

The background features a vibrant, multi-colored gradient. It starts with a dark blue on the left, transitions through purple and magenta, and then into bright orange and yellow towards the right. A diagonal line separates the darker blue on the left from the lighter colors on the right.

AWS
re:Invent

A M Z 3 0 1

Amazon.com: Enterprise database migration at scale

Thomas Park

Sr. Manager Software Development
Amazon.com

Doug Booth

Principal Business Development Manager
Amazon Web Services

Agenda

Amazon's legacy database infrastructure

Reasons to migrate databases

Migration Strategy and lessons learned

Benefits realized


Review of Documented Cased Studies

Video



Amazon's Legacy Database Infrastructure

In the beginning



Welcome to Amazon.com Books!

*One million titles,
consistently low prices.*

(If you explore just one thing, make it our personal notification service. We think it's very cool!)

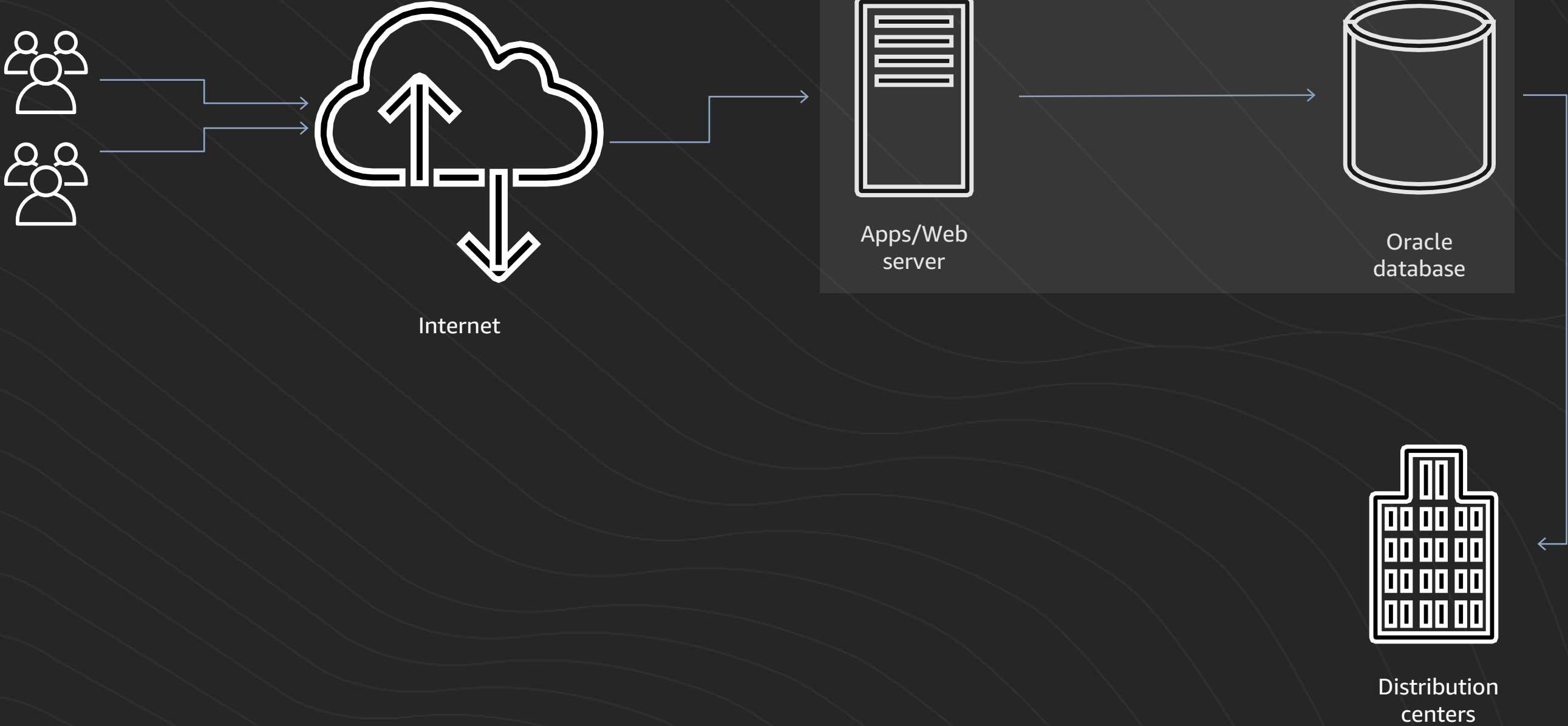
SPOTLIGHT! -- AUGUST 16TH

These are the books we love, offered at Amazon.com low prices. The spotlight moves **EVERY** day so please come often.

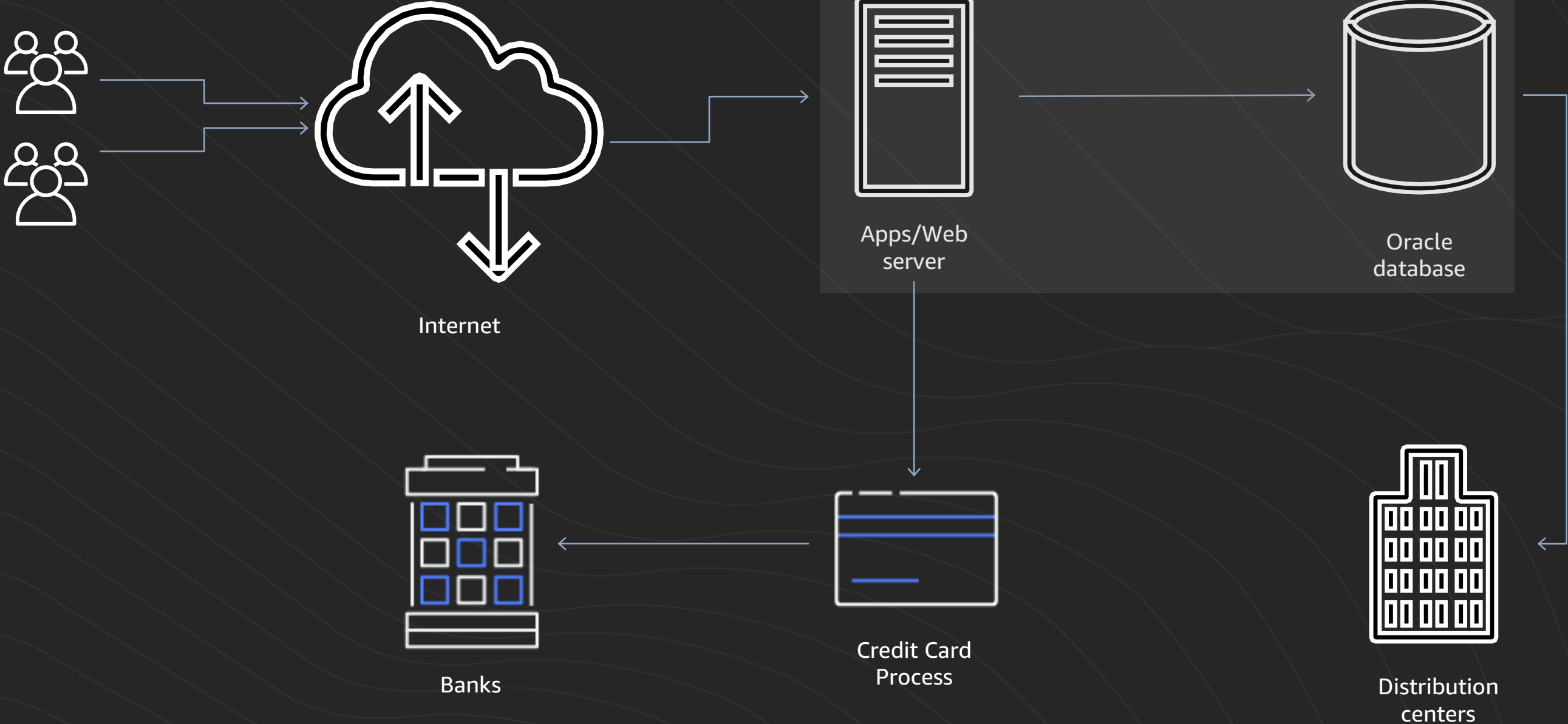
ONE MILLION TITLES

Search Amazon.com's [million title catalog](#) by author, subject, title, keyword, and more... Or take a look at the [books we recommend](#) in over 20 categories... Check out our [customer reviews](#) and the [award winners](#) from the Hugo and Nebula to the Pulitzer and Nobel... and [bestsellers](#) are 30% off the publishers list...

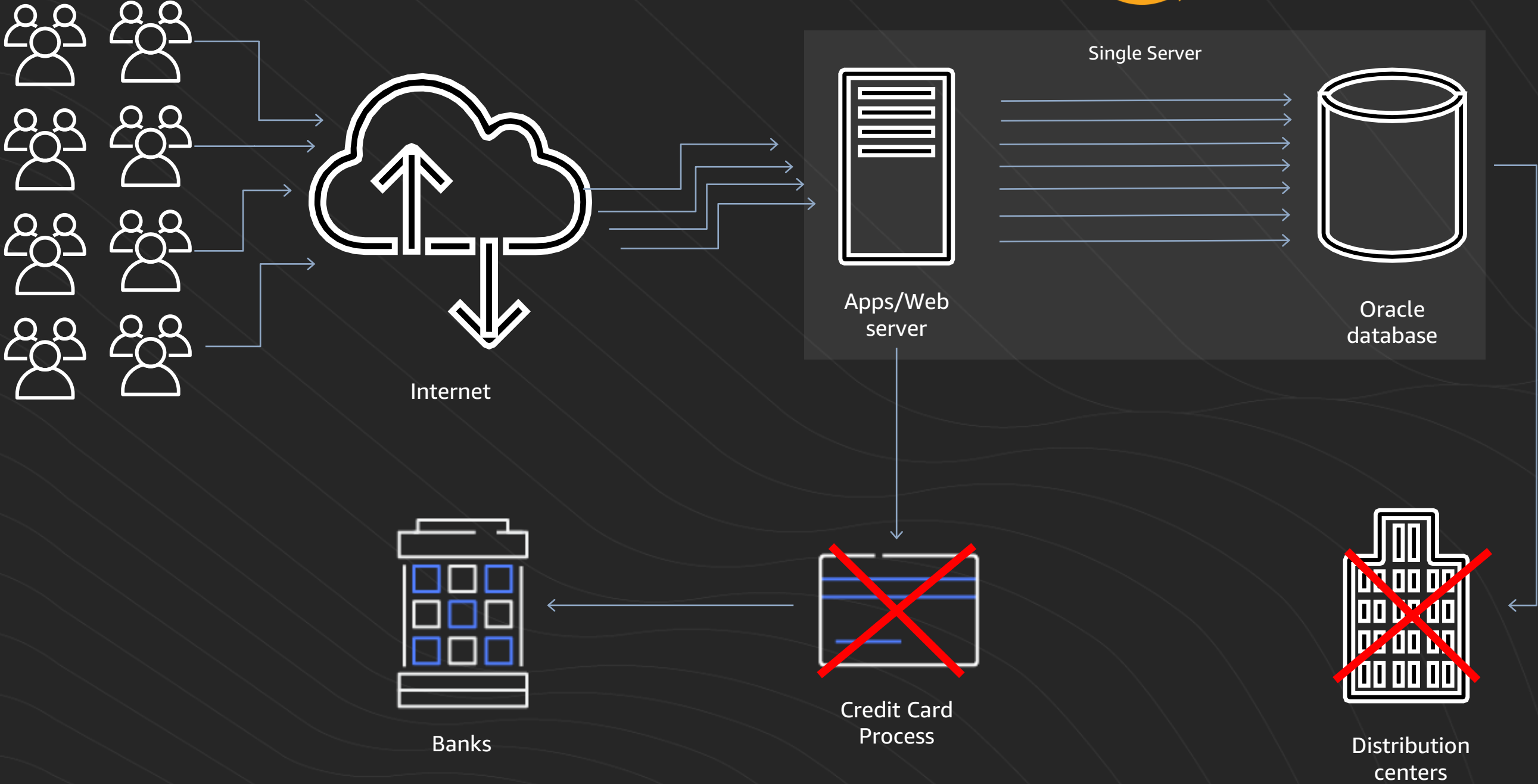
amazon

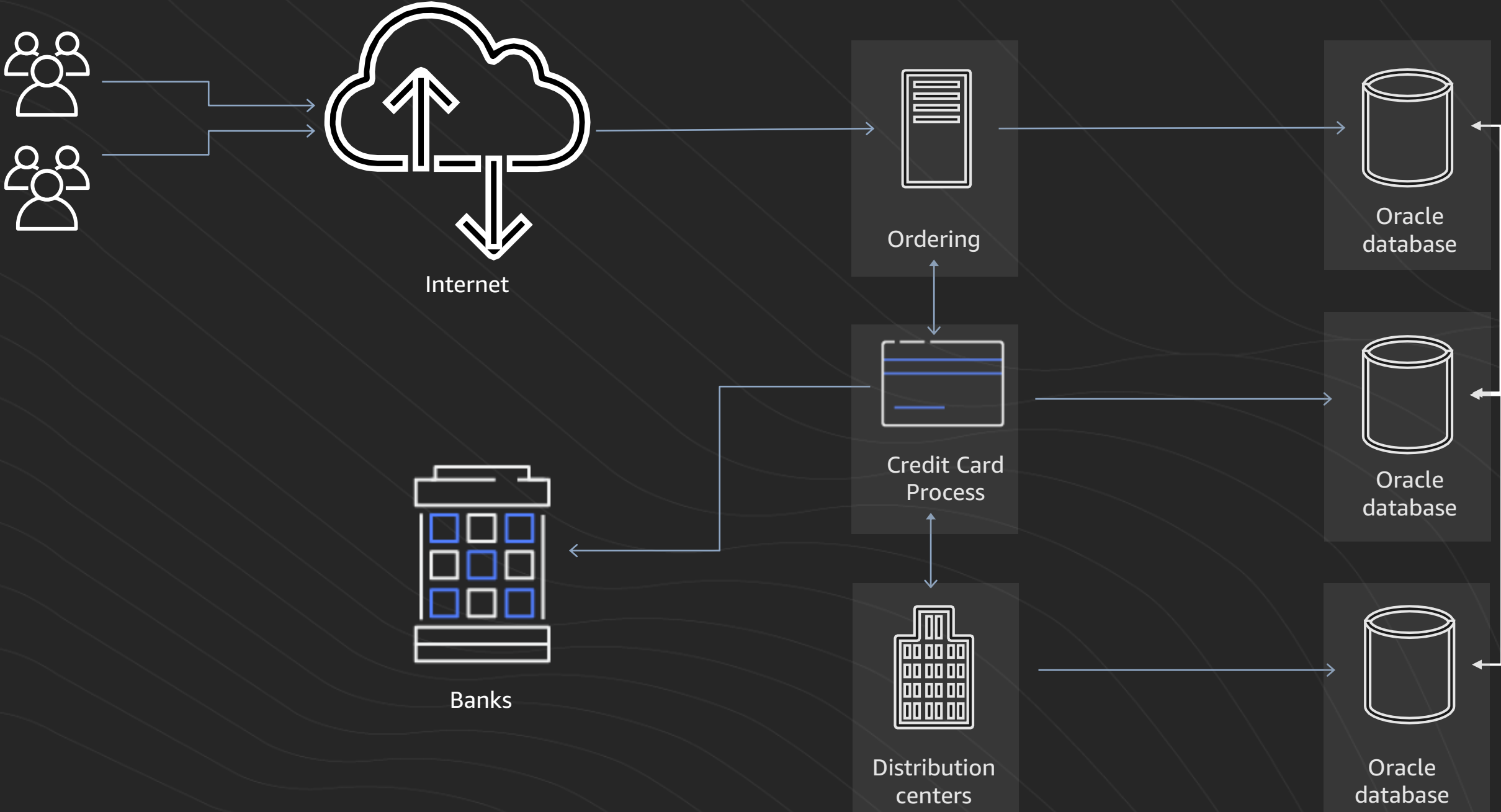


amazon

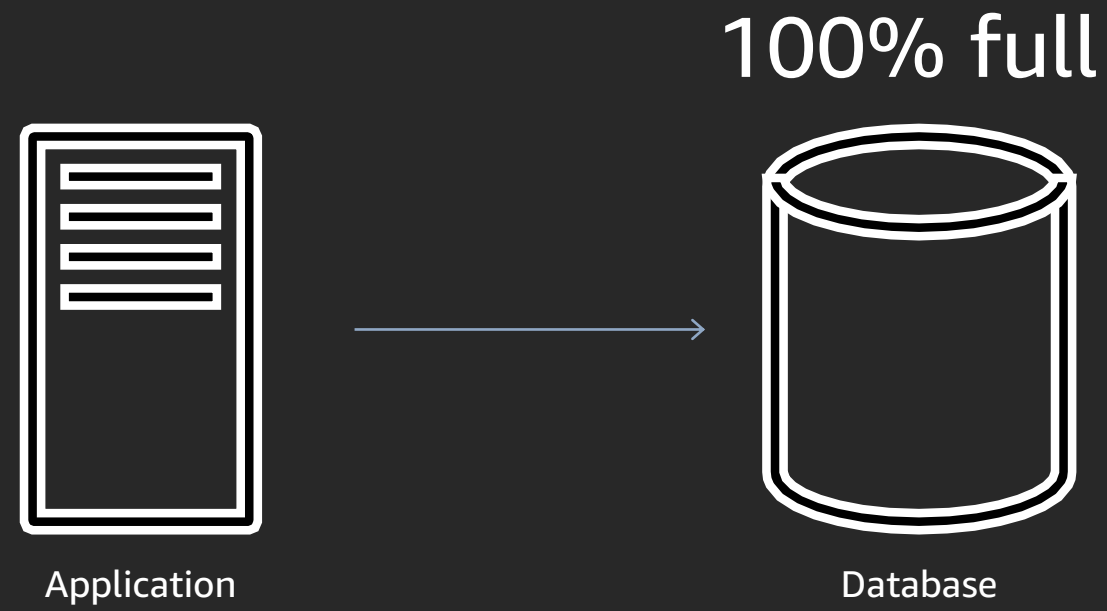


amazon

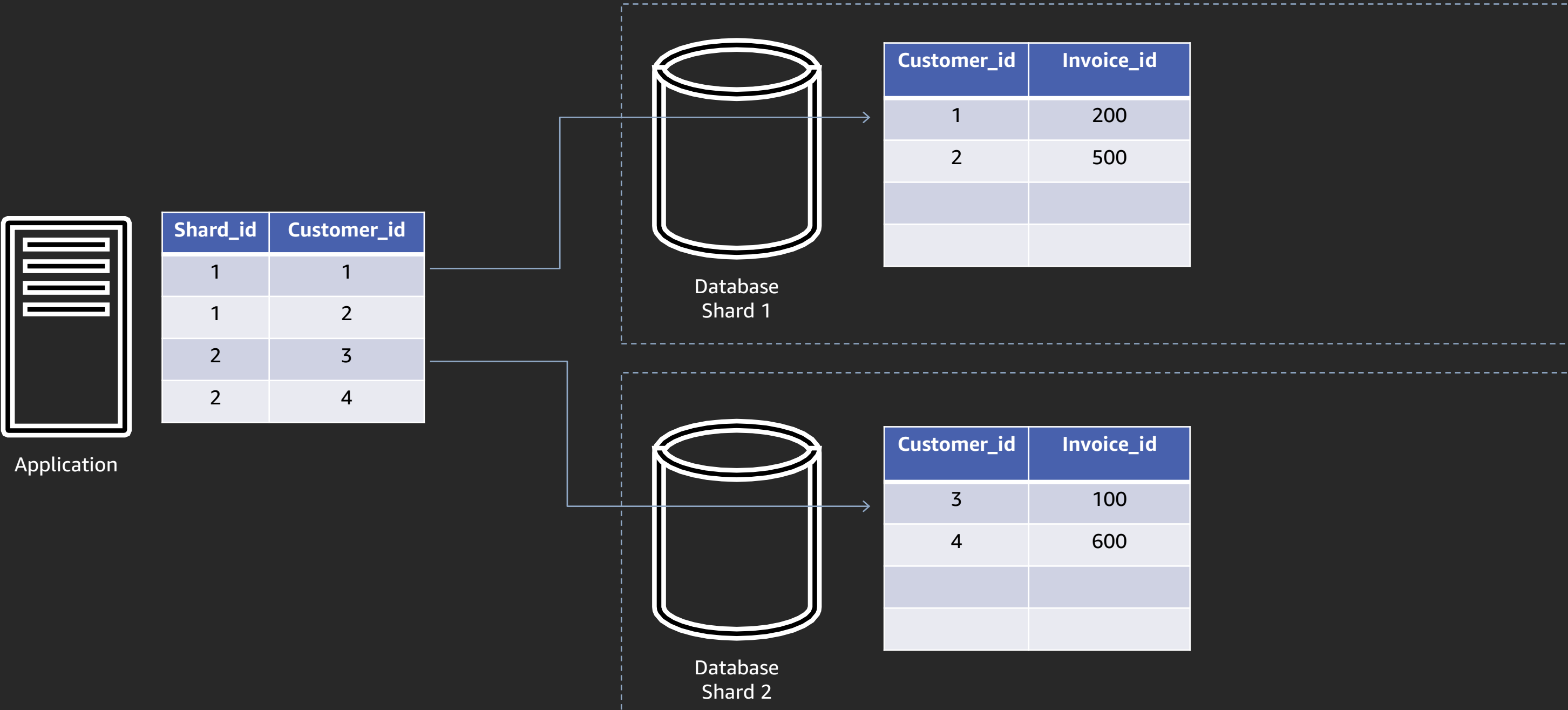




DB Sharding in relational databases



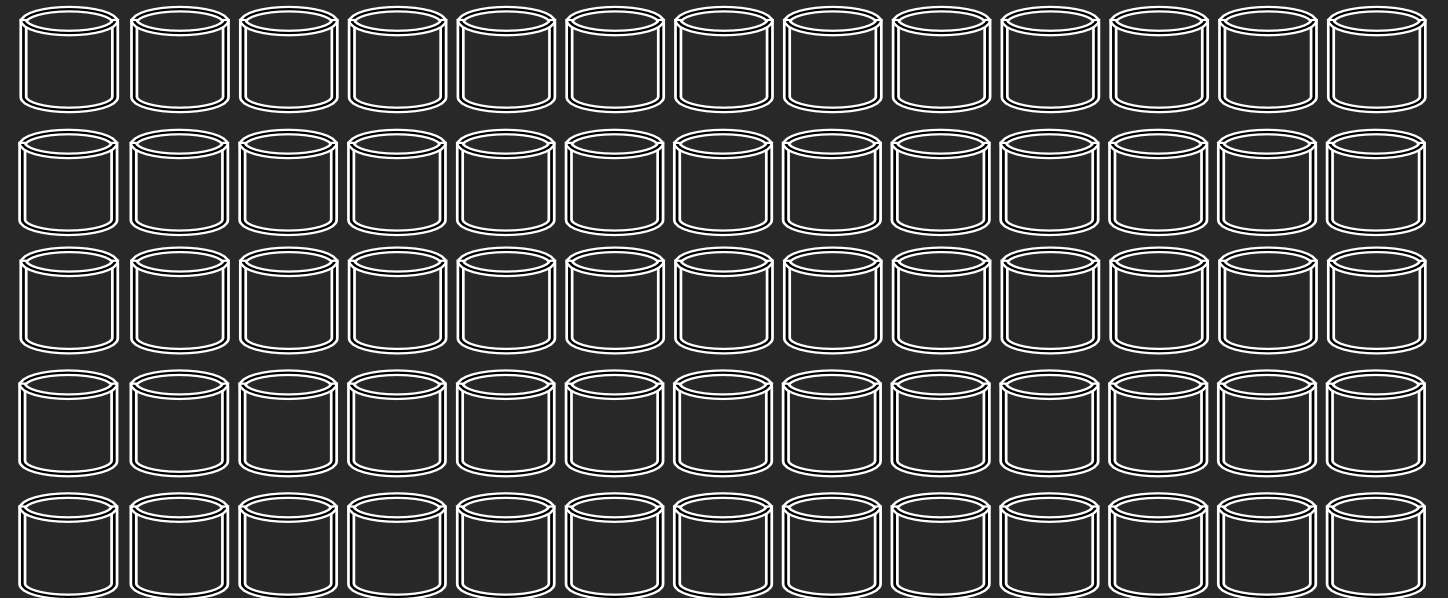
Database sharding in relational databases



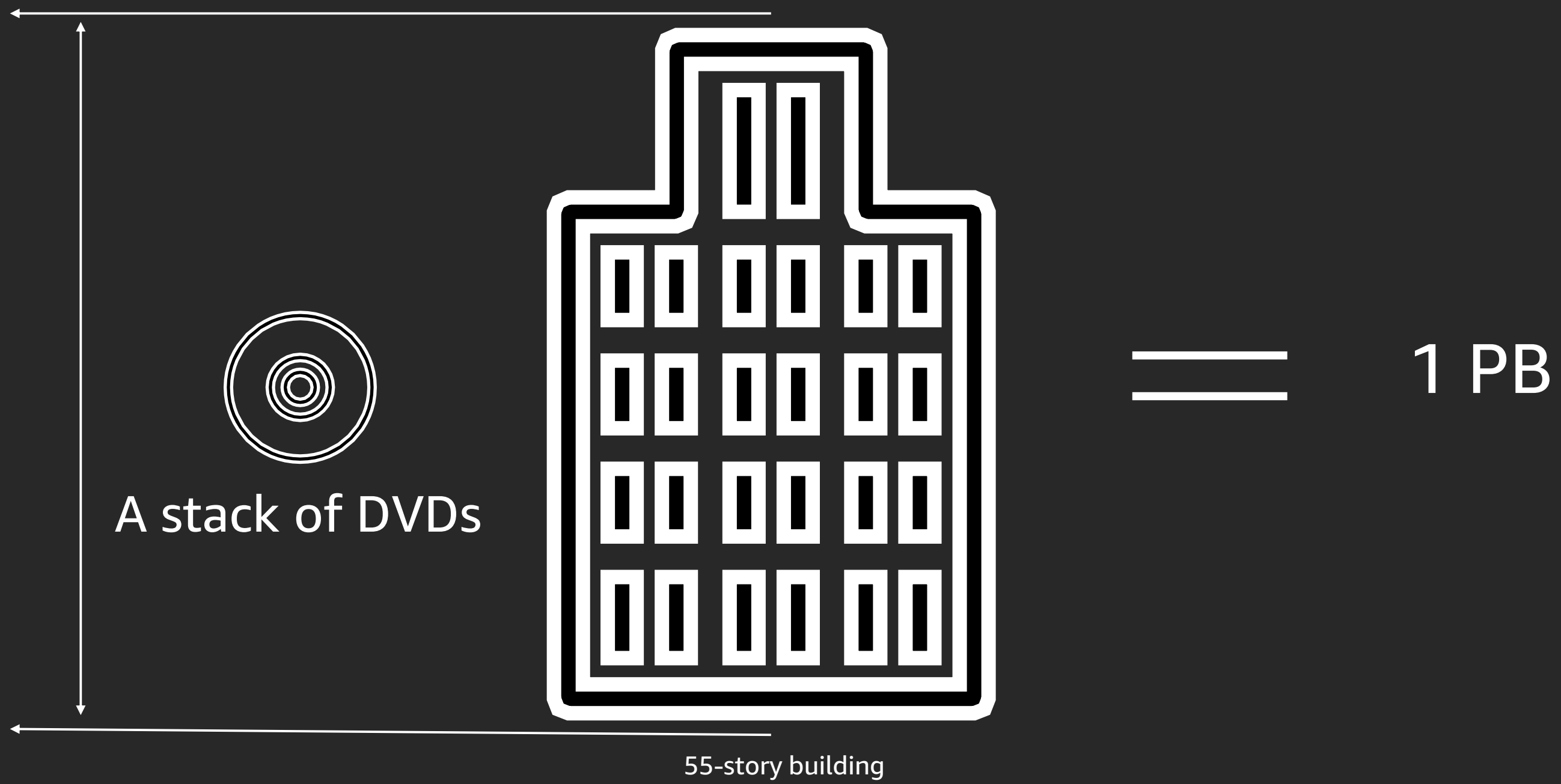
Legacy database infrastructure in 2017

OLTP Databases

- 75PB of data
- Nearly 7,500 OLTP databases
- Over 100 teams with 1000s of applications
- Processing millions of orders, payments, and fulfillments around the world



How big is 1PB of data?



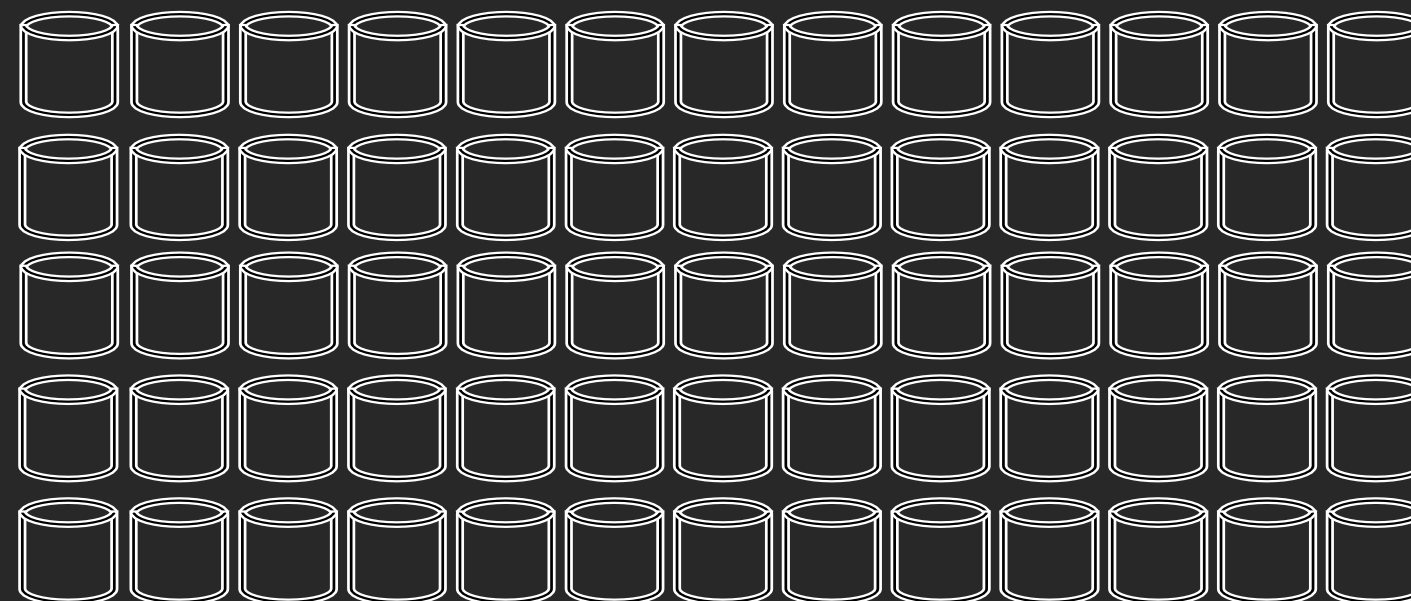
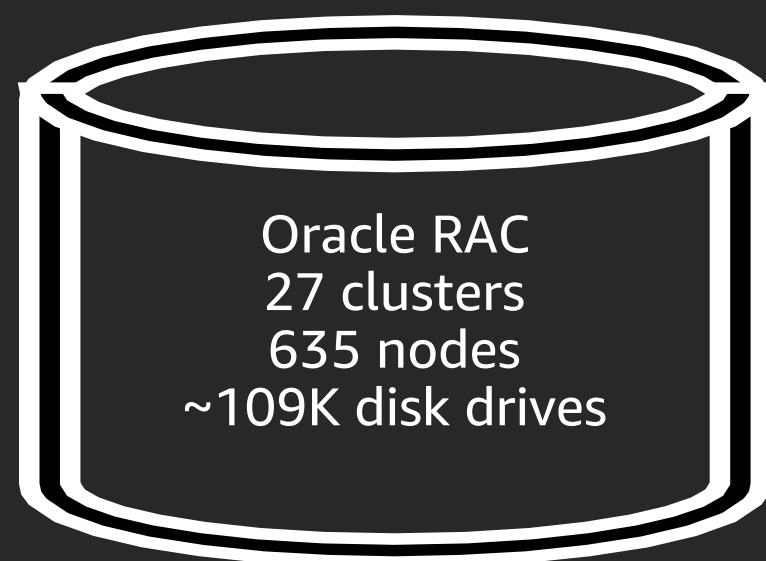
Legacy database infrastructure in 2017

Data Warehouse Databases

- 50PB of data
- 75,000 DW tables
- 600,000 user analytics jobs / day
- 1,800 teams publishing data / day
- 3,300 teams consuming data / day

OLTP Databases

- 75PB of data
- Nearly 7,500 OLTP databases
- Over 100 teams with 1000s of applications
- Processing millions of orders, payments, and fulfillments around the world



Reasons to migrate databases

Key migration requirements

Before we talk about reasons why, here are the requirements



Service availability

Needs to be near-100% available and operate at full scale



Data availability

All data needs to be available and consistent at all times



Migration time

Needed to be completed in a defined time frame with little to no downtime

Reasons to migrate

1. Scalability risks due to data volume increase and global market expansion
2. Latency risks due to data volume and transaction rates increase
3. Cost risks due to HW/SW costs increase
4. Availability risks due to legacy code/architecture
5. Operation risks due to HW provisioning/management time/resources

Managing scaling issues

Some teams faced 60 weeks to scale for peak events like PrimeDay

Even with best third-party support, high risk of failure due to scale

Punitive licensing tactics lacking customer obsession



Transaction rates

Climbed as
business expanded

Improve latency and reduce operations costs

Amazon and our customers are latency sensitive (retail, video, Alexa) and saw 40% improvements at 2X-4X the load

Managing 75PB of data growing at an exponential rate is expensive and time consuming. Some teams saw 90% reduction in operations costs with AWS-managed database services.

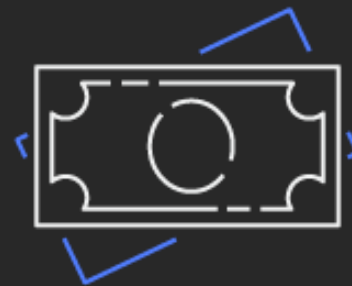


Latency

Improved by 40%

Cost risks due to hardware and software pricing

- Hardware pricing and availability were costly to manage, and timing could be unreliable
- Software pricing and vendor tactics were difficult and costly to manage (renewals, sitting on hardware, cost of DV with HW sitting, managing capacity, planning)



CapEx

HW/SW costs



Software

License management

Availability and operations risks

- The ability to scale databases BOTH up and down at scale was needed
- Hundreds of people were managing and scaling legacy systems
- Operations needed to be reliable, predictable, and measureable
- Data needed to be designed for 100% availability
- Data needed to be consistent 100% of the time



Availability & Operation

HW provisioning
and management

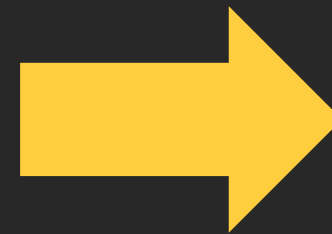
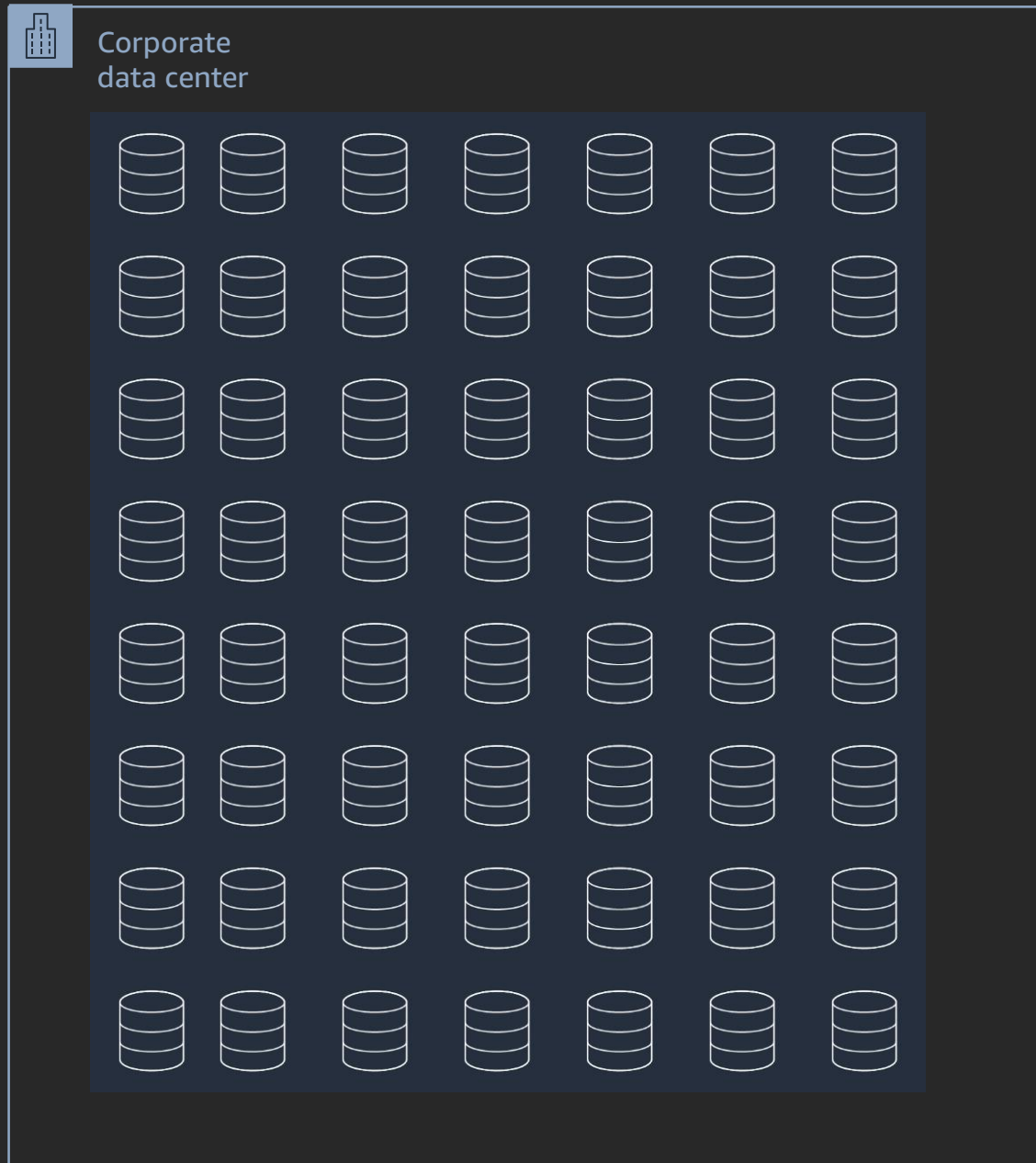
PRIMEDAY drives need for massive scalability

- Adding PrimeDay created a situation where we no longer had one year to plan for peak. Now with a Peak event every 6 months, Amazon had half the time to plan and scale.
- In 2015 there were 34.4M products ordered with 9 countries participating
- In 2019 there were 175M products ordered with 18 countries participating. It was the largest shopping event in Amazon history.

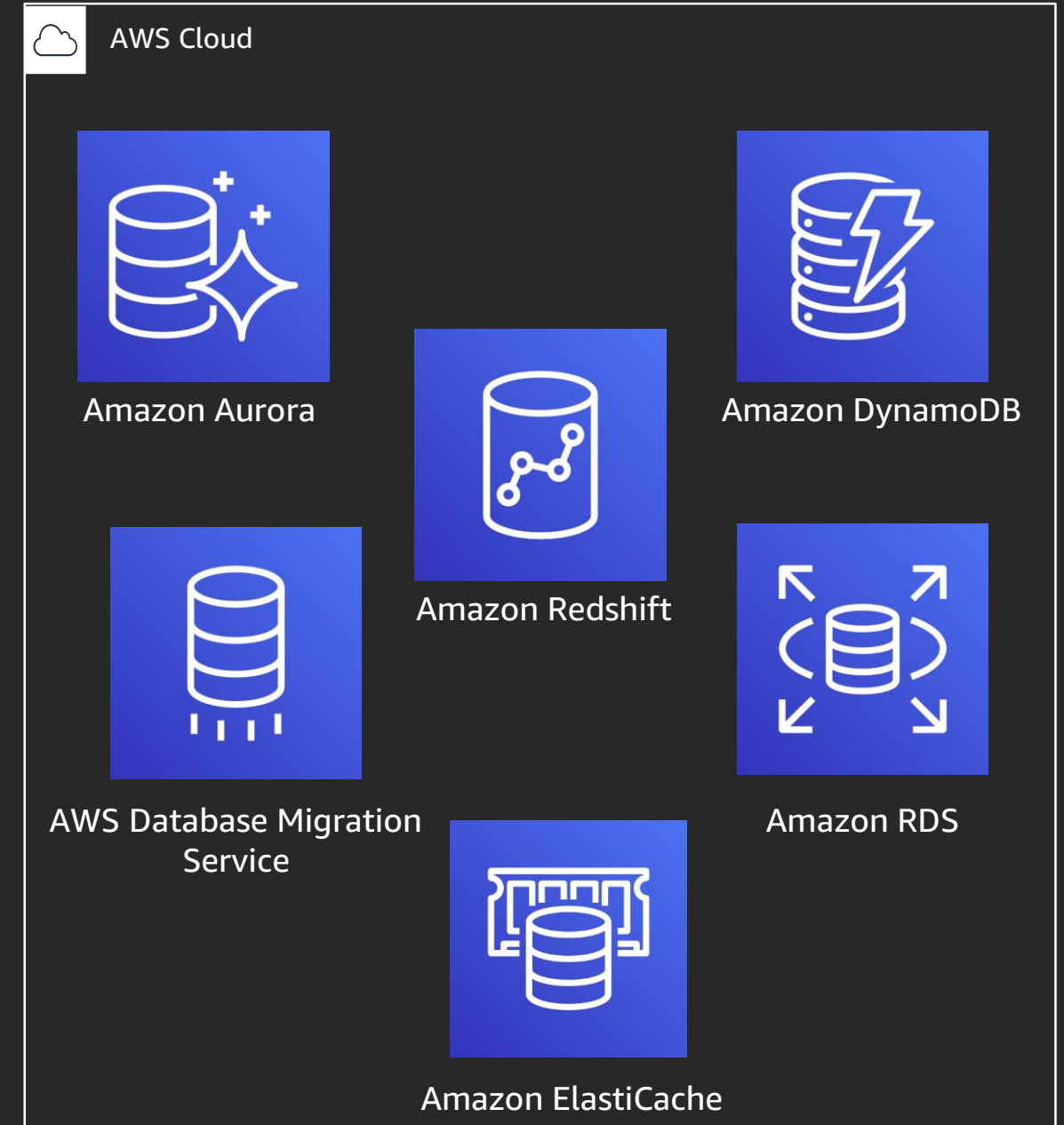
Migration Strategy & Lessons Learned

Database migration

Oracle databases



AWS purpose-built databases



Strategy #1: Increase your visibility

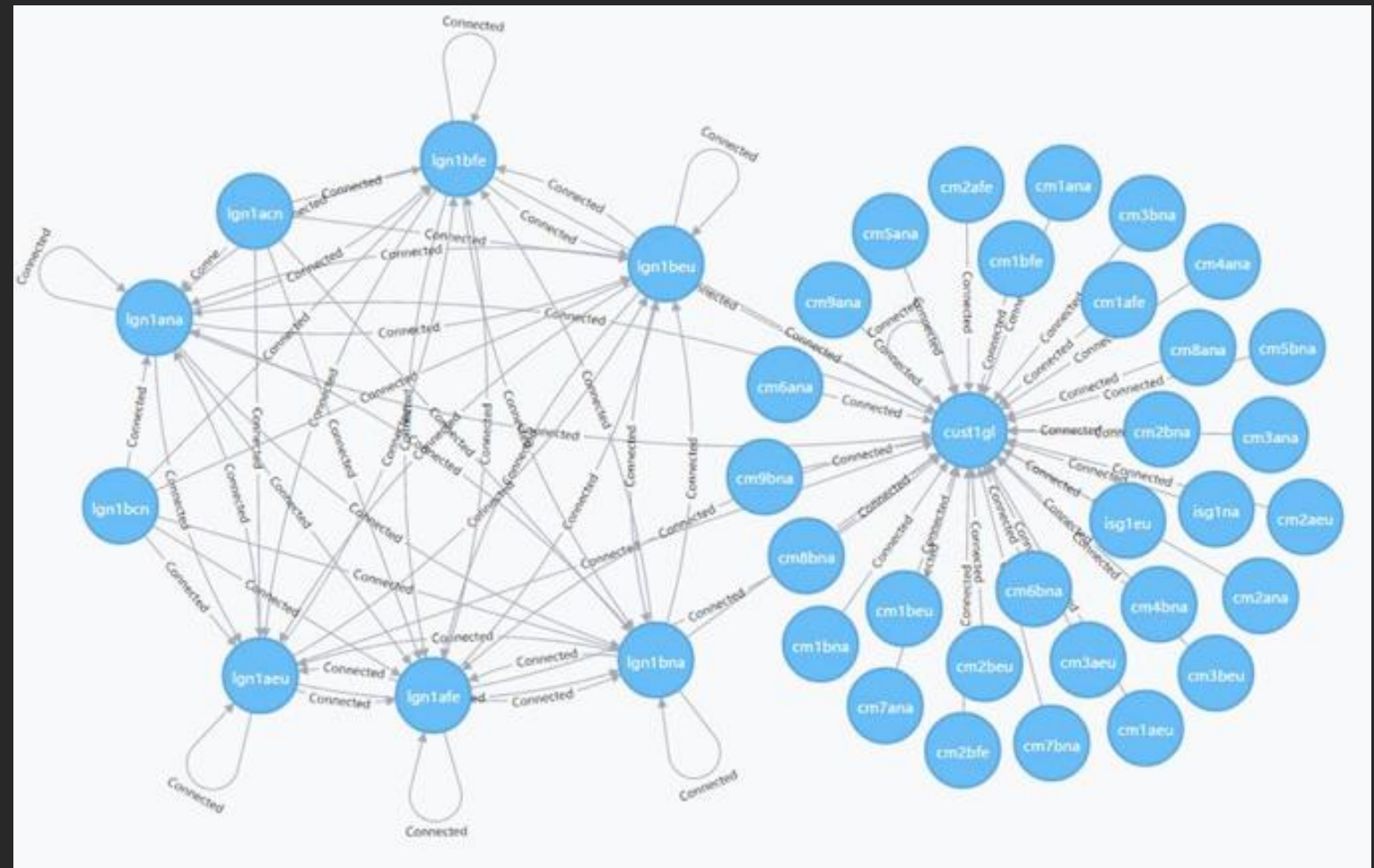
Well-defined KPIs and success metrics will allow you to track migration progress

- How many Oracle databases do you have today?
- How many new Oracle databases did you launch today?
- How many Oracle databases did you migrate/decommission today?
- Who owns the databases?

Strategy #1: Increase your visibility

Well-defined KPIs and success metrics will allow you to track migration progress

- How many Oracle databases do you have today?
- How many new Oracle databases did you launch today?
- How many Oracle databases did you migrate/decommission today?
- Who owns the databases?

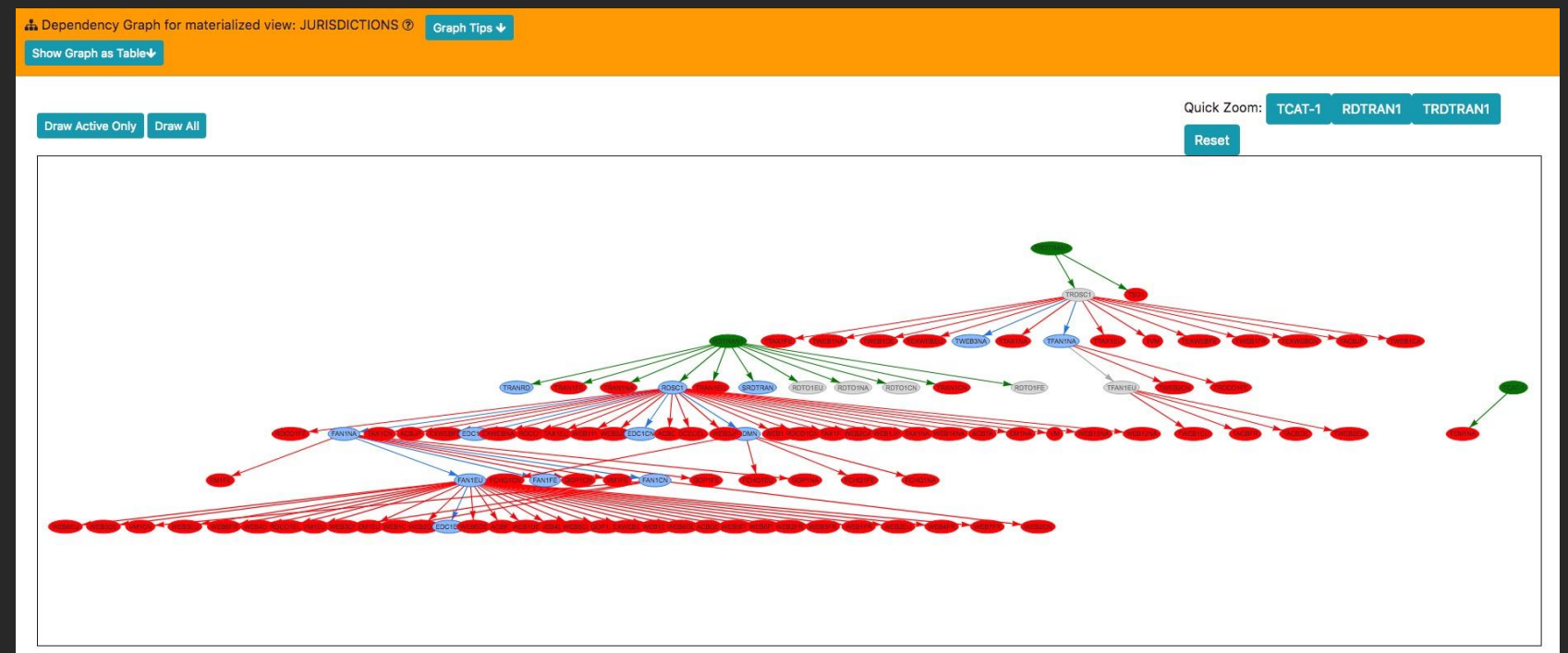


Strategy #1: Increase your visibility

- Oracle database inventory collector
- Database module dependencies
- Database CRUD operations reports
- Database object owner/type reports
- DW jobs and DB connections reports
- Advanced Search
- DB burndown chart
- Project status
- Database migration status

Strategy #1: Increase your visibility

- Oracle database inventory collector
- Database module dependencies
- Database CRUD operations reports
- Database object owner/type reports
- DW jobs and DB connections reports
- Advanced Search
- DB burndown chart
- Project status
- Database migration status

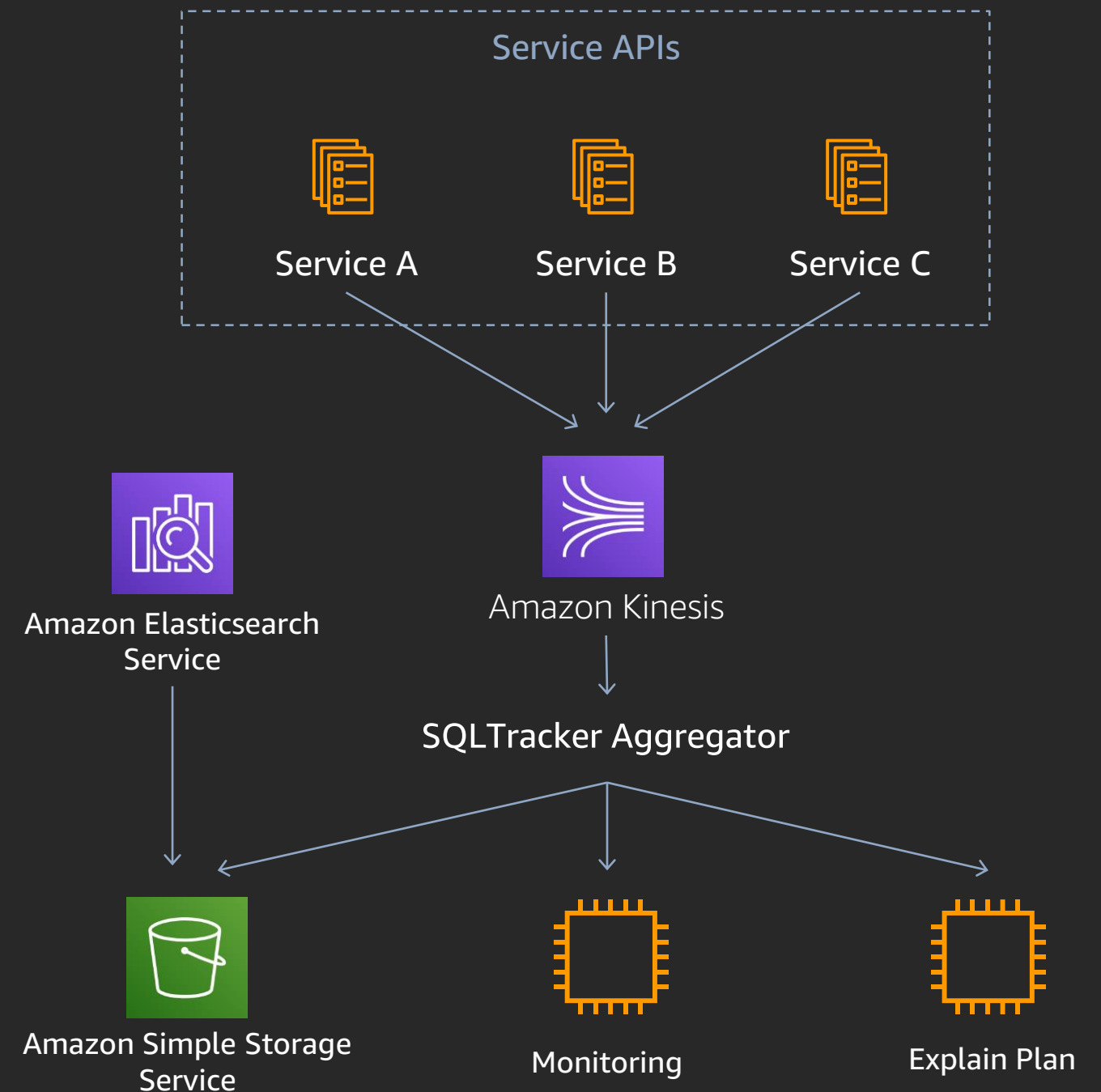


Strategy #1: Increase your visibility

- SQLTracker is a java library to track all SQL execution of DB queries for centralized analysis
- Wrappers for `java.sql.Connection` and `javax.sql.DataSource` instances

Strategy #1: Increase your visibility

- SQLTracker is a java library to track all SQL execution of DB queries for centralized analysis
- Wrappers for `java.sql.Connection` and `javax.sql.DataSource` instances



Strategy #2: Accelerate executive support

- Executive Support
- AWS Partnership with Account Team + Aurora/RDS, DynamoDB, DMS and Redshift teams
- Internal AWS Advocates and people interested in doing it a new way
- Reporting structure to track progress and drive accountability

Strategy #2: Accelerate executive support

- Executive Support
- AWS Partnership with Account Team + Aurora/RDS, DynamoDB, DMS and Redshift teams
- Internal AWS Advocates and people interested in doing it a new way
- Reporting structure to track progress and drive accountability

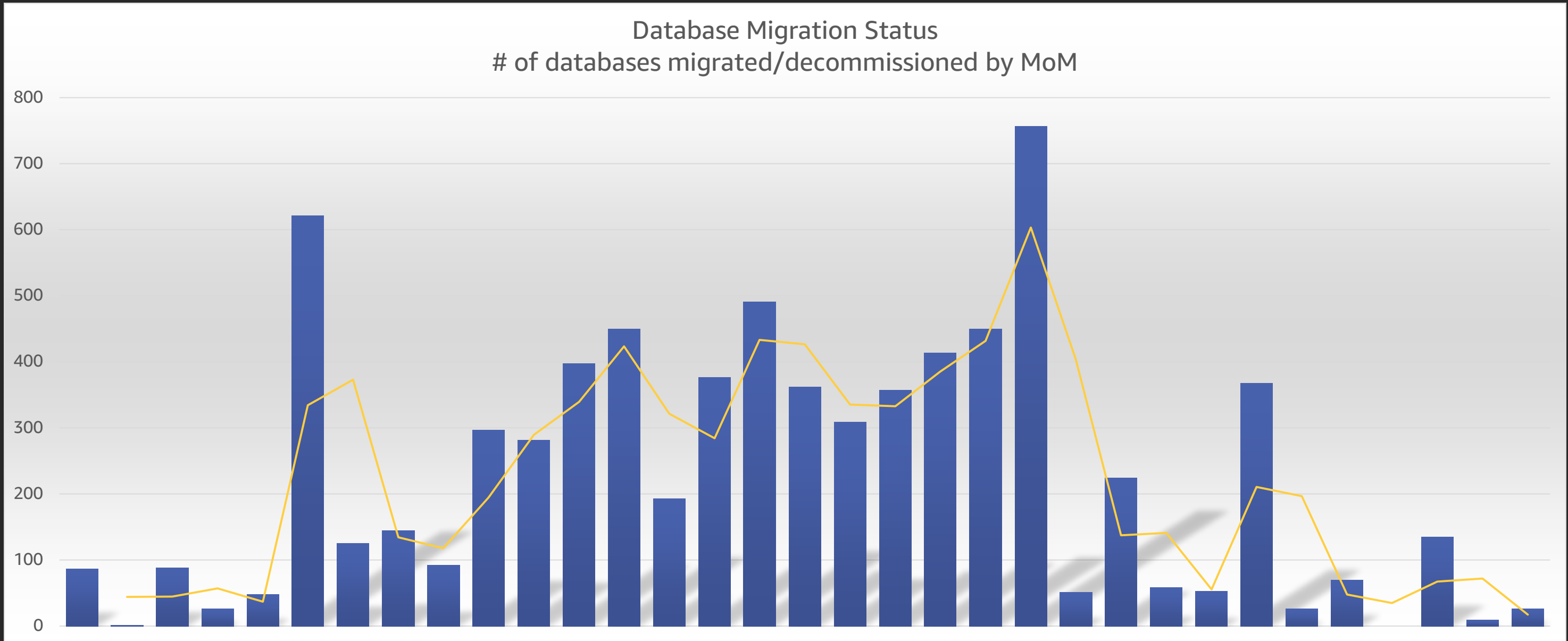
The screenshot displays an AWS dashboard for user 'joonpark@'. At the top, it shows 'Active Instances: 0' and 'Ownership Discrepancies: 0'. Below this, there are two summary tables. The first table, titled 'Self', shows 0 active instances and 0 discrepancies for the user 'joonpark'. The second table, titled 'Direct Reports', shows 0 active instances and 0 discrepancies for three different aliases. Below these tables is a section titled 'Active Instances Burndown for joonpark@'. This section includes a table with columns for 'Tracking Org', 'Tracking Org Leader', 'Current Instance Count', and quarterly data from 2017 Q1 to 2019 Q4. The table currently shows 'No Active Databases Found' and 'Showing 0 to 0 of 0 entries'. There are also buttons for 'Copy', 'Excel', 'CSV', 'PDF', and 'Print', along with a search bar and pagination controls.

Self	Active Instances	Discrepancies
joonpark	0	0

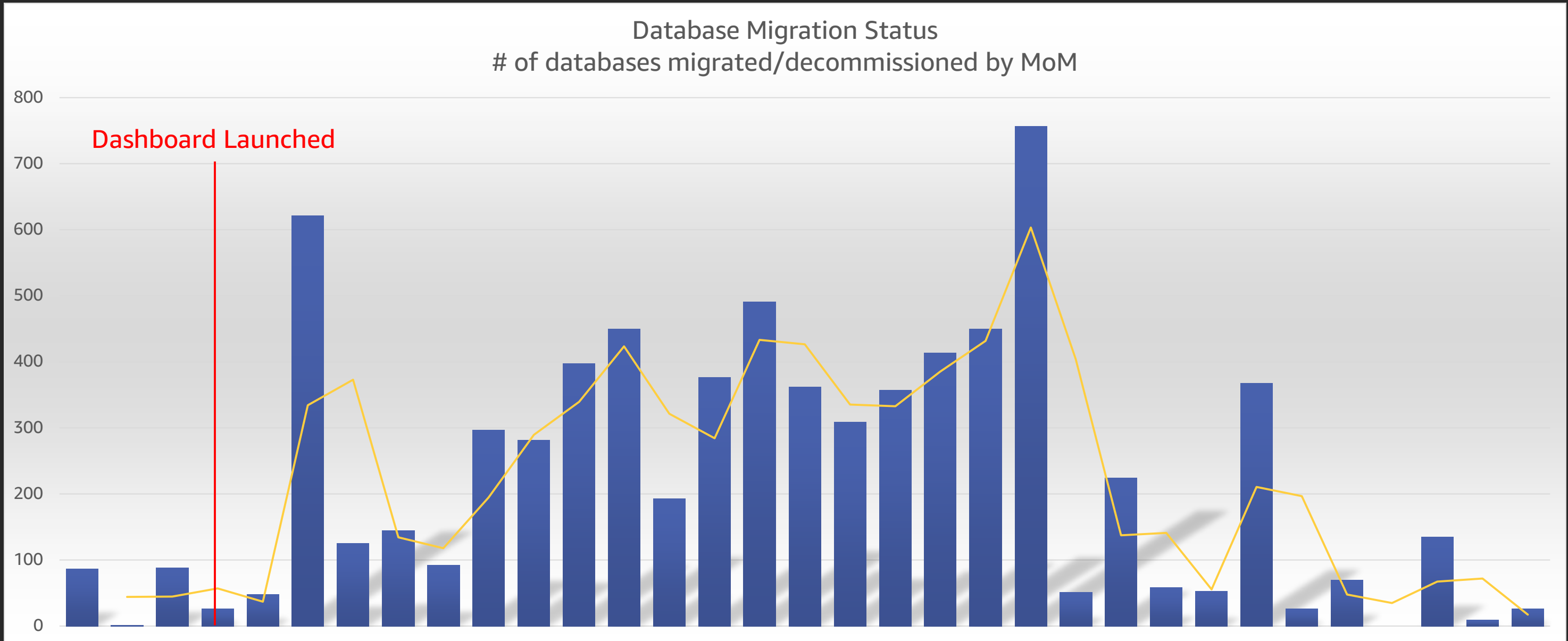
Direct Reports	Active Instances	Discrepancies
[Redacted]	0	0
[Redacted]	0	0
[Redacted]	0	0

Tracking Org	Tracking Org Leader	Current Instance Count	2017 Q1	2017 Q2	2017 Q3	2017 Q4	2018 Q1	2018 Q2	2018 Q3	2018 Q4	2019 Q1	2019 Q2	2019 Q3	2019 Q4
No Active Databases Found														

Strategy #2: Accelerate executive support



Strategy #2: Accelerate executive support

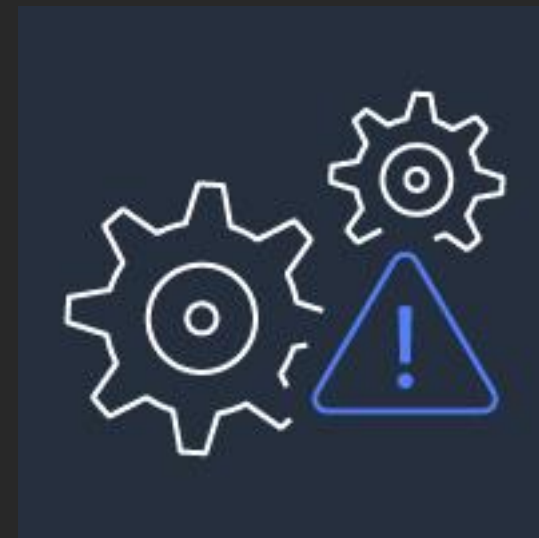


Strategy #3: Address FUD and real technical issues

- There will be many reasons “why not”
- Think about what is on the other side for this investment
- Realize the team has not been doing this for 20 years in this way and need time to adapt

Fear Uncertainty Doubt

vs



Real technical issues

Get Oracle DBAs support

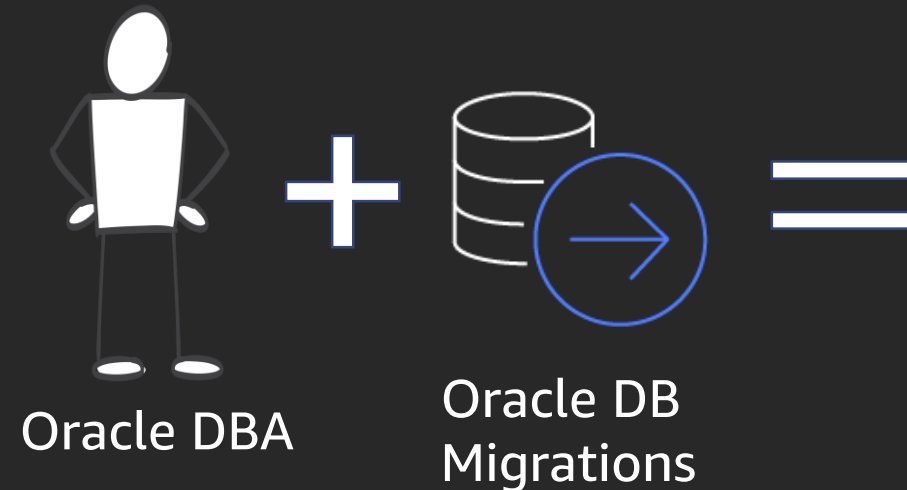
- DBA support is critical for mitigating risks and identifying dependencies for DB migrations
- Provide DBA career path and training opportunities (e.g., Database Engineer, AWS Certifications, SA, System Development Engineers, DB Migration Specialists, etc.)



Oracle DBA

Get Oracle DBAs support

- DBA support is critical for mitigating risks and identifying dependencies for DB migrations
- Provide DBA career path and training opportunities (e.g., Database Engineer, AWS Certifications, SA, System Development Engineers, DB Migration Specialists, etc.)

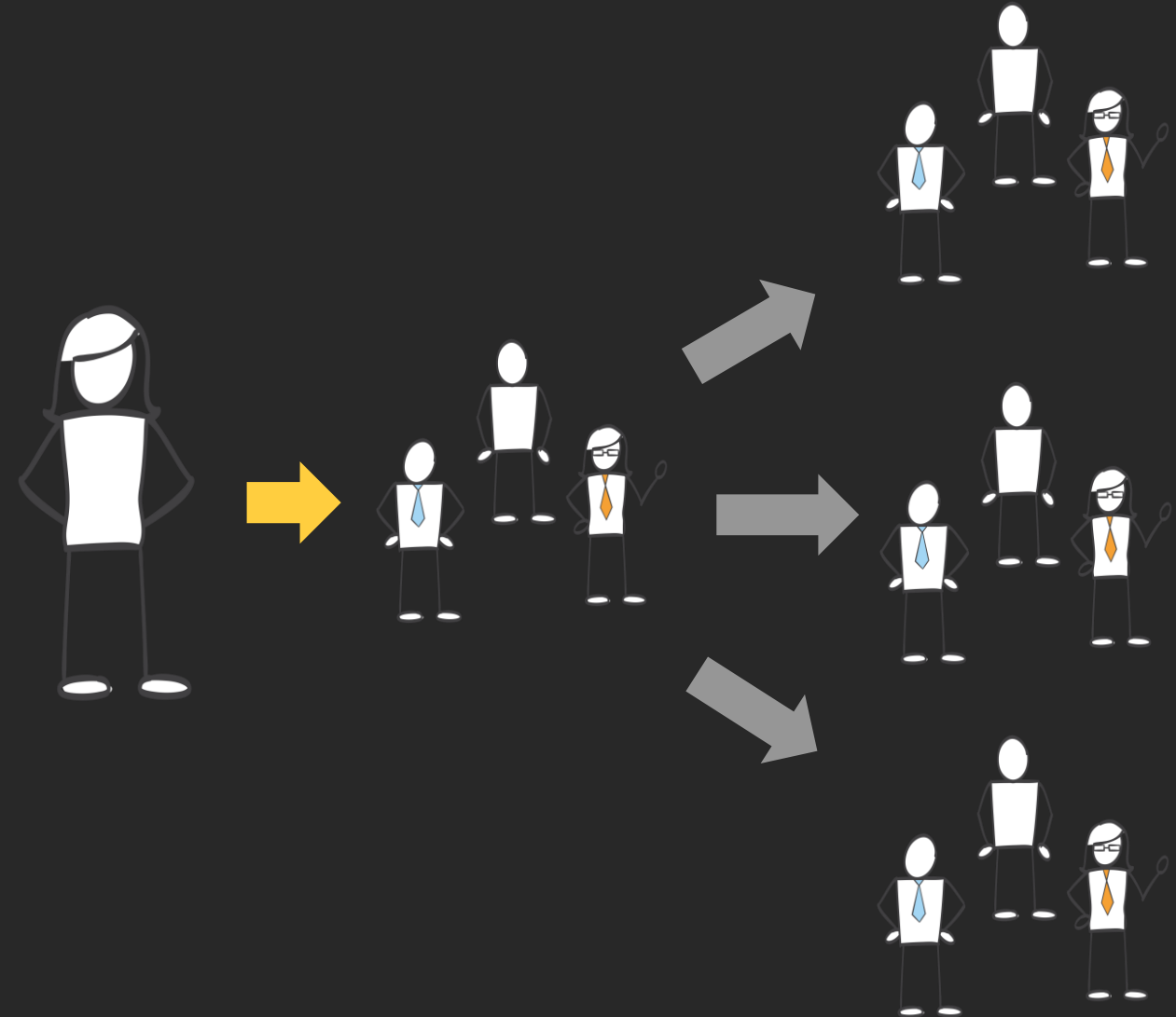


Strategy #4: Create force multipliers

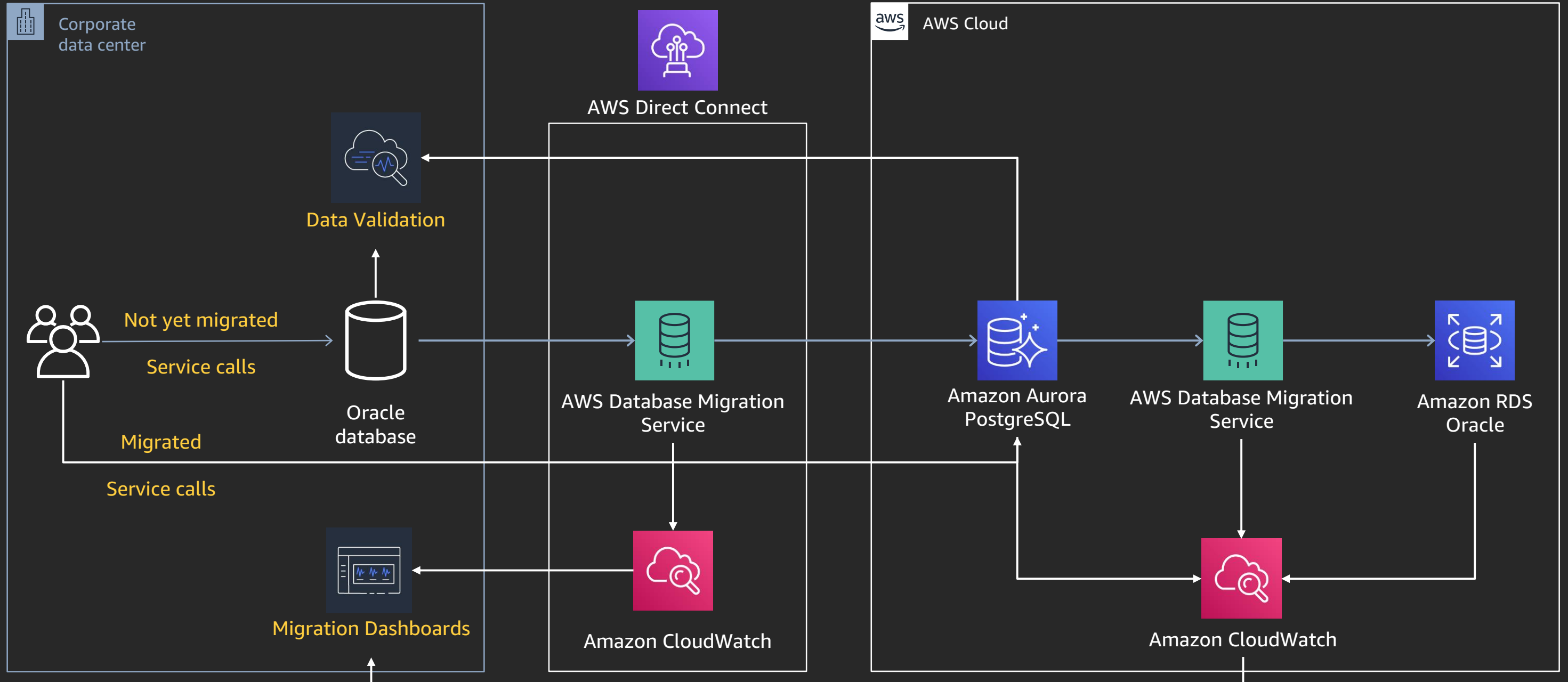
- Create force multipliers – Enable the enablers
- Add routine cadence of Learning Sessions on AWS (e.g., Aurora/RDS, Redshift, DynamoDB, DMS, etc.)
- Share reference architectures, best practices, design patterns, and lessons learned

Strategy #4: Create force multipliers

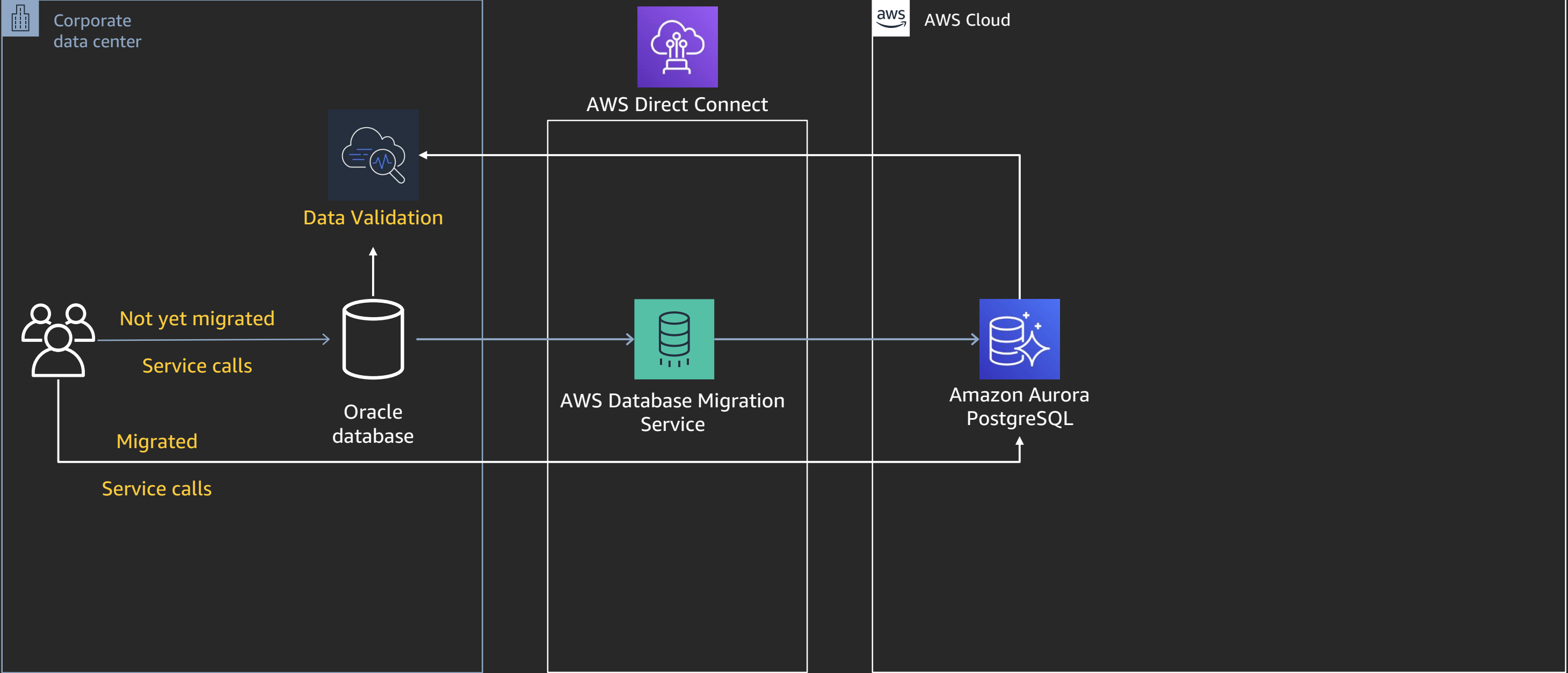
- Create force multipliers – Enable the enablers
- Add routine cadence of Learning Sessions on AWS (e.g., Aurora/RDS, Redshift, DynamoDB, DMS, etc.)
- Share reference architectures, best practices, design patterns, and lessons learned



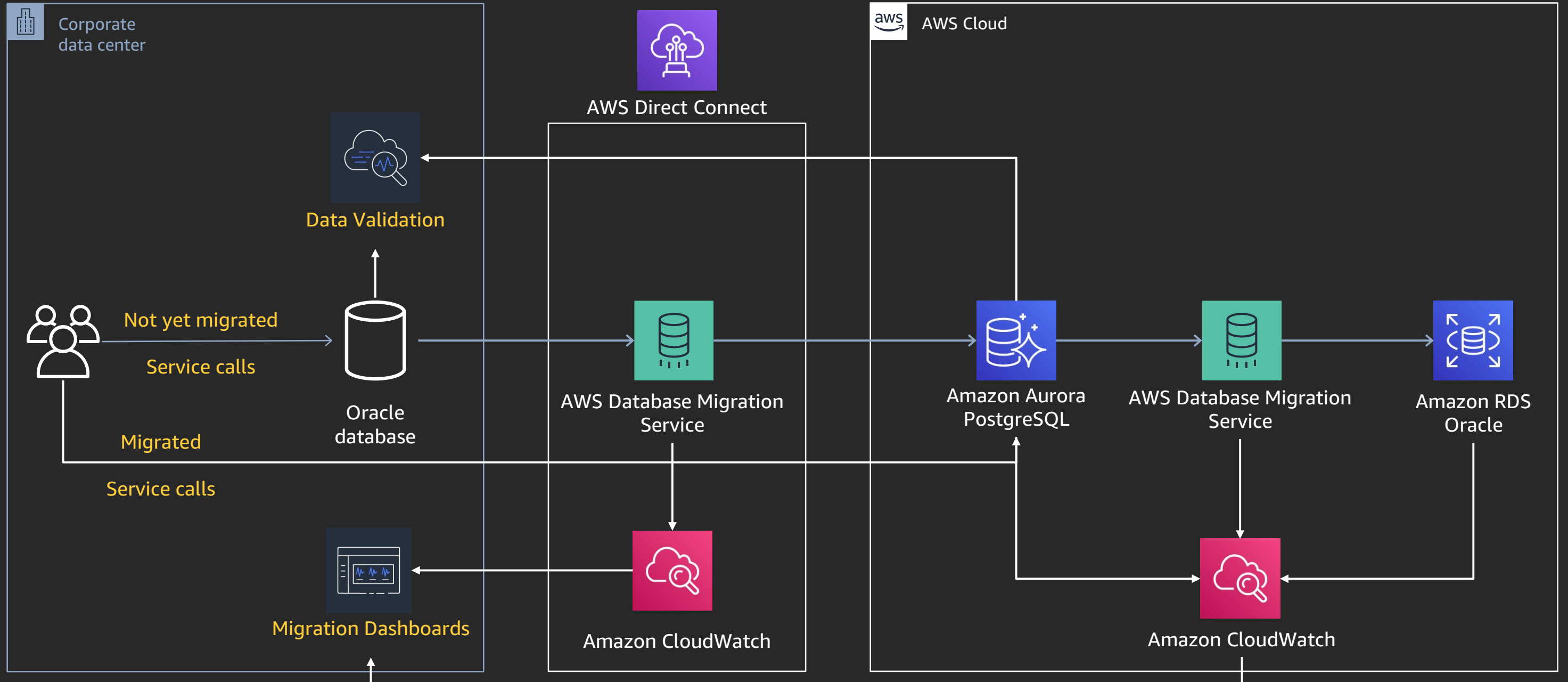
Oracle to Amazon RDS/Aurora PostgreSQL



Oracle to Amazon RDS/Aurora PostgreSQL



Oracle to Amazon RDS/Aurora PostgreSQL



AWS Blog: “How to solve some common challenges faced while migrating from Oracle to PostgreSQL”

- ORA-01555 during extracting data from the source
- Data type conversion Issues
- Empty string vs null in PostgreSQL
- Null behavior in composite unique index
- Choosing numeric data type in PostgreSQL (numeric vs BIGINT)
- SEQUENCE cache behavior in PostgreSQL
- Performance issues with plaintext searches in PostgreSQL

Empty string vs null in PostgreSQL

In Oracle

```
create table emp(name varchar(10), dept int);  
Table created.
```

```
insert into emp(name,dept) values(null, 1);  
1 row created.
```

```
select * from emp where name is null;  
NAME      DEPT
```

```
-----  
|          | 1
```

##Lets inster an empty string now

```
insert into emp(name,dept) values('', 2);  
1 row created.
```

Lets select all values where name is null.

#As you can see Oracle auto converts the empty string to Null values.

```
select * from emp where name is null;
```

```
NAME      DEPT
```

```
-----  
|          | 1
```

```
|          | 2
```

#Lets see what happens if we compare empty string using "=" operator

```
select * from emp where name = '';
```

```
no rows selected
```

#As you can see no rows found, oracle stores the empty strings as Null.

Oracle

Empty string == null

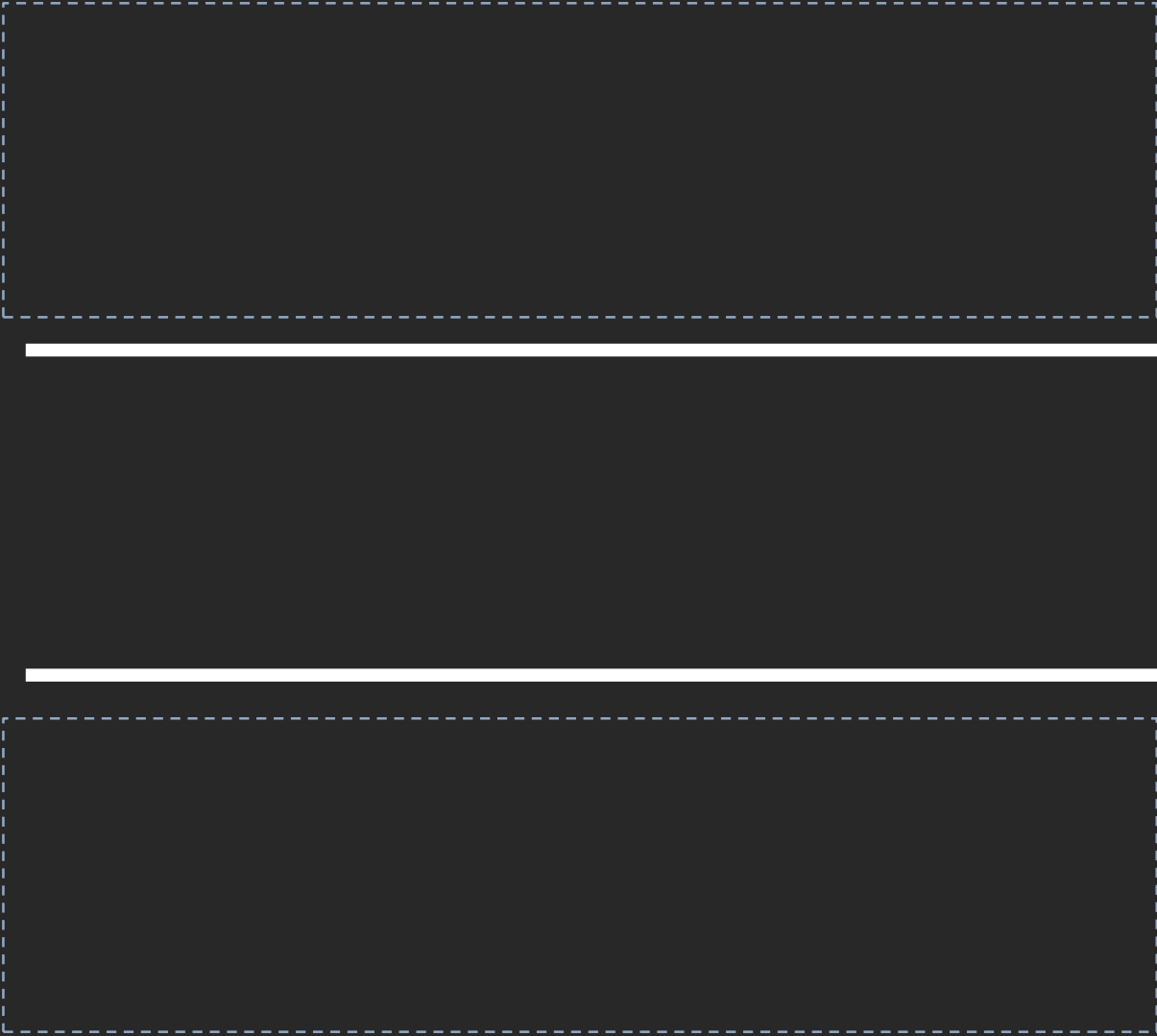
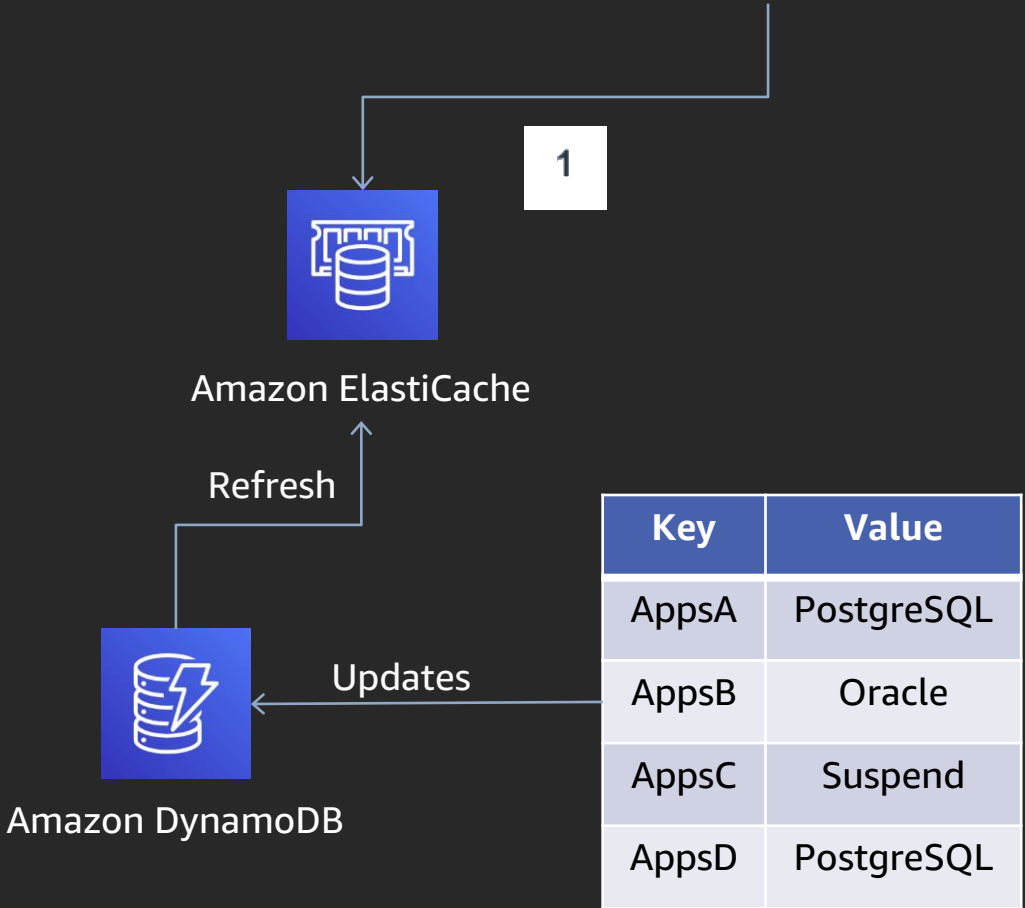
Empty string vs null in PostgreSQL

```
1 In PostgreSQL
2 -----
3 create table emp (name character varying(10), dept numeric);
4 CREATE TABLE
5
6 insert into emp(name,dept) values(null, 1);
7
8 #No surprise here, null is treated as null in PostgreSQL
9 select * from emp where name is null;
10 name | dept
11 -----
12      |    1
13
14 #Now lets inster the empty string
15 insert into emp(name,dept) values('', 2);
16
17 # PostgreSQL only displays the record which was intered as null.
18 select * from emp where name is null;
19 name | dept
20 -----
21      |    1
22
23 #Lets query the empty string using the "=" operator
24 select * from emp where name = '';
25 name | dept
26 -----
27      |    2
28
29 # Here we go, we have a empty string displayed now.
```

PostgreSQL

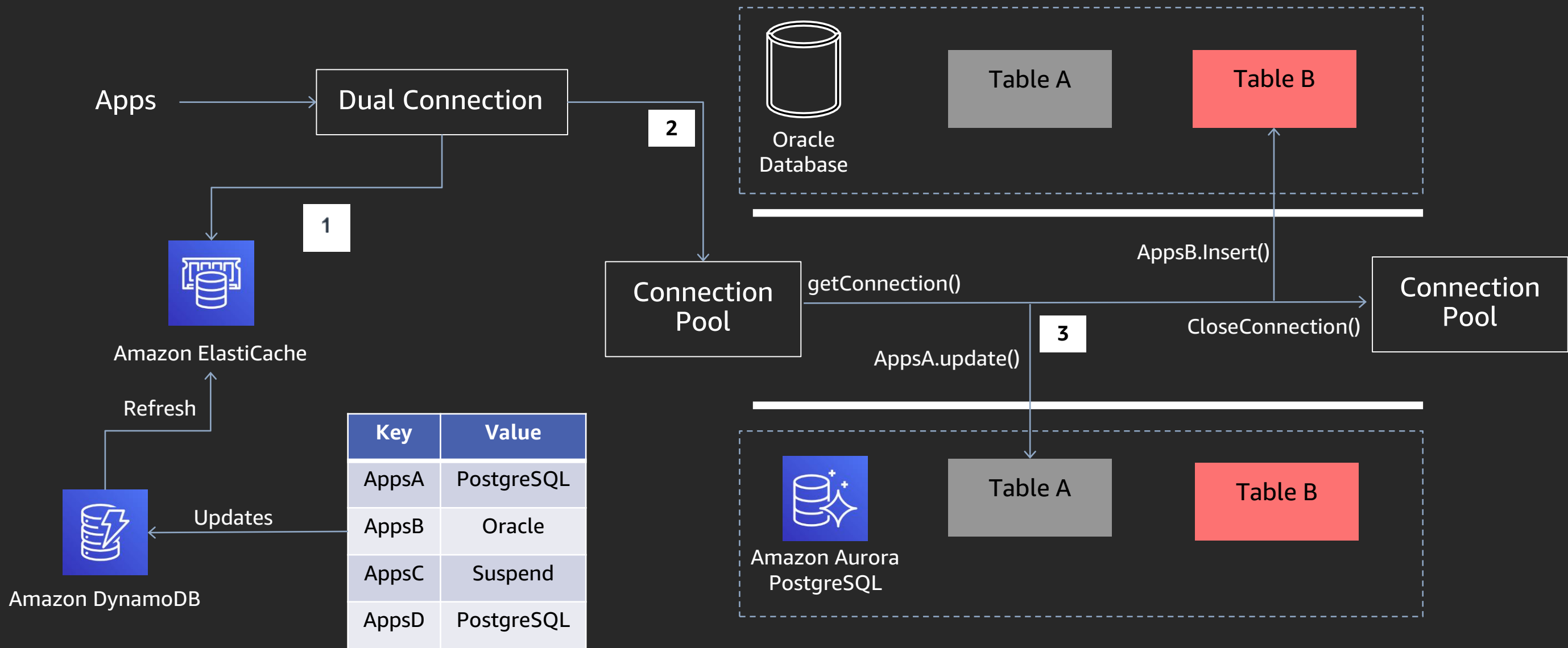
Empty string != null

Dynamic Connection



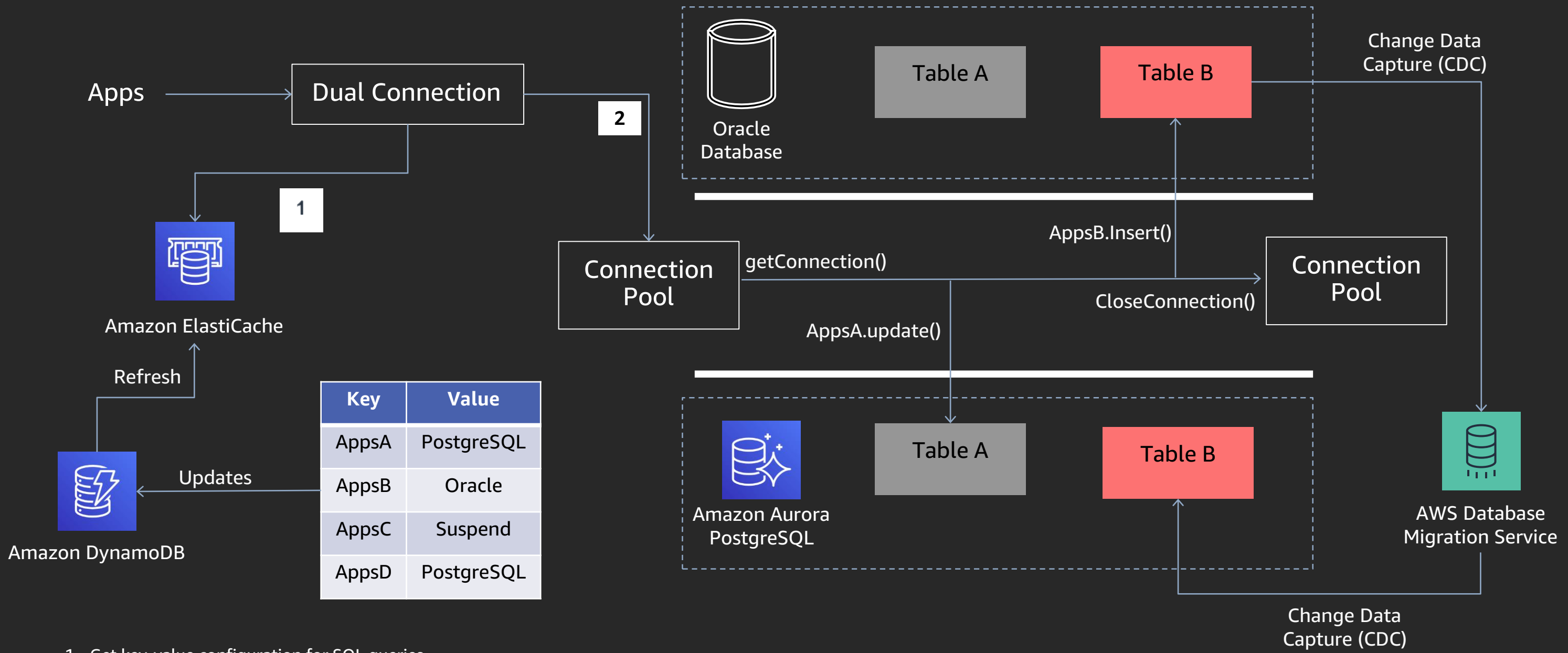
1. Get key-value configuration for SQL queries
2. Determine which connection to be used
3. Execute SQL queries based on the configuration value

Dynamic Connection



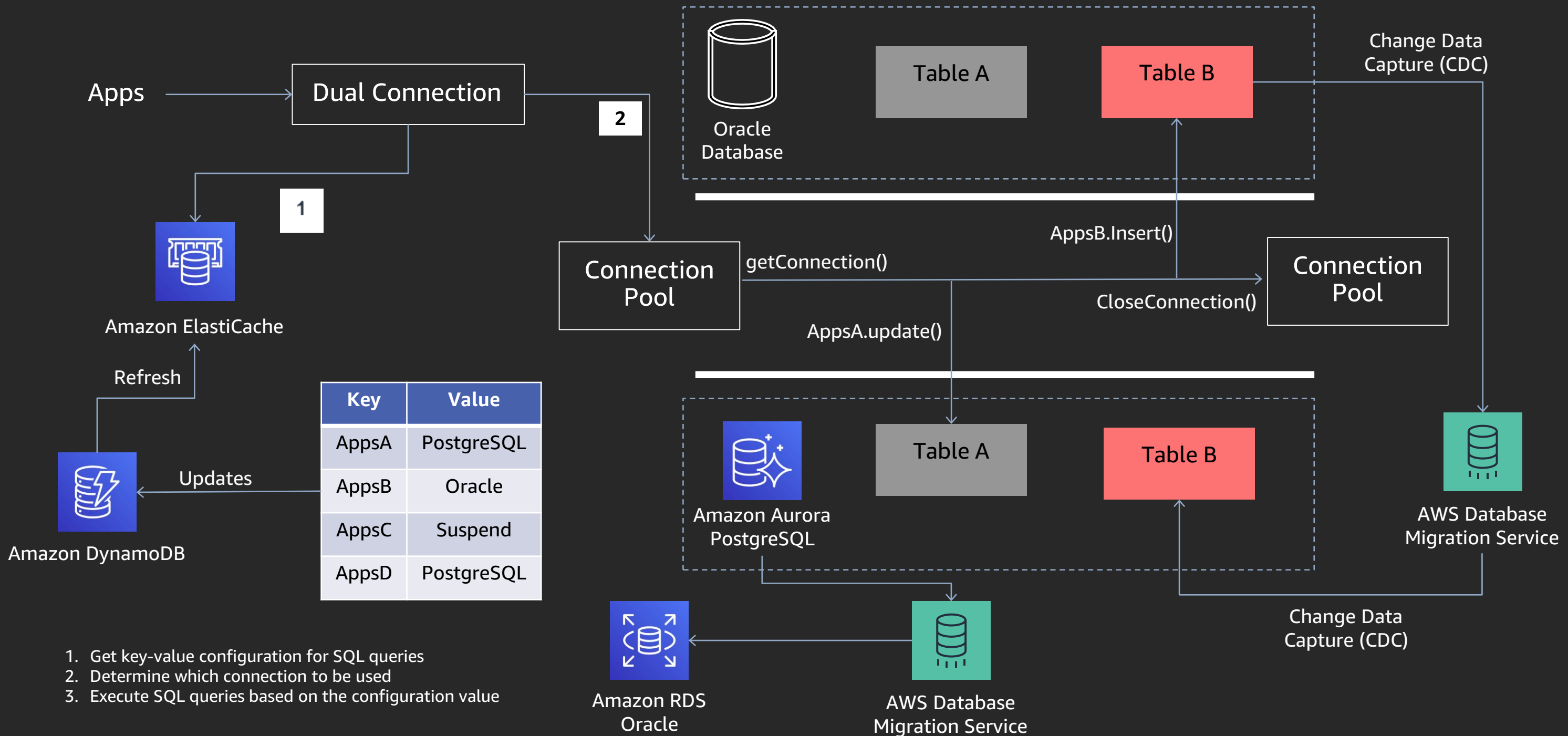
1. Get key-value configuration for SQL queries
2. Determine which connection to be used
3. Execute SQL queries based on the configuration value

Dynamic Connection



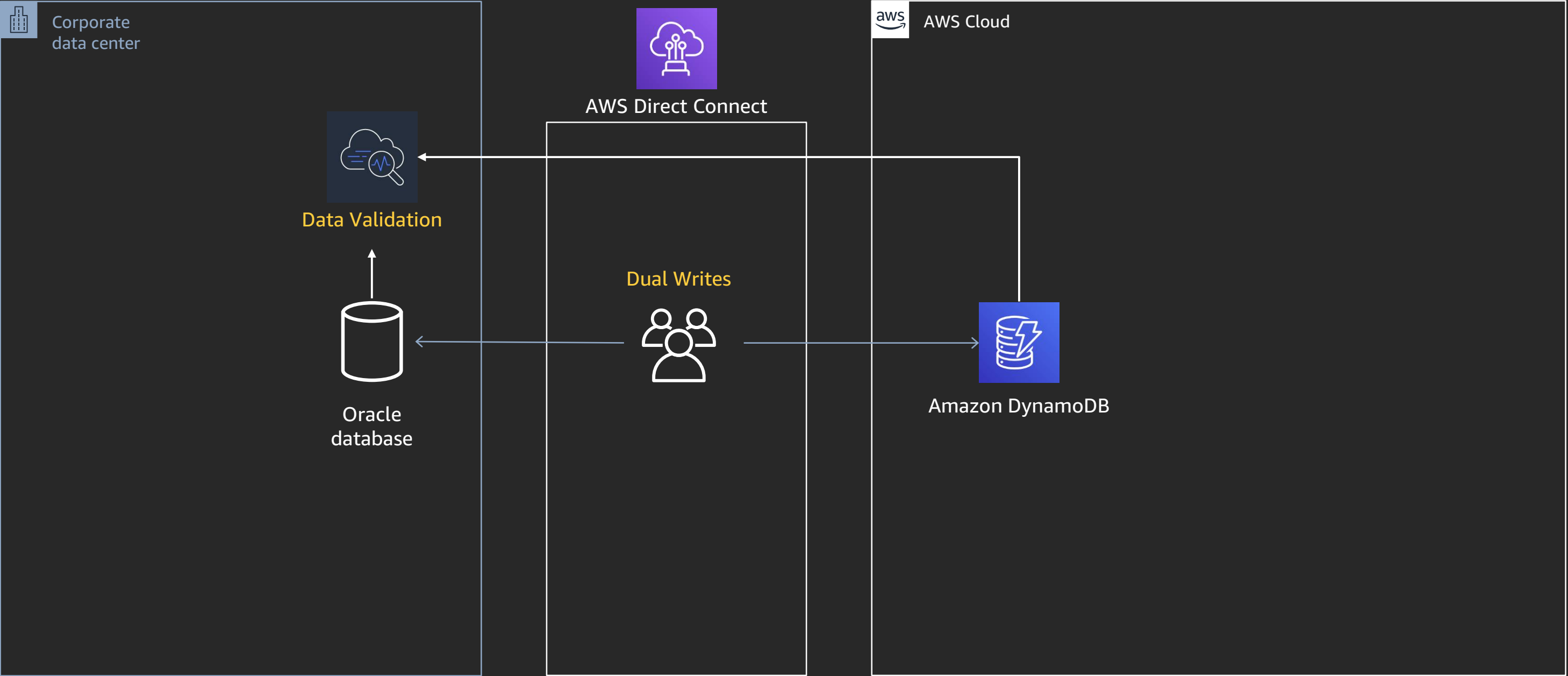
1. Get key-value configuration for SQL queries
2. Determine which connection to be used
3. Execute SQL queries based on the configuration value

Dynamic Connection

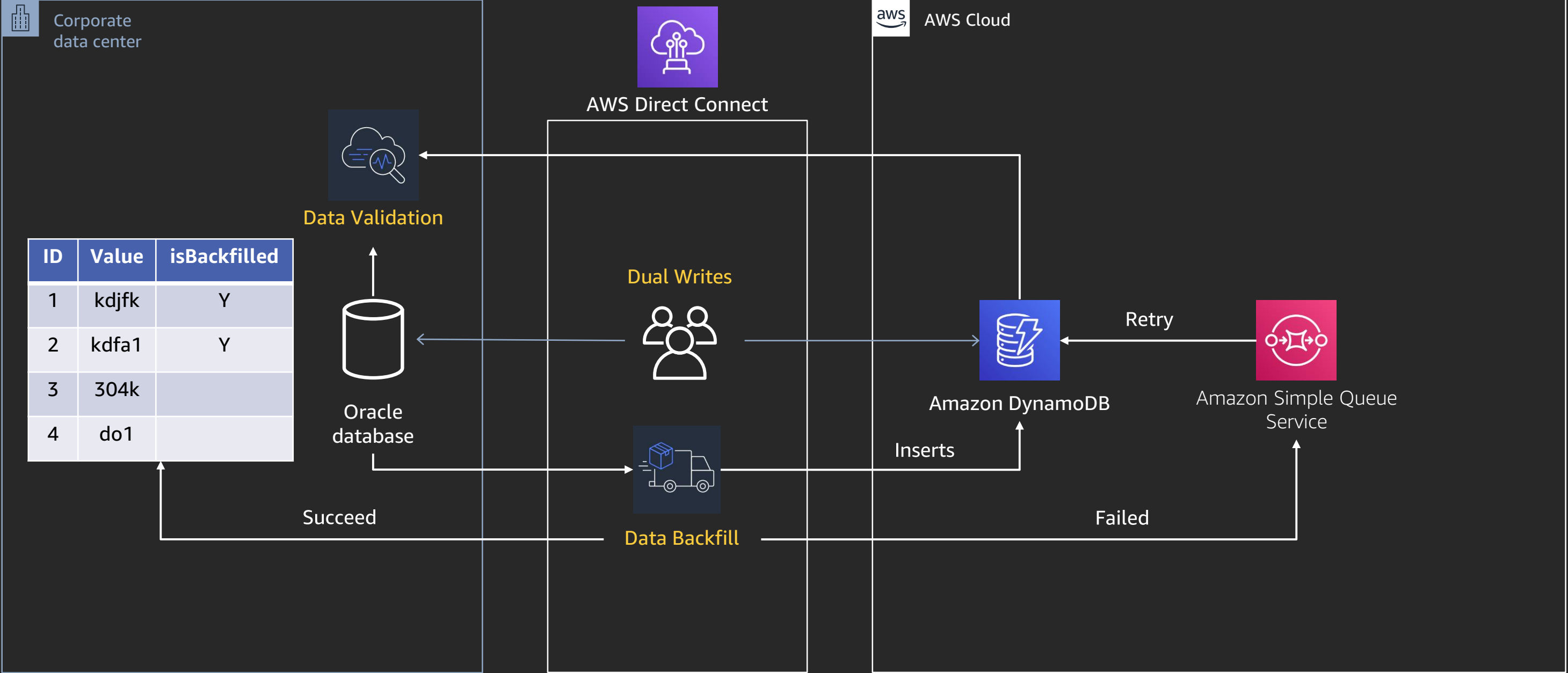


1. Get key-value configuration for SQL queries
2. Determine which connection to be used
3. Execute SQL queries based on the configuration value

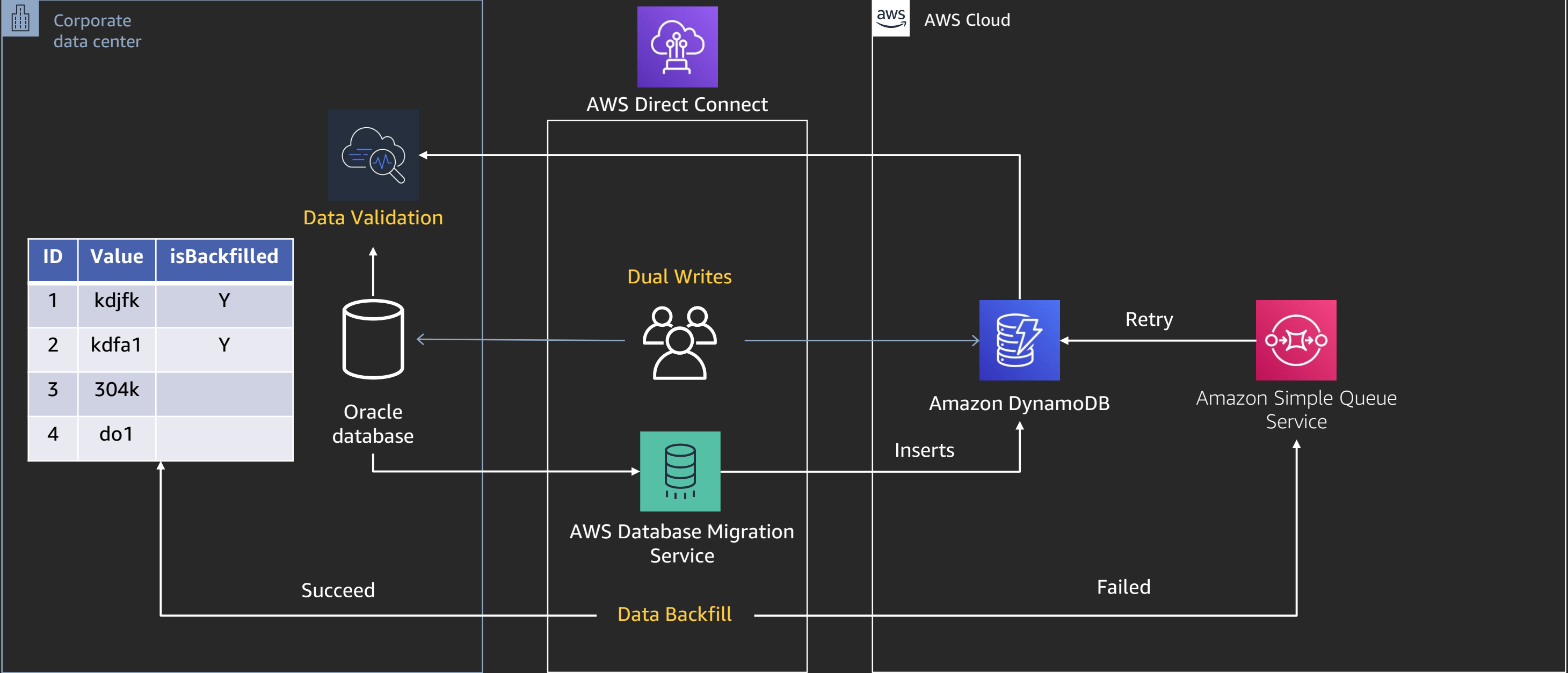
Oracle to Amazon DynamoDB



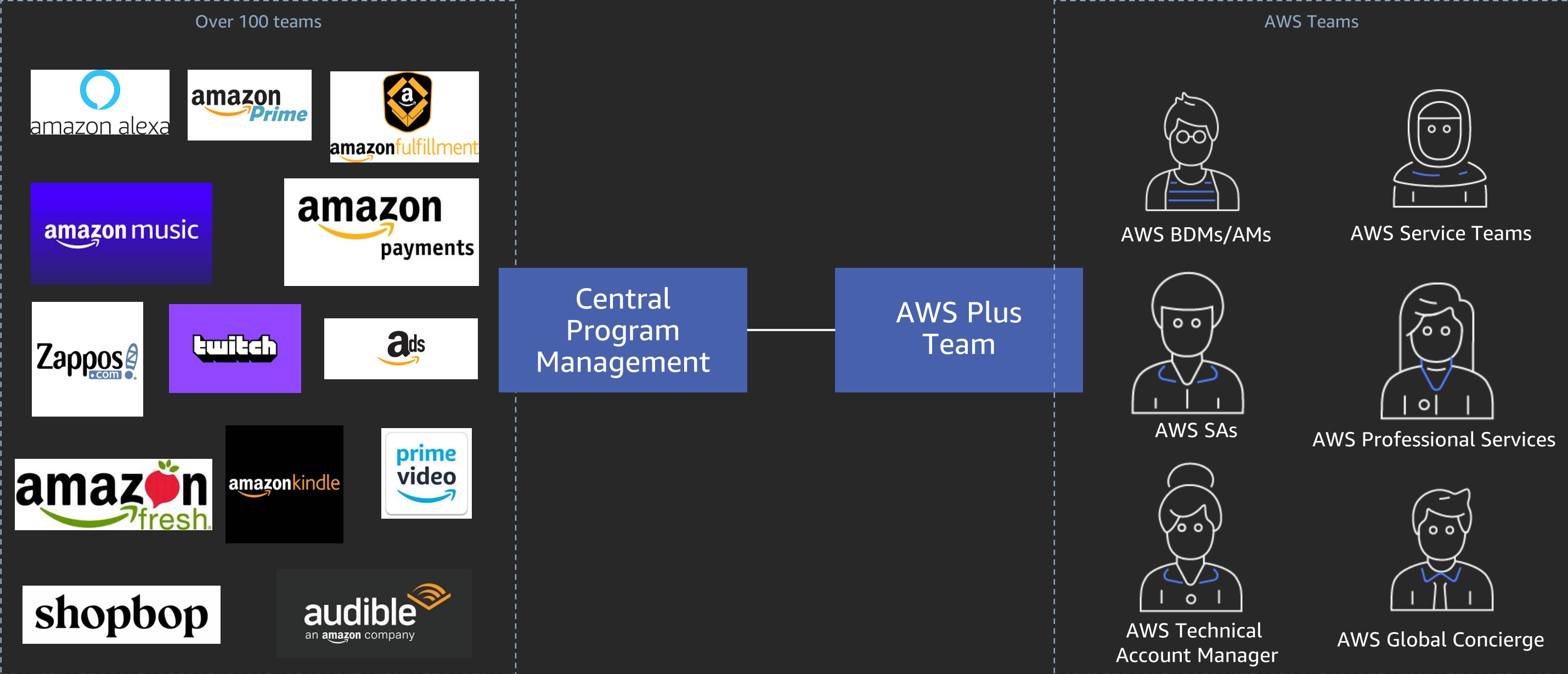
Oracle to Amazon DynamoDB



Oracle to Amazon DynamoDB with AWS DMS



Strategy #5 Work with AWS team



Benefits realized

Outcome

Amazon migrated nearly 7,500 databases from Oracle to Amazon RDS and Aurora

Amazon migrated over 303 business-critical services to DynamoDB

Reduced operations costs

Improved latency and increased innovation

Reduced scaling efforts for peak events like PrimeDay

The infographic is divided into three main sections: Analytics, Databases, and Data movement. Each section contains several AWS services with their respective icons and descriptions.

Analytics	Databases
Redshift Data warehousing	Aurora MySQL, PostgreSQL
EMR Hadoop + Spark	DynamoDB Key value, Document
Athena Interactive analytics	RDS MySQL, PostgreSQL, MariaDB, Oracle, SQL Server
Elasticsearch Service Operational Analytics	ElastiCache Redis, Memcached

Data movement

Database Migration Service | Schema Conversion Tool | Snowball | Snowmobile | Kinesis Data Firehose | Kinesis Data Streams

Benefits realized

Purpose-built databases allowed Amazon.com to choose the right tool for the right reasons: One size for all no longer fits



Cost reduction

Teams reported 40%-90%
operational cost savings



Performance improvements

Teams reported latency improvements
of 40% at 2X-4X the load



Administrative overhead

Reduced peak scaling effort 10X and
administrative overhead with
AWS managed database services

Keys to our success (and yours)

- Purpose-built databases and managed database services: Use the right tools for the job
- Full Leadership Support and reporting on status by VP Org.
- Training and resources — self-paced and class based
- Relationship with AWS and Amazon.com working as one
- Internal community passionate about change
- Documenting the wins and evangelizing internally
 - AWS innovators – Amazon.com

<https://pages.awscloud.com/AWS-Innovators-Amazon.html>

Case studies



AFT

Amazon delights customers with a vast selection of products and fast, often free, shipping. However, to deliver a simple, seamless shopping experience requires a massive amount of infrastructure and technology behind the scenes

[Read more »](#)



Analytics

Amazon builds and operates thousands of microservices to serve millions of customers. These include catalog browsing, order placement, transaction processing, delivery scheduling, video services, and Prime registration

[Read more »](#)



Amazon Database Migration (Business Data Technologies)

When people hear that Amazon is on the verge of concluding an enterprise-level, multiyear initiative to move the company's data from Oracle databases onto Amazon Web Services (AWS), this question might come to mind

[Read more »](#)

<https://pages.awscloud.com/AWS-Innovators-Amazon.html>

Case studies



Wallet

E-commerce buyers want to pay for goods and services easily and securely. Amazon.com enables such payments with Amazon Wallet Service (Wallet), which every Amazon customer around the world interacts with when making purchases on Amazon or visiting the My Wallet page.

[Read more »](#)



Prime Video

Tens of millions of people watch movies and TV shows on Prime Video, the video-streaming service from Amazon that also allows viewers... purchase content and download it for offline viewing. Customers want instant access to videos

[Read more »](#)



Advertising

People performing product searches on the Amazon.com website or mobile app represent a very specific—and valuable—category of potential customer. Because they are searching on a retail site, they are closer to the purchase decision than someone performing the same search on... general-interest web search engine.

[Read more »](#)

<https://pages.awscloud.com/AWS-Innovators-Amazon.html>

Case studies



Buyer Fraud

As one of the world's largest online retailers, Amazon is also one of the world's largest targets for online fraud," says Balachandra Krishnamurthy... software development manager on the Amazon Transaction Risk Management Services (TRMS) team.

[Read more »](#)



Items & Offers

The Item Master Service (IMS) is responsible for the ingestion and processing of data from various suppliers and production of an authoritative catalog. IMS publishes catalog data to the Amazon website, search indexes

[Read more »](#)



FLASH

Across its many business entities, Amazon processes more than 20 billion financial transactions each month, including accounts receivable, accounts payable, royalties, amortizations, remittances, payments, and cash

[Read more »](#)

<https://pages.awscloud.com/AWS-Innovators-Amazon.html>

AWS Migration Playbooks

MIGRATION PLAYBOOK

Migrate from Microsoft SQL Server to Amazon Aurora MySQL

Migrate your Microsoft SQL Server Database to Amazon Aurora MySQL with minimal downtime.

AWS Database Migration Service, AWS Schema Conversion Tool, Amazon Aurora, Amazon RDS for SQL Server

MIGRATION PLAYBOOK

Migrate from Microsoft SQL Server to Amazon Aurora PostgreSQL

Migrate your Microsoft SQL Server Database to Amazon Aurora PostgreSQL with minimal downtime.

AWS Database Migration Service, AWS Schema Conversion Tool, Amazon Aurora

MIGRATION PLAYBOOK

Migrate from Oracle to Amazon Aurora PostgreSQL

Migrate your Oracle Database to Amazon Aurora PostgreSQL with minimal downtime.

AWS Database Migration Service, AWS Schema Conversion Tool, Amazon Aurora

11 STEPS

Migrate from Oracle to Amazon Aurora MySQL

Migrate your Oracle Database to Amazon Aurora MySQL with minimal downtime.

AWS Database Migration Service, AWS Schema Conversion Tool, Amazon Aurora, Amazon RDS for Oracle

11 STEPS

Migrate from Oracle to Amazon Redshift

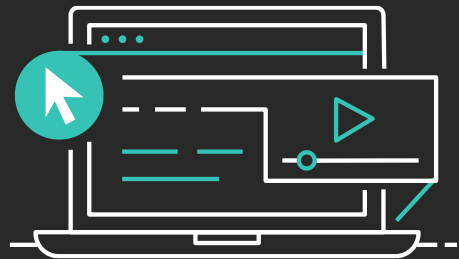
Migrate your Oracle Data Warehouse to Amazon Redshift with minimal downtime.

AWS Database Migration Service, AWS Schema Conversion Tool, Amazon Redshift, Amazon RDS for Oracle

<https://aws.amazon.com/dms/resources/>

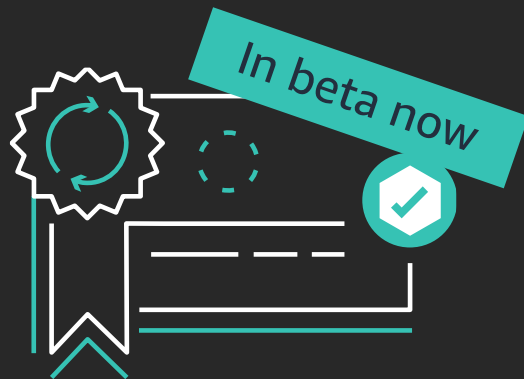
Learn databases with AWS Training and Certification

Resources created by the experts at AWS to help you build and validate database skills



25+ free digital training courses cover topics and services related to databases, including:

- Amazon Aurora
- Amazon Neptune
- Amazon DocumentDB
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Redshift
- Amazon RDS



Validate expertise with the new **AWS Certified Database - Specialty** beta exam

Visit aws.training

Thank you!



Please complete the session survey in the mobile app.