



aws INNOVATE

DATA EDITION

23 August, 2022

Modernise your data infrastructure with fully managed, purpose-built databases

William Wong

Senior Specialist Solutions Architect, Databases

Amazon Web Services



Agenda

- Why modernize?
- Identify the different stages of modernization?
- Common tools and service used at each stage:
 - Moved to managed
 - Break free
 - Modernizing with purpose-built

A modern data strategy



Modernize



Unify



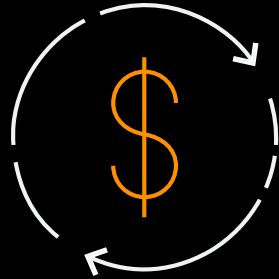
Innovate

Why modernize?

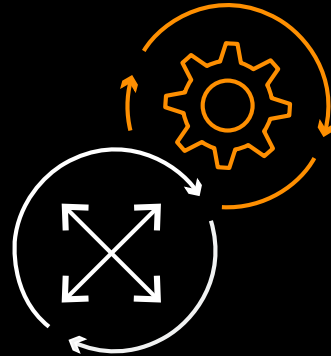
Customers are looking to



Modernize
infrastructure



Reduce
CAPEX

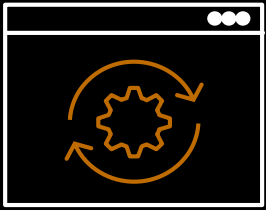


Automate
heavy lifting



Embrace purpose-
built tools

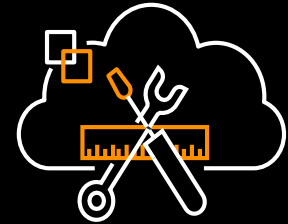
Modernize your data infrastructure



Moving to managed
data services

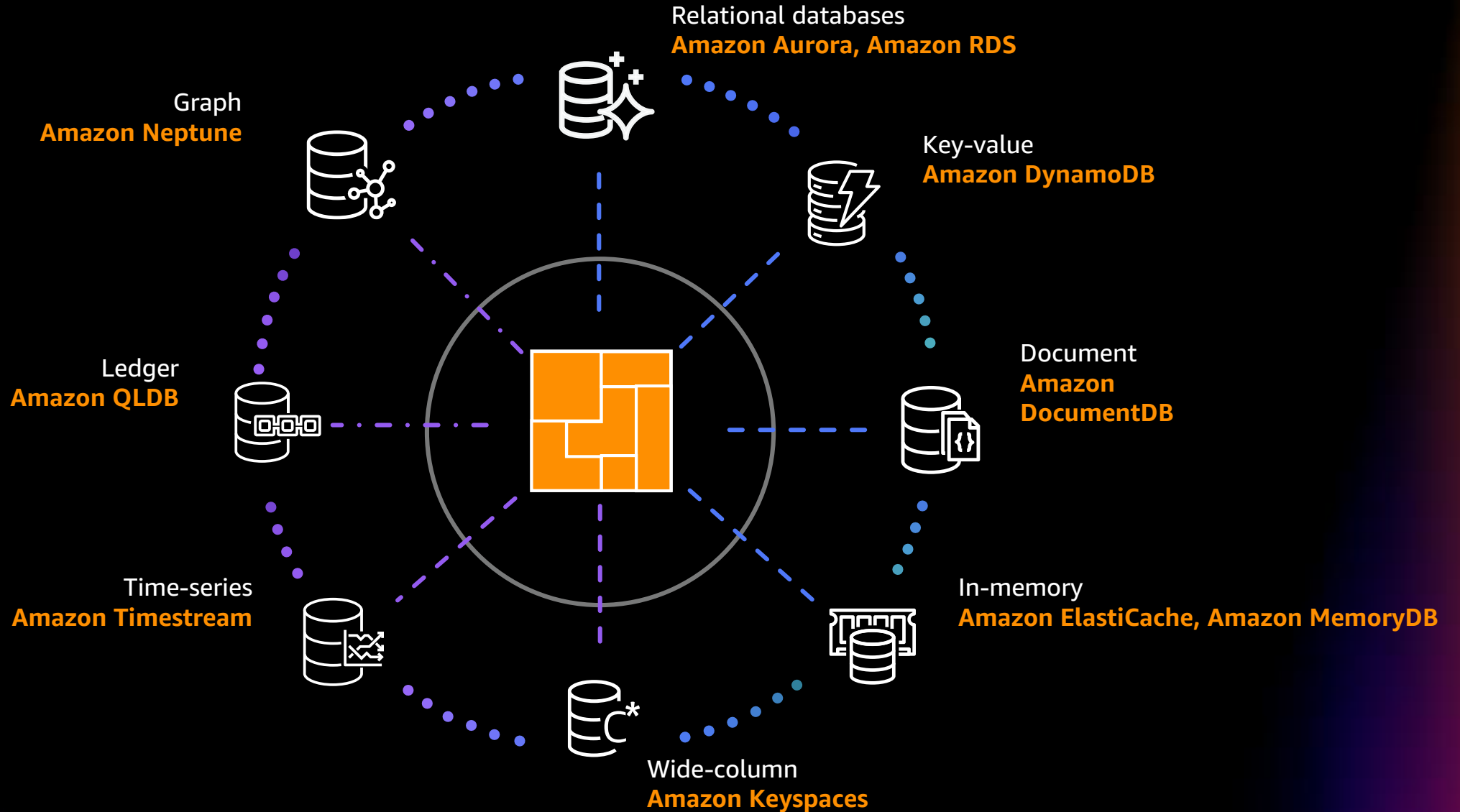


Breaking free from
legacy databases

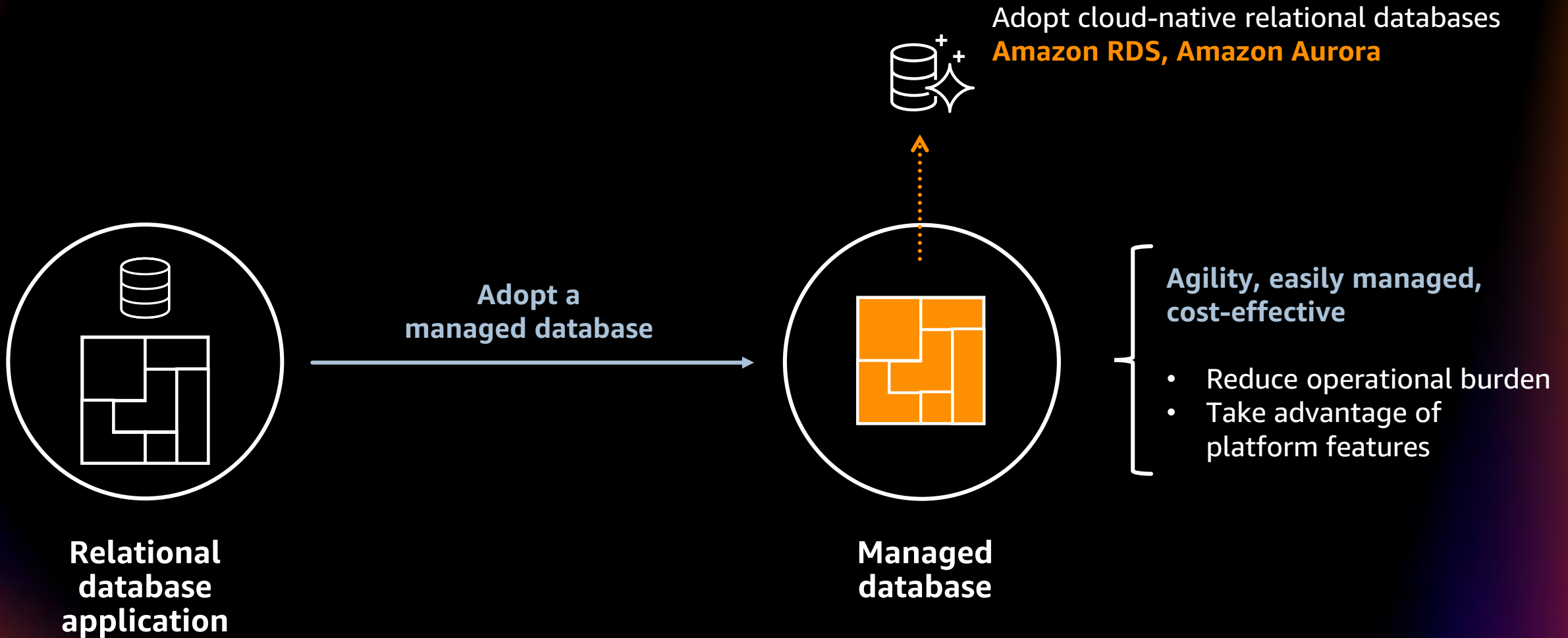


Modernizing with
purpose-built tools

The right tool for the job



Step 1: Move to a managed database



Amazon Relational Database Service (Amazon RDS)

MANAGED RELATIONAL DATABASE SERVICE WITH A CHOICE OF POPULAR DATABASE ENGINES



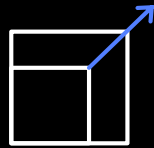
Microsoft
SQL Server

ORACLE®



Easy to administer

Easily deploy and maintain hardware, OS, and DB software; built-in monitoring



Performant & scalable

Scale compute and storage with a few clicks; minimal downtime for your application



Available & durable

Automatic multi-AZ data replication; automated backup, snapshots, and failover



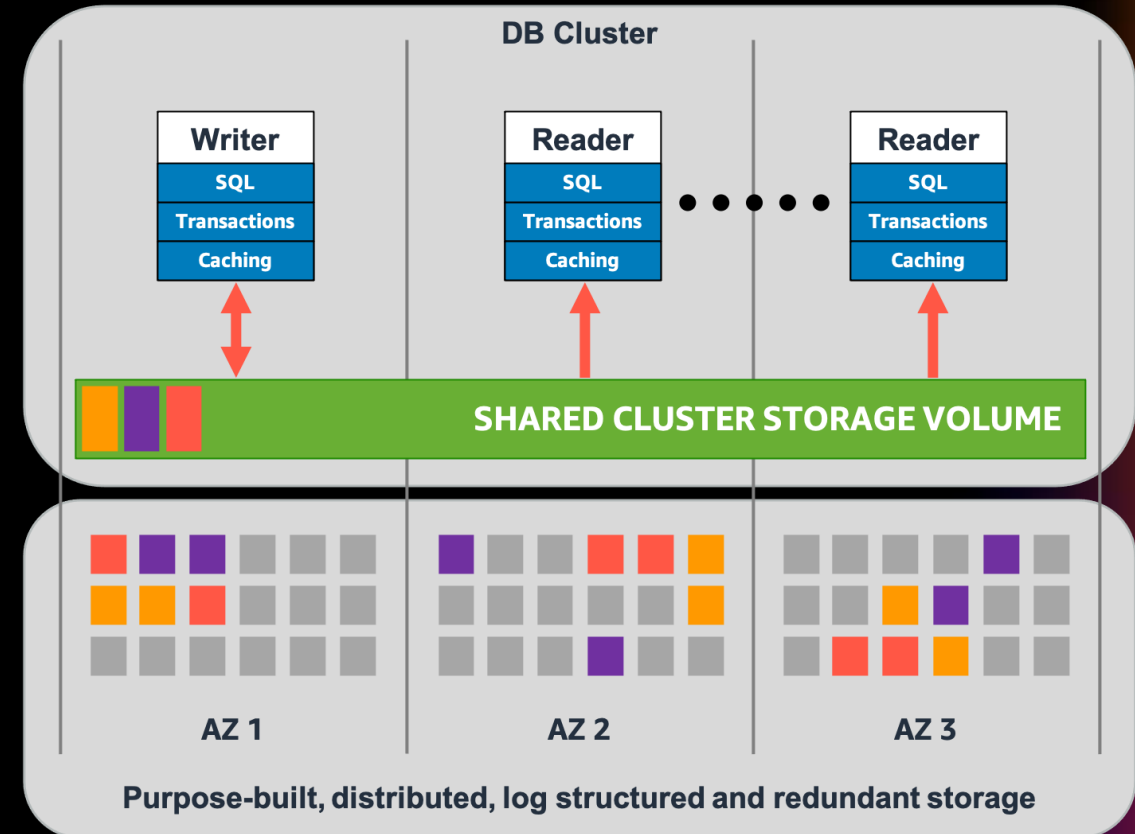
Secure & compliant

Data encryption at rest and in transit; assists customers to help them achieve their compliance goals through tooling

Step 2: Break free with Amazon Aurora

ENTERPRISE DATABASE AT OPEN-SOURCE PRICE, DELIVERED AS A MANAGED SERVICE

- Drop-in compatibility with MySQL and PostgreSQL
- Simplicity and cost-effectiveness of open-source databases
- Throughput and availability of commercial databases
- Simple pay-as-you-go pricing
- Global database with deployment in multiple Regions

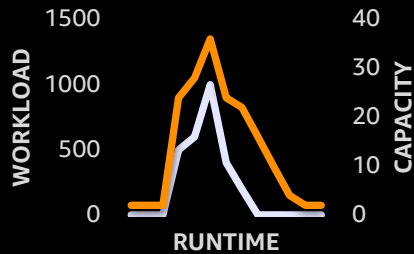


<https://aws.amazon.com/rds/aurora/>

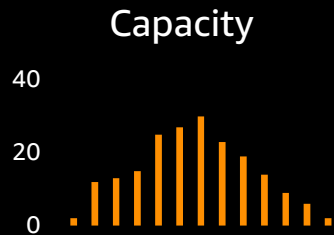
Introducing Amazon Aurora Serverless v2



An auto-scaling configuration for Amazon Aurora that now supports even the most demanding applications and database workloads



Scale instantly, from hundreds to **hundreds-of-thousands of transactions**, in a fraction of a second



Scale in **fine-grained increments** to provide just the right amount of database capacity



Full breadth of Amazon Aurora capabilities including multi-AZ, global database



Up to 90% cost savings when compared to provisioning for peak load

<https://aws.amazon.com/rds/aurora/serverless/>



Amdocs brings innovation and cloud benefits to RevenueONE by leveraging Amazon Aurora

Challenge

Amdocs needed to re-architect its RevenueONE solution to leverage the full benefits of the cloud and deliver the flexibility and scalability needed to support advanced monetization requirements in the 5G era

Solution

Amdocs migrated RevenueONE from a legacy commercial database to Amazon Aurora PostgreSQL to deliver a cloud-native solution for their customers and derive the benefits of a fully managed, pay-as-you-go database service

Results and benefits

- **Scalability** and cloud-native architecture to accelerate the launch of new 5G services
- **Reduced licensing costs** in favor of a pay-as-you-go model
- **Reduced operational overhead** by moving from self-managed to fully managed databases

“Prepping and getting a database cluster up and running took three weeks for installing, confirming network, testing for latency. On Amazon Aurora we’re now able to effectively do it in a day.”

Jay Deen
CTO, Amdocs Media

[Learn more](#)



Break Free Tools

HOMOGENEOUS AND HETEROGENEOUS



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

AWS Database Migration Service (AWS DMS) easily and securely migrates and/or replicates your databases **and** data warehouses to AWS



Babelfish for Aurora PostgreSQL

RUN SQL SERVER APPLICATIONS ON POSTGRESQL WITH LITTLE TO NO CODE CHANGE



Reduce migration time and risk

Migrate at your own pace

Source code available on GitHub

Modernizing with purpose-built tools

DEVELOPERS WANT THE RIGHT DATABASE TO MEET APPLICATION'S UNIQUE REQUIREMENTS

Lack of scalability



Lack of developer flexibility



Performance issues



Inability to integrate all data types



One-size-fits-all approach of using a relational database for everything is **no longer working**

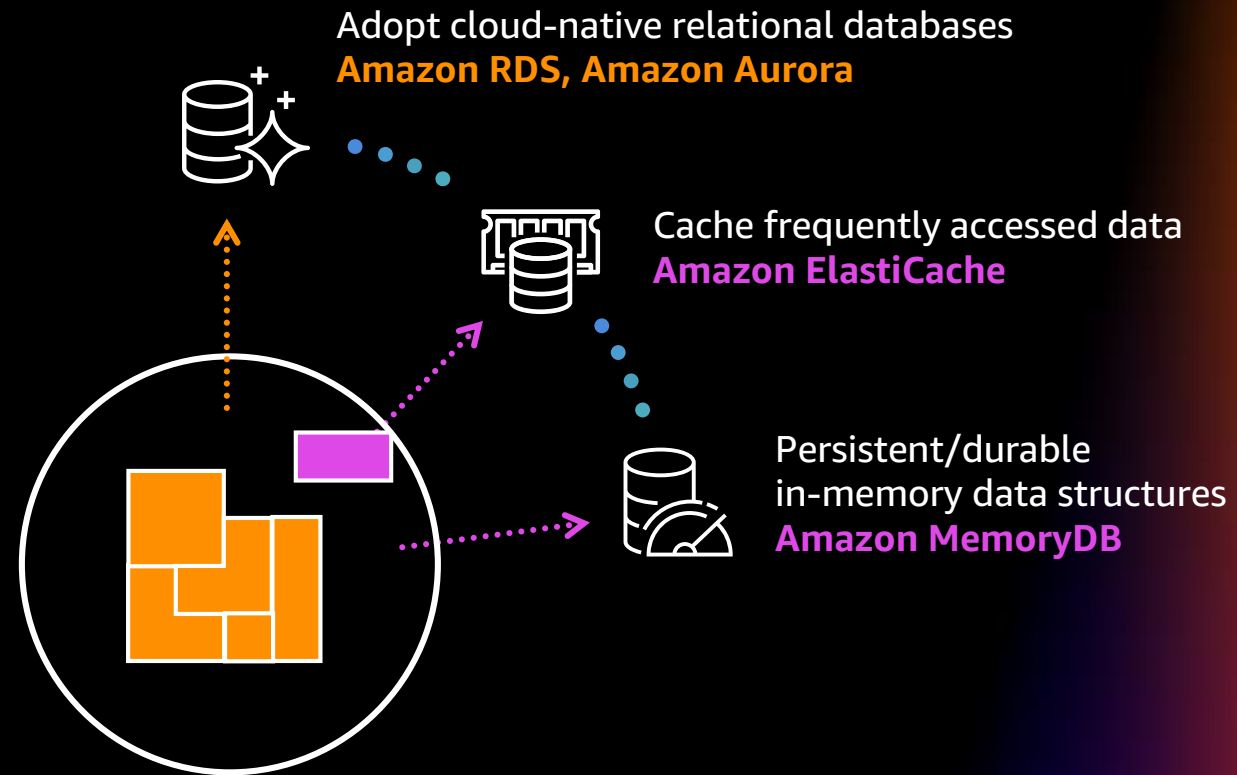
Step 3: Cache and persistent in-memory data

Address scaling pain

- Low-hanging fruit, easy to adopt
- Offload repetitive read operations
- Offload session state
- Caching

Increase performance

- Frequent counters
- Fast-changing rankings
- Submillisecond data access



Amazon ElastiCache

MANAGED, REDIS, OR MEMCACHED-COMPATIBLE IN-MEMORY DATA STORE



Fully managed



**Consistent high
performance**



Unlimited scale

Amazon MemoryDB for Redis

REDIS-COMPATIBLE, DURABLE, IN-MEMORY DATABASE SERVICE



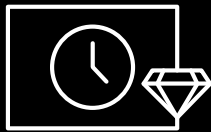
Fully managed



Ultra-fast performance



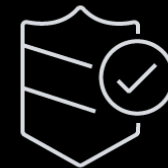
Redis compatibility



**Durability and high
availability**



High scalability



Security



PayPay Corporation is one of Japan's leading mobile payments applications, with more than 30 million users completing millions of cashless transactions every day. PayPay has been using Amazon ElastiCache for Redis both as a distributed cache and a low-latency persistent data store.

Challenge

The team was asked to build the QR Code Payment Service in a short period of 3 months.

Solution

PayPay adopted a microservices architecture on AWS designed to support the growth of the business. Each microservice team can choose the right database solution according to the application requirements.

“We chose ElastiCache because we needed a fully managed and highly scalable distributed Redis cluster for numerous microservices. For example, our authentication service is based on access tokens stored in ElastiCache for Redis. Because every API call requires authentication, this service handles a huge amount of traffic per second and requires submillisecond performance. ElastiCache for Redis makes this possible.”

Shilei Long
Senior Software Architect
PayPay Corporation

<https://aws.amazon.com/solutions/case-studies/paypay/>



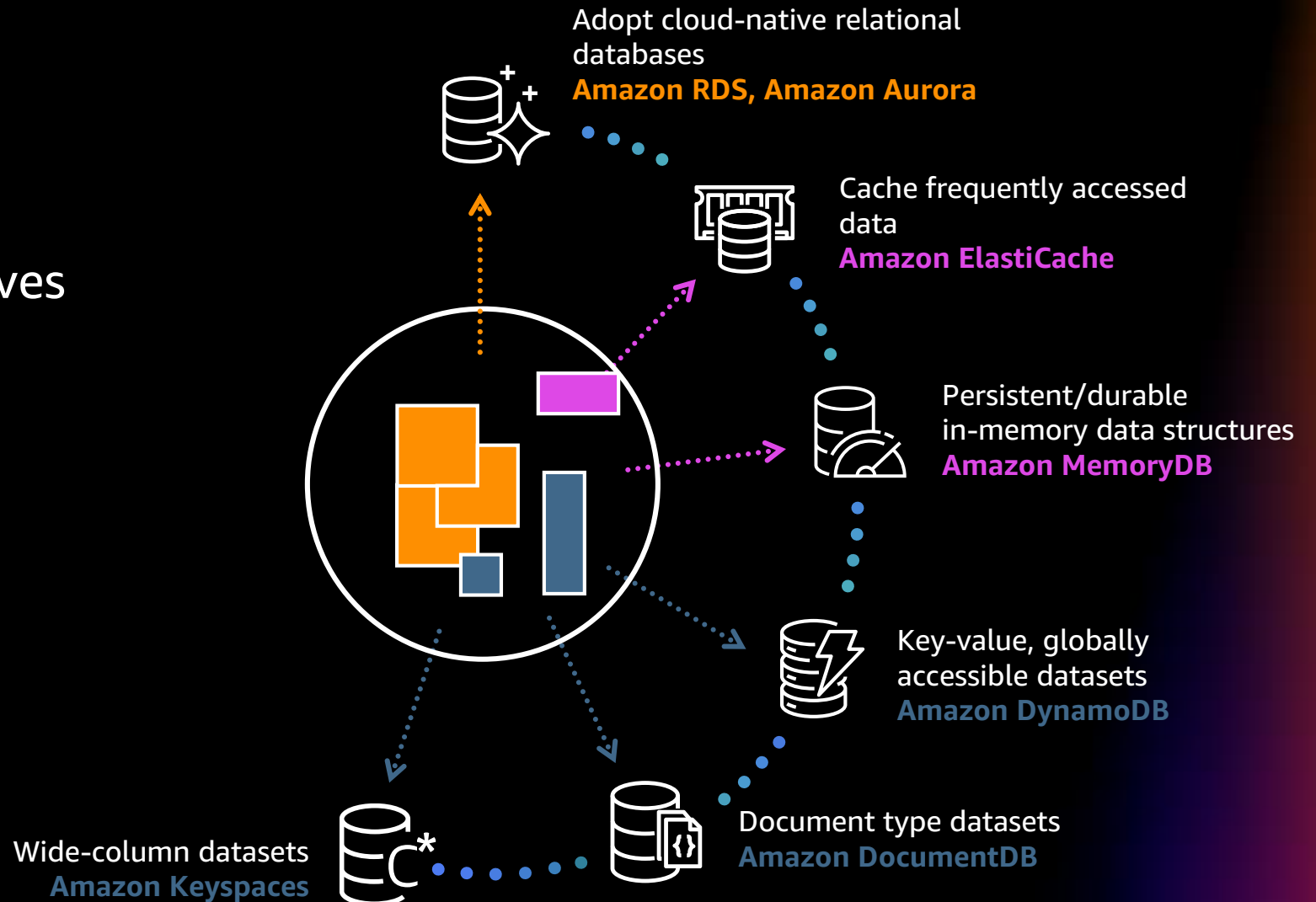
Step 4: Nonrelational query patterns

Breaking down the monolith

- Right tool for the job
- Microservice access pattern drives data store decision

Benefits

- Agility
- Scalability
- Performance



Key-value database

- Simple key-value pairs
- Partitioned by keys
- Resilient to failure
- High-throughput, low-latency reads and writes
- Consistent performance at scale

```
PUT {
  TableName: "Gamers",
  Item: {
    "GamerTag": "Hammer57",
    "Level": 21,
    "Points": 4050,
    "Score": 483610,
    "Plays": 1722
  }
}
```

```
GET {
  TableName: "Gamers",
  Key: {
    "GamerTag": "Hammer57",
    "ProjectionExpression": "Points"
  }
}
```

Gamers				
Primary key	Attributes			
Gamer tag	Level	Points	High score	Plays

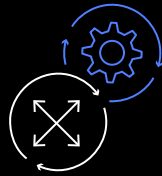
Amazon DynamoDB

FAST AND FLEXIBLE NOSQL DATABASE AT ANY SCALE



Performance at scale

- Delivers consistent, single-digit millisecond latency
- Handles millions of requests per second



No servers to manage

- Maintenance free
- Auto-scaling
- On-demand capacity mode
- Up to 99.999% SLA



Enterprise-ready

- ACID transactions
- Encryption at rest
- Continuous backups, on demand backup and restore
- Integration with other AWS services



Global tables

- Build global applications
- Get fast access to local data
- Automated global replication



Zoom Video Communications, Inc. provides videotelephony and online chat services through a cloud-based peer-to-peer software platform and is used for teleconferencing, telecommuting, distance education, and social relations.

Challenge

Beginning in early 2020, Zoom's software usage saw a significant global increase after quarantine measures were adopted in response to the COVID-19 pandemic.

Solution

Amazon DynamoDB global tables using on-demand capacity allowed Zoom to scale and meet that global increase in demand for their services.

<https://aws.amazon.com/dynamodb/customers/>



“ On the backend, we were able to manage this surge with Amazon DynamoDB for Zoom Meetings. Using DynamoDB global tables in conjunction with on-demand mode enabled us to scale nearly infinitely with no performance issues, even with our sudden spike in usage. ”

Yasin Mohammed

Engineering Manager, Cloud Operations
Zoom Video Communications, Inc.

Why document databases?



Documents map naturally to how humans model data



Documents (objects/JSON) are common application data models



Document databases store JSON-like documents



Document databases provide flexible schema and indexing



Ad hoc querying and aggregations

```
#Ad hoc queries
db.col.find( { "name.last": "Smith" } )

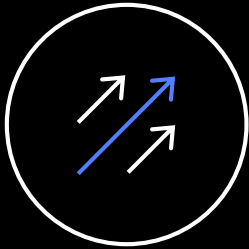
#Count of documents
db.col.find({ product: "xyz", region: "msp" }).count()

#Return distinct fields
db.col.distinct( "region" )

#Power aggregations over documents
db.col.aggregate([
    { $match: { product: "xyz" } },
    { $group: { _id: "$c_id", rev: { $sum: "$price" } } },
    { $sort: { rev: -1 } }
])
```

Amazon DocumentDB

FAST, SCALABLE, AND FULLY MANAGED MONGODB-COMPATIBLE DATABASE SERVICE



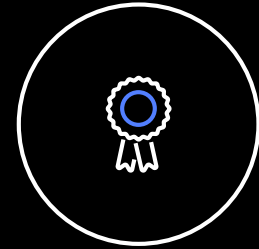
Fast and scalable

- Scale compute in minutes
- Storage and I/O autoscaling
- Storage scales to 64 TB
- Scale out to 15 replicas for millions of reads
- Globally distributed



Enterprise-ready

- Built-in high availability
- Backups enabled by default
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting

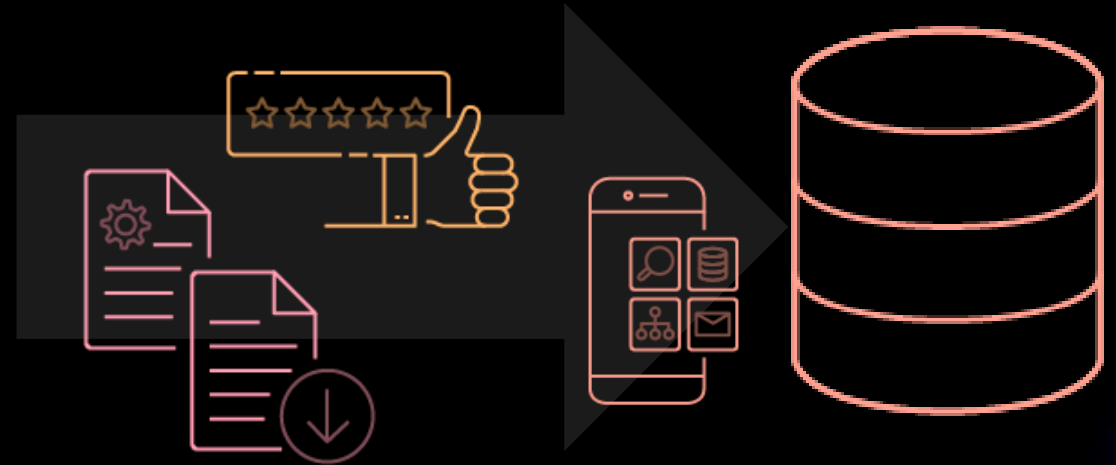


MongoDB-compatible

- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

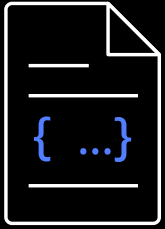
Wide column: Apache Cassandra

- Open-source, wide-column data store
- Cassandra Query Language (CQL)
- Large-scale applications that require fast read and write performance



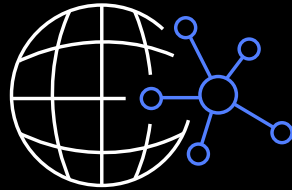
Amazon Keyspaces

FULLY MANAGED APACHE CASSANDRA-COMPATIBLE DATABASE SERVICE



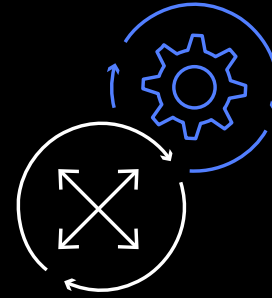
Compatible with Apache Cassandra

- Existing applications
- New applications
- Developer skill sets



Performance at any scale

- Handles millions of requests per second
- Delivers single-digit-millisecond latency



Serverless

- No compaction
- Fully managed
- Autoscaling



Enterprise-ready

- Point-in-time recovery backups
- Encryption at rest
- IAM integration

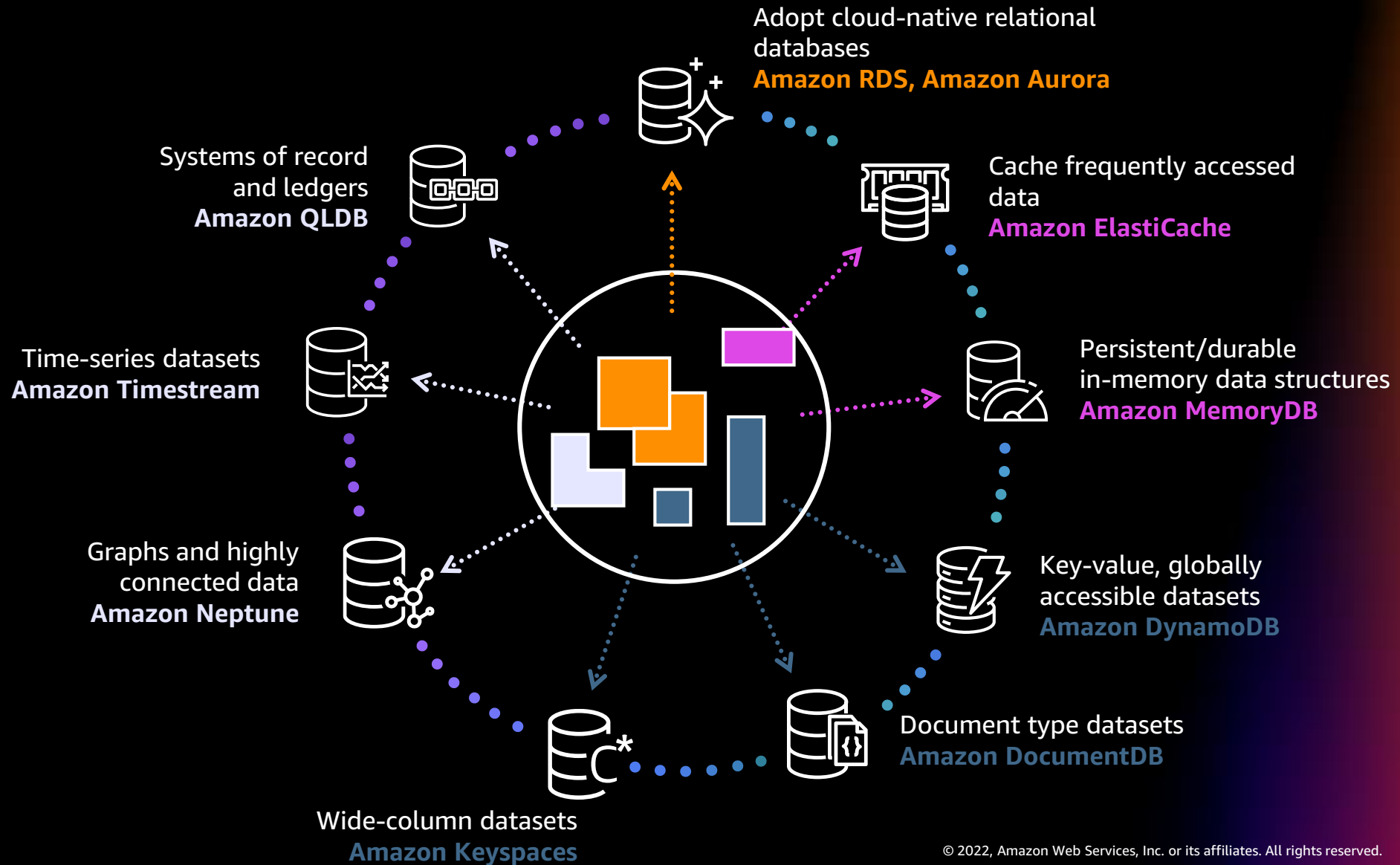
Step 5: Specialized data interactions

Specialized datasets

- Social graphs
- Recommendation engines
- Measurements
- Systems of record
- Digital records

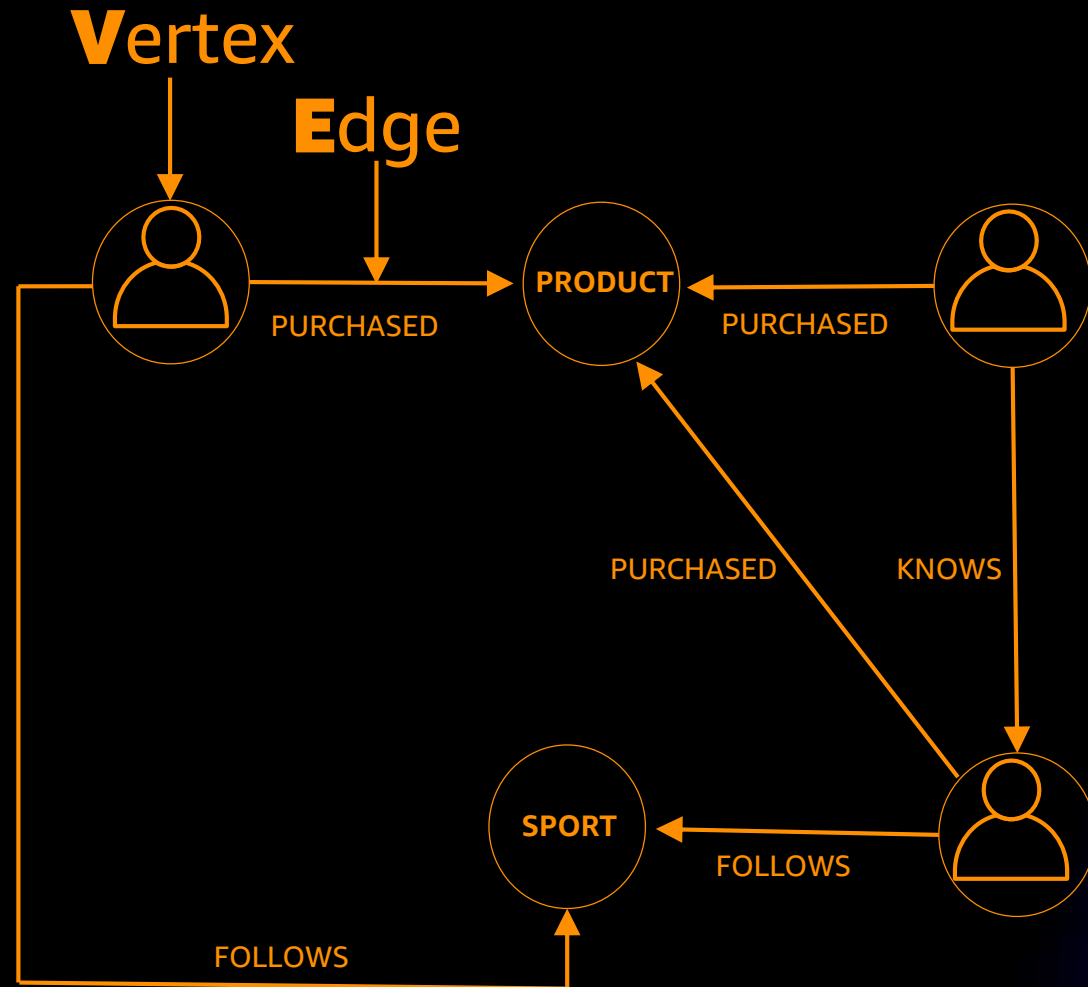
Benefits

- Agility
- Scalability
- Performance



Graph data

- Relationships are first-class objects
- Vertices connected by Edges



Amazon Neptune

FAST RELIABLE GRAPH DATABASE BUILT FOR THE CLOUD

Open



Supports Property Graphs &
W3C RDF graph models

Fast



Query billions of relationships
with millisecond latency

Reliable



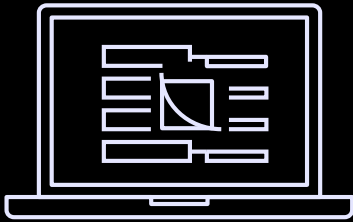
6 replicas of your data across
3 AZs with full backup and
restore

Easy



Build powerful queries easily
with Gremlin and SPARQL

Time-series use cases



Clickstream
analytics



Monitoring and
maintenance



Transaction
analytics

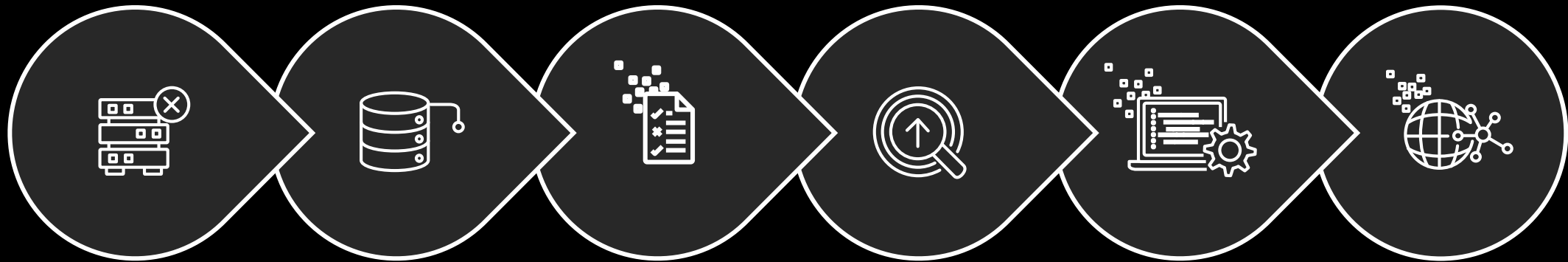


DevOps
analysis

Amazon Timestream

SERVERLESS, FULLY MANAGED TIME SERIES DATABASE

Collect, store, and process data sequenced by time **at any scale**



Serverless auto
scaling for ingesting,
storing, and querying

Data lifecycle
management

No predefined
schema

Access data across
storage tiers with
adaptive query engine

Built-in
time series
analytics in SQL

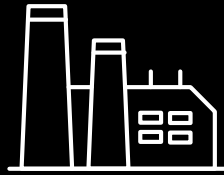
Integrates with
ML, analytics,
data collectors

Ledger databases



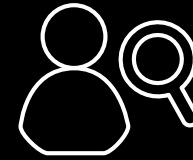
Banking and finance

Keeping track of transactions, trades, and accounts



Manufacturing

Recording components used in manufacturing



Ownership

Maintaining records of asset ownership

Amazon QLDB

FULLY MANAGED LEDGER DATABASE SERVICE

IMMUTABLE



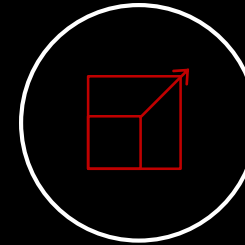
Append-only, immutable journal **tracks history of all changes**

VERIFIABLE



All changes are **cryptographically chained** and **verifiable**

SERVERLESS



Scalable and highly available, no need to provision storage and I/O in advance

EASY TO USE



PartiQL offers familiar SQL operators, ION data model is rich and expressive

The Zerobase Foundation offers a free, open-source, privacy-first contact tracing platform to help stop the spread of COVID-19

Challenge:

Zerobase needed a secure and performant database solution to support its privacy-first contact tracing app, helping to track the spread of COVID-19 while keeping user data private

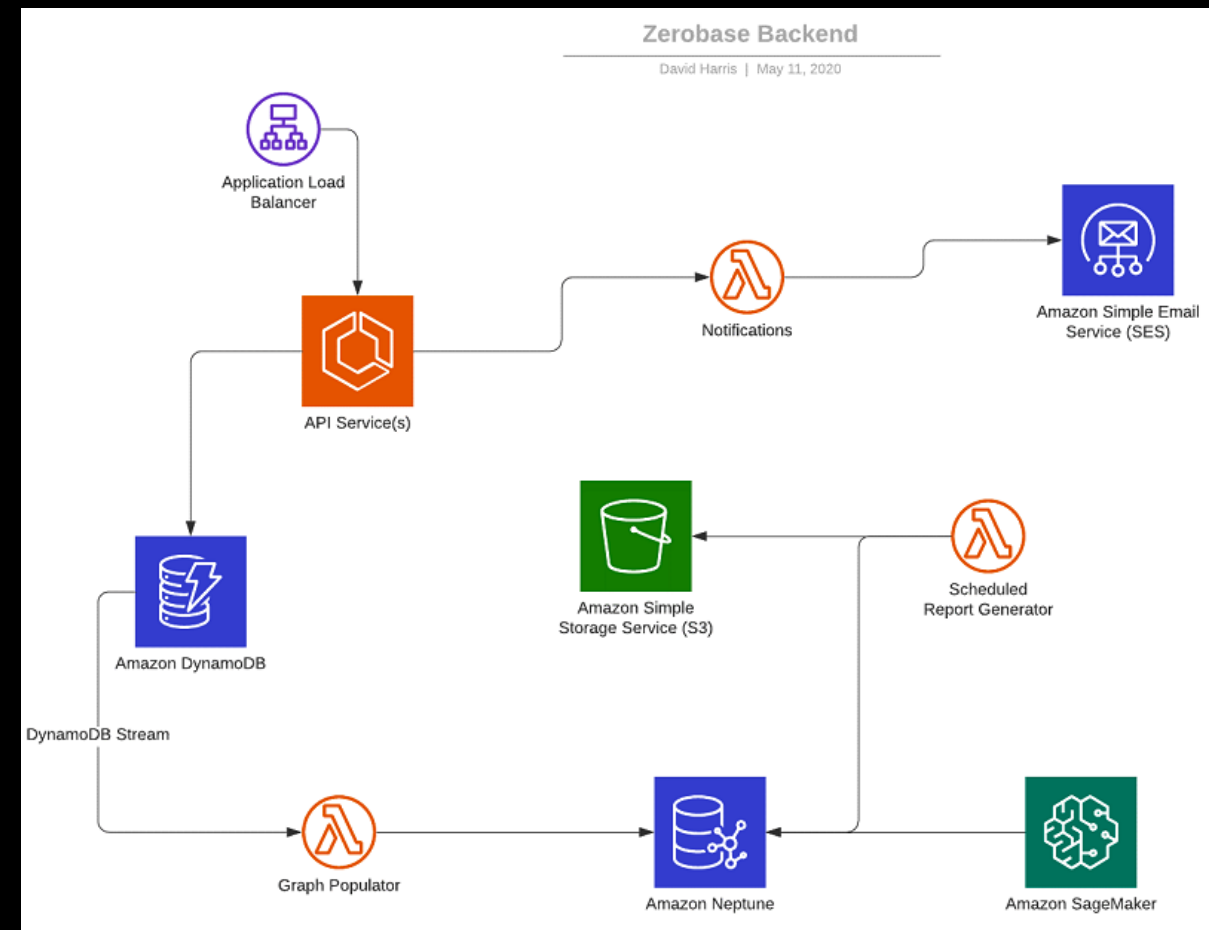
Solution:

Zerobase opted for an event-driven architecture using Amazon DynamoDB as the primary data store backing its API and Amazon Neptune to power its contact graph for tracking potential exposures.

Results:

With AWS purpose-built databases, Zerobase is able to maintain high availability, consistency, low latency, and security for its contact tracing solution, while only paying for the compute they require

<https://aws.amazon.com/blogs/database/zerobase-creates-private-secure-and-automated-contact-tracing-using-amazon-neptune/>

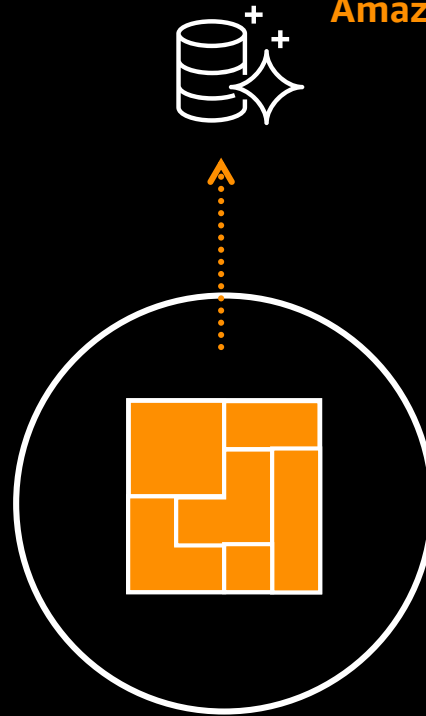


To sum it all up

1. Move workloads to managed DB services

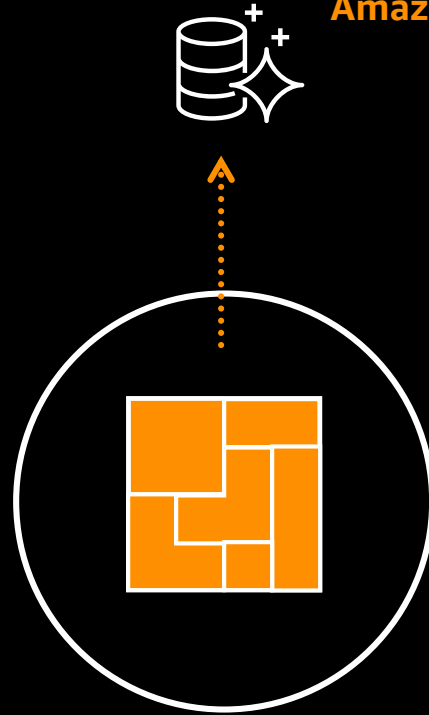
Adopt cloud-native relational databases

Amazon RDS

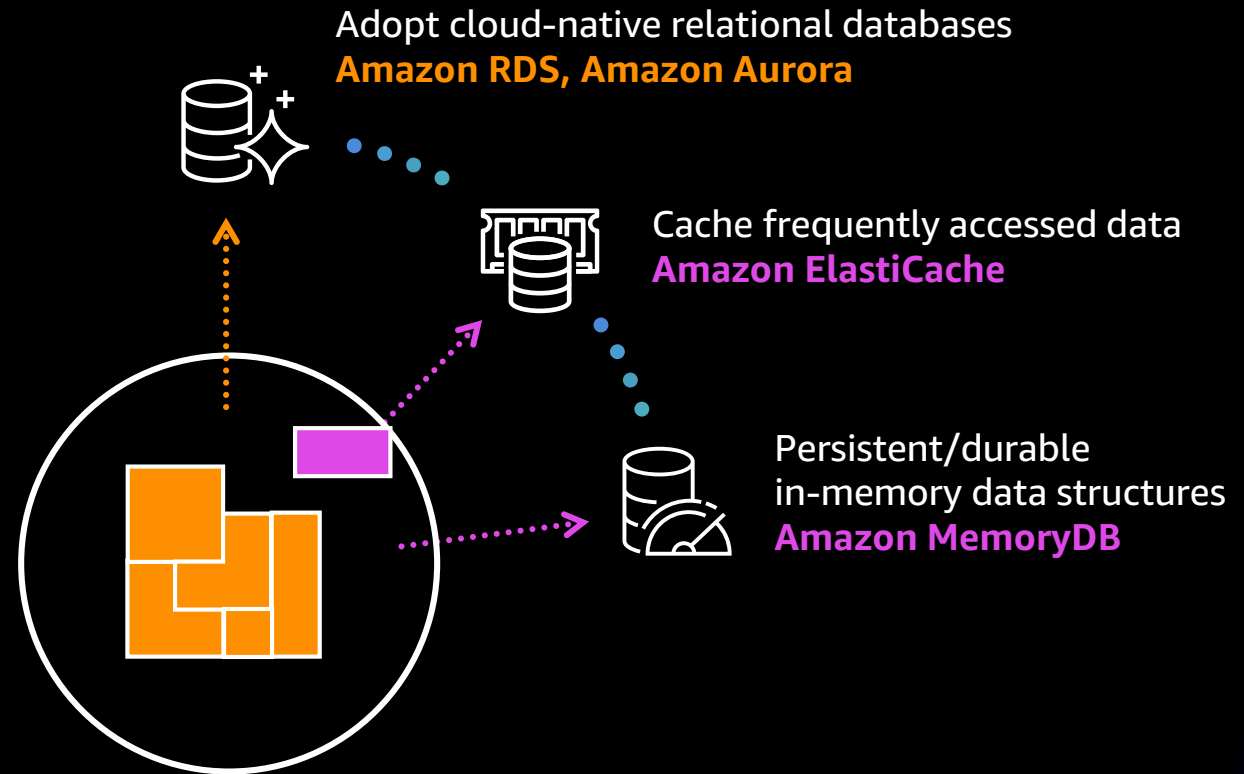


2. Break free with Open Source Databases

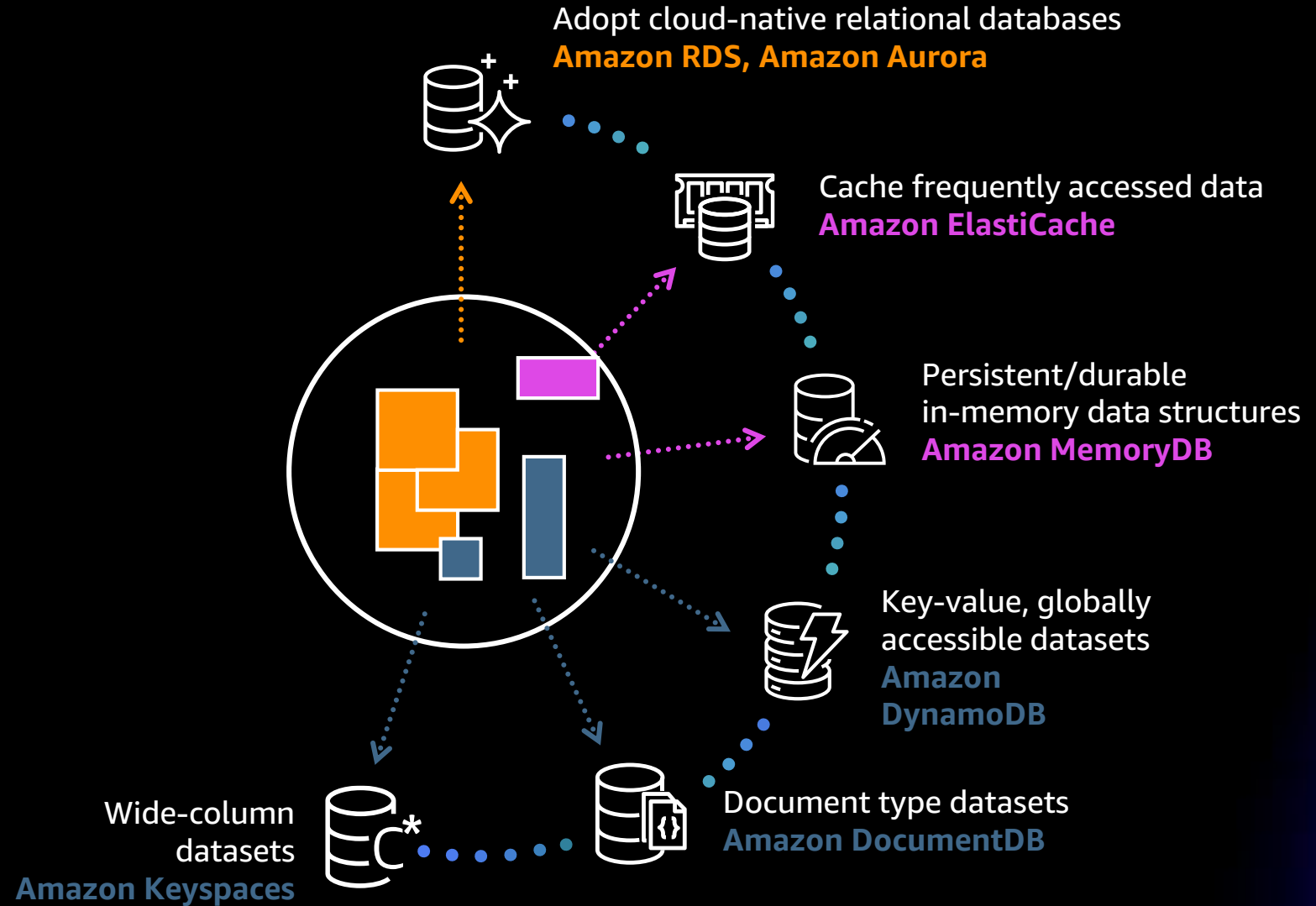
Adopt cloud-native relational databases
Amazon RDS, Amazon Aurora



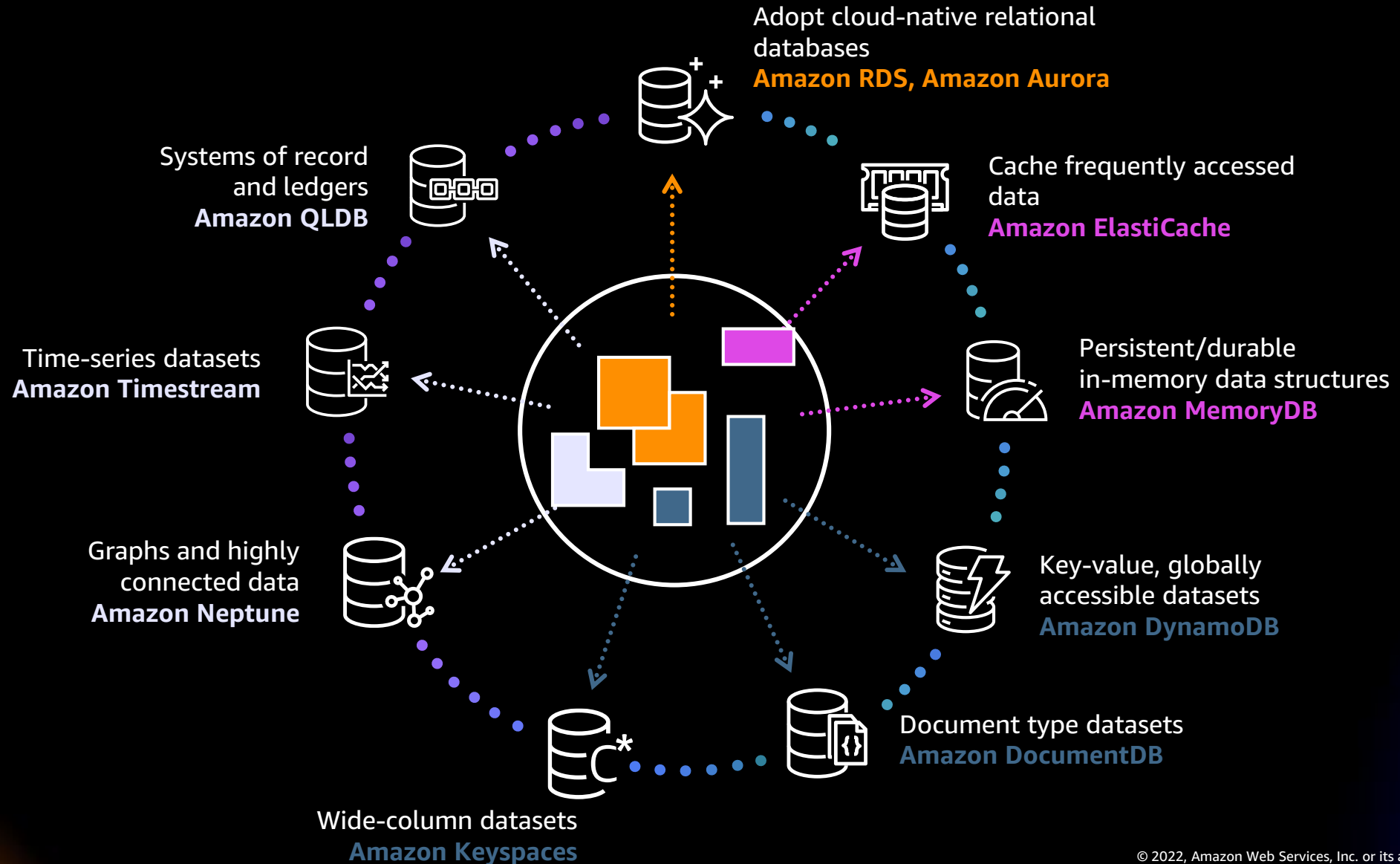
3. Cache and offload in-memory workloads



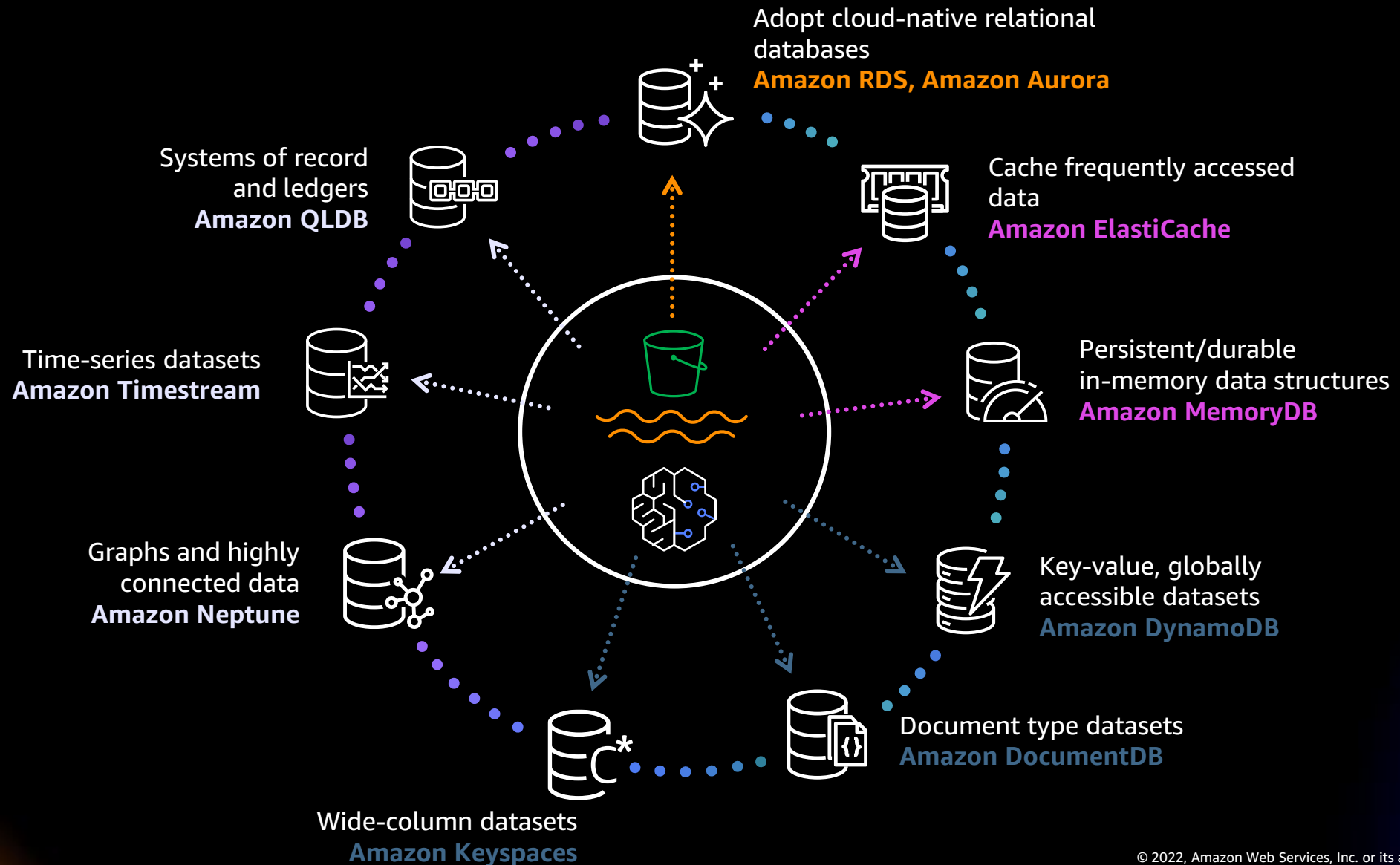
4. Move nonrelational query patterns to NoSQL



5. Move specialized dataset to specialized DBs



Unify using data lakes, and innovate using ML



Visit the AWS Data resource hub

A modern data strategy can help you manage, act on, and react to your data so you can make better decisions, respond faster, and uncover new opportunities. Dive deeper with these resources today.

- Harness data to reinvent your organization
- In unpredictable times, a data strategy is key
- Make data a strategic asset
- Rewiring your culture to be data-driven
- Put your data to work with a modern analytics approach
- ... and more!



<https://tinyurl.com/data-hub-aws>

[Visit resource hub](#)

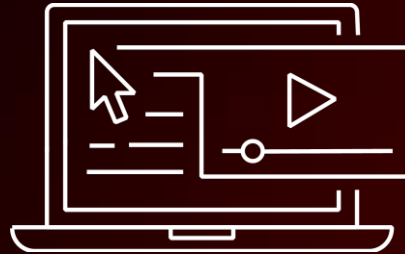
AWS Training and Certification for Data and Analytics



AWS Data & Analytics FREE Training Resources

Discover how to harness data, one of the world's most valuable resources, and innovate at scale.

<https://bit.ly/3Ntlhy7>



AWS Data Analytics Learning Plan

This learning plan expose you to the fastest way to get answers from all your data to all your users. It can also help prepare you for the AWS Certified Data Analytics - Specialty certification exam.

<https://bit.ly/3wBVjD1>



AWS Certified Data Analytics - Specialty

Earning AWS Certified Data Analytics – Specialty validates expertise in using AWS data lakes and analytics services.

<https://go.aws/3lwFORR>

Thank you for attending AWS Innovate – Data Edition

We hope you found it interesting! A kind reminder to **complete the survey**.
Let us know what you thought of today's event and how we can improve the event experience for you in the future.



aws-apj-marketing@amazon.com



twitter.com/AWSCloud



facebook.com/AmazonWebServices



youtube.com/user/AmazonWebServices



slideshare.net/AmazonWebServices



twitch.tv/aws

Thank you!