ĐIỆN TOÁN ĐÁM MÂY
(Cloud Computing)

PRACTICES
Practice 6 – Cloud Database

Presenter: **Dr. Nguyen Dinh Long**
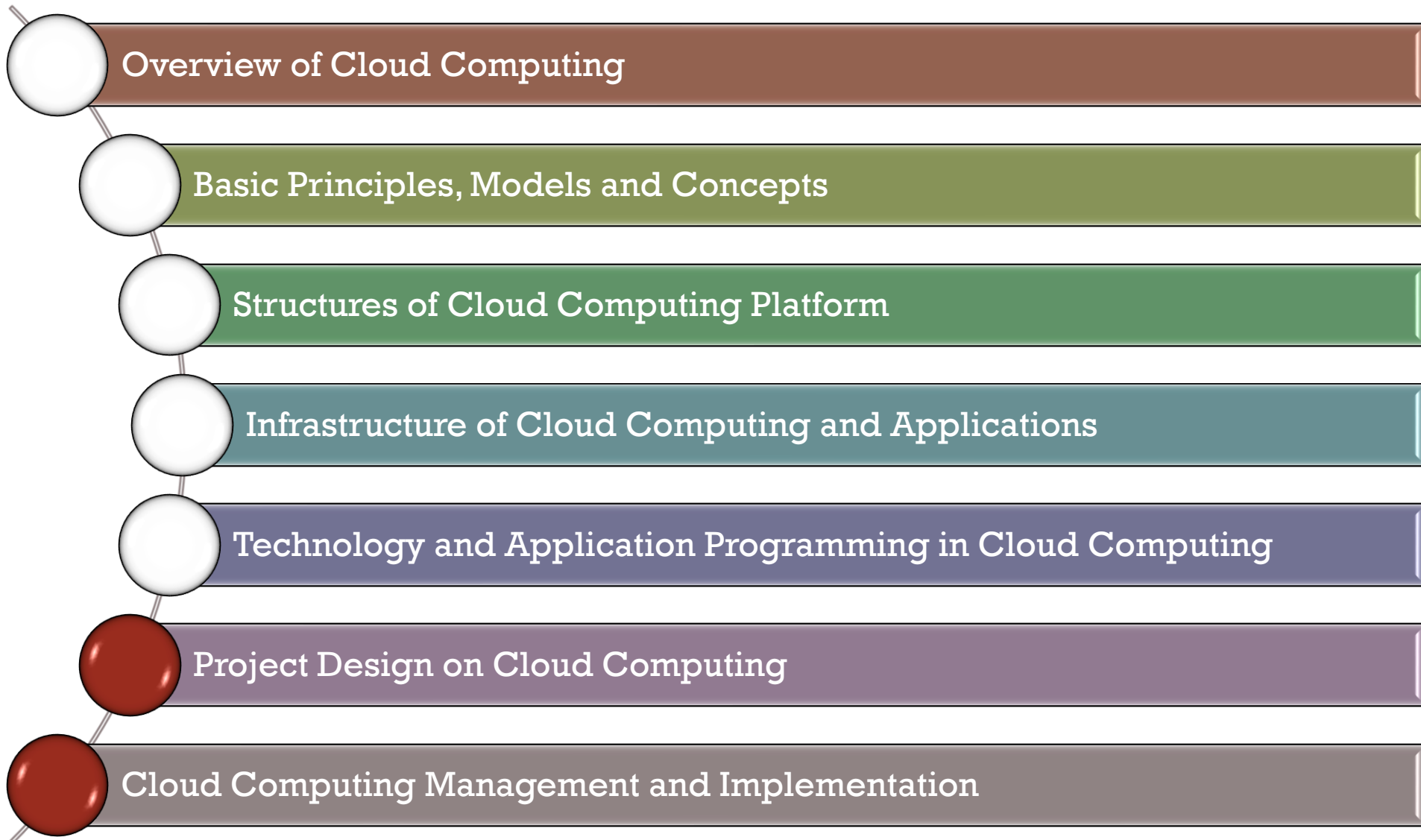
Email: dinhlonghcmut@gmail.com

Phone: +84 947 229 599

Google-site: https://sites.google.com/view/long-dinh-nguyen

Nov. 2022

Outline

- Overview of Cloud Computing
- Basic Principles, Models and Concepts
- Structures of Cloud Computing Platform
- Infrastructure of Cloud Computing and Applications
- Technology and Application Programming in Cloud Computing
- Project Design on Cloud Computing
- Cloud Computing Management and Implementation

References

Main:

- Thomas Erl, Zaigham Mahmood, and Ricardo Puttini. 2013. *Cloud Computing Concepts, Technology & Architecture*. Prentice Hall.

- Michael J. Kavis. 2014. *Architecting the Cloud: Design Decisions for Cloud Computing Service Models*. Wiley

- Arshdeep Bahga, and Vijay Madisetti. 2013. *Cloud Computing: A Hands-On Approach*. CreateSpace Independent Publishing Platform

More:

- Rajkuma Buyya, Jame Broberg and Andrzej Goscinski. 2011. *Cloud Computing –Principles and paradigms*, Wiley

- Nick Antonopoulos, and Lee Gillam. 2010. *Cloud Computing - Principles, Systems and Applications*, Springer-Verlag London Limited.

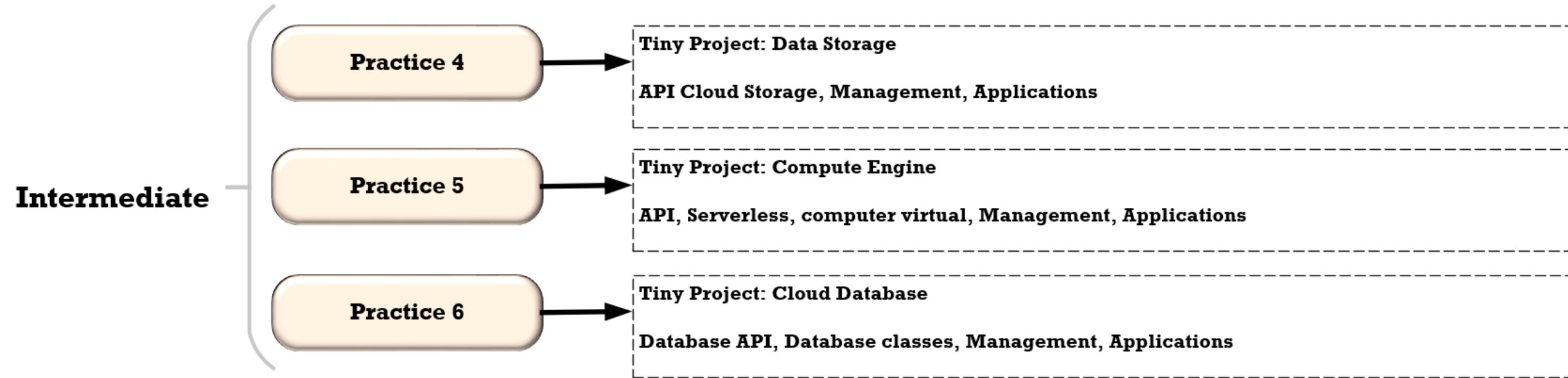- Slides here are modified from several sources in Universities and Internet.

Cloud Computing: Practices

Levels: Beginning (3 weeks) – **Intermediate (3 weeks)** – Advanced (3 weeks)

Groups: 9 with 5 person/group

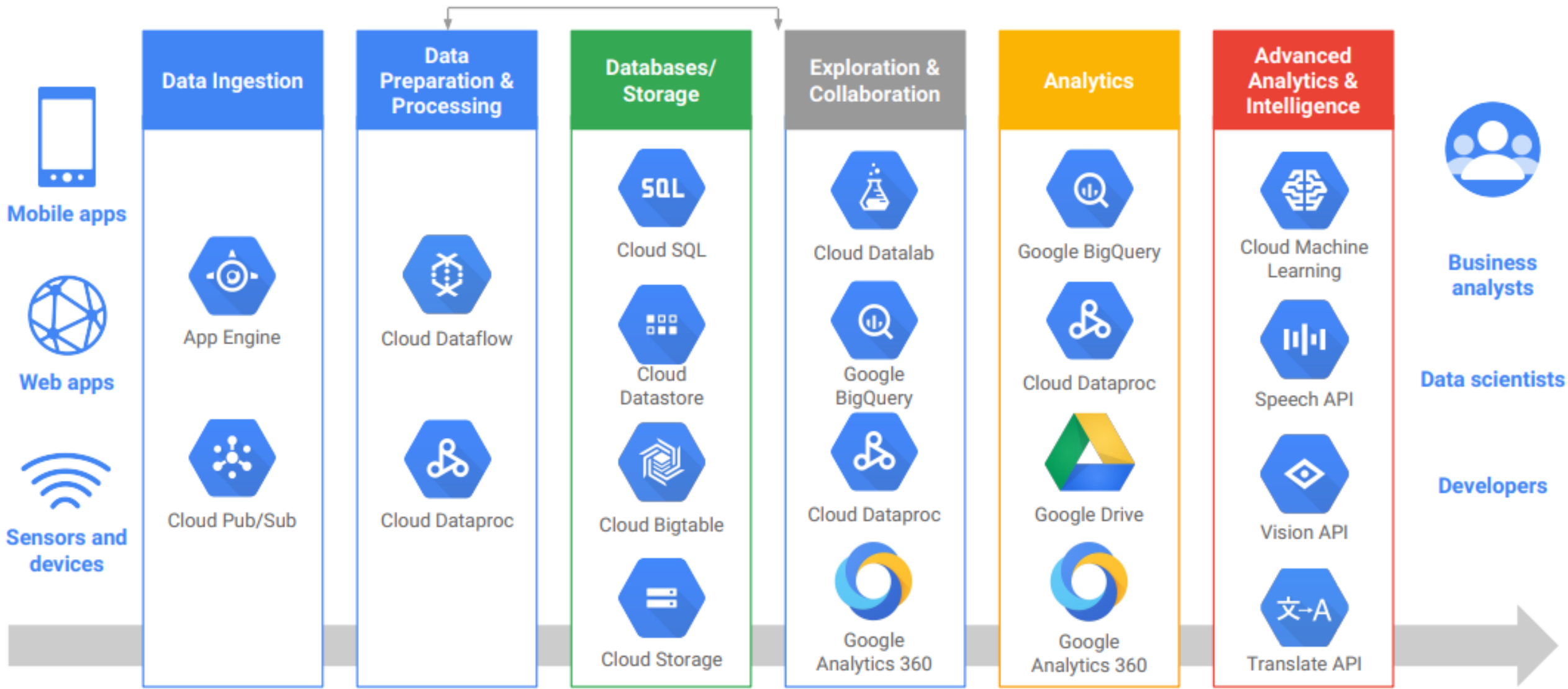Practice: submit a report for each group, submit to our Google Classroom

Cloud Computing: Practices

.

Intermediate

Practice 4 ➝ Tiny Project: Data Storage

API Cloud Storage, Management, Applications

Practice 5 ➝ Tiny Project: Compute Engine

API, Serverless, computer virtual, Management, Applications

Practice 6 ➝ Tiny Project: Cloud Database

Database API, Database classes, Management, Applications

Content of Practice 6

1. Introduction of Database services in Google Cloud

2. Type of Google Cloud

3. Considerations for Cloud Database

4. Cloud Database management

5. Applications

6

Google Cloud Database

Mobile apps

Web apps

Sensors and devices

Data Ingestion	Data Preparation & Processing	Databases/ Storage	Exploration & Collaboration	Analytics	Advanced Analytics & Intelligence
App Engine	Cloud Dataflow	Cloud SQL	Cloud Datalab	Google BigQuery	Cloud Machine Learning
Cloud Pub/Sub	Cloud Dataproc	Cloud Datastore	Google BigQuery	Cloud Dataproc	Speech API
		Cloud Bigtable	Cloud Dataproc	Google Drive	Vision API
		Cloud Storage	Google Analytics 360	Google Analytics 360	Translate API

Business analysts

Data scientists

Developers

7

Google Cloud Database

❑ **What is a Cloud Database?**

▪ A cloud database is a database that is deployed, delivered, and accessed in the cloud. Cloud databases organize and store structured, unstructured, and semi-structured data just like traditional on-premises databases. However, they also provide many of the same benefits of cloud computing, including speed, scalability, agility, and reduced costs.

Cloud database defined

A cloud database is a database built to run in a public or hybrid cloud environment to help organize, store, and manage data within an organization. Cloud databases can be offered as a managed database-as-a-service (DBaaS) or deployed on a cloud-based virtual machine (VM) and self-managed by an in-house IT team.

Cơ sở dữ liệu đám mây là cơ sở dữ liệu được xây dựng để chạy trong môi trường đám mây công cộng hoặc đám mây phức hợp nhằm giúp tổ chức, lưu trữ và quản lý dữ liệu trong một tổ chức/dự án. Cơ sở dữ liệu đám mây có thể được cung cấp dưới dạng dịch vụ cơ sở dữ liệu được quản lý (DBaaS) hoặc được triển khai trên máy ảo (VM) dựa trên đám mây và do nhóm CNTT nội bộ tự quản lý.
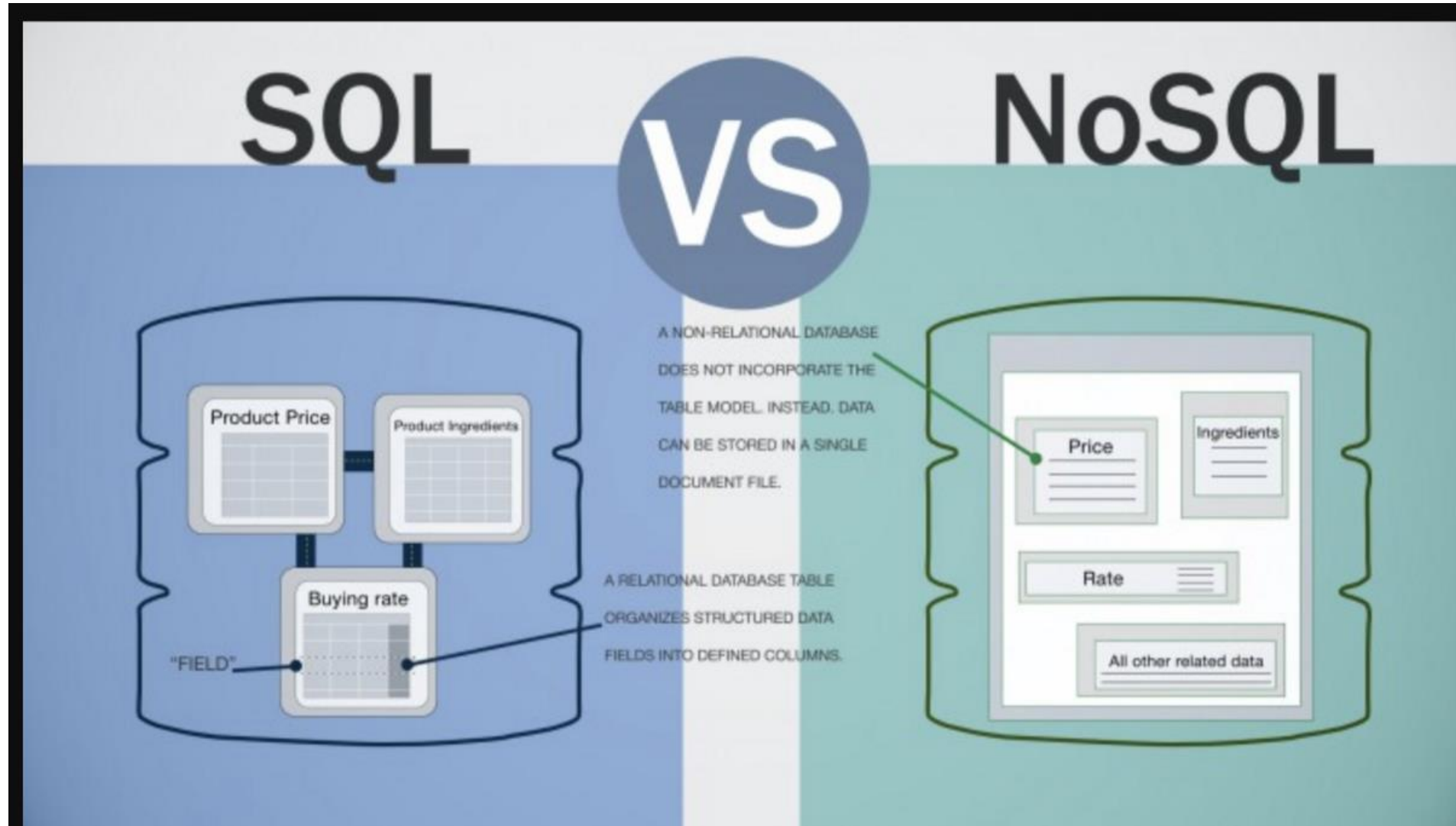
Google Cloud Database

❑ **Why use a cloud database?**

▪ The amount of data generated and collected today is growing exponentially. It's not only more varied, but also wildly disparate. Data can now reside across on-premises databases and distributed cloud applications and services, making it difficult to integrate using traditional approaches. In addition, real-time data processing is becoming essential to business success—delays and lags in data delivery to mission-critical applications could have catastrophic consequences.

▪ As cloud adoption accelerates and the way we use data continues to evolve, legacy databases face significant challenges.

▪ Cloud databases provide flexibility, reliability, security, affordability and more. Providing a solid foundation for building modern business applications. In particular, they can rapidly adapt to changing workloads and demands without increasing the workload of already overburdened teams.
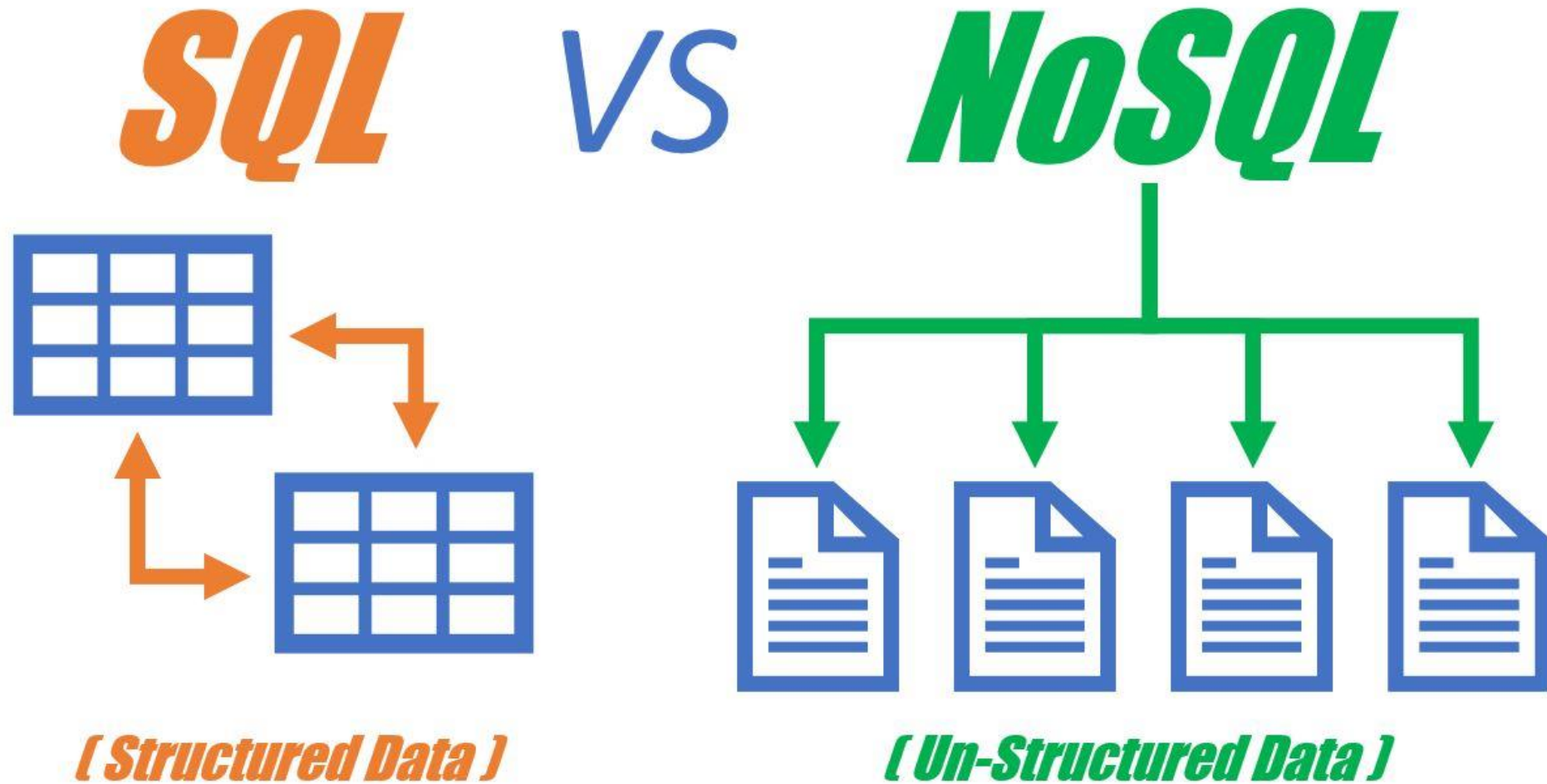
Google Cloud Database

❑ **Why use a cloud database?**

Google Cloud Database

□ **Why use a cloud database?**

SQL VS NoSQL

(Structured Data)   (Un-Structured Data)

Google Cloud Database

❑ **Advantages of cloud databases:**

.

Reduced operational overhead
Cloud databases eliminate the management and maintenance of any physical infrastructure. Your cloud provider is responsible for provisioning, updating, and maintaining all the hardware, operating systems, and database software.

Lower total cost of ownership (TCO)
The cloud service provider owns and operates infrastructure allowing teams to focus on building applications. In addition, pay-as-you-go options lets you provision what you need, when you need it, and scale up or down depending on your usage.

Safe, secure data
Cloud providers invest in the best technologies and experts to offer multiple layers of protection and centralized security policies that can help protect customer data without slowing innovation.

Improved agility and scalability
You can launch a new cloud database or decommission one in minutes. This allows you to test, operationalize, and validate new ideas faster. Plus, cloud databases can dynamically scale as your applications grow and deliver consistent performance under high load.

Flexible database options
You can choose purpose-built cloud databases with the capabilities and performance that match your specific use case and application needs.

Better reliability
Cloud platforms, including cloud databases, come with a host of built-in features designed to maintain constant connectivity and fulfill SLAs, including high availability, automated backups, and robust disaster recovery.

Google Cloud Database

❑ **Advantages of cloud databases:**

.

Reduced ope... ...**d scalability**
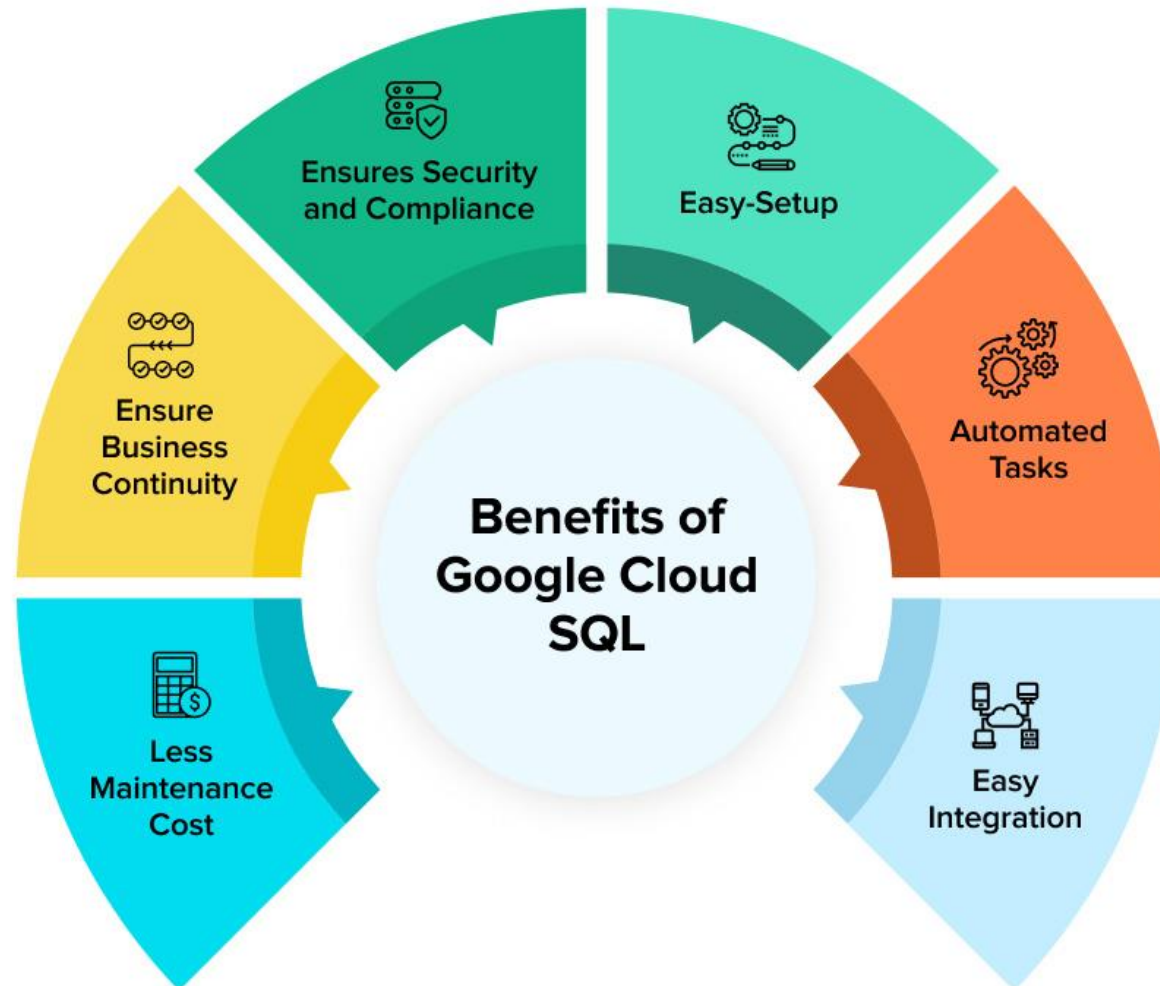
Cloud databases eliminat... ...ase or decommission one
maintenance of any physi... ..., operationalize, and
provider is responsible fo... ...ud databases can
maintaining all the hardw... ...ions grow and deliver
database software. ...load.

Lower total co... ...**e options**

The cloud service provide... ...d databases with the
allowing teams to focus o... ...match your specific use
addition, pay-as-you-go o...
you need, when you need...
depending on your usage.

Safe... ...**ility**

Cloud providers invest in... ...databases, come with a
to offer multiple layers of... ...o maintain constant
security policies that can... ...uding high availability,
without slowing innovati... ...aster recovery.

Benefits of Google Cloud SQL: Ensures Security and Compliance, Easy-Setup, Automated Tasks, Easy Integration, Less Maintenance Cost, Ensure Business Continuity

net solutions

13

Google Cloud Database

❑ **Considerations for cloud databases:**

While the benefits of cloud databases can help organizations address many modern obstacles that impede growth and digital transformation, there are some common considerations of cloud databases to keep in mind as you plan your migration to the cloud.

- Vendor lock-in

- Difficulty integrating data with other systems

- Complex and lengthy migrations

- Underestimating cloud costs

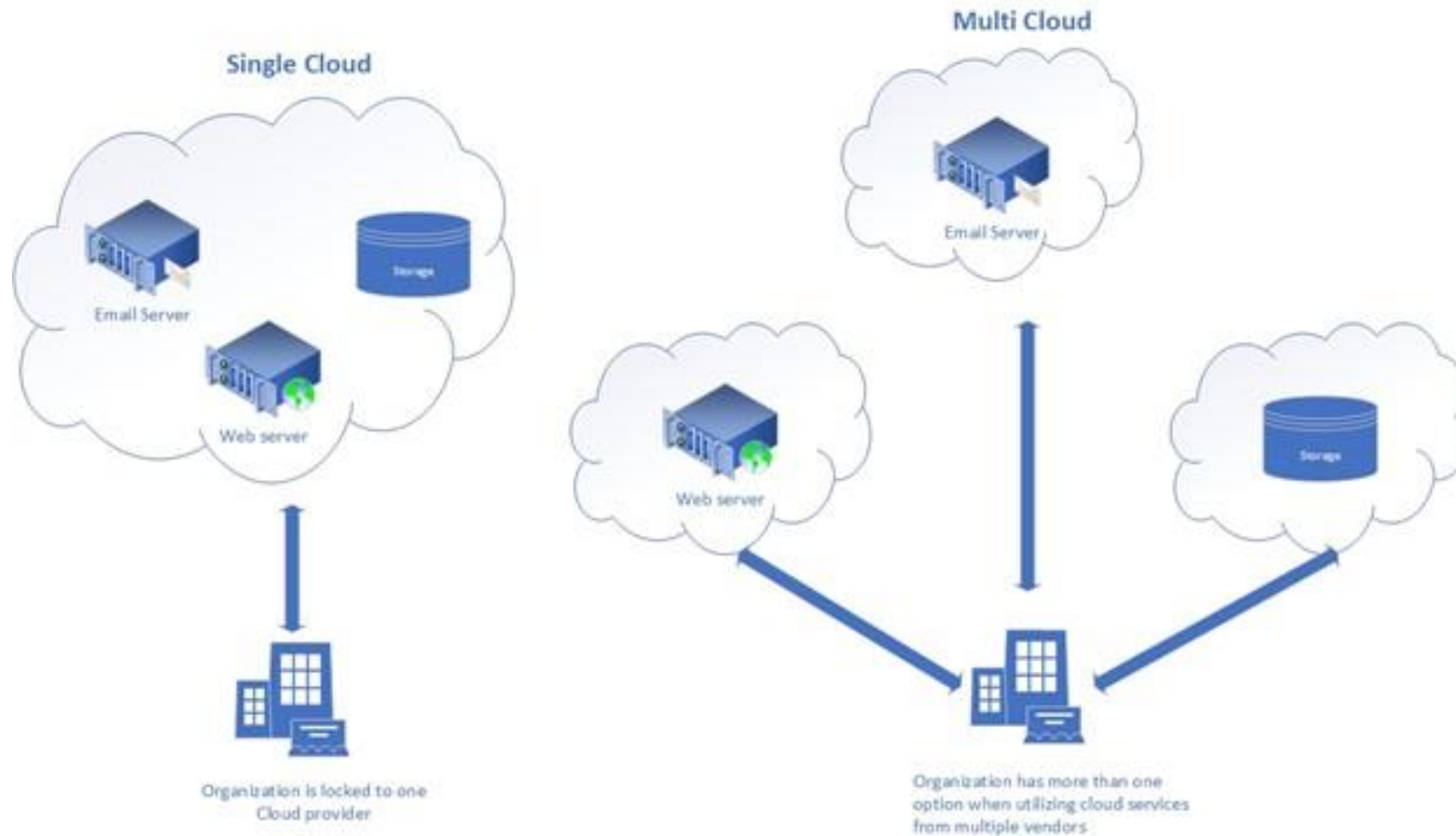- Possibility of connection downtime

- Cloud security concerns

The above isn't necessarily a list of cloud database disadvantages, but rather factors that should be carefully weighed when selecting a cloud service provider. You'll need to take the time to evaluate the best cloud database and service model that meets your specific business needs.

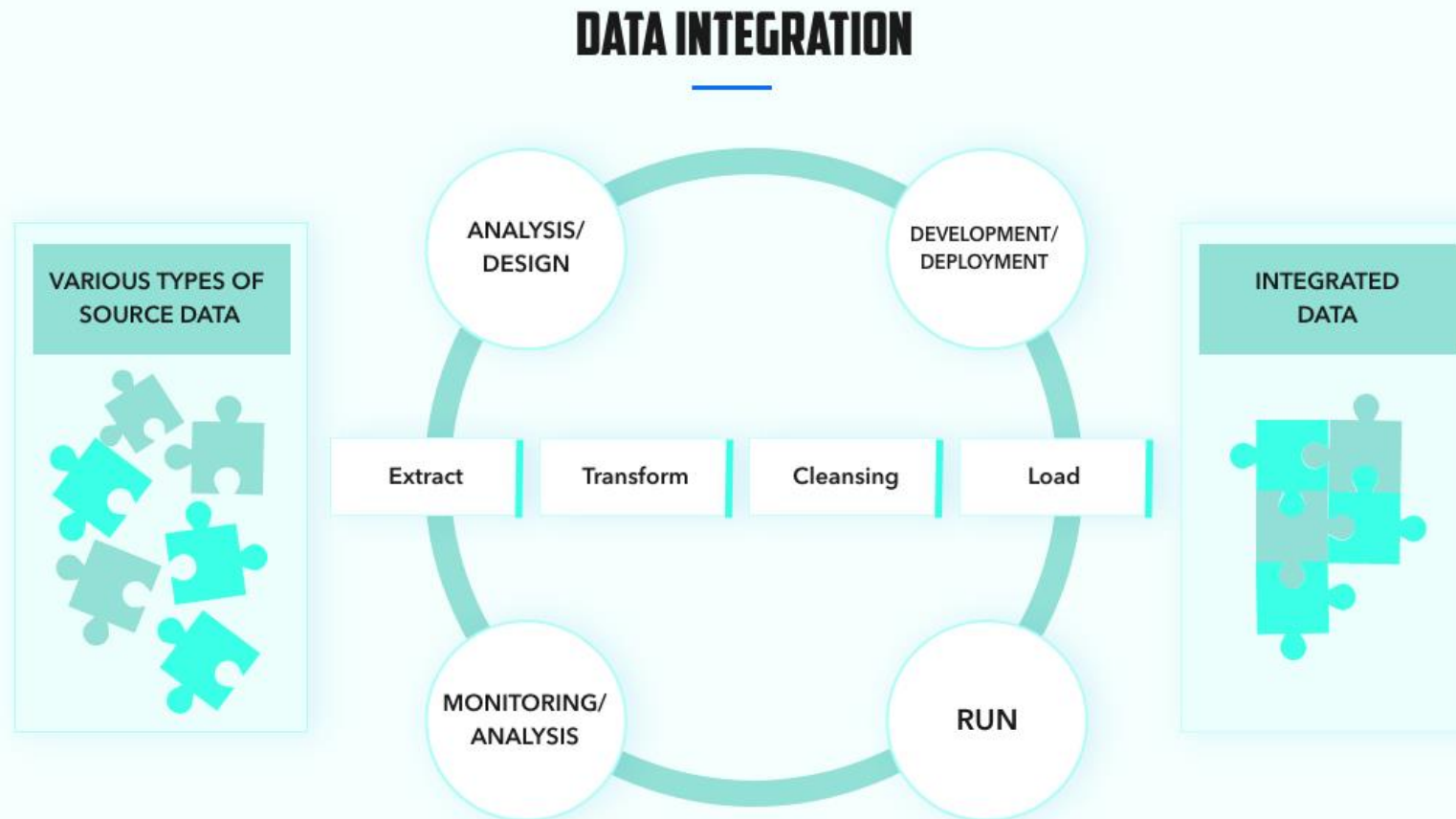# Google Cloud Database

❑ **Considerations for cloud databases:**

▪ Vendor lock-in

Single Cloud

Email Server

Storage

Web server

Organization is locked to one
Cloud provider

Multi Cloud

Email Server

Web server

Storage

Organization has more than one
option when utilizing cloud services
from multiple vendors

Google Cloud Database

❑ **Considerations for cloud databases:**

▪ Difficulty integrating data with other systems

DATA INTEGRATION

VARIOUS TYPES OF SOURCE DATA

ANALYSIS/ DESIGN

DEVELOPMENT/ DEPLOYMENT

INTEGRATED DATA

Extract | Transform | Cleansing | Load
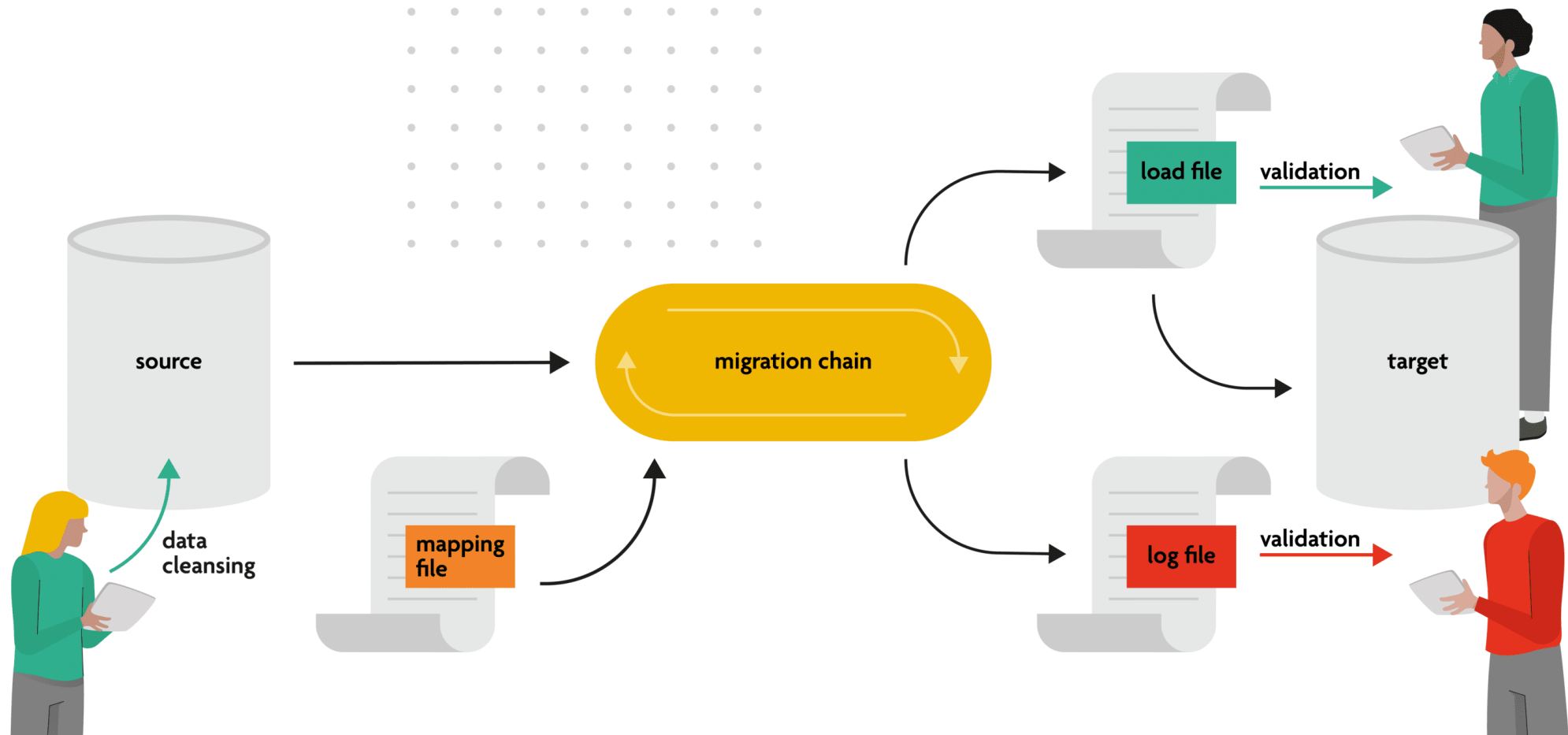
MONITORING/ ANALYSIS

RUN

Google Cloud Database

❑ **Considerations for cloud databases:**

▪ Complex and lengthy migrations

Google Cloud Database

❑ **Considerations for cloud databases:**

▪ Underestimating cloud costs

Hybrid Cloud Cost Considerations

- Management and Support Fees

- Data Transfer

- Storage Fees

- Hybrid Cloud Software and Server Fees

- Customization and Integration Fees

- Compliance Fees

Google Cloud Database

❑ **Considerations for cloud databases:**

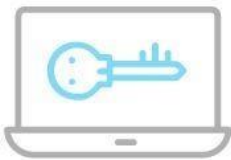▪ Possibility of connection downtime

Google Cloud Database

☐ **Considerations for cloud databases:**

▪ Cloud security concerns

68% Misconfiguration of the cloud platform/ wrong setup

58% Unauthorized access

52% Insecure interfaces /APIs

50% Hijacking of accounts, services or traffic

43% External sharing of data

36% Malicious insiders

33% Foreign state-sponsored cyber attacks

28% Denial of service attacks

Data Breaches — 1

Inadequate Identity and Access Management — 2

Insecure APIs — 3

System Vulnerabilities — 4

Account or Service Hijacking — 5

Malicious Insider Threats — 6

Advanced Persistent Threats (APTs) — 7

Malware Injection — 8

Data Loss — 9

Insufficient Due Diligence — 10

Poor IP Protection — 11

Abuse of Cloud Services — 12

DoS Attacks — 13

Vulnerabilities Caused by Shared Technology — 14

Communication with CSPs — 15

VERITIS

Google Cloud Database

❏ **Types of cloud databases:**

Like a traditional on-premises database, cloud databases can be classified into relational databases and non-relational databases.

▪ **Relational cloud databases** consist of one or more tables of columns and rows and allow you to organize data in predefined relationships to understand how data is logically related. These databases typically use a fixed data schema, and you can use structured query language (SQL) to query and manipulate data. They are highly consistent, reliable, and best suited to dealing with large amounts of structured data.

Examples of relational databases include SQL Server, Oracle, MySQL, PostgreSQL, Spanner, and Cloud SQL.

▪ **Non-relational cloud databases** store and manage unstructured data, such as email and mobile message text, documents, surveys, rich media files, and sensor data. They don't follow a clearly-defined schema like relational databases and allow you to save and organize information regardless of its format.

Examples of non-relational databases include MongoDB, Redis, Cassandra, Hbase, and Cloud Bigtable..

Which Database should I use?

RELATIONAL

Cloud SQL
Managed MySQL, PostgreSQL, SQL Server

Cloud Spanner
Cloud-native with large scale, consistency, 99.999% availability

Bare Metal
Lift and shift Oracle workloads to Google Cloud

Good For:

General purpose SQL DB

RDBMS+ scale, HA, HTAP

RDBMS+ scale, HA, HTAP

Use Case:

- Web frameworks
- ERP
- CRM
- Ecommerce and web
- SaaS application

- Gaming
- Global financial ledger
- Supply chain/ inventory management

- Legacy applications
- Data center retirement

NON-RELATIONAL (NO SQL)

DOCUMENT

Firestore
Cloud Native, serverless, NoSQL document database, backend-as-a-service, global strong consistency, 99.999% SLA

KEY VALUE

Cloud Bigtable
Cloud-native NoSQL wide-column store for large scale, low-latency workloads

Good For:

Large scale, complex hierarchical data

Heavy read + write, events

Use Case:

- Mobile/web/ IoT applications
- Real-time sync
- Offline sync
- Personalized apps

- Personalization
- Adtech
- Recommendation engines
- Fraud detection

IN MEMORY

Memory Store
Fully managed Redis and Memcached for sub-millisecond data access

Good For:

In-memory and Key-value store

Use Case:

- Caching
- Gaming
- Leaderboard
- Social chat or news feed

- Session store
- Personalization
- Adtech

Google Cloud Database

Cloud SQL

❑ **Cloud SQL:**

Fully managed relational database service for MySQL, PostgreSQL, and SQL Server with rich extension collections, configuration flags, and developer ecosystems.

- Reduce maintenance costs with fully managed MySQL, PostgreSQL and SQL Server databases
- Ensure business continuity with reliable and secure services backed by 24/7 SRE team
- Automate database provisioning, storage capacity management, and other time-consuming tasks
- Database observability made easy for developers with Cloud SQL Insights
- Easy integration with existing apps and Google Cloud services like GKE and BigQuery

Secure and compliant	Scale as you go	Set up in minutes
Data encryption at rest and in transit. Private connectivity with Virtual Private Cloud and user-controlled network access with firewall protection. Compliant with SSAE 16, ISO 27001, PCI DSS, and HIPAA.	Scale your instances effortlessly with a single API call whether you start with simple testing or you need a highly available database in production.	Standard connection drivers and built-in migration tools allow you to create and connect to your first database in just a few minutes.

Google Cloud Database

❑ **Cloud Spanner:**

Fully managed relational database with unlimited scale, strong consistency, and up to 99.999% availability.

- Get all the benefits of relational semantics and SQL with unlimited scale
- Start at any size and scale with no limits as your needs grow. Get started with a free trial.
- Enjoy high availability with zero scheduled downtime and online schema changes
- Deliver high-performance ACID transactions with strong consistency over regions and continents
- Focus on innovation, automating cumbersome tasks such as manual database sharing

Scale as needed with no limits	Make your database highly available anywhere	Do fewer thankless tasks with a simpler experience
Globally distributed, ACID-compliant cloud database that automatically handles replicas, sharding, and transaction processing, so you can quickly scale to meet any usage pattern and ensure success of products.	Cloud Spanner is built on Google's dedicated network and battle tested by Google services used by billions. It offers up to 99.999% availability with zero downtime for planned maintenance and schema changes.	IT admins and DBAs are inundated with operating databases. With Cloud Spanner, creating or scaling a globally replicated database now takes a handful of clicks and reduces your cost of maintaining databases.

Google Cloud Database

❏ **BigQuery:** *Cloud data warehouse to power your data-driven innovation*

BigQuery is a completely serverless and cost-effective enterprise data warehouse.

It has built-in machine learning and BI that works across clouds, and scales with your data.

BigQuery's serverless architecture lets you use SQL queries to analyze your data. You can store and analyze your data within BigQuery or use BigQuery to assess your data where it lives. To test how it works for yourself, query data—without a credit card—using the BigQuery sandbox.

Upload data files from local sources	Automate data movement into BigQuery with DTS	Deploy data pipelines into BigQuery with Data Fusion

Google Cloud Database

❑ **Cloud Bigtable:**

A fully managed, scalable NoSQL database service for large analytical and operational workloads with up to 99.999% availability.

- Build responsive applications with consistent, single-digit millisecond latency
- Seamlessly scale to match your storage and throughput needs; no downtime during reconfiguration
- Ideal for use cases such as personalization, fraud detection, real-time analytics, and IoT
- Ensure high availability with multi-primary replication in up to 8 regions
- Easily connect to Google Cloud services such as BigQuery or the Apache ecosystem

Google Bigtable

Fast and performant

Use Cloud Bigtable as the storage engine that grows with you from your first gigabyte to petabyte-scale for low-latency applications as well as high-throughput data processing and analytics.

Seamless scaling and replication

Start with a single node per cluster, and scale to hundreds of nodes dynamically supporting peak demand at low latency. Replication also adds high availability and workload isolation for live serving apps.

Simple and integrated

Fully managed service that integrates easily with big data tools like Hadoop, Dataflow, and Dataproc. Plus, support for the open source HBase API standard makes it easy for development teams to get started.

Google Cloud Database

❑ **Firestore:**

Easily develop rich applications using a fully managed, scalable, and serverless document database.

- Serverless document database that effortlessly scales to meet any demand, with no maintenance

- Accelerate development of mobile, web, and IoT apps with direct connectivity to the database

- Built-in live synchronization and offline mode makes it easy to develop real-time applications

- Fully customizable security and data validation rules to ensure the data is always protected

- Seamless integration with Firebase and Google Cloud services like Cloud Functions and BigQuery

Launch applications and features faster	Effortlessly scale to meet unpredictable demand	Simple and flexible with pay as you go
Firestore offers a great developer experience with built-in live synchronization, offline support, and ACID transactions. These features are available across a robust set of client and server-side libraries.	Firestore automatically scales up and down based on demand. It requires no maintenance, and provides high availability of 99.99–99.999% achieved through strongly consistent data replication.	No-ops database lets you pay only for what you use—no up-front expenditure or underutilized resources. Simplified architecture lets your apps talk directly to Firestore from your mobile or web clients.

27

Google Cloud Database

❑ **Memorystore:**

Reduce latency with scalable, secure, and highly available in-memory service

for Redis and Memcached.

- Build application caches that provide sub-millisecond data access
- 100% compatible with open source Redis and Memcached
- Migrate your caching layer to cloud with zero code change

Focus on building great apps

Memorystore automates complex tasks for open source Redis and Memcached like enabling high availability, failover, patching, and monitoring so you can spend more time coding.

Scale as needed

Scale as your application grows. You can scale reads to over a million QPS with Redis 6 and Read Replicas. Memorystore for Memcached provides clusters as large as 5 TB to meet your caching needs.

Highly available

Memorystore for Redis standard tier instances supports up to five read replicas which are replicated across zones and provide a 99.9% availability SLA, resulting in minimal disruption of applications.

Google Cloud Database

❑ **Cloud database management:**

Cloud databases follow two deployment models: traditional self-managed and managed database as a service (DBaaS).

▪ Traditional self-managed cloud databases: The cloud database is installed and hosted on a virtual machine in a traditional deployment model. With this model, you maintain more control over your database, and your in-house IT team is still responsible for database management and maintenance.

▪ Managed database service: The cloud database is accessed as a service. It runs on the cloud service provider's physical infrastructure, where the provider is responsible for most of the operational, maintenance, and administrative database management tasks. Typically, it includes automated provisioning, scaling, security, upgrade, and health monitoring—everything needed to keep your database consistently available and high performance.

Google Cloud Database

❑ **Cloud database management:**

Cloud databases follow two deployment models: traditional self-managed and managed database as a service (DBaaS).

▪ Traditional self-managed cloud databases: The cloud database is installed and hosted on a virtual machine in a traditional deployment model. With this model, you maintain more control over your database, and your in-house IT team is still responsible for database management and maintenance.
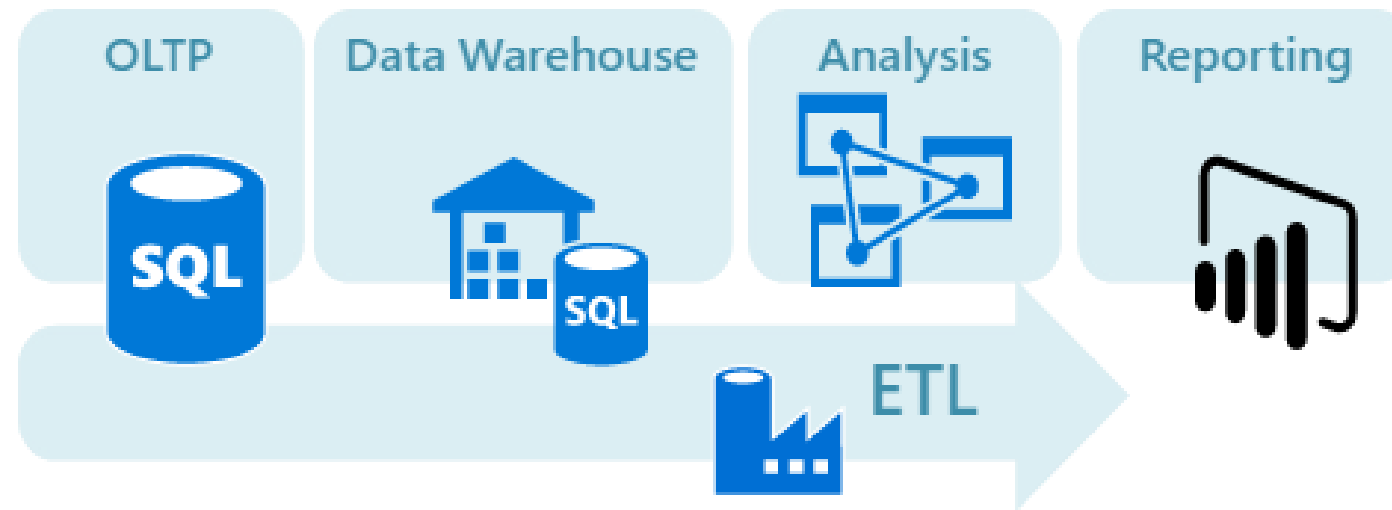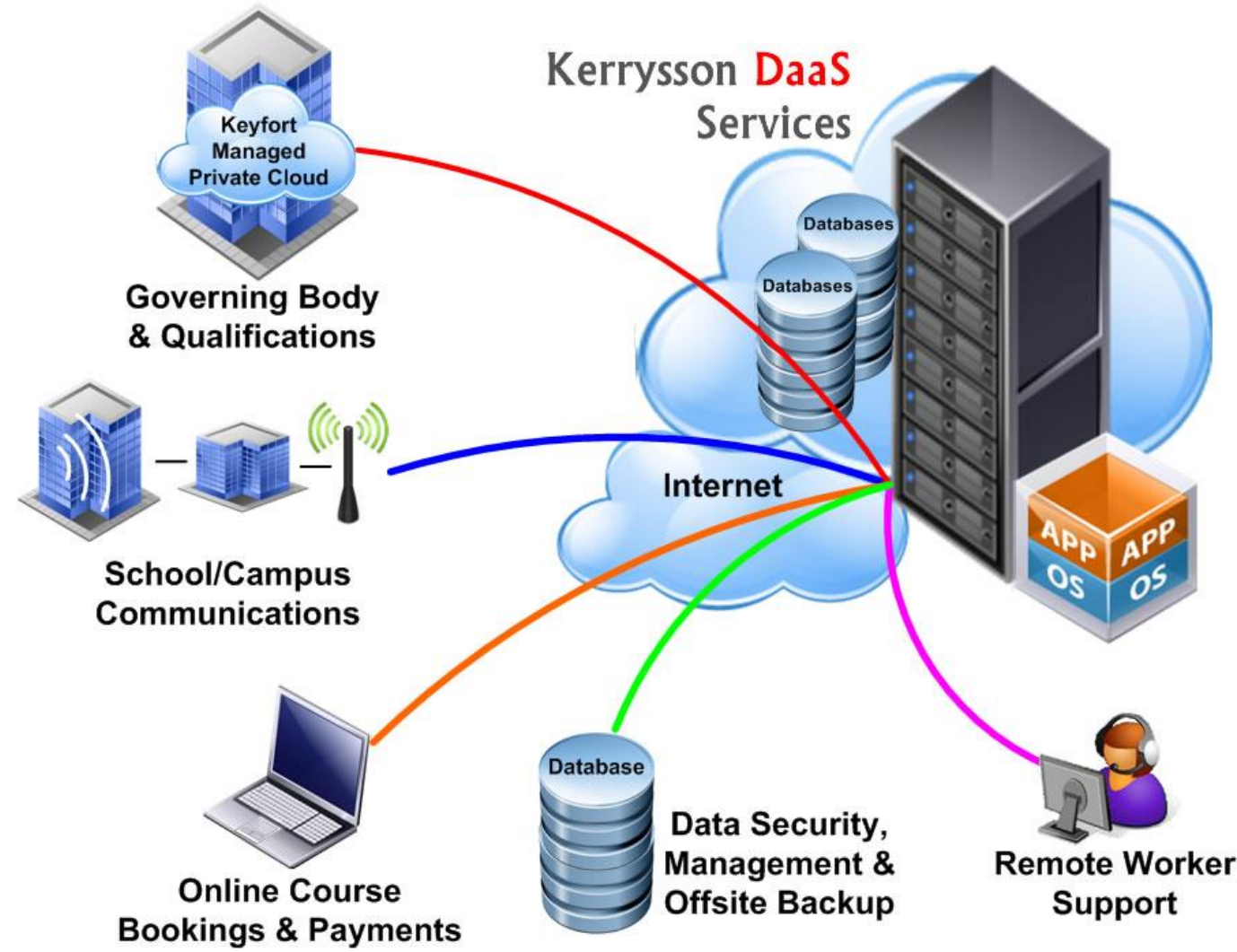
Google Cloud Database

❑ **Cloud database management:**

▪ Managed database service: The cloud database is accessed as a service. It runs on the cloud service provider's physical infrastructure, where the provider is responsible for most of the operational, maintenance, and administrative database management tasks.

Google Cloud Database

❑ **Cloud SQL pricing**

Pricing for Cloud SQL depends on your instance type:

- MySQL and PostgreSQL

- SQL Server

❑ **MySQL and PostgreSQL pricing**

Cloud SQL pricing is composed of the following charges:

- CPU and memory pricing

- Storage and networking pricing

- Instance pricing

Google Cloud Database - Report

TASK I: Benefits of Cloud Database

https://cloud.google.com/learn/what-is-a-cloud-database

➢ **Phân tích lợi ích của Cloud Database**

➢ **Những vấn đề gì cần xem xét khi ứng dụng Cloud Database**

(*bằng Tiếng Việt !!!*)

TASK II: Cloud Database Services

https://cloud.google.com/learn/what-is-a-cloud-database

➢ **Phân tích các tính năng (key features), lợi ích (benefits) của các dịch vụ Cloud Database**

Cloud SQL

Cloud Spanner

BigQuery

Cloud Bigtable

Firestore

Memorystore

(bằng Tiếng Việt !!!)

Google Cloud Database - Report

TASK III: Cloud Database Pricing (cont.)

https://cloud.google.com/sql/pricing

https://cloud.google.com/bigquery/pricing

➢ **Phân tích và tính toán chi phí dịch vụ Cloud Database (Singapore (asia-southeast1), Monthly)**

MySQL and PostgreSQL pricing	
CPU and memory pricing	2 vCPU, 8 GB memory, 1 HA vCPU, 4 GB HA memory

Storage and networking pricing
   o Traditional Storage
   o HA Storage
   o Network
	
	20 GB SSD storage, 100 GB HDD storage, 20 GB backups
	10 GB SSD storage, 60 GB HDD storage, 10 GB backups
	IPv4 addresses (12h/day)

Instance pricing	db-g1-small

Google Cloud Database - Report

TASK III: Cloud Database Pricing

https://cloud.google.com/sql/pricing

https://cloud.google.com/bigquery/pricing

➢ **Phân tích và tính toán chi phí dịch vụ Cloud Database (Jakarta (asia-southeast2), Monthly)**

SQL Server pricing	
CPU and memory pricing	3 vCPU, 16 GB memory, 2 HA vCPU, 8 GB HA memory

Storage and networking pricing
  o Traditional Storage
  o HA Storage
  o Network
	
	20 GB SSD storage, 100 GB HDD storage, 20 GB backups
	10 GB SSD storage, 60 GB HDD storage, 10 GB backups
	IPv4 addresses (12h/day)

Licensing	Standard (12 core hours/day) and Web (12 core hours/day)

Google Cloud Database - Report

TASK III: Cloud Database Pricing

https://cloud.google.com/sql/pricing

https://cloud.google.com/bigquery/pricing

➢ **Phân tích và tính toán chi phí dịch vụ Cloud Database (Hong Kong (asia-east2), Monthly)**

BigQuery pricing	
Analysis pricing ○ On-demand analysis pricing ○ Flat-rate analysis pricing	28 TB 5 slots (Monthly flat-rate commitments)
Storage pricing ○ Active storage ○ Long-term storage	Active logical storage (100 GB), Active physical storage (80 GB) Long-term logical storage (500 GB), Long-term physical storage (250GB)

TASK IV: Cloud Database – BigQuery services
(Enable the BigQuery sandbox)

https://cloud.google.com/bigquery/docs/sandbox

➢ **Khởi tạo dịch vụ BigQuery Sandbox**

☺ Tiến hành theo các bước hướng dẫn (link cung cấp)

☺ Trình tự và ảnh chụp minh chứng

☺ Show các quotas và limits của BigQuery sandbox

Google Cloud Database - Report

TASK V: Cloud Database – BigQuery services

(Enable the BigQuery with Cloud Console)

https://cloud.google.com/bigquery/docs/quickstarts/query-public-dataset-console

➢ **Khởi tạo dịch vụ BigQuery with Cloud console**

☺ Tiến hành theo các bước hướng dẫn (link cung cấp)

☺ Query public data

☺ Load and Query data

☺ Explore the console with bigquery