

Purpose-built databases for modern applications

김일호 Lead Solutions Architect for Emerging Tech. 15-Sep-2020



Agenda



What's a modern application

Why consider purpose-built databases



AWS Databases: The right tool for the right job



Unprecedented Data Growth Drives Innovation

Explosion of data



Data grows 10x every 5 years driven by networkconnected smart devices Micro-services changes data and analytics requirements



Micro-services architecture decreases need for one-size-fitsall databases and increases need for real-time monitoring and analytics Rapid rate of change driven by DevOps



Transition from IT to DevOps increases rate of change

What's a modern application?





Characteristics of a Modern Digital Business





Technology decisions drive your business strategy



Modernization changes how you work



Start from scratch Goal is innovation Deploy, debug, & monitor Goal is low risk & high

Operators

reliability/availability



App architectures & patterns have evolved over the years...



Traditional three-tier application architecture



Modern application requirements Requires more performance, scale, and availability











E-Commerce

Media streaming

Social Online media gaming

Shared economy

Users	1M+
Data volume	Terabytes—petabytes
Locality	Global
Performance	Microsecond latency
Request rate	Millions per second
Access	Mobile, IoT, devices
Scale	Virtually unlimited
Economics	Pay-as-you-go
Developer access	Instance API access
Development	Apps and storage are decoupled

A modern three-tier application architecture





Decoupling data into each service improves resilience





A simple MSA architecture



Note: https://aws.amazon.com/getting-started/hands-on/break-monolith-app-microservices-ecs-docker-ec2/module-three/

Instead of a monolithic application,



build microservices with purpose-built tools



Internet-scale e-commerce



The world's largest e-commerce business, Amazon.com, migrated entirely to purpose-built AWS databases because of:

- Cost savings
- Performance improvements
- Administrative overhead reductions



Elasticity & Scalability matter



Amazon Prime Day 2019 – Powered by AWS

- Amazon S3: 63 PB added
- EC2: 372,000 at start. Scaled to 426,000
- Amazon DynamoDB: 7.11 trillion calls
- Amazon Aurora: 148 billion transactions
- Amazon S3: 185 petabytes transferred

Note: https://aws.amazon.com/blogs/aws/amazon-prime-day-2019-powered-by-aws/



Why consider purpose-built databases?







Airbnb uses AWS databases for their internetscale cloud-based applications

User search history: Amazon DynamoDB

- Massive data volume
- Need quick lookups for personalized search

Session state: Amazon ElastiCache

• In-memory store for submillisecond site rendering

Relational data: Amazon RDS (MySQL)

- Referential integrity
- Primary transactional database





Duolingo uses AWS databases to serve up over **31 billion items** for 80 language courses with **high performance** and **scalability**

Primary database: Amazon DynamoDB

- 24,000 reads and 3,000 writes per second
- Personalize lessons for users taking 6bn exercises per month

In-memory caching: Amazon ElastiCache

Instance access to common words and phrases

Transactional data: Amazon Aurora

• Maintain user data



Expedia uses AWS databases for global inventory and pricing analytics that require **consistent performance** at scale

Real-time stream processing: Amazon ElastiCache

- Multi-stream union and self-join
- 24-hour look-back window

Transactional data: Amazon Aurora

• Operational queries

Data warehousing: Amazon Redshift

• Analytical queries





Capital One migrated its monolithic mainframe to **highly available** AWS databases for their microservices-based applications

Transactional data: Amazon RDS

• State management

Analytics: Amazon Redshift

• Web logs

Consistent low latency: Amazon DynamoDB

• User data and mobile app





FINRA uses different databases based on the purpose

Market data search: Amazon DynamoDB

- Massive data volume
- Need quick lookups for searches

Session state: Amazon ElastiCache

• In-memory store for sub-millisecond site rendering

Data collection: Amazon DocumentDB

Fraud detection: Amazon Neptune

Relational data: Amazon RDS, Amazon Aurora

- Referential integrity
- Primary transactional database



AWS purpose-built databases: the right tool for the right job



Our portfolio Broad and deep portfolio, purpose-built for builders

Business Intelligence & Machine Learning



Purpose-built databases





Amazon Aurora MySQL and PostgreSQL-compatible relational database built for the cloud



Performance & scalability

5x throughput of standard MySQL and 3x of standard PostgreSQL; scale-out up to15 read replicas



Availability & durability

Fault-tolerant, self-healing storage; six copies of data across three AZs; continuous backup to S3



Highly secure

Network isolation, encryption at rest/transit



Fully managed

Managed by RDS: no server provisioning, software patching, setup, configuration, or backups



Amazon DynamoDB Fast and flexible key value database service for any scale





Performance at scale

Consistent, single-digit millisecond response times at any scale; build applications with virtually unlimited throughput

Serverless architecture

No hardware provisioning, software patching, or upgrades; scales up or down automatically; continuously backs up your data



Enterprise security

Encrypts all data by default and fully integrates with AWS Identity and Access Management for robust security



Global replication

Build global applications with fast access to local data by easily replicating tables across multiple AWS Regions



Amazon DocumentDB



Fast, scalable, highly available MongoDB-compatible database service



Millions of requests per second, millisecond latency



Same code, drivers, and tools you use with MongoDB



Simple and fully managed



Secure and compliant



2x throughput of managed MongoDB services



Deeply integrated with AWS services



© 2020, Amazon Web Services, Inc. or its Affiliates. All rights reserved.

Unlimited scale

Read scaling with

replicas. Write and memory

scaling with sharding.

Non disruptive scaling

Amazon ElastiCache

Managed, Redis or Memcached-compatible in-memory data store

Consistent high performance

In-memory data store and cache for sub-millisecond response times AWS manages all hardware and software setup, configuration, monitoring

Fully managed









Amazon Neptune

Fast, reliable graph database built for the cloud



Supports Apache TinkerPop & W3C RDF graph models

Query billions of relationships with millisecond latency 6 replicas of your data across 3 AZs with full backup and restore

Build powerful queries easily with Gremlin and SPARQL



Amazon Timestream Fast, scalable, fully managed time series database



millions of inserts per second (10M/second)

engine maintains steady, predictable performance interpolation, smoothing, and approximation

configuration, server provisioning, software patching



Amazon Quantum Ledger Database



Fully managed ledger database Track and verify history of all changes made to your application's data





Amazon Keyspaces (for Apache Cassandra)

Scalable, highly available, and managed Apache Cassandra-compatible database service





A migration is difficult?

Migration Tools

AWS Database Migration Service (DMS)

DMS easily and securely migrates and/or replicate your databases *and* data warehouses to AWS

Schema Conversion Tool (SCT)

SCT converts your commercial database and data warehouse schemas to open-source engines or AWS-native services, such as Amazon Aurora and Redshift

Flexible, powerful migration tooling Most sources and targets, higher conversion automation

Database Migration Service (DMS)

Schema Conversion Tool (SCT)





For heterogeneous migrations (e.g., Oracle to Aurora PostgreSQL or Aurora MySQL, and Cassandra to Amazon DynamoDB)



SAMSUNG

Samsung Migrates 1.1 Billion Users across Three Continents from Oracle to Amazon Aurora with AWS Database Migration Service

2020

The cloud, as it turns out, is big enough for everyone. It's even big enough for Samsung Electronics, the world's second-largest IT company by revenue. As users of Samsung Account—the company's certification and authorization service—ballooned to over 1.1 billion, the electronics giant decided it was time to move its massive database away from its monolithic legacy Oracle internet data center (IDC) solution, which was expensive and made it difficult to scale to accommodate growing traffic. Looking to the cloud, Samsung Electronics found Amazon Web Services (AWS) a perfect fit for this shift. In less than 18 months, Samsung Electronics migrated its global Samsung Account data to Amazon Aurora, a MySQL- and PostgreSQL-compatible relational database built for the cloud. This migration from Oracle resulted in a profound cost reduction as Samsung was able to free itself from the restrictive licensing costs of the on-premises legacy database. With the migration to cloud-native Aurora, Samsung is able to focus on innovating and positioning itself for the company's expected growth.





The scalability of Amazon Aurora is the best benefit—especially if we focus on the cost."

Salva Jung Principal Architect and Engineering Manager

Purpose-built databases





Our approach aws



Get started

Purpose-built Databases aws.amazon.com/products/databases/

Database Migration Service aws.amazon.com/dms/resources/

Database Freedom aws.amazon.com/solutions/databasemigrations/database-freedom/

AWS Database Migration Partners aws.amazon.com/dms/partners/

Migration Resources Library aws.amazon.com/solutions/database-migrations/resources/

