Database Systems:

Module 14, Lecture 2 – Cloud Databases – The Big Three

Instructor: Alan Paradise

Cloud Databases

LESSON OBJECTIVES

- Describe the "Big Three" in the Cloud Computing marketplace
- Differentiate the types and styles of databases available in the cloud

Cloud Databases

Let's look at the "Big Three" Cloud Providers' database offerings

- Amazon Web Services
- Microsoft Azure
- Google Cloud Platform

Cloud Databases

AWS

- Lets you choose from 15 different database engines including relational, key-value, document, in-memory, graph, time series, and ledger (blockchain secured) databases
- High-availability & Scaling: AWS provides continuous monitoring of your clusters to keep your workloads up and running with self-healing storage and automated scaling
- You don’t need to worry about data management tasks like software patching, backup, recovery, setup.

Cloud Databases

RELATIONAL database that stores and provides access to data points that are related to one another. Uses SQL.	Traditional Applications ERP / CRM E-Commerce	Amazon Aurora Amazon RDS Amazon RedShift
KEY-VALUE designed for storing, retrieving, managing associative arrays more commonly known today as a dictionary or hash table.	High Traffic Web Apps E-Commerce Systems Game Apps	Amazon DynamoDB
IN-MEMORY system that primarily relies on main memory for computer data storage. In the event of a power loss, data stored in volatile RAM is lost.	Caching Session Management Gaming Leaderboards Geospatial Apps	Amazon Elasticache for Memcashed Amazon Elasticache for Redis
DOCUMENT type of nonrelational database that is designed to store and query data as JSON-like documents.	Content Management Catalogs User Profiles	Amazon DocumentDB

Cloud Databases

WIDE-COLUMN NoSQL databases that works well for storing enormous amounts of data that can be collected. Its architecture uses persistent, sparse matrix, multi-dimensional mapping.	Industrial apps Equipment maintenance Fleet management Route optimization	Amazon Managed Apache-Cassandra Service
GRAPH uses graph structures for semantic queries with nodes, edges, and properties to represent and store data	Fraud detection Social networking Recommendation engines	Amazon Neptune
TIME SERIES software system optimized for storing and serving time series through associated pairs of times and values	IoT applications DevOps Industrial telemetry	Amazon Timestream
LEDGER A NoSQL database that provides an immutable, transparent, and cryptographically verifiable transaction log owned by a central authority.	Systems of record Supply chain Registrations Banking transactions	Amazon Quantum Ledger Database

RELATIONAL:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon Aurora built for the cloud, that combines the performance/availability of traditional enterprise databases with simplicity and cost-effectiveness of open source databases. **Five times faster** than standard MySQL databases and **three times faster** than standard PostgreSQL databases	High Performance/Scalability High Availability/Durability Highly Secure PostgreSQL/MySQL Compatible Migration Support Fully Managed	Capital One Verizon United Nations Arizona State University
Amazon RDS Built for cloud, available on several database instance types - **optimized for memory**, performance or I/O - and provides you with six familiar database engines to choose from, including Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database, and SQL Server.	Easy to Administer Highly Scalable Available and Durable Fast Secure Inexpensive	GE Appliances Netflix Expedia Intuit Blackboard Unilever
Amazon Redshift **powers analytical workloads**, you can query petabytes of structured and semi-structured data across **data warehouse** and your data lake using standard SQL	Inexpensive Optimal Storage Diverse Workloads Managed Storage	Lyft Comcast Yelp McDonalds

KEY-VALUE:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon DynamoDB key-value/ document DB delivering single-digit millisecond performance at any scale. Fully managed, multi-region, durable database with built-in security, backup, in-memory caching for internet-scale applications. **It can handle more than 10 trillion requests per day** and **can support peaks of more than 20 million requests per second.**	Performance at Scale No Servers to Manage Enterprise Ready Internet Scale Maintains Low Latency Maintains Concurrency	Nike Samsung Snapchat Airbnb Tinder GumGum

IN-MEMORY:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon ElastiCache for Memcached Memcached-compatible in-memory key-value store service that can be used as a cache or a data store, ideal for cases where frequently accessed data must be in-memory.	Extreme Performance Secure and Hardened Memcached-Compatible Auto Discovery Easily Scalable Fully Managed	Major League Baseball Adobe AdRoll
Amazon ElastiCache for Redis **Blazing fast in-memory data store** that provides **sub-millisecond latency** to power internet-scale real-time applications. Built on open-source Redis and compatible with the Redis APIs, ElastiCache for Redis works with your Redis clients and uses the open Redis data format to store your data	Redis-Compatible Extreme Performance Fully Managed and Hardened Highly Available and Reliable Easily Scalable Secure and Compliant	Zynga Grab Dream11 Coffee Meets Bagel

DOCUMENT:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon DocumentDB Fully managed document database service that supports MongoDB workloads. **Designed for 99.99% availability** and replicates six copies of your data across three AWS Availability Zones (AZs).	MongoDB-Compatible Fully Managed Performance at Scale	The Washington Post Freshop FINRA

WIDE-COLUMN:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon Managed Apache Cassandra Services Can run your Cassandra workloads on AWS using the same Cassandra application code and developer tools that you use today.	Apache-Cassandra Compatible No Servers to Manage Performance at Scale Highly Available and Secure	Pegasystems Reltio Adobe McDonalds

GRAPH:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon Neptune Built to store and navigate relationships. They have advantages over relational databases for use cases like social networking, recommendation engines, and fraud detection, where you need to create relationships between data and quickly query these relationships	Supports open graph API's High Performance/Scalability High Availability and Durability Highly Secure Fully Managed	Amazon Alexa Siemens Pearson Blackfynn PaySense Thomson Reuters

Cloud Databases

TIME SERIES:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon Timestream Database service for **IoT and operational applications** that makes it **easy to store and analyze trillions of events per day at 1/10th the cost of relational databases**. Driven by the rise of IoT devices, IT systems, and smart industrial machines, time-series data — data that measures how things change over time — is one of the fastest growing data types.	Extremely Fast. Processes trillions of events per day. Up to 1,000X faster query performance. Built-In Analytics Serverless Optimized Query-Processing Engine Manage Patching, Setup, Configurations	Enterprise Sporting Goods Market Research Accounting Construction

LEDGER:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon Quantum Ledger Database Provides a transparent, immutable, and cryptographically verifiable **transaction log** owned by a central trusted authority. Amazon QLDB **tracks every application data** change and maintains a complete and **verifiable history of changes** over time.	Immutable and Transparent Cryptographically Verifiable Performant and Highly Scalable Serverless Easy to Use Highly Available	Splunk Zilliant Realm Digital Asset Health Direct Driver & Vehicle Licensing Agency

Cloud Databases

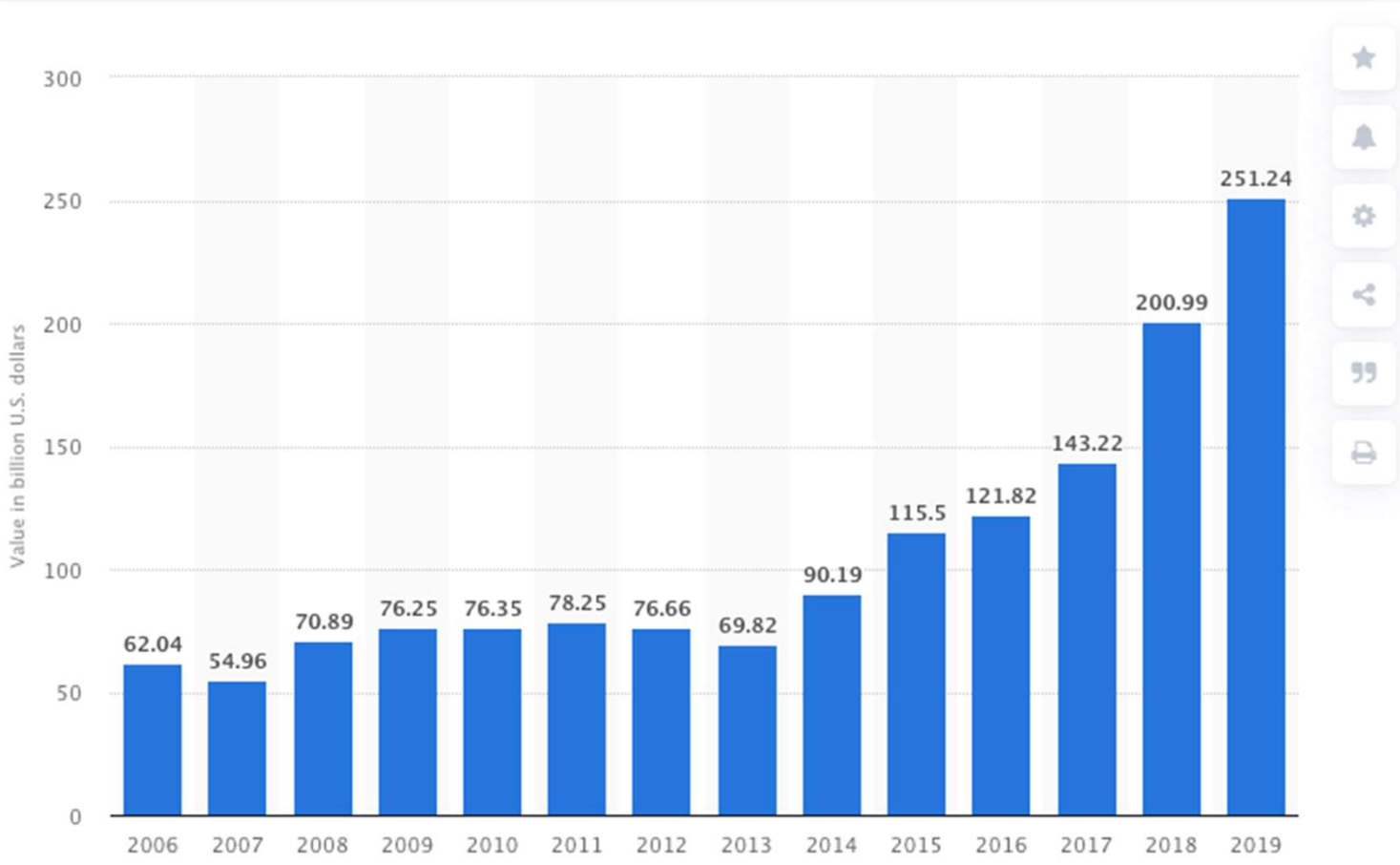
Microsoft Azure

NOTE: Microsoft already has a HUGE market share

- Windows PCs
- Windows Servers (Mail, Database, App, Web)
- MS Office

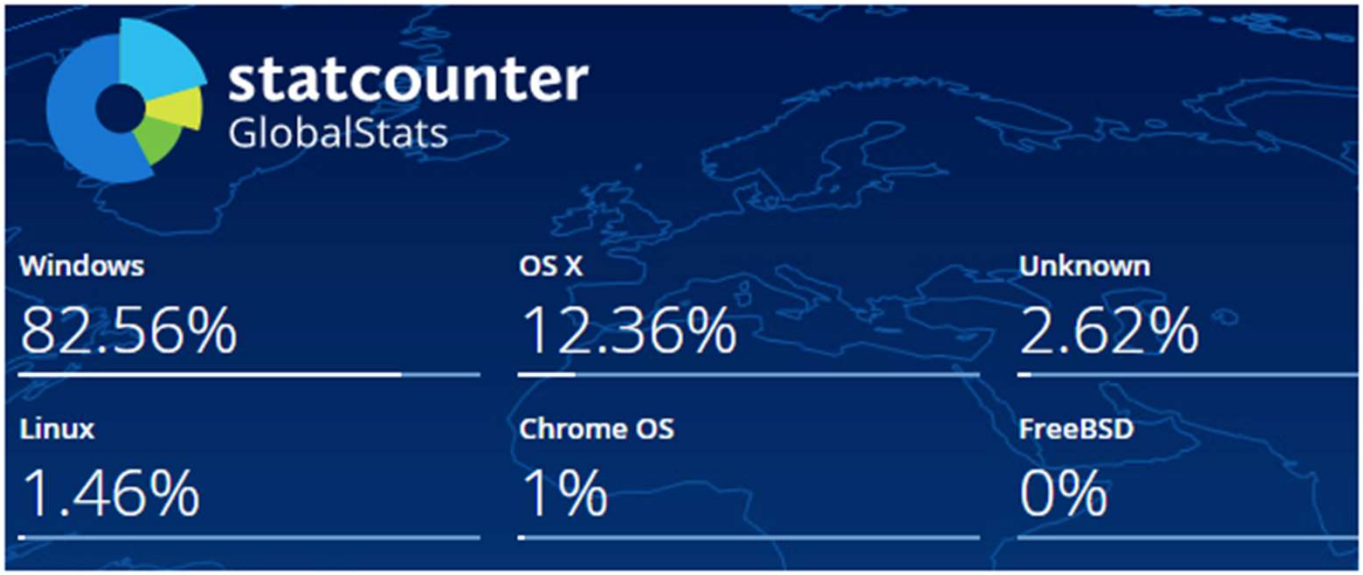
Microsoft's global brand value from 2006 to 2019

(in billion U.S. dollars)



Cloud Databases

Desktop & Laptop OS



Cloud Databases

RELATIONAL database that stores and provides access to data points that are related to one another. Uses SQL.	Azure SQL Database Azure Database for MySQL Azure Database for PostgreSQL SQL Server on VMs Azure Database for MariaDB Azure Synapse Analytics Azure Data Explorer Azure Database Migration Service
NO-SQL/NON RELATIONAL designed for storing, retrieving, managing associative arrays more commonly known today as a dictionary or hash table.	Table Storage Azure Cosmos DB Azure Cache for Redis

RELATIONAL:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Azure SQL Database Migrate your SQL Server applications, with no code changes, to experience the benefits of a fully managed and intelligent service. Or build for future app growth and scale up to 100 TB with Hyperscale.	Easily Migrate Data Built-In Machine Learning Scalability/Availability Advanced Data Security	AccuWeather PAYCHEX Allscripts ABB
Azure Database for MySQL Deliver high availability and elastic scaling to open-source mobile and web apps with a managed community MySQL database service, or migrate MySQL workloads to the cloud.	Flexible Pricing Database Protection Language/Framework of choice Scalability/Availability	FundWorks GeekWire School District 42
Azure Database for PostgreSQL Build scalable and secure enterprise-ready apps on community PostgreSQL, scale out single node PostgreSQL with high performance, or migrate PostgreSQL and Oracle workloads to the cloud.	Integration with JSONB Indexing/Extensions High Performance Scaling Intelligent Performance Recommendations	Enlyft Sivantos Higher Ed Profiles Somerset
SQL Server on VMs Run your SQL Server apps in the cloud with seamless scaling and pay-per-minute pricing, or migrate SQL Server or Oracle workloads to the cloud.	**High-Performance VMs** Best TCO Built-In Security Manageable	RedHat SUSE Ubuntu Windows Server
Azure Database for MariaDB Deliver high availability and elastic scaling to open-source mobile and web apps with a managed community MariaDB database service.	Easily deploy Applications Achieve Business Continuity Flexible Pricing Unparalleled Security Gaming Digital Marketing Financial Management Retail E-Commerce	N/A

Cloud Databases

NON- RELATIONAL:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Table Storage Rapidly develop with massing semi-structured datasets using a NoSQL key-value store.	Petabytes of structured data Supports flexible Data Schema Made for Enterprise Designed for Developers	XBOX Carnegie Mellon University CSA GreenButton
Azure Cache for Redis Power fast, scalable applications with an open-source-compatible in-memory data store.	High Performance Fully-Managed Service Built-In Reliability Flexible Scalability Open Source Compatible	CSA ISO CJIS ITAR
Azure Cosmos DB Build applications with guaranteed low latency and high availability anywhere, at any scale, or migrate Cassandra, MongoDB, and other NoSQL workloads to the cloud.	Global Distribution Millisecond Latency Elastic, Automatic Scaling Multi-Model	American Cancer Society ExxonMobil Symantec asos

Cloud Databases

Google Cloud Platform

Google is the "underdog" versus AWS and Azure

How to take marketshare?

- Lower Prices
- Be "App centric" versus "server centric"
- Simpler approach
- Friendlier help, easily available

Cloud Databases

RELATIONAL database that stores and provides access to data points that are related to one another. Uses SQL.	Cloud Spanner Cloud SQL
NO-SQL/NON RELATIONAL designed for storing, retrieving, managing associative arrays more commonly known today as a dictionary or hash table.	Cloud Bigtable Cloud Firestore Firebase Realtime Database Cloud Memorystore

Cloud Databases

RELATIONAL:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Cloud Spanner Cloud Spanner helps future-proof your database backend. It can scale to arbitrarily large database sizes to help avoid rewrites and migrations. You get	Global Scale Fully Managed Relational Semantics Multi-Language Support	Streak Dragon Ball Legends Optiva The Next Platform
best of relational database structure with non relational database scaling and performance with **strong consistency across rows.**	Transactional Consistency Enterprise Security Highly Available	
Cloud SQL Cloud SQL is fully compatible with applications using MySQL, PostgreSQL, and SQL Server. You can connect with nearly any application, anywhere in the world. Cloud SQL automates backups, replication, and failover to ensure your database is reliable, highly available, and flexible to your performance needs.	Fully Managed Integrated Reliable High Performance Secure Access Availability Protection Scalaility	Descartes Labs Signify WideOrbit

Cloud Databases

Cloud MemoryStore	Open Source Redis	Descartes Labs
Memorystore for Redis provides a fully managed in-memory data store service built on scalable, secure, and highly available infrastructure managed by Google. Use Memorystore to build application caches that provides sub-millisecond data access. Memorystore is compatible with the Redis protocol, allowing easy migration with zero code changes.	High Availability Patching/Monitoring Scalability Google Grade Security Easy Lift/Shift Migration	

Cloud Databases

In Summary:

Industry is *gradually* shifting

- Away from relational to NoSQL
- Away from internal hosting to cloud

As a database expert, you need to understand these trends, and advise your organization/clients

Cloud Databases

Next Topic: Careers in Database