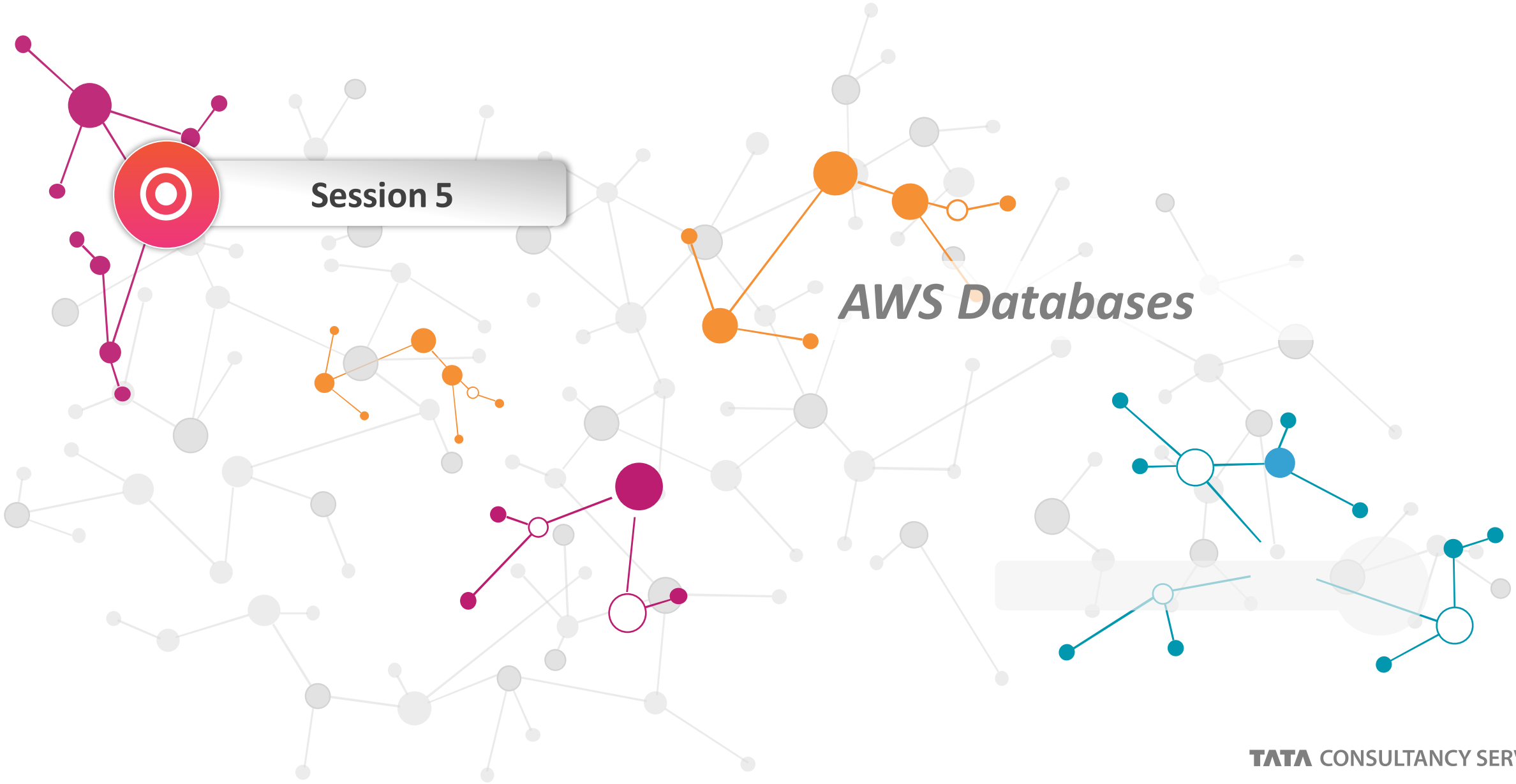
TATA CONSULTANCY SERVICES

Building on belief

AWS Academy

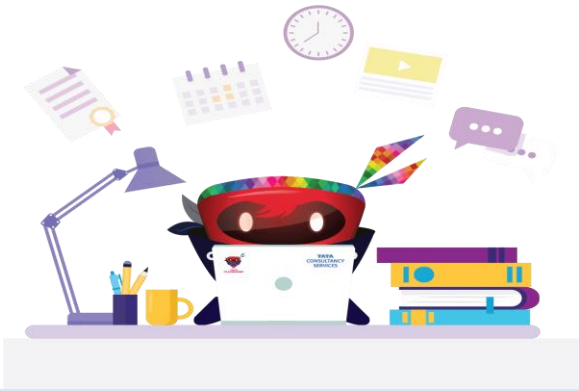
North America
Talent Development Team

Date


Session 5
AWS Databases

- *AWS Database Services*
- *Traditional vs AWS Data services model*
- *Amazon RDS*
- *Redshift*
- *DynamoDB*
- *ElastiCache*
- *Neptune*

Module # 1



AWS Databases

AWS Database Services

AWS Database Services

Amazon RDS

Managed Relational Database Service

Amazon Redshift

Petabyte-scale Data Warehouse Service

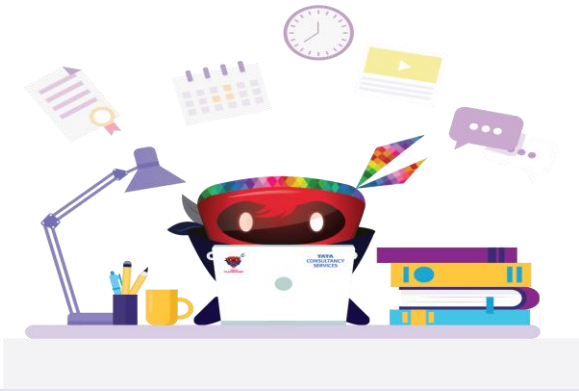
DynamoDB

Fast & Scalable NoSQL Service

ElastiCache

In-memory Caching Service

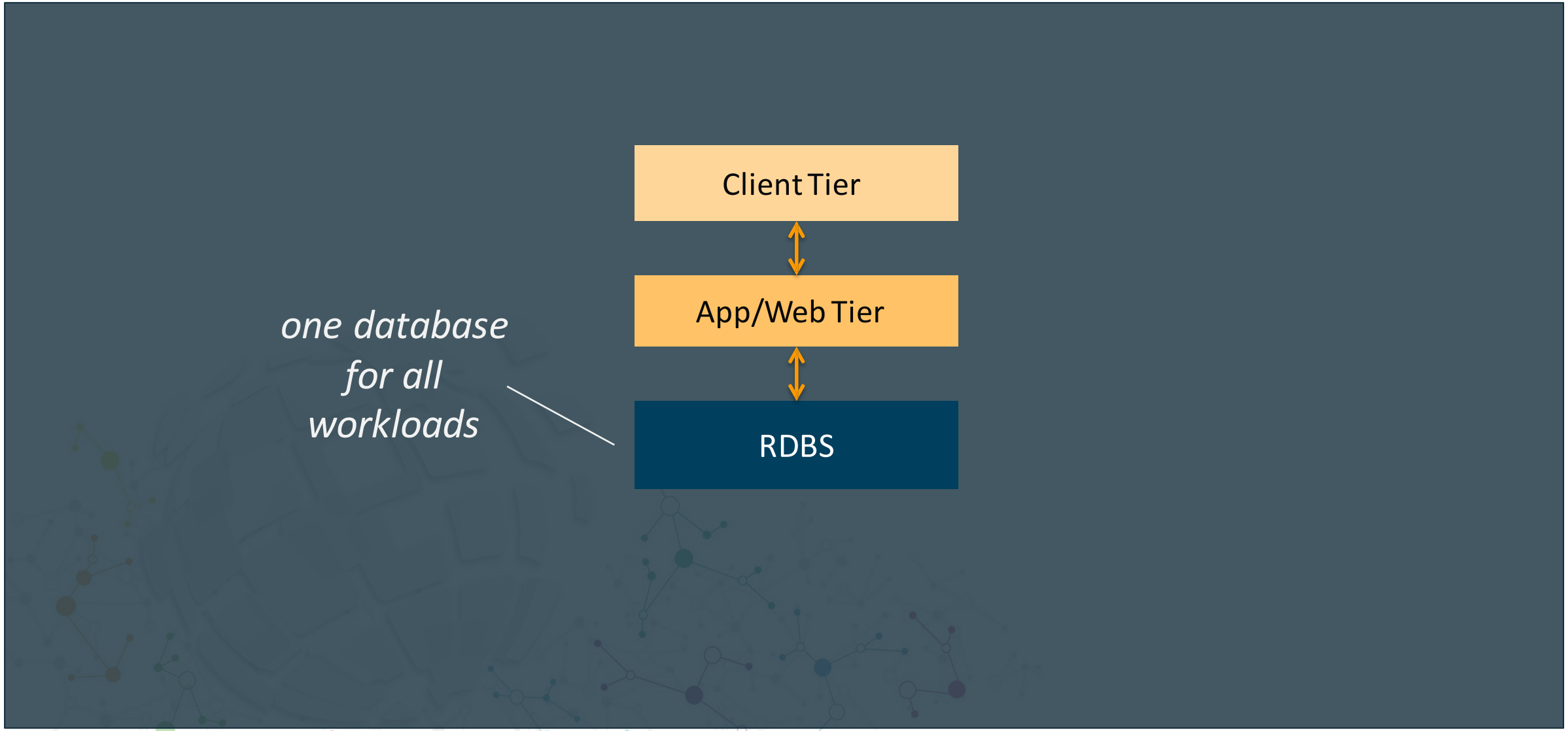
Module # 2



AWS Databases

Traditional vs AWS Data services model

Traditional Database Architecture

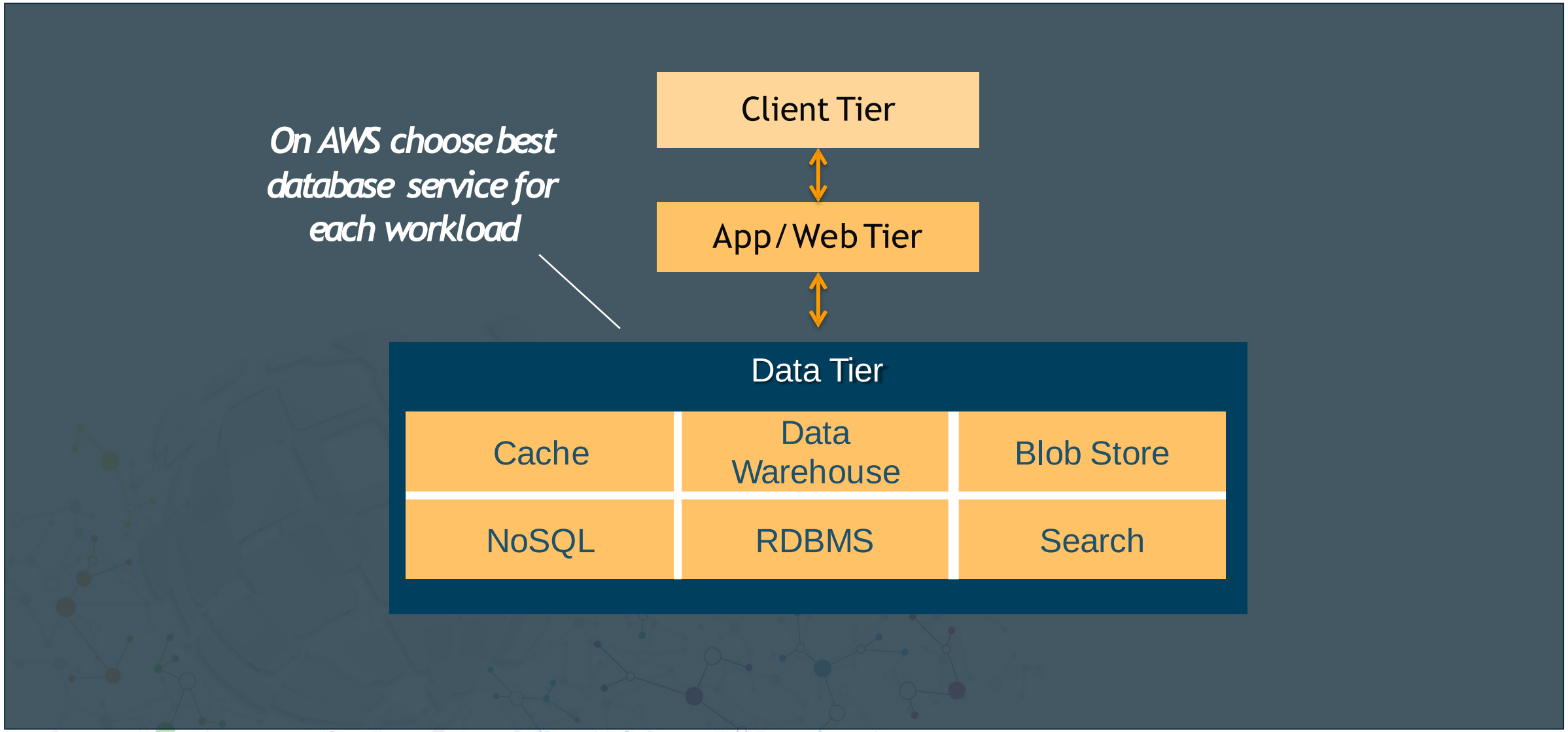


Traditional Database Architecture

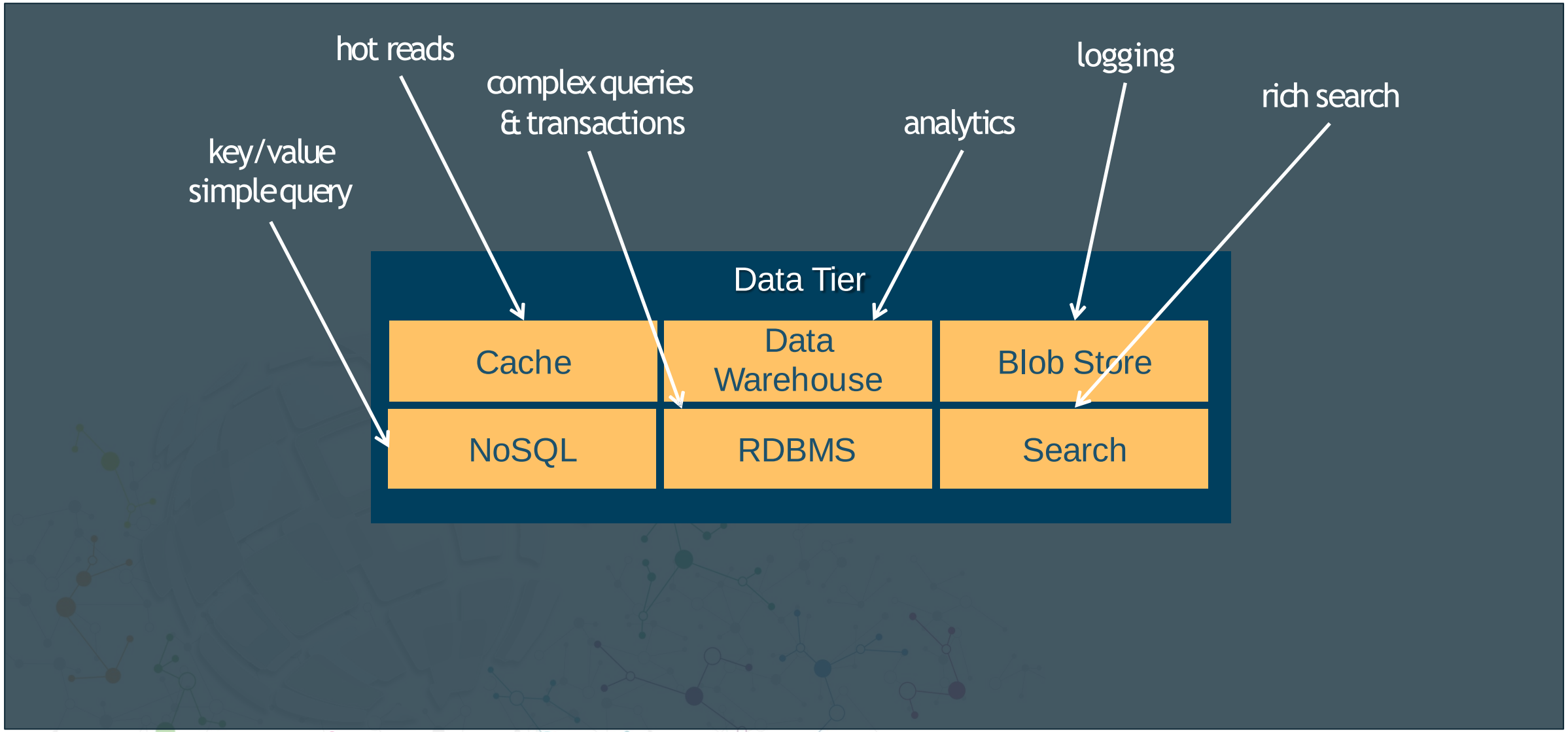
Key-value access
Complex queries
OLAP transactions
Analytics

All forced into the relational database

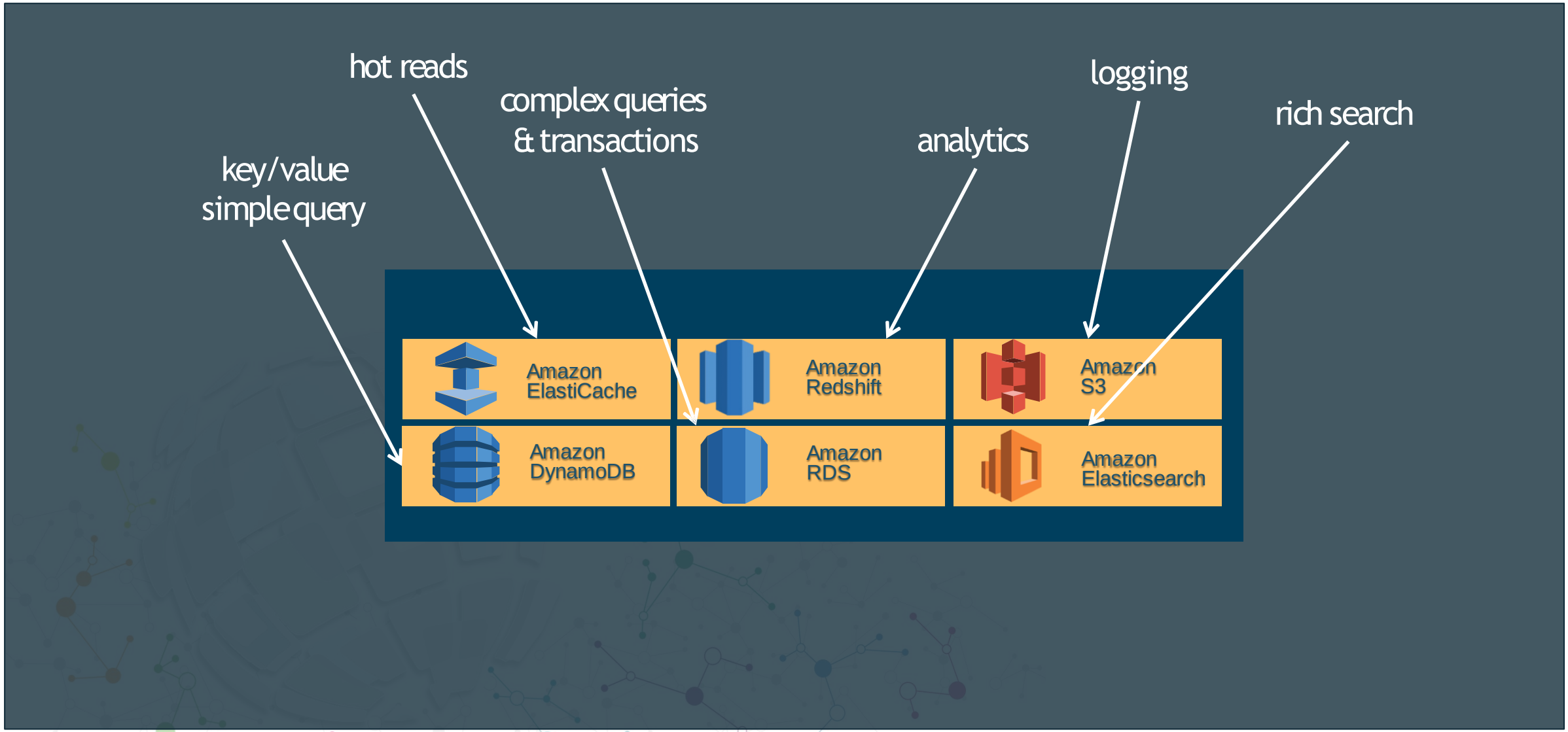
AWS Data Tier Architecture



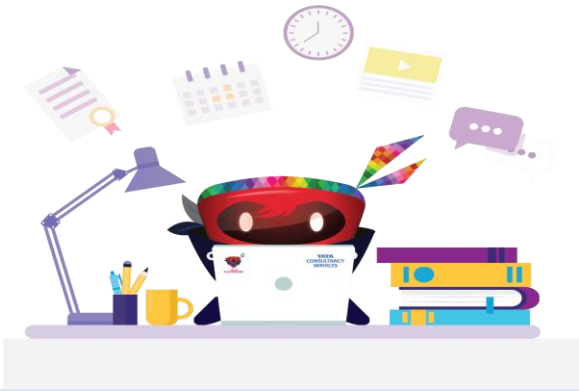
Workload Driven Data Store Selection



AWS Database Services for the Data Tier



Module # 3



AWS Databases

Amazon RDS

Amazon RDS

Amazon RDS

Relational Databases

Amazon Aurora, MySQL, PostgreSQL, Oracle, SQL Server, MariaDB

Fully managed; zero admin

If you host your databases on-premises

- App optimization
- Scaling
- High availability
- Database backups
- DB s/w patches
- DB s/w installs
- OS patches
- OS installation
- Server maintenance
- Rack & stack
- Power, HVAC, net

aws

If you host your databases in Amazon EC2

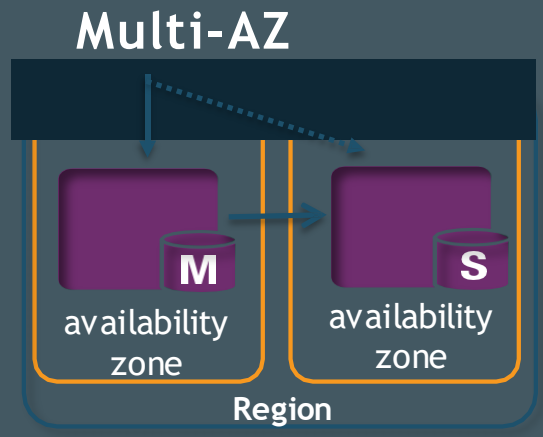
If you choose Amazon RDS

The Self-managed vs. AWS-managed decision

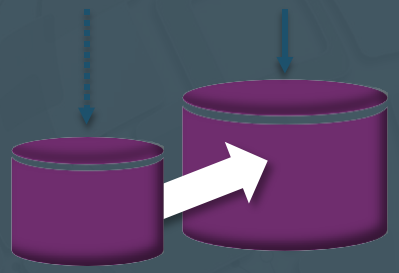
Self-managed database	AWS-managed database
You have full responsibility for upgrades and backup	AWS provides upgrades, backup, and failover as a service
You have full responsibility for security	AWS provides high infrastructure security, certifications; gives you tools to ensure DB security
Full control over parameters of server, OS, and database	Database is a managed appliance, so you can easily automate
Replication is expensive, complex, and requires a lot of engineering	AWS provides failover as a packaged service

Key Amazon Amazon RDS Features

Amazon RDS Configuration	Improve Availability	Increase Throughput	Reduce Latency
Push-Button Scaling		✓	
Multi AZ	✓		
Read Replicas		✓	
Provisioned IOPS		✓	✓

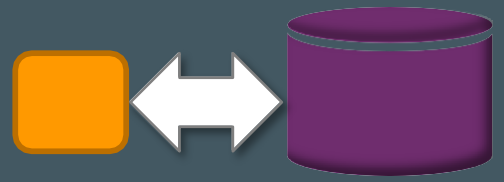


Push-Button Scaling



Read Replicas

Provisioned IOPS

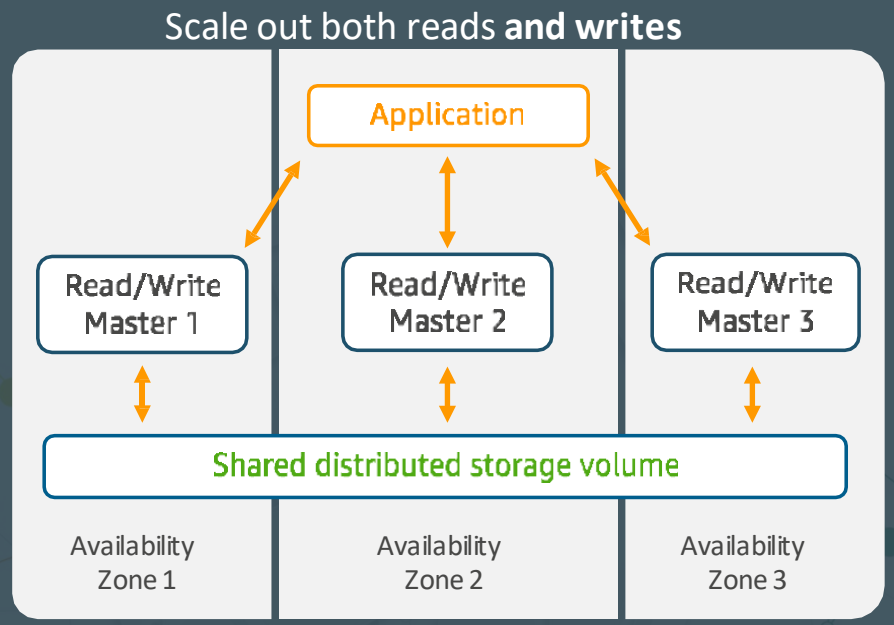


Amazon RDS -- Aurora

- Amazon Aurora is a MySQL and PostgreSQL compatible relational database engine.
- Aurora MySQL is compatible with MySQL version 5.6 and above.
- Aurora PostgreSQL is compatible with PostgreSQL version 9.6.8 and above.
- Aurora supports up to 15 low latency read replicas.
- Aurora Replicas share the same data volume as the primary instance in the same AWS Region.
- Amazon Aurora is designed to offer greater than 99.99% availability, replicating 6 copies of your data across 3 Availability Zones.
- Aurora underlying storage grows automatically as needed, up to 128 tebibytes (TiB) for Aurora MySQL and 64 TiB for Aurora PostgreSQL

Aurora | Multi- Master

First relational DB service with scale–out reads and writes, across multiple datacenters



Zero application downtime from ANY instance failure

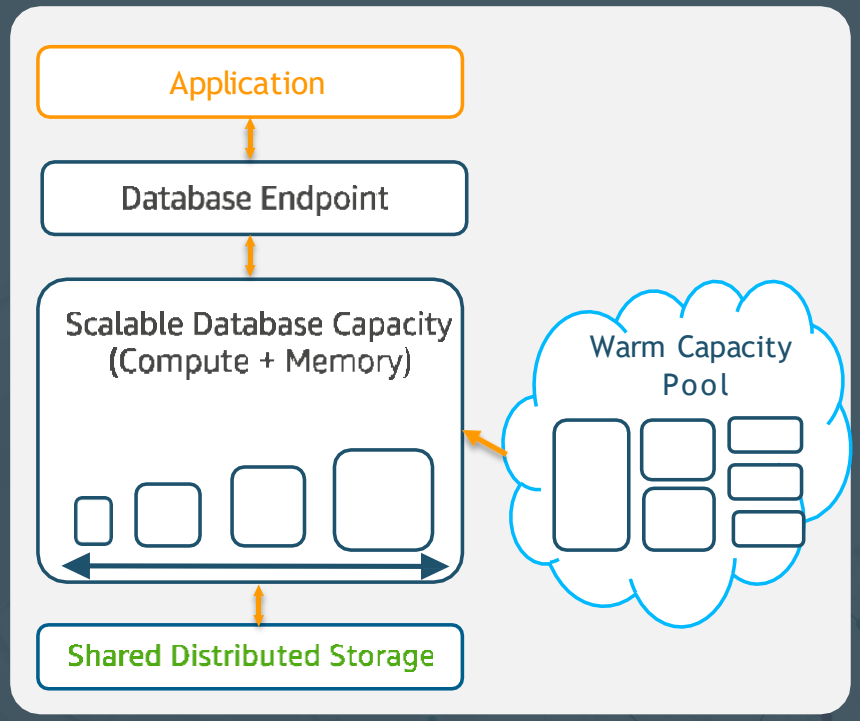
Zero application downtime from ANY AZ failure

Faster write performance and higher scale

Aurora Serverless

Aurora Serverless

On-demand, auto-scaling database for applications with variable workloads

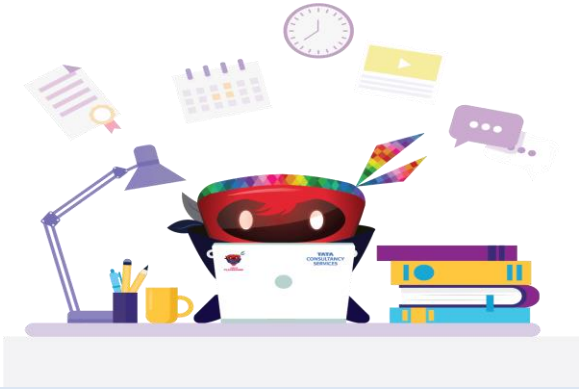


Starts up on demand, shuts down when not in use

Automatically scales with no instances to manage

Pay per second for the database capacity you use

Module # 3



AWS Databases

Amazon Redshift

Amazon Redshift

Amazon Redshift

Petabyte scale

for as low as $999/TB per year

Massively parallel

Columnar Store

Relational data warehouse

Fully managed = no admin

Amazon Redshift

Fast and powerful, petabyte-scale data warehouse

- Fully managed Relational Database
- Highly-parallel
- Columnar Data Store

Data warehouse-type queries

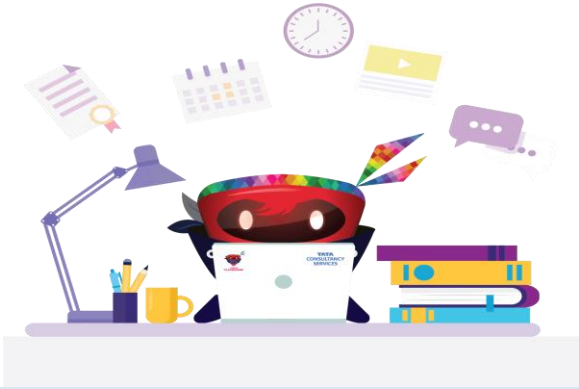
- Aggregations, historical analysis
- BI Tool integration

Grow with your data

- 160 GB ➔ 1.6 PB

Deepest integration with your data lake and AWS services

Module # 3



AWS Databases

Amazon DynamoDB

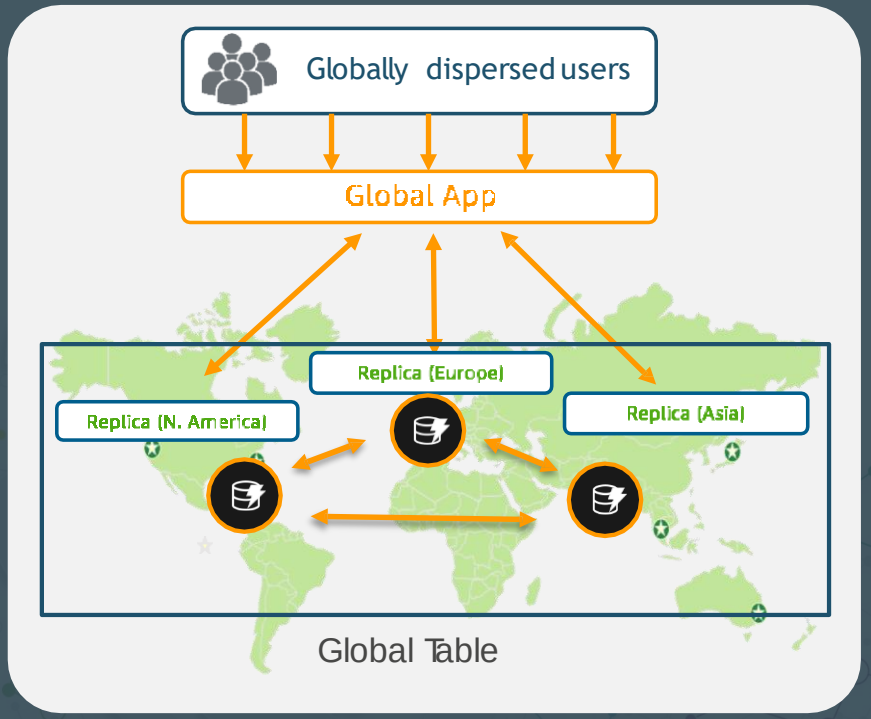
Amazon DynamoDB

Amazon DynamoDB

- NoSQL Database
- Seamless scalability
- Zero admin
- Single digit millisecond latency
- Multi-Master
- Multi-Region

DynamoDB Global Tables

First fully managed, multi-master, multi-region database



- Build high performance, globally distributed applications
- Low latency reads & writes to locally available tables
- Disaster proof with multi-region redundancy
- Easy to setup and no application re-writes required

Amazon DynamoDB

- Fully managed NoSQL database service
- Massively scalable, distributed key/value store
- Fast and predictable
- Built-in fault tolerance
- In-memory caching

NoSQL vs. SQL for a new app: how to choose?

Want simplest possible DB management?

Want app to manage DB integrity?

Amazon DynamoDB

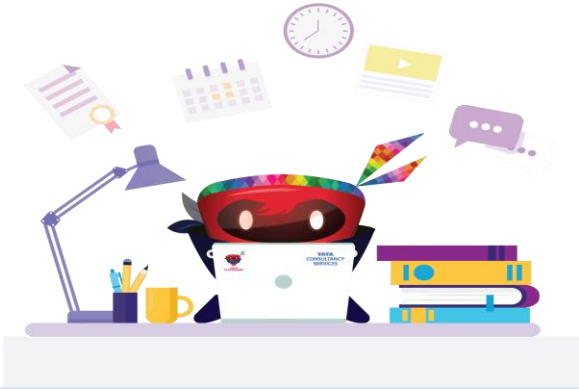
Need joins, transactions, frequent table scans?

Want DB engine to manage DB integrity?

Team has SQL skills?

Amazon RDS

Module # 3



AWS Databases

Amazon ElastiCache

Amazon ElastiCache

In-memory cache in the cloud

Improve latency and throughput for read-heavy workloads

Supports open-source caching engines

- Memcached
- Redis

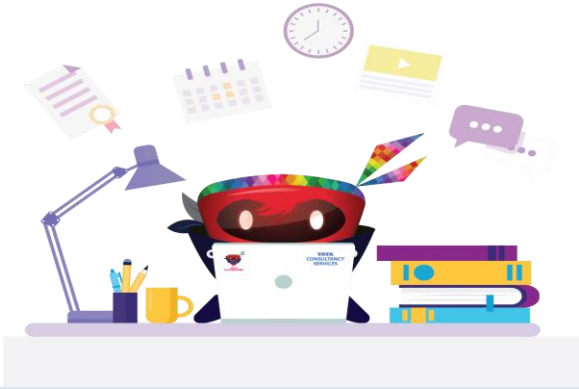
Fully managed

Multi-AZ

Examples

- Caching of MySQL database query results
- Caching of complex query post-processing results

Module # 3



AWS Databases

Amazon Neptune

Amazon Neptune

Amazon Neptune

Fully managed graph database for highly connected data

Open

Supports Apache TinkerPop™ & W3C RDF graph models

Fast & Scalable

Store billions of relationships; query 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

It's all about

choice

Performance-oriented
Cost-oriented

Enabling Certainty

Any Questions?

Thank You!

Enabling Certainty

TATA CONSULTANCY SERVICES