

CLOUD DATABASES: AWS, GOOGLE CLOUD & AZURE



SUNNY THAKOR : 240160510052

HARDIK MEKHIYA : 240160510027

AKSHIT SONANI : 240160510049

INTRODUCTION TO CLOUD DATABASES

❑ What are Cloud Databases?

A cloud database is a database that runs on a cloud platform instead of on your own computer or servers. It lets you store, manage, and access data through the internet.

Examples: Amazon RDS, Google Cloud SQL, Azure SQL, MongoDB Atlas.

❑ Why They Matter in Modern Applications?

- **Scalability** – Apps like social media or e-commerce can support millions of users without slowing down.
- **Accessibility** – Data is available from anywhere, which is useful for global teams and remote users.
- **Cost efficiency** – Companies pay only for the storage and resources they use.
- **Reliability & Security** – Cloud providers ensure backups, disaster recovery, and protection against data loss.
- **Speed of development** – Developers can build apps faster without worrying about database setup.

Examples: Netflix uses cloud databases to recommend shows to millions of users.



AWS DATABASES OVERVIEW

Amazon Web Services (AWS) offers many cloud database services to match different needs. Each one is designed for specific types of applications.

❑ Amazon RDS (Relational Database Service)

- A managed service for relational databases like MySQL, PostgreSQL, Oracle, and SQL Server.
- AWS handles updates, backups, and scaling.
- Good for apps that need traditional relational data storage (e.g., e-commerce, ERP).

❑ Amazon DynamoDB (NoSQL)

- A fully managed **NoSQL database**.
- Stores data in key-value and document format.
- Extremely fast and scalable, used in apps with millions of requests per second.
- Example: Gaming, IoT, mobile apps, real-time chat apps.

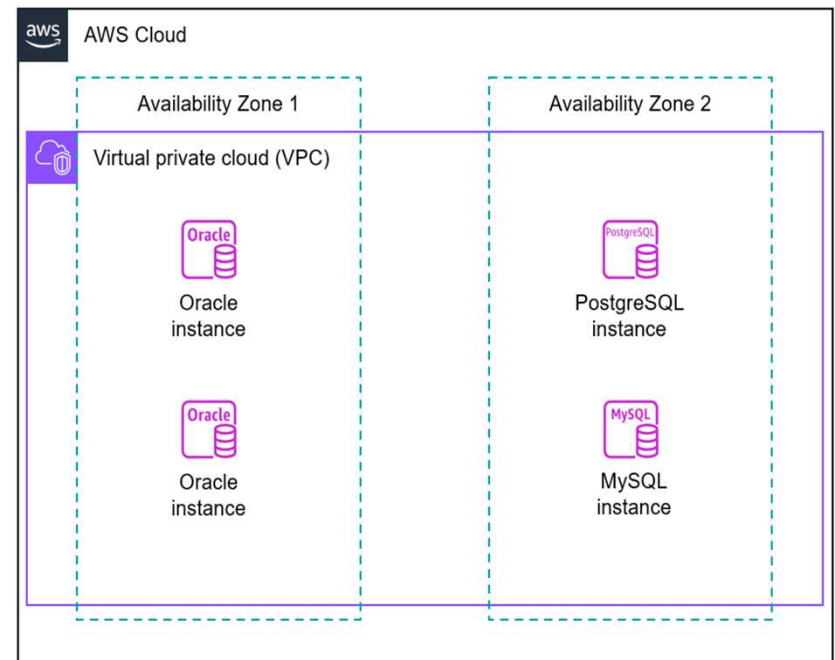


AWS - RDS (RELATIONAL DATABASE)

Amazon Relational Database Service (Amazon RDS) is a web service that makes it easier to set up, operate, and scale a relational database in the AWS Cloud. It provides cost-efficient, resizable capacity for an industry-standard relational database and manages common database administration tasks.

❑ Amazon RDS DB instances

- **VPC (Virtual Private Cloud):** A private network inside AWS where the database instances run securely.
- **Availability Zone 1:** Hosts two Oracle database instances for redundancy or load distribution.
- **Availability Zone 2:** Hosts one PostgreSQL instance and one MySQL instance, giving flexibility to run different database engines.
- **Multi-AZ setup:** By placing databases in different zones, the architecture ensures high availability and disaster recovery (if one AZ goes down, the other still works).





AWS - DYNAMODB(NOSQL)

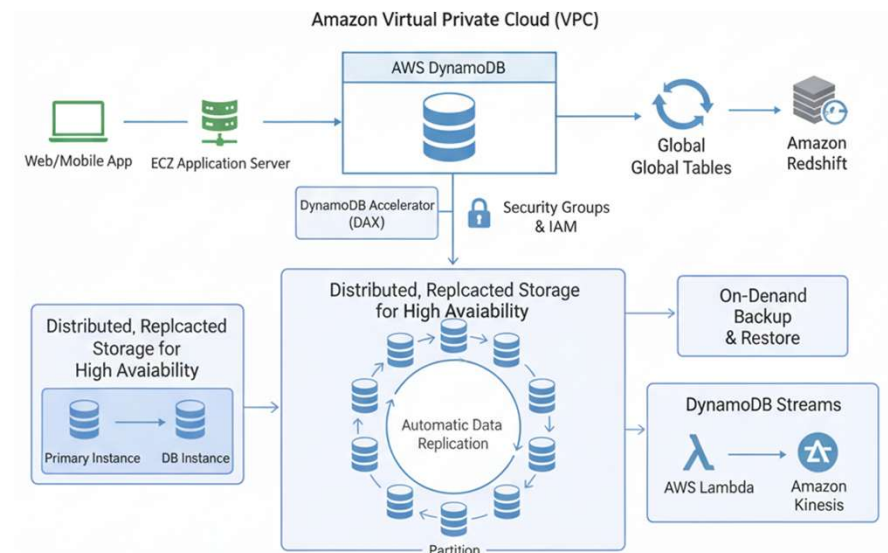
Amazon DynamoDB is a fully managed NoSQL database service that provides fast and predictable performance at any scale.

It is serverless, automatically handling scaling, replication, and backups.

DynamoDB is ideal for applications needing low-latency access, such as gaming, IoT, and real-time analytics.

❑ DynamoDB:

- Data is stored in distributed, replicated storage across multiple partitions and Availability Zones for high availability.
- Automatic data replication ensures fault.
- On-demand backup & restore protects data.
- DynamoDB Streams capture real-time data changes and can trigger AWS Lambda or feed into Amazon Kinesis.
- Integration with Amazon Redshift for analytics.





GOOGLE CLOUD DATABASES OVERVIEW

All Google Cloud databases run on **Google's global infrastructure** (data centers in multiple regions).
Databases are provided as **managed services** — no need to manage physical servers.
High availability through **replication** across zones/regions.

❑ Relational Databases (SQL)

○ Cloud SQL

- Managed relational DB service
- Supports **MySQL, PostgreSQL, SQL Server**
- Use case: traditional apps, ERP, finance

○ Cloud Spanner

- **Relational + globally distributed**
- Horizontal scaling + strong consistency
- Use case: banking, retail, large-scale systems

❑ Non-Relational Databases (NoSQL)

○ Firestore

- Document-based NoSQL
- Real-time sync (great for mobile/web apps)

○ Bigtable

- Wide-column NoSQL
- Petabyte-scale, high throughput
- Use case: IoT, analytics, time-series data



GOOGLE CLOUD - CLOUD SQL

Google Cloud SQL is a fully managed relational database service that supports MySQL, PostgreSQL, and SQL Server. It handles backups, replication, updates, and failover automatically so you don't need to manage servers. It integrates with other Google Cloud services like BigQuery and Cloud Storage for analytics and data management.

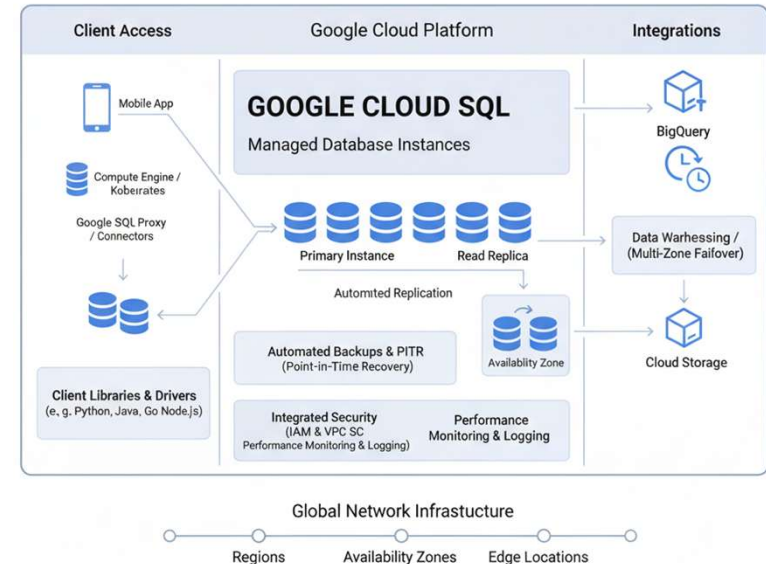
❑ Google Cloud SQL:

- Provides managed database instances (MySQL, PostgreSQL, SQL Server).
- Supports primary instance with read replicas for scaling reads.
- Automated backups & point-in-time recovery.
- High availability with multi-zone failover.

❑ Integrations:

- Works with BigQuery for analytics.
- Triggers event/data pipelines.
- Stores backups or exports in Cloud Storage

GOOGLE CLOUD SQL ARCHITECTURE





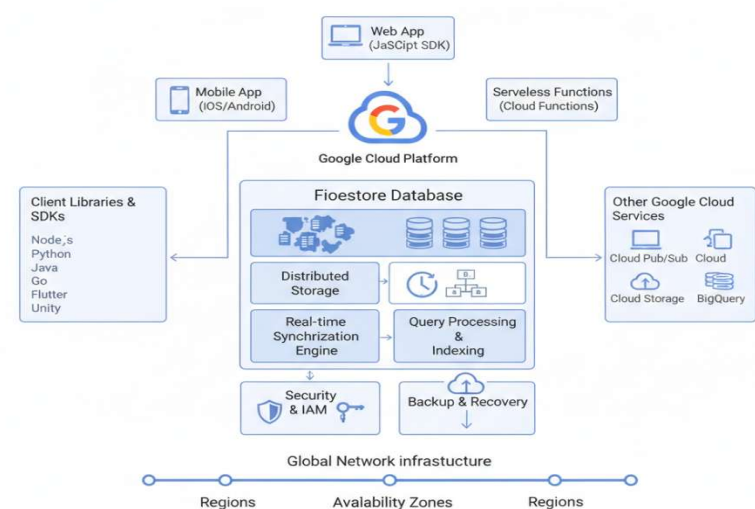
GOOGLE CLOUD - FIRESTORE

Google Cloud Firestore is a NoSQL document database for web and mobile apps. It offers real-time synchronization and offline support, so data stays updated across devices instantly. It's scalable, serverless, and secure, with easy integration into Google Cloud services.

❑ Firestore Database Core:

- **Distributed Storage** stores documents and collections.
- **Real-time Synchronization Engine** keeps data updated instantly across clients.
- **Query Processing & Indexing** supports fast searches.
- **Security & IAM** manages authentication and access.
- **Backup & Recovery** ensures reliability.
- **Global Infrastructure:** Data is replicated across regions and availability zones for performance and fault tolerance.

GOOGLE CLOUD FIRESTORE ARCHITECTURE





Which Database should I use?

#GCPSketchnotes @PVERGADIA THECLOUDGIRL.DEV
03.22.2023



RELATIONAL



Cloud SQL

Managed MySQL, PostgreSQL, SQL Server
SLA: 99.95%



AlloyDB

Managed PostgreSQL-compatible, with 100x faster analytics & 4x faster transaction queries
SLA: 99.99%

RELATIONAL & NON-RELATIONAL



Spanner

Multi-dialect (PostgreSQL, GoogleSQL) database, unlimited scaling
SLA: 99.999%

NON-RELATIONAL (NO SQL)



Firestore

Serverless, document database with built-in cross-client sync and offline caching
SLA: 99.999%



Bigtable

Key-value store for large scale, low latency workloads, Supports HBase APIs
SLA: 99.999%

IN MEMORY



Memorystore

Managed Redis & Memcached
In memory cache & key-value store
SLA: 99.9%

BARE METAL



Oracle

Lift & shift oracle workloads to Google Cloud, run on certified, pre-configured and tuned infrastructure high performance and availability

Fully managed experience for low-touch administration with improved availability, security & governance, support

Good For:

- Fastest lift & shift migrations - OLTP workloads
- Managed experience
- Common API & control plane across database engines

- PostgreSQL workloads that need high performance & high availability or scale
- HTAP (hybrid transaction/analytical processing)
- Migrations off commercial databases

Highest scale and availability requirements without compromising on SQL capabilities, consolidating multiple databases for cost savings

- Rapid and cost-efficient application development (no need for middle tier)
- Easy aggregation from multiple data sources

- Cost-sensitive applications
- Need high throughput & consistent single-digit millisecond latencies irrespective of scale

- Sub-millisecond latencies for reads & writes
- Cache to improve app performance & throughput and reduce costs

Oracle applications and databases migrating to cloud

Use Case:

Web Frameworks



ERP



CRM



Ecommerce & Web



Operational Analytics



ERP, CRM



Financial Services



Ecommerce & Web



SaaS



Order & Inventory Management



Online Banking, payments & ledger



Gaming: player profiles & gameplay data



Electronic Medical Records



Catalog Metadata Management



Web & Mobile Apps



News Feeds, Social Chat, Influencer Engagements



Game Saves, Player Profiles



Retail Catalogs, Point of Sales



Real-Time Analytics (Personalization, Fraud detection)



IoT/Clickstream/Time Series



Feature Stores, Operational Data Hubs & Data Fabrics



Batch Unstructured Data Processing



Financial Markets, Crypto Ledgers



Database caching



Session Store



Jobs and Queues



Leaderboards



Fast Data Ingestion



Legacy Applications



Datacenter Exits



AZURE DATABASES OVERVIEW

Microsoft Azure provides a range of managed database services for different application needs.

❑ **Azure SQL Database**

- Fully managed relational database based on Microsoft SQL Server.
- Scales automatically and offers strong security.
- Great for enterprise apps, finance, and business systems.

❑ **Azure Cosmos DB**

- Globally distributed **NoSQL database**.
- Supports multiple data models: key-value, document, graph, and column.
- Provides very fast performance worldwide.
- Ideal for apps needing global reach, like social media or gaming.

❑ **Azure Database for PostgreSQL**

- Managed PostgreSQL service.
- Handles open-source workloads with built-in high availability.
- Commonly used for analytics, SaaS apps, and web backends.

❑ **Azure Database for MySQL**

- Managed MySQL service in the cloud.
- Automatic scaling, backups, and patching.
- Popular for content management systems (CMS) and e-commerce websites.



AZURE - SQL DATABASE

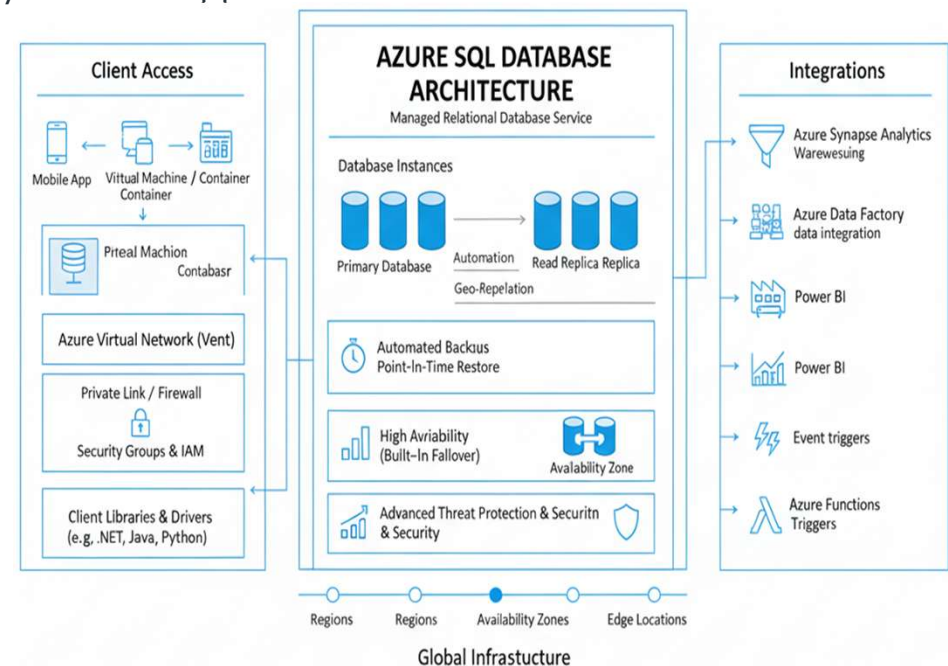
Azure SQL Database is a relational database (RDBMS) service provided by **Microsoft Azure** that is widely used by developers when creating new applications in the cloud. It is managed completely by Microsoft and is a highly scalable platform-as-a-service (PaaS) designed especially for cloud applications.

SQL Azure Architecture:

- **Database Instances** Supports automated replication, read replicas, and geo-replication.
- **Automated Backups** → Point-in-time restore.
- **High Availability** → Built-in failover for reliability.
- **Performance Monitoring & Tuning** → Optimizes queries and workloads.

Integrations:

- **Azure Synapse Analytics** → Data warehousing.
- **Azure Data Factory** → Data integration.
- **Power BI** → Data visualization.
- **Event Triggers** → React to database events.
- **Azure Functions** → Trigger serverless workflows.





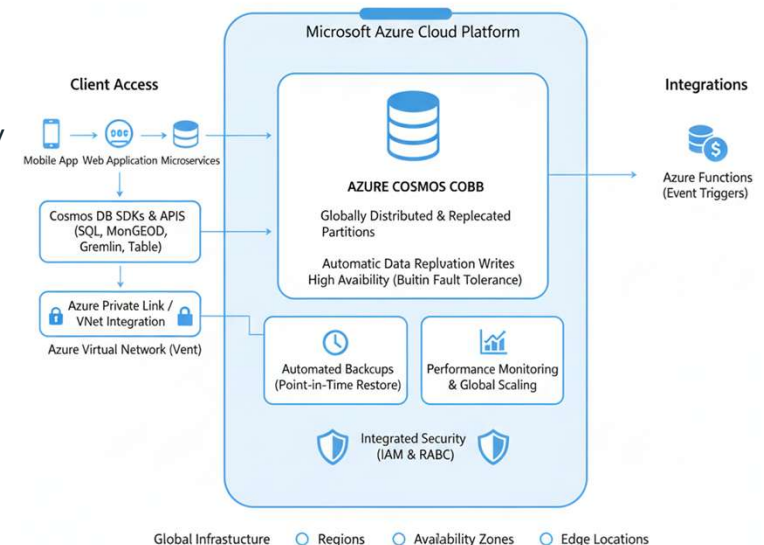
AZURE - COSMOS DB

Azure Cosmos DB is a fully managed, globally distributed NoSQL database service by Microsoft. It offers multi-model support (SQL, MongoDB, Gremlin, Table, Cassandra), automatic scaling, and low-latency access worldwide. With built-in high availability, security, and replication, it's ideal for modern, mission-critical applications.

❑ Azure Cosmos DB

- **Globally Distributed & Replicated Partitions** → Data automatically replicated across regions.
- **Automatic Data Replication Writes** → Ensures fault tolerance and availability.
- **High Availability (Built-in Fault Tolerance)** → Continuous availability across multiple regions.
- **Automated Backups (Point-in-Time Restore)** → Data recovery at any point in time.
- **Performance Monitoring & Global Scaling** → Monitors workloads and scales automatically worldwide.
- **Integrated Security (IAM & RBAC)** → Access control with identity and role-based security.

AZURE COSMOS DB ARCHITECTURE



A handy new Google Cloud, AWS, and Azure product map

Filter by entering a product name or characteristic.

Service category ▼	Service type	Google Cloud product	AWS offering	Azure offering
App Modernization	CI/CD	Cloud Build	AWS CodeBuild, AWS CodeDeploy, AWS CodePipeline	Azure DevOps
Compute	VMware connectivity	VMware Engine	VMware Cloud on AWS	Azure VMware Solution
Management tools	API management	Apigee API Management	Amazon API Gateway	Azure API Management
Machine learning (ML)	ML platform	TensorFlow Enterprise	Tensorflow on AWS	Azure Databricks

THANK YOU !

