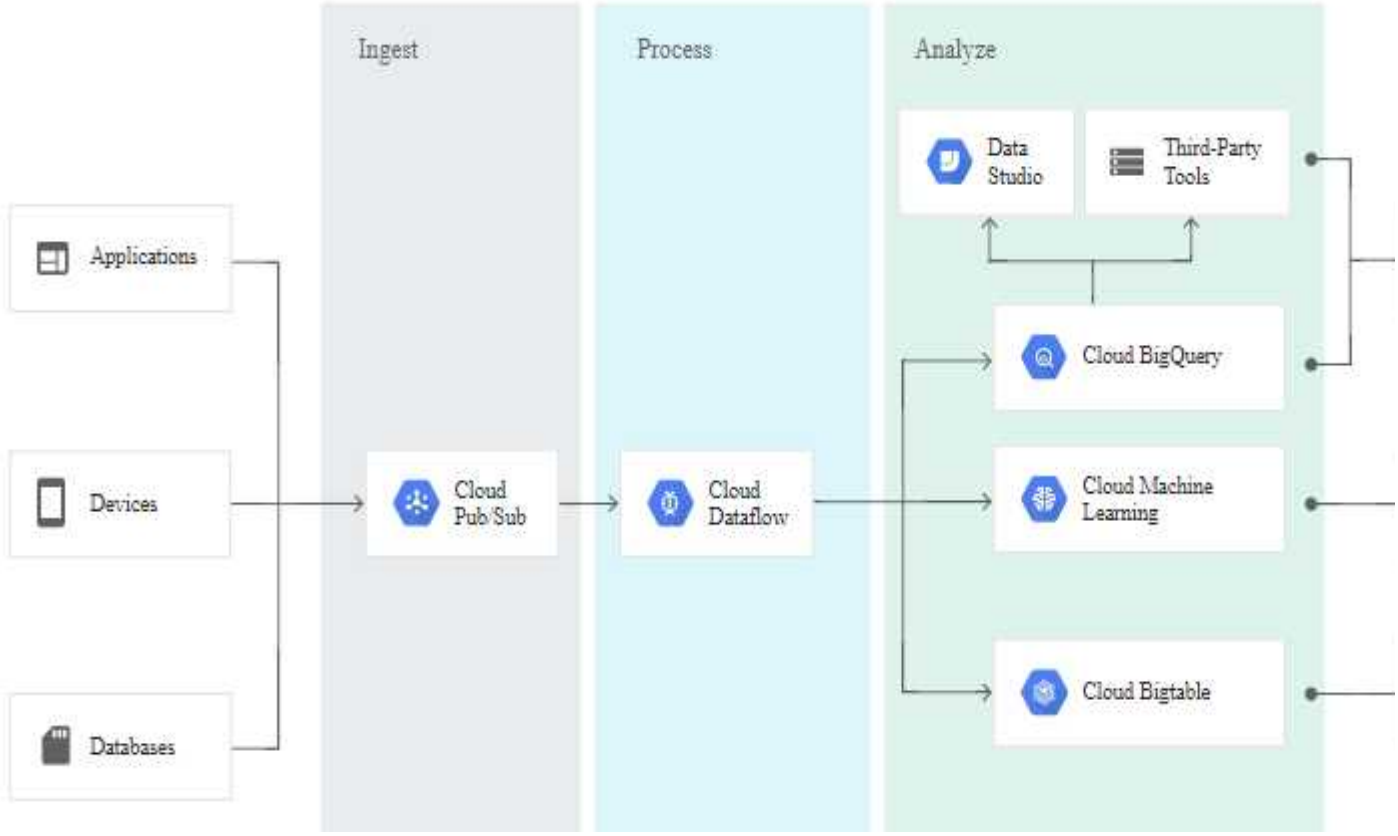
http://www.eweek.com/big-data-and-analytics/google-touts-value-of-cloud-iot-core-for-analyzing-connected-car-data

https://cloud.google.com/solutions/iot/

The push endpoint can be a load balancer.

A container cluster can be used.

Cloud Pub/Sub for Stream Analytics



Reference: https://cloud.google.com/pubsub/

https://cloud.google.com/solutions/iot/

https://cloud.google.com/solutions/designing-connected-vehicle-platform

https://cloud.google.com/solutions/designing-connected-vehicle-platform#data_ingestion

http://www.eweek.com/big-data-and-analytics/google-touts-value-of-cloud-iot-core-for-analyzing-connected-car-data

https://cloud.google.com/solutions/iot/

Question: 8

For this question, refer to the TerramEarth case study.

The TerramEarth development team wants to create an API to meet the company's business requirements. You want the development team to focus their development effort on business value versus creating a custom framework. Which method should they use?

A. Use Google App Engine with Google Cloud Endpoints. Focus on an API for dealers and partners.

B. Use Google App Engine with a JAX-RS Jersey Java-based framework. Focus on an API for the public.

C. Use Google App Engine with the Swagger (open API Specification) framework. Focus on an API for the public.

D. Use Google Container Engine with a Django Python container. Focus on an API for the public.

E. Use Google Container Engine with a Tomcat container with the Swagger (Open API Specification) framework. Focus on an API for dealers and partners.

Answer: A

Explanation:

https://cloud.google.com/endpoints/docs/openapi/about-cloud-endpoints?hl=en_US&_ga=2.21787131.-1712523161.1522785064

https://cloud.google.com/endpoints/docs/openapi/architecture-overview

https://cloud.google.com/storage/docs/gsutil/commands/test

Develop, deploy, protect and monitor your APIs with Google Cloud Endpoints. Using an Open API Specification or one of our API frameworks, Cloud Endpoints gives you the tools you need for every phase of API development.

From scenario:

Business Requirements

Decrease unplanned vehicle downtime to less than 1 week, without increasing the cost of carrying surplus inventory

Support the dealer network with more data on how their customers use their equipment to better position new products and services

Have the ability to partner with different companies – especially with seed and fertilizer suppliers in the fast-growing agricultural business – to create compelling joint offerings for their customers.

Reference: https://cloud.google.com/certification/guides/cloud-architect/casestudy-terramearth

Question: 9

For this question, refer to the TerramEarth case study

Your development team has created a structured API to retrieve vehicle data. They want to allow third parties to develop tools for dealerships that use this vehicle event data. You want to support delegated authorization against this data. What should you do?

A. Build or leverage an OAuth-compatible access control system.

B. Build SAML 2.0 SSO compatibility into your authentication system.

C. Restrict data access based on the source IP address of the partner systems.

D. Create secondary credentials for each dealer that can be given to the trusted third party.

Answer: A

Explanation:

https://cloud.google.com/appengine/docs/flexible/go/authorizing-apps

https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#delegate_application_authorization_with_oauth2

Delegate application authorization with OAuth2

Cloud Platform APIs support OAuth 2.0, and scopes provide granular authorization over the methods that are supported. Cloud Platform supports both service-account and user-account OAuth, also called three-legged OAuth.

Reference: https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#delegate_application_authorization_with_oauth2

https://cloud.google.com/appengine/docs/flexible/go/authorizing-apps

Question: 10

For this question, refer to the TerramEarth case study.

TerramEarth plans to connect all 20 million vehicles in the field to the cloud. This increases the volume to 20 million 600 byte records a second for 40 TB an hour. How should you design the data ingestion?

A. Vehicles write data directly to GCS.

B. Vehicles write data directly to Google Cloud Pub/Sub.

C. Vehicles stream data directly to Google BigQuery.

D. Vehicles continue to write data using the existing system (FTP).

Answer: B

Explanation:

https://cloud.google.com/solutions/data-lifecycle-cloud-platform

https://cloud.google.com/solutions/designing-connected-vehicle-platform

Question: 11

For this question, refer to the TerramEarth case study

You analyzed TerramEarth's business requirement to reduce downtime, and found that they can achieve a majority of time saving by reducing customers' wait time for parts You decided to focus on reduction of the 3 weeks aggregate reporting time Which modifications to the company's processes should you recommend?

A. Migrate from CSV to binary format, migrate from FTP to SFTP transport, and develop machine learning analysis of metrics.

B. Migrate from FTP to streaming transport, migrate from CSV to binary format, and develop machine learning analysis of metrics.

C. Increase fleet cellular connectivity to 80%, migrate from FTP to streaming transport, and develop machine learning analysis of metrics.

D. Migrate from FTP to SFTP transport, develop machine learning analysis of metrics, and increase dealer local inventory by a fixed factor.

Answer: C

Explanation:

The Avro binary format is the preferred format for loading compressed data. Avro data is faster to load because the data can be read in parallel, even when the data blocks are compressed.

Cloud Storage supports streaming transfers with the gsutil tool or boto library, based on HTTP chunked transfer encoding. Streaming data lets you stream data to and from your Cloud Storage account as soon as it becomes available without requiring that the data be first saved to a separate file. Streaming transfers are useful if you have a process that generates data and you do not want to buffer it locally before uploading it, or if you want to send the result from a computational pipeline directly into Cloud Storage.

Reference: https://cloud.google.com/storage/docs/streaming

https://cloud.google.com/bigquery/docs/loading-data

Question: 12

For this question refer to the TerramEarth case study.

Which of TerramEarth's legacy enterprise processes will experience significant change as a result of increased Google Cloud Platform adoption.

A. Opex/capex allocation, LAN changes, capacity planning

B. Capacity planning, TCO calculations, opex/capex allocation

C. Capacity planning, utilization measurement, data center expansion

D. Data Center expansion, TCO calculations, utilization measurement

Answer: B

Explanation:

Capacity planning, TCO calculations, opex/capex allocation From the case study, it can conclude that Management (CXO) all concern rapid provision of resources (infrastructure) for growing as well as cost management, such as Cost optimization in Infrastructure, trade up front capital expenditures (Capex) for ongoing operating expenditures (Opex), and Total cost of ownership (TCO)

Question: 13

For this question, refer to the TerramEarth case study.

To speed up data retrieval, more vehicles will be upgraded to cellular connections and be able to transmit data to the ETL process. The current FTP process is error-prone and restarts the data transfer from the start of the file when connections fail, which happens often. You want to improve the reliability of the solution and minimize data transfer time on the cellular connections. What should you do?

A. Use one Google Container Engine cluster of FTP servers. Save the data to a Multi-Regional bucket. Run the ETL process using data in the bucket.

B. Use multiple Google Container Engine clusters running FTP servers located in different regions. Save the data to Multi-Regional buckets in us, eu, and asia. Run the ETL process using the data in the bucket.

C. Directly transfer the files to different Google Cloud Multi-Regional Storage bucket locations in us, eu, and asia using Google APIs over HTTP(S). Run the ETL process using the data in the bucket.

D. Directly transfer the files to a different Google Cloud Regional Storage bucket location in us, eu, and asia using Google APIs over HTTP(S). Run the ETL process to retrieve the data from each Regional bucket.

Answer: D

Explanation:

https://cloud.google.com/storage/docs/locations

Question: 14

For this question, refer to the TerramEarth case study.

TerramEarth's 20 million vehicles are scattered around the world. Based on the vehicle's location its telemetry data is stored in a Google Cloud Storage (GCS) regional bucket (US. Europe, or Asia). The CTO has asked you to run a report on the raw telemetry data to determine why vehicles are breaking down after 100 K miles. You want to run this job on all the data. What is the most cost-effective way to run this job?

A. Move all the data into 1 zone, then launch a Cloud Dataproc cluster to run the job.

B. Move all the data into 1 region, then launch a Google Cloud Dataproc cluster to run the job.

C. Launch a cluster in each region to preprocess and compress the raw data, then move the data into a multi region bucket and use a Dataproc cluster to finish the job.

D. Launch a cluster in each region to preprocess and compress the raw data, then move the data into a region bucket and use a Cloud Dataproc cluster to finish the jo

Answer: D

Explanation:

Storageguarantees 2 replicates which are geo diverse (100 miles apart) which can get better remote latency and availability.

More importantly, is that multiregional heavily leverages Edge caching and CDNs to provide the content to the end users.

All this redundancy and caching means that Multiregional comes with overhead to sync and ensure

consistency between geo-diverse areas. As such, it's much better for write-once-read-many scenarios. This means frequently accessed (e.g. "hot" objects) around the world, such as website content, streaming videos, gaming or mobile applications.

Reference: https://medium.com/google-cloud/google-cloud-storage-what-bucket-class-for-the-best-performance-5c847ac8f9f2

Question: 15

For this question, refer to the TerramEarth case study.

TerramEarth has equipped unconnected trucks with servers and sensors to collet telemetry data. Next year they want to use the data to train machine learning models. They want to store this data in the cloud while reducing costs. What should they do?

A. Have the vehicle' computer compress the data in hourly snapshots, and store it in a Google Cloud storage (GCS) Nearline bucket.

B. Push the telemetry data in Real-time to a streaming dataflow job that compresses the data, and store it in Google BigQuery.

C. Push the telemetry data in real-time to a streaming dataflow job that compresses the data, and store it in Cloud Bigtable.

D. Have the vehicle's computer compress the data in hourly snapshots, a Store it in a GCS Coldline bucket.

Answer: D

Explanation:

Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example:

Cold Data Storage - Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

Disaster recovery - In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

Reference: https://cloud.google.com/storage/docs/storage-classes

Question: 16

For this question refer to the TerramEarth case study

Operational parameters such as oil pressure are adjustable on each of TerramEarth's vehicles to increase their efficiency, depending on their environmental conditions. Your primary goal is to increase the operating efficiency of all 20 million cellular and unconnected vehicles in the field How can you accomplish this goal?

A. Have your engineers inspect the data for patterns, and then create an algorithm with rules that make operational adjustments automatically.

B. Capture all operating data, train machine learning models that identify ideal operations, and run locally to make operational adjustments automatically.

C. Implement a Google Cloud Dataflow streaming job with a sliding window, and use Google Cloud Messaging (GCM) to make operational adjustments automatically.

D. Capture all operating data, train machine learning models that identify ideal operations, and host in Google Cloud Machine Learning (ML) Platform to make operational adjustments automatically.

Answer: B

Explanation:

Question: 17

Your agricultural division is experimenting with fully autonomous vehicles.

You want your architecture to promote strong security during vehicle operation.

Which two architecture should you consider?

Choose 2 answers: