

# *Cloud Computing*

---

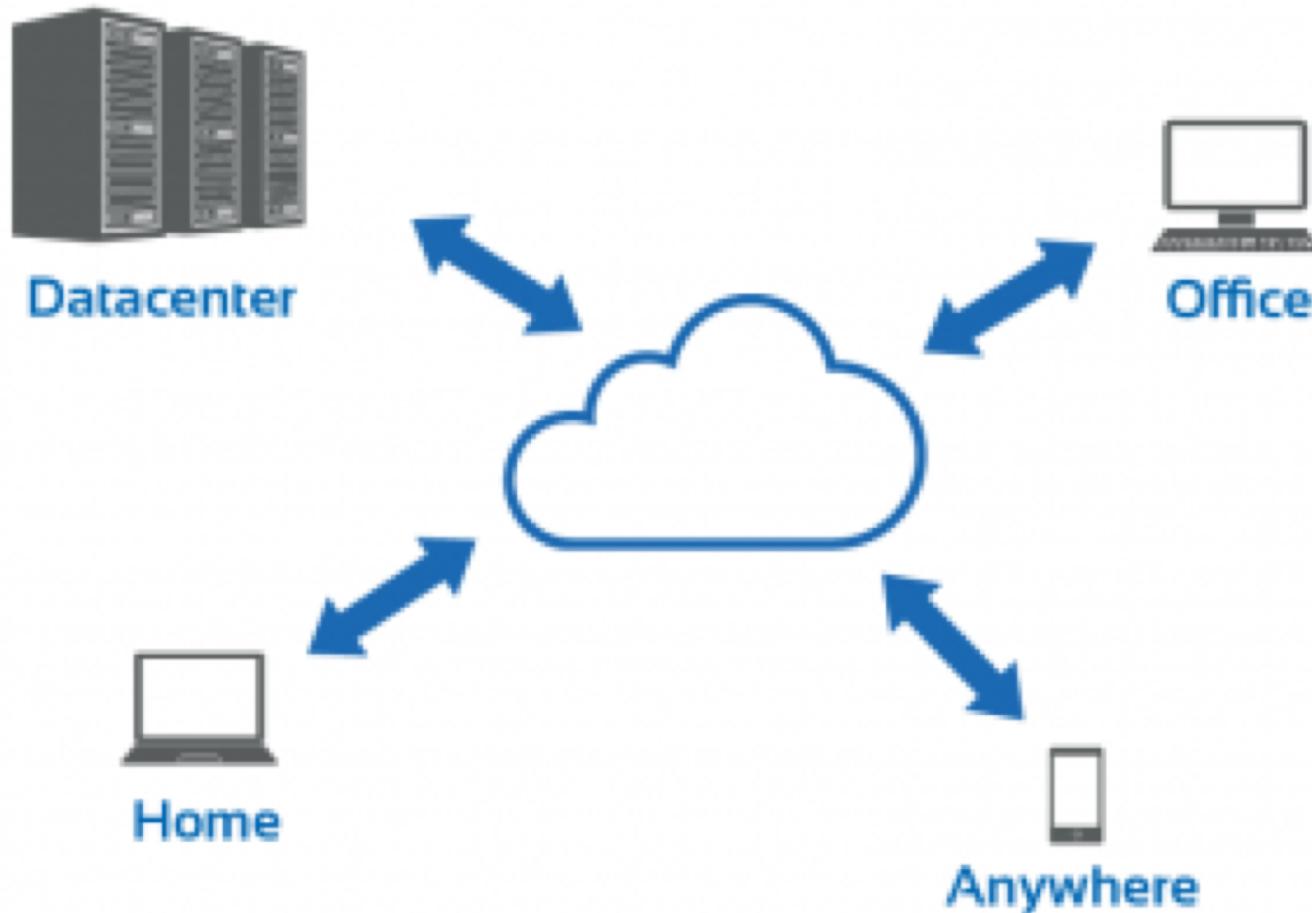
**A little history . . . Life in the Corporate “Data Center”**



## A little history . . .

- Things to think about:
  - Scaling, Lead time
  - Floor space
  - Power management
  - Heating/Cooling
  - Redundancy -- SPOF
  - Fire Suppression
  - Battery Backup -- UPS
  - Network Wiring
  - Data & Server Backups
  - Network Switching
  - 7X24 support
  - Alerts/Alarms

# *Cloud Computing*



## Welcome, “Cloud Computing”

- **Private** – my private cloud in my own data center
- **Public** – a shared environment hosted by a provider

## A metaphor

- The electrical grid
  - You don't know where it comes from
  - It's there when you need it, just plug in
  - Use what you want
  - Need more? Just take it.
  - Pay for what you use

## **Definition . . .**

- “**Computing Services and Solutions are delivered and consumed in real time over the internet.**”
- **Characteristics of Public Cloud Services**
  - Offsite hosting
  - Pay per use (setup/initial, plus ongoing)
  - Shared space
  - Massively Scalable
  - On-Demand Provisioning
  - Rapid Deployment
  - Lowers innovation barriers
  - Leading edge architecture

## **Advantages of Cloud Computing**

- Ubiquitous (available from anywhere with an internet connection)
- Automated change management
- Massively Scalable
- On-Demand Provisioning
- Rapid Deployment
- Lowers innovation barriers
- Leading edge architecture
- Lower Cost

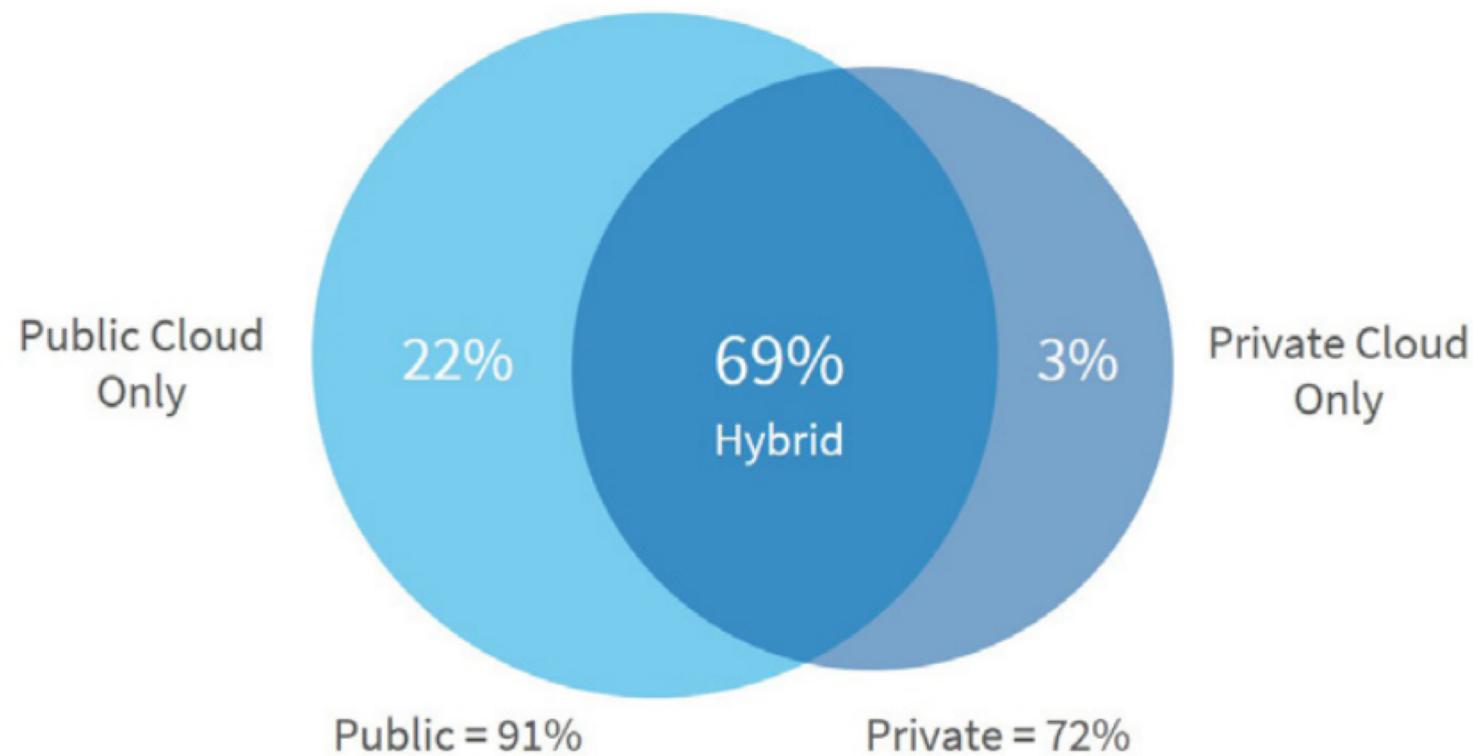
## **Disadvantages of Cloud Computing**

- Surrender Control
- Less Robust Monitoring
- Requires Large Network Pipe/Capacity
- Less Secure – multi-tenancy, DDOS

## **Private versus Public Cloud**

- Private: Leverage the advantages, with few disadvantages
  - Massively Scalable
  - On-Demand Provisioning
  - Rapid Deployment
  - More secure
  - Better Monitoring
  - BUT → Still requires significant internal infrastructure in your own corporate data center

## **94% of Respondents Using Cloud**



*Source: RightScale 2019 State of the Cloud Report from Flexera*

# *Cloud Computing*

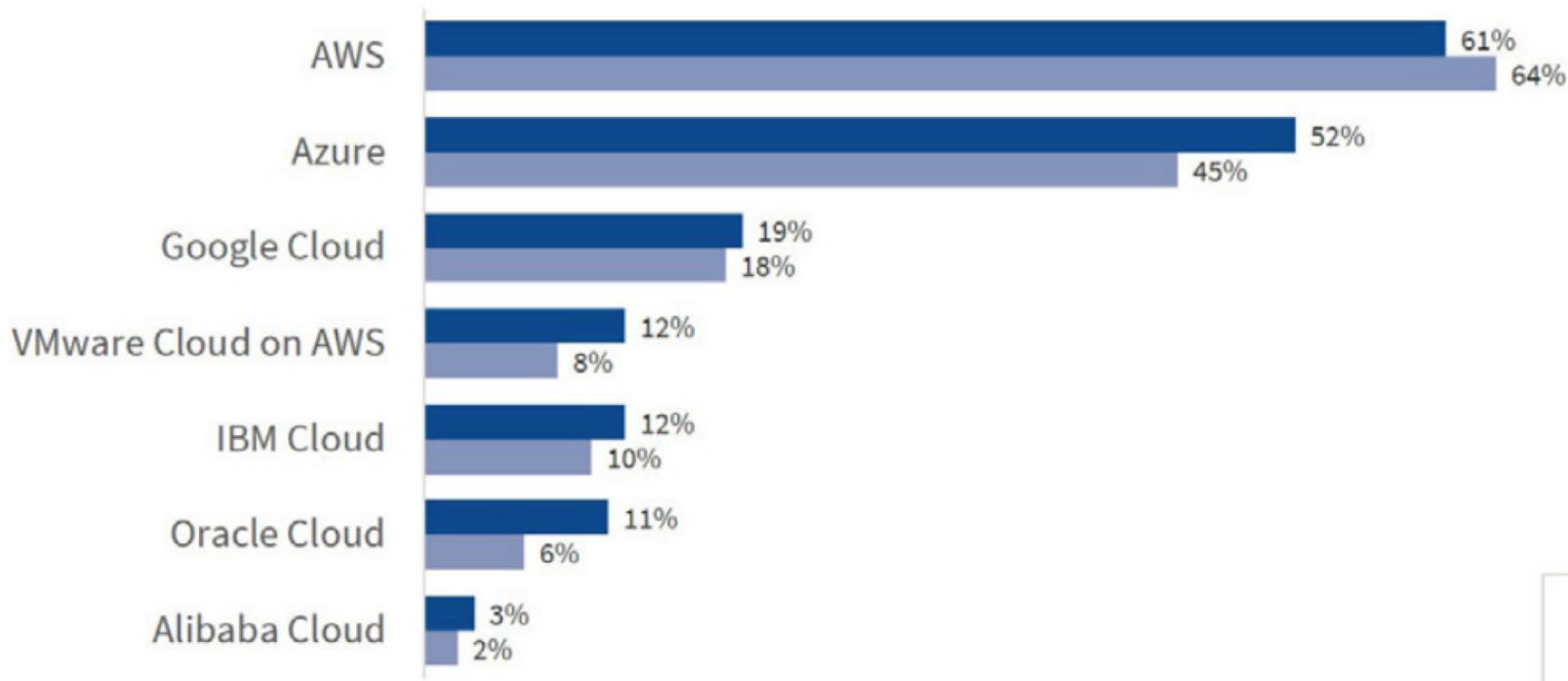
Worldwide Cloud IT Infrastructure Market Forecast by Deployment Type, 2016 - 2022 (shares based on Value)



# *Cloud Computing*

## Public Cloud Adoption

% of All Respondents



Source: RightScale 2019 State of the Cloud Report from Flexera

## Trends in Cloud Computing for 2020

1. Businesses increasingly favor a multi-cloud strategy.
2. Cloud security has become a major area of concern.
3. Cloud cost management is still a struggle.

Read more at: <https://www.stratospherenetworks.com/blog/top-3-cloud-computing-trends-to-watch-in-2020/>

# *Databases In The Cloud*

---

- **IaaS - Infrastructure-as-a-Service**
  - A cloud service providing infrastructure - computers, networking resources, storage.

I can pay for cloud-based servers and storage, and spin up database servers according to my own architecture/design, using my chosen DBMS software. For example, Oracle's "BYOL" offering.

(see article posted to this week's moodle page)

# *Databases In The Cloud*

---

- **SaaS - Software-as-a-Service**
  - A cloud service providing users access to software in a self-service, on-demand fashion.

Top cloud providers offer MANY, MANY database alternatives. Their equipment: their servers, their storage, their software licenses.

- I pay for storage by the byte.
- I pay for data movement into and out of my databases by the byte.
- I pay for network bandwidth consumption by the byte.
- I pay for CPU cycles as I burn them.

# *Databases In The Cloud*

---

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

Amazon Web Services

Microsoft Azure

Google Cloud Platform

A BIG **SHOUT OUT:**

To **Muntaha Pasha**, CompSci Major, whose independent research project provided me with great info to share with you on this topic.

## 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.

# *Databases In The Cloud*

---

<b>RELATIONAL</b> 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
<b>KEY-VALUE</b> 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
<b>IN-MEMORY</b> 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 Memcached Amazon Elasticache for Redis
<b>DOCUMENT</b> type of nonrelational database that is designed to store and query data as JSON-like documents.	Content Management Catalogs User Profiles	Amazon DocumentDB

# *Databases In The Cloud*

---

<b>WIDE-COLUMN</b> 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
<b>GRAPH</b> uses graph structures for semantic queries with nodes, edges, and properties to represent and store data	Fraud detection Social networking Recommendation engines	Amazon Neptune
<b>TIME SERIES</b> software system optimized for storing and serving time series through associated pairs of times and values	IoT applications DevOps Industrial telemetry	Amazon Timestream
<b>LEDGER</b> 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

# *Databases In The Cloud*

## 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. <b>Five times faster than standard MySQL databases and three times faster than standard PostgreSQL databases</b>	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 - <b>optimized for memory</b> , performance or I/O - and provides you with six familiar database engines to choose from, including <u>Amazon Aurora</u> , <u>PostgreSQL</u> , <u>MySQL</u> , MariaDB, <u>Oracle Database</u> , and SQL Server.	Easy to Administer  Highly Scalable  Available and Durable  Fast  Secure  Inexpensive	GE Appliances  Netflix  Expedia  Intuit  Blackboard  Unilever
Amazon Redshift  <b>powers analytical workloads</b> , you can query petabytes of structured and semi-structured data across <b>data warehouse</b> and your data lake <b>using standard SQL</b>	Inexpensive  Optimal Storage  Diverse Workloads  Managed Storage	Lyft  Comcast  Yelp  McDonalds

## KEY-VALUE:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
Amazon DynamoDB <u>key-value/ document DB</u> 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  <u>Blazing fast in-memory data store that provides sub-millisecond latency to power internet-scale real-time applications.</u> 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 <a href="#">MongoDB workloads</a> . 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 <a href="#">Cassandra workloads</a> 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

# *Databases In The Cloud*

## 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 — <u>data that measures how things change over time</u> — 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

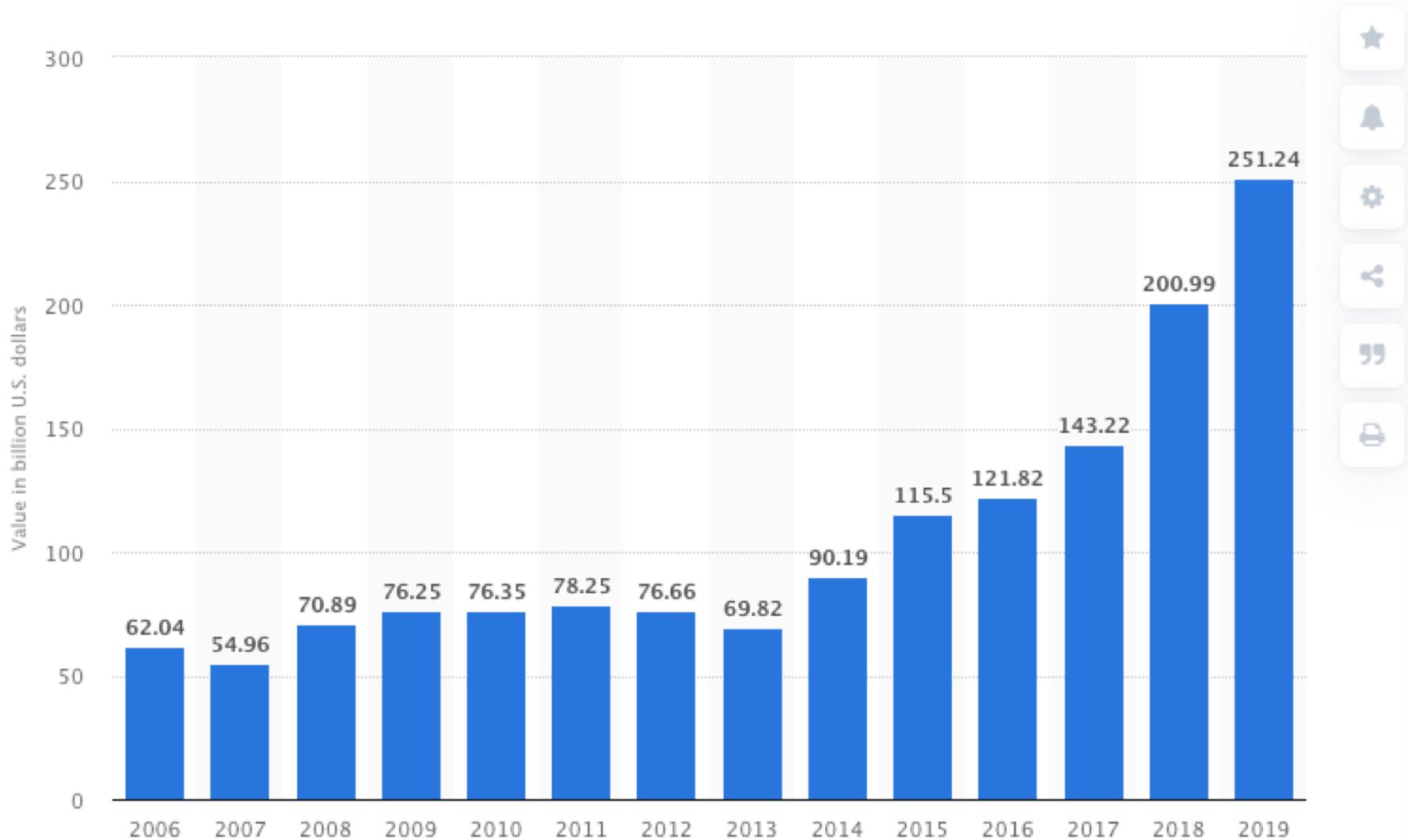
## LEDGER:

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.  <u>Amazon QLDB tracks every application data change and maintains a complete and verifiable history of changes over time.</u>	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

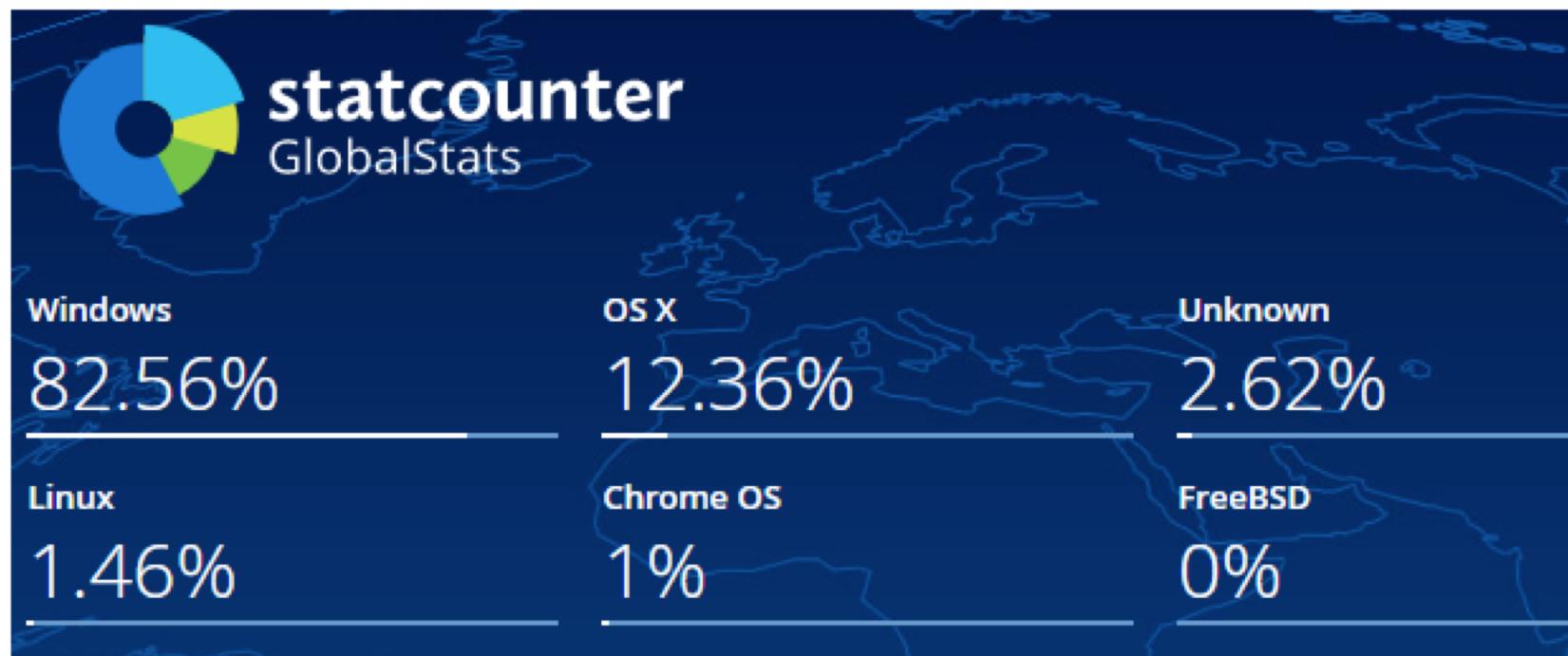
## Microsoft Azure

- 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)*



## Desktop & Laptop OS



350 systems in ranking, January 2020

Rank			DBMS	Database Model	Score		
Jan 2020	Dec 2019	Jan 2019			Jan 2020	Dec 2019	Jan 2019
1.	1.	1.	Oracle	Relational, Multi-model	1346.68	+0.29	+77.85
2.	2.	2.	MySQL	Relational, Multi-model	1274.65	-1.01	+120.39
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1098.55	+2.35	+58.29
4.	4.	4.	PostgreSQL	Relational, Multi-model	507.19	+3.82	+41.08
5.	5.	5.	MongoDB	Document, Multi-model	426.97	+5.85	+39.78
6.	6.	6.	IBM Db2	Relational, Multi-model	168.70	-2.65	-11.15
7.	7.	↑ 8.	Elasticsearch	Search engine, Multi-model	151.44	+1.19	+8.00
8.	8.	↓ 7.	Redis	Key-value, Multi-model	148.75	+2.51	-0.27
9.	9.	9.	Microsoft Access	Relational	128.58	-0.89	-13.04
10.	↑ 11.	10.	SQLite	Relational	122.14	+1.78	-4.66
11.	↓ 10.	11.	Cassandra	Wide column	120.66	-0.04	-2.32
12.	12.	12.	Splunk	Search engine	88.67	-1.85	+7.25
13.	13.	13.	MariaDB	Relational, Multi-model	87.45	+0.66	+8.63
14.	14.	↑ 15.	Hive	Relational	84.24	-1.81	+14.33
15.	15.	↓ 14.	Teradata	Relational, Multi-model	78.29	-0.21	+2.10
16.	16.	↑ 20.	Amazon DynamoDB	Multi-model	62.02	+0.39	+6.93
17.	17.	↓ 16.	Solr	Search engine	56.57	-0.65	-4.92
18.	↑ 19.	18.	FileMaker	Relational	55.11	-0.03	-2.05
19.	↑ 20.	19.	SAP HANA	Relational, Multi-model	54.69	+0.52	-1.95
20.	↓ 18.	↑ 21.	SAP Adaptive Server	Relational	54.59	-0.96	-0.45
21.	21.	↓ 17.	HBase	Wide column	53.34	-0.12	-7.05
22.	22.	22.	Neo4j	Graph	51.66	+1.10	+4.86

# *Databases In The Cloud*

---

<b>RELATIONAL</b> 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
<b>NO-SQL/NON RELATIONAL</b> 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

# *Databases In The Cloud*

## NON- RELATIONAL:

<b>DATABASE NAME</b>	<b>BENEFITS</b>	<b>CUSTOMERS USING IT</b>
<b>Table Storage</b> 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 <u>GreenButton</u>
<b>Azure Cache for Redis</b> 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
<b>Azure Cosmos DB</b> 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 <u>asos</u>

## **Google Cloud Platform**

Google is the "underdog" versus AWS and Azure

How to take marketshare?

- Lower Prices
- App centric versus server centric  
(Like Kubernetes to orchestrate app dev)
- Simpler approach
- Friendlier help, easily available

# *Databases In The Cloud*

---

<b>RELATIONAL</b> database that stores and provides access to data points that are related to one another. Uses SQL.	Cloud Spanner Cloud SQL
<b>NO-SQL/NON RELATIONAL</b> 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

# *Databases In The Cloud*

## Google Cloud Platform

### 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 <b>strong consistency across rows.</b>	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  Scalability	Descartes Labs  Signify  WideOrbit

## NON - RELATIONAL:

DATABASE NAME	BENEFITS	CUSTOMERS USING IT
<p>Cloud Bigtable Ideal for ad tech, fintech, and IoT, Cloud Bigtable offers consistent sub-10ms latency. Replication provides higher availability, higher durability, and resilience in the face of zonal failures.</p> <p><u><a href="#">Cloud Bigtable is designed with a storage engine for machine learning applications and provides easy integration with open source big data tools.</a></u></p>	<p>Low Latency Massively Scalable Fast Performance Simple Integration Integrates easily with Hadoop, and other popular Big-Data tools. Fully Managed</p>	<p>Beam BigQuery TensorFlow JanusGraph OpenTSDB</p>
<p>Cloud Firestore Cloud Firestore is a fast, fully managed, serverless, cloud-native NoSQL <u>document database</u> that simplifies storing, syncing, and querying data for your mobile, web, and IoT apps at global scale. Its client libraries provide live synchronization and offline support, while its security features and integrations with Firebase and Google Cloud Platform (GCP) accelerate building truly serverless apps.</p>	<p>Automatic Upgrades Security Datastore Mode Powerful Query Engine Multi-Region Replication Built for Cloud-Specific Apps</p>	<p>Hawkin Dynamics</p>
<p>Firebase Realtime Database The Firebase Realtime Database is a cloud-hosted NoSQL database that lets you store and sync data between your users in real time. Firebase helps you develop high-quality apps, grow your user base, and earn more money. Each feature works independently, and they work even better together.</p>	<p>Collaborate across devices with ease Serverless Apps Optimized for Offline Usage Strong User-Based Security</p>	<p>GreenLion Soft</p>

# *Databases In The Cloud*

---

<p><b>Cloud MemoryStore</b></p> <p>Memorystore <a href="#">for Redis</a> 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.</p>	<p>Open Source Redis</p> <p>High Availability</p> <p>Patching/Monitoring</p> <p>Scalability</p> <p>Google Grade Security</p> <p>Easy Lift/Shift</p> <p>Migration</p>	<p>Descartes Labs</p>
--	--	-----------------------

## Sample Pricing for Cloud Firestore

FEATURE	PRICE
Stored data	\$0.18/GB
Bandwidth	<a href="#">Google Cloud Pricing</a>
Document writes	\$0.18/100K
Document reads	\$0.06/100K
Document deletes	\$0.02/100K

# *Google Cloud Platform Pricing*

## 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