PUBLIC SECTOR SYMPOSIUM

BRUSSELS | MARCH 28, 2023

BAT303

Minimize Downtime During Database Migrations to AWS

Rotem Agmon (he/him)
Solutions Architect
AWS

Agenda

- Why migrate?
- Introduction to AWS Database Migration Service (AWS DMS)
- AWS DMS - Deep dive
- Post Migration Optimization

Why migrate?

Drivers for migration

Infrastructure

- Hardware end of life
- Lack of redundancy or automation
- Failure at different layers

Scalability

- Change instance sizes as required
- Scale across Regions as required

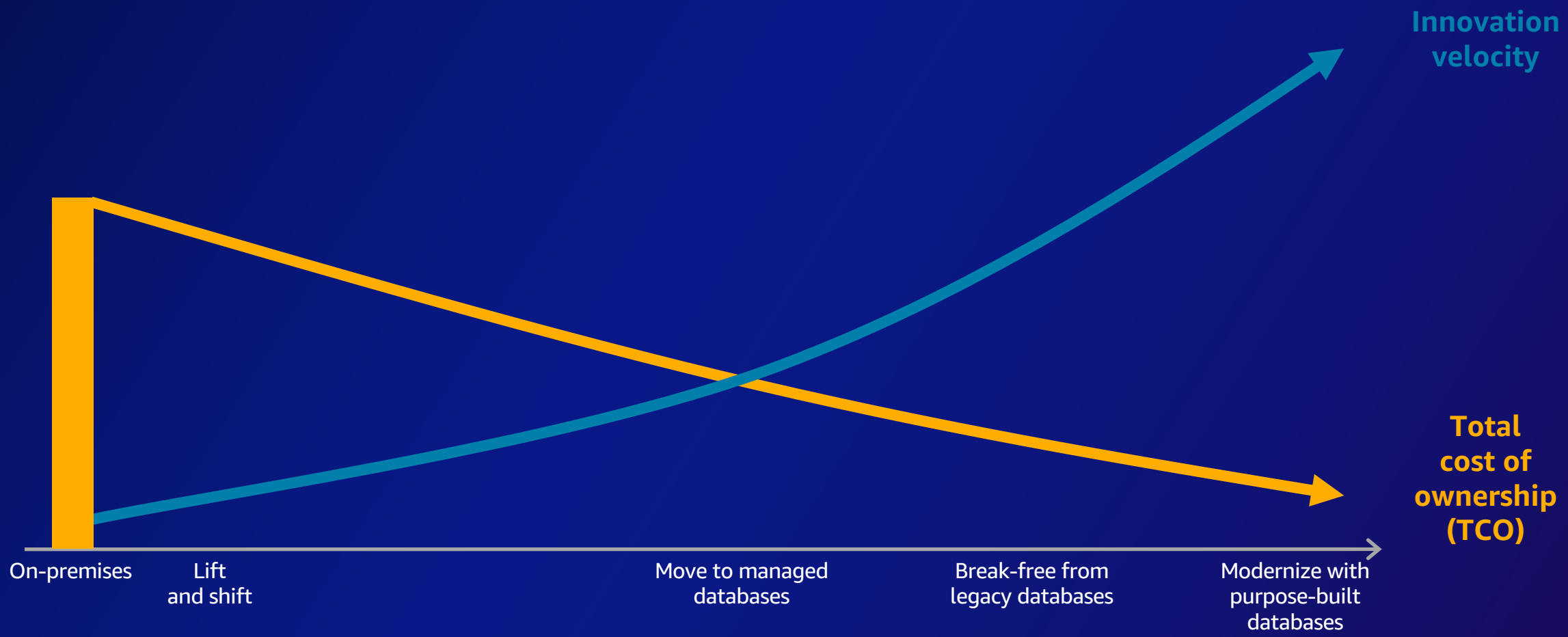
Cost optimization

- Pay for what you use
- Open-source options, license savings
- Fully managed databases

Innovation

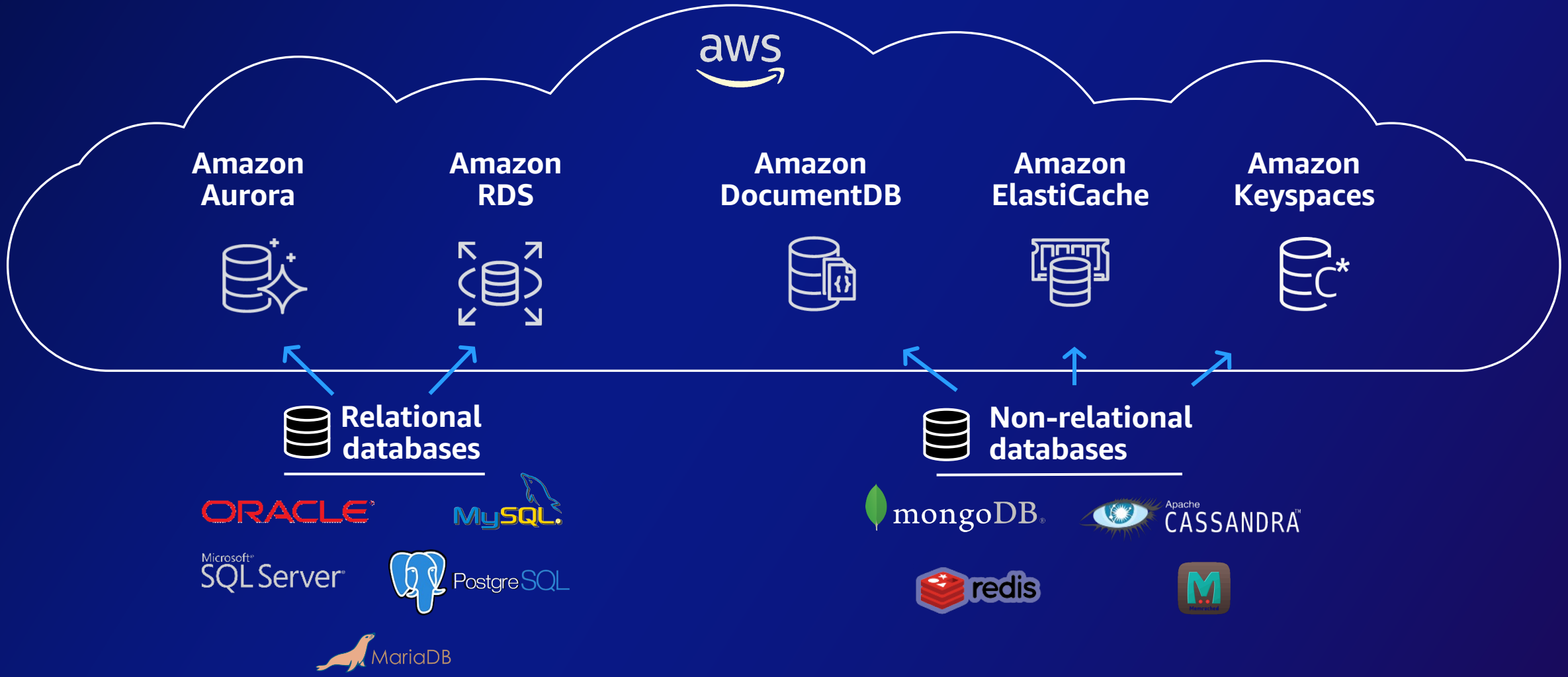
- Modernize your database tier
- Choose the right database for your workload

Modernizing leads to maximum innovation velocity and optimal total cost of ownership



Move to fully managed databases

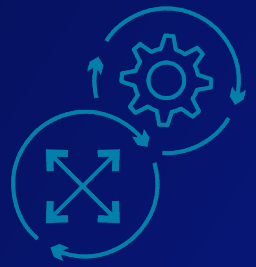
Migrate on-premises or self-managed databases to fully managed services



Introduction to AWS DMS

AWS DMS

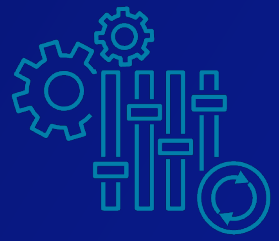
Securely migrate your databases and data warehouses to AWS with ease and minimal downtime



Managed Migration

Automated migration service will discover, assess, convert and migrate your database and analytics workloads to AWS

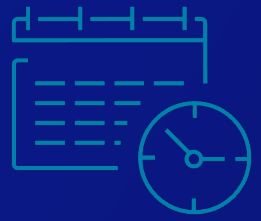
Easy to use with a few clicks to start the process



Breadth of Options

Migrate data to and from most widely used commercial and open-source databases

Supports SQL, NoSQL, text-based and data warehouse targets



High availability

Multi-AZ option; highly resilient and self–healing with continuous replication and monitoring for minimal downtime.

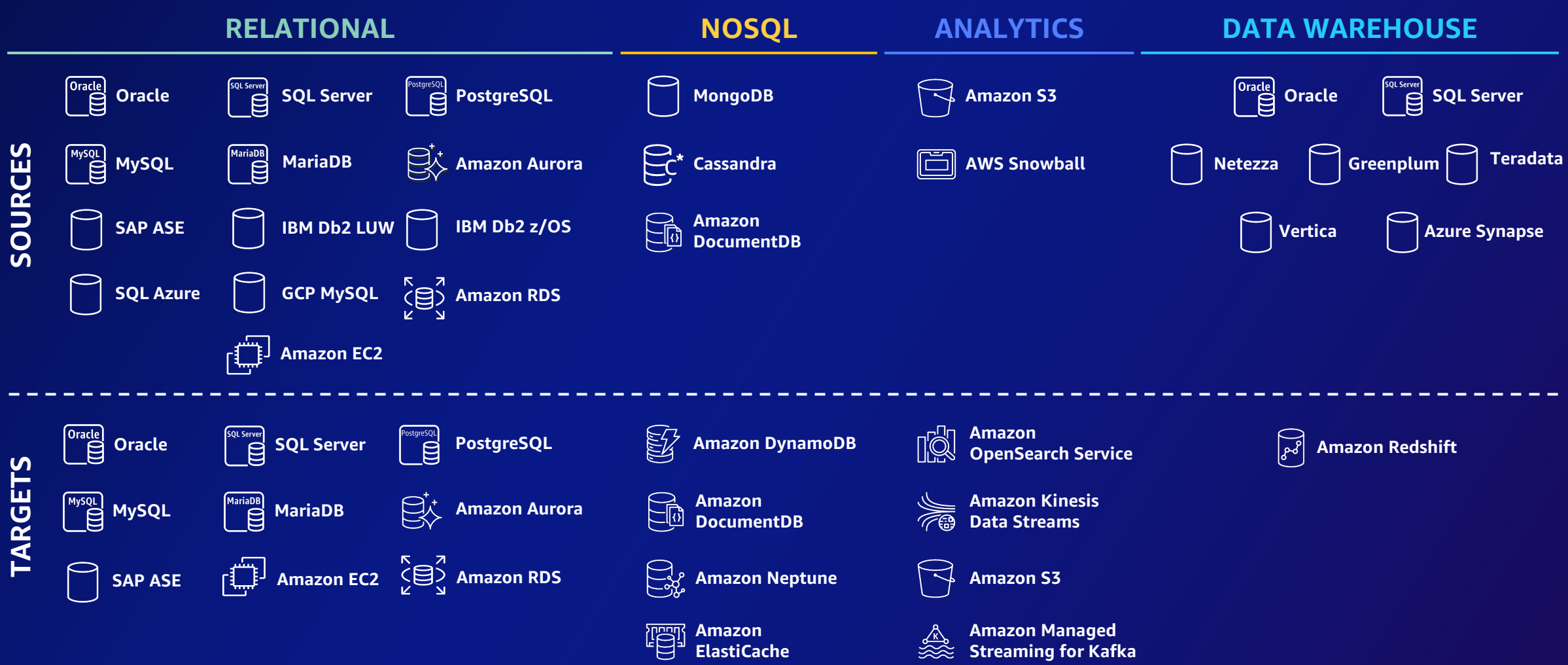
Low Cost

Self-service tool

Only pay for compute resources and additional log storage used

800,000+ databases migrated using DMS

AWS DMS supported sources and targets



Migration planning is a pain (we know)

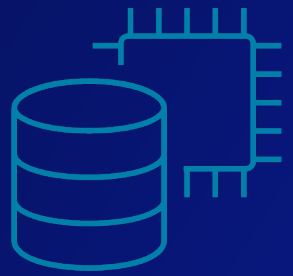
Manual, time-consuming fleet discovery

Complex and error-prone process

Expensive to bring in consultants

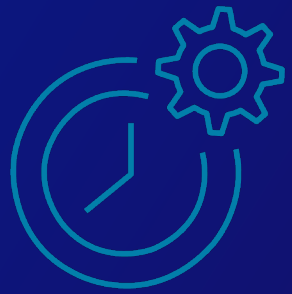
AWS DMS Fleet Advisor

Automatically inventories and assesses on-premises database and analytics server fleets



Identify databases to migrate at scale with minimal effort

Save on costs associated with migration planning and migrating workloads



Build a migration plan in hours instead of weeks

AWS DMS Fleet Advisor: How it works

Accelerate database and analytics migrations with automated inventory and migration recommendations

AWS DMS Fleet Advisor: Getting started

Accelerate database and analytics migrations with automated inventory and migration recommendations

Analyzes key data important for planning a migration, including database version, database size, number of queries per second, and IOPS.

Discovers and analyzes the same source and target databases as AWS DMS (see full list here)

- **Example sources: Oracle, Microsoft SQL Server, MySQL, MongoDB, Amazon EC2**
- **Example targets: Amazon Aurora, Amazon DynamoDB, Amazon Redshift, Amazon EC2**

Export Fleet Advisor's findings into a .csv report to quickly share with stakeholders to get alignment

Fleet Advisor is a free capability of AWS DMS

Get started with just a few clicks in the AWS Management Console

When to choose native migration tools

- Target supports native replication
- Moving ALL the data
- No transformations required
- Downtime window is large enough

When to use backup and restore

- Database is small
- Downtime window is large enough
- Moving all the data
- No transformations required

Larger database migration

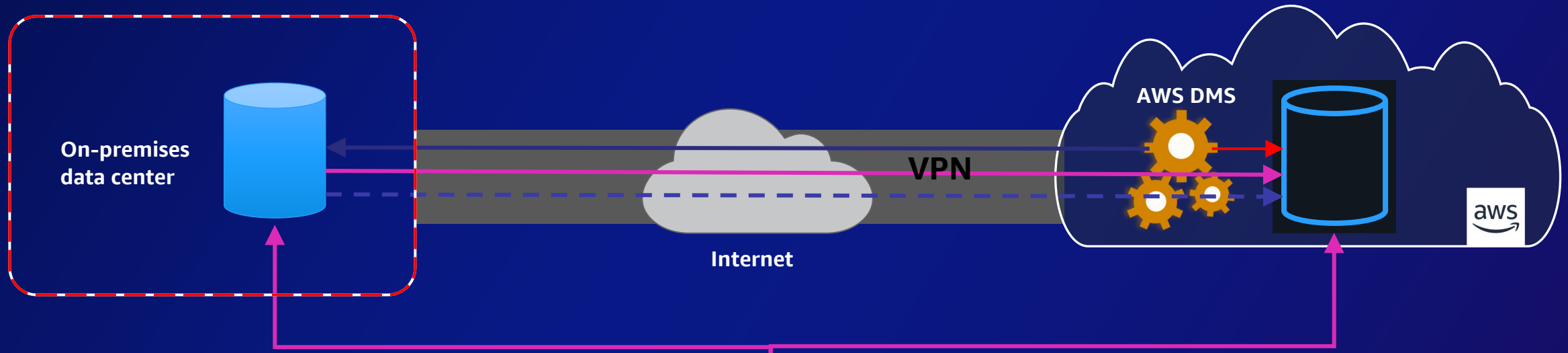
- **What is large?**
  - Moving all the data takes longer than you are willing to wait
  - > 6 TB is generally considered large

- **What can I do?**
  - Use AWS Snowball
  - Let AWS Professional Services help you

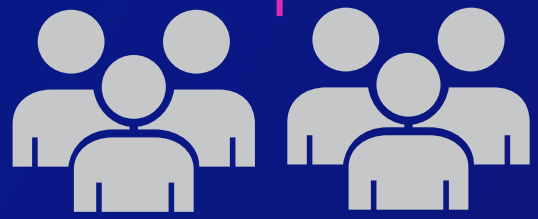
AWS DMS - Deep dive

The data migration process

The data migration process

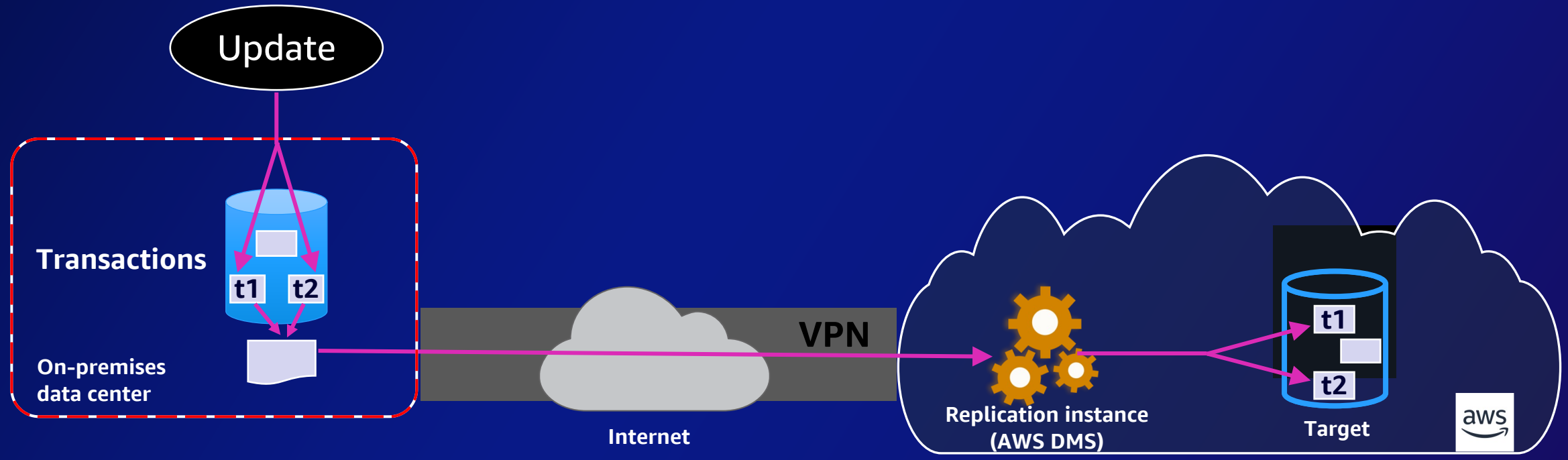


- Start a replication instance
- Connect to source and target databases
- Select tables, schemas, or databases



- Let AWS DMS load data and keep them in sync
- Switch applications over to the target after they are in sync at your convenience
- AWS DMS captures ongoing changes after initial migration using CDC

Change data capture (CDC)

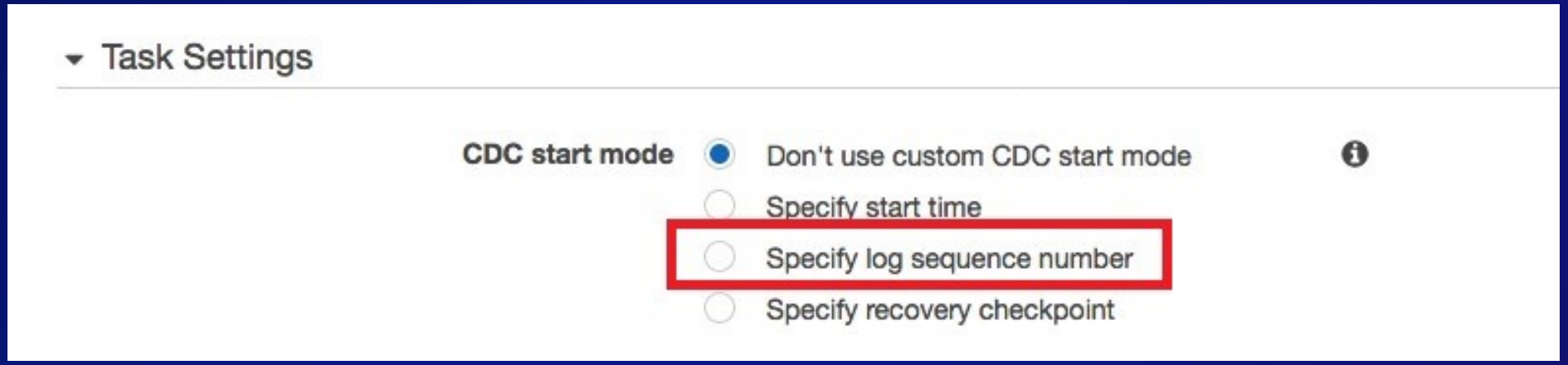


- CDC for continuous data replication
- Tracks data changes by reading transaction logs
- Continuously streams changes for downstream applications to consume

Change apply after bulk load

AWS DMS Tasks

- Configure Tasks in any of the below modes –
  - Full Load
  - Full Load and Ongoing Replication
  - Replicate Data Changes Only
- Replicate from specific LSN or Recovery Point
- Stop Replication at Server Stop Time or Commit Stop Time



Database Migration Options

- **Load is table by table**
  - Configurable number of tables in parallel
- **One time load or Change Data Capture (CDC)**
  - Read from database log on the source and apply to the target
- **Filter criteria available for selective loading**
  - Select only a few tables or a subset of data in your tables
- **Multiple sources and targets. Mix and match.**
  - One side of the migration must be in AWS
- **Ongoing replication support**
  - Keep your replication going until your application is ready to cutover

Database Migration Options

https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Source.SQLServer.html

The following lists SQL Server editions you can use as a source with on-premises databases.

SQL Server version	Full load	Ongoing replication (CDC)
2005, 2008, 2008R2, 2012, 2014, 2016, 2017, 2019	Enterprise Standard Workgroup Developer Web	Enterprise Edition Developer Standard Edition (version 2016 SP1 and later)

When using SQL Server 2005 as a source, only Full Load is supported.

Oracle Database as Source

- On Amazon EC2 - Oracle 10.2+, 11g and up to 12.2, 18c, and 19c
- On Amazon RDS – 11g (versions 11.2.0.4 and later) and up to 12.2, 18c, and 19c
- SSL is supported
- AWS DMS can read using LogMiner or BinaryReader from Oracle DB (only required for CDC)
- For high volume DBs BinaryReader reduces the load, but LogMiner support encryption & compression options
- Source DB must be running in ARCHIVELOG mode and backup retention period must be > 1 days (for RDS)
- Tables with LOBs must have a primary key to use CDC
- Oracle Active Data Guard Stand-by database can be a source

https://docs.amazonaws.cn/en_us/dms/latest/userguide/CHAP_Source.Oracle.html

DMS Best Practices

- Performance of an AWS DMS Migration depends on –
  - Resource availability on the source
  - The available network throughput
  - The resource capacity of the replication server
  - The ability of the target to ingest changes
  - The type and distribution of source data
  - The number of objects to be migrated
- Load multiple tables in parallel (default 8 – up to 50 per task) – careful!
- Disable/drop primary key indexes, secondary indexes, referential integrity constraints and triggers
- Disable backups and transaction logging on target until ready for cutover

Post Migration Optimization

AWS Well-Architected Framework – Performance Efficiency

Design considerations: performance efficiency

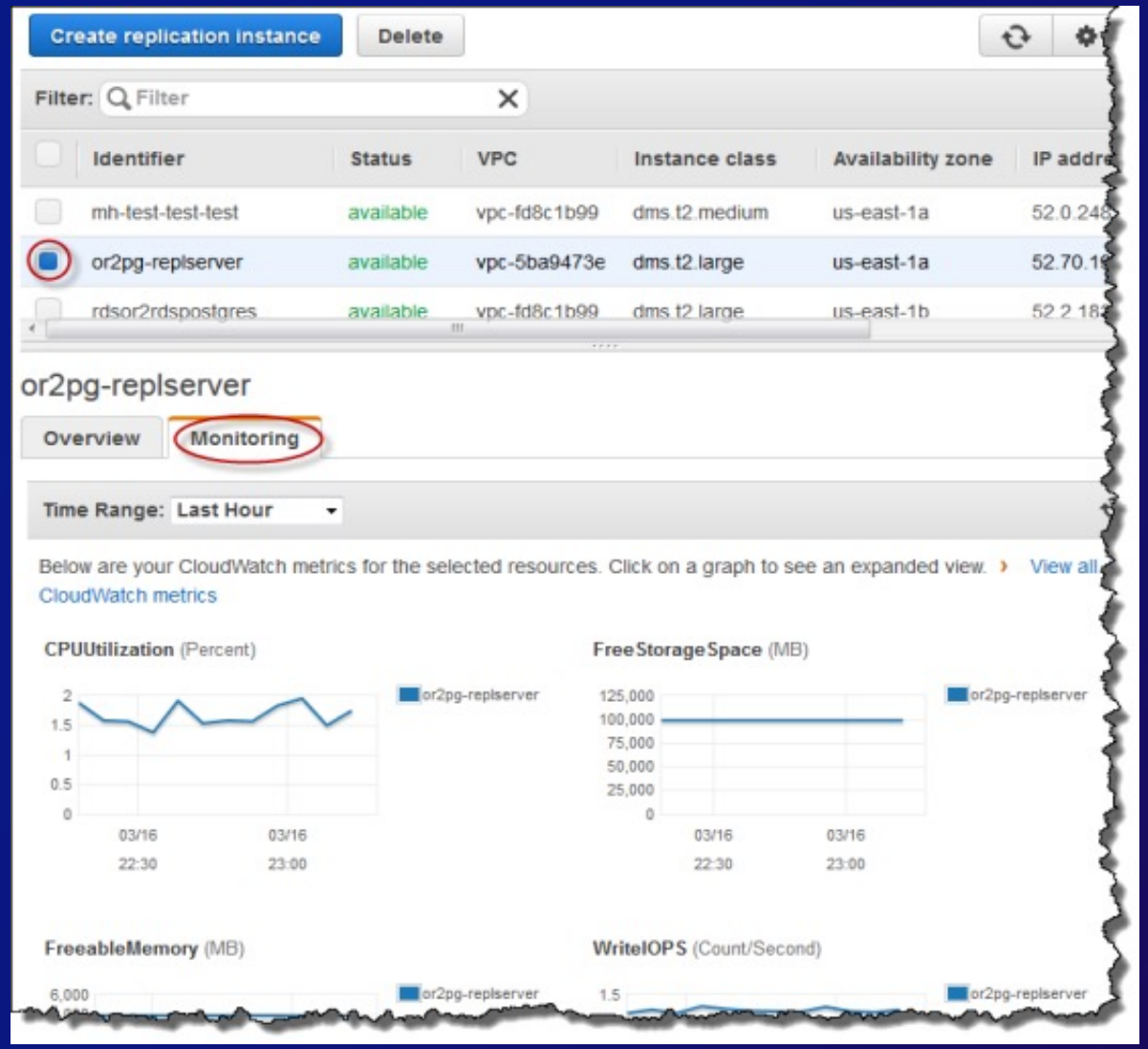
- Appropriately size your replication instance
- Understand source workload pattern and effects on replica lag
- Appropriately size EBS volume for performance
- Understand impacts of multi-AZ deployment
- Deploy replication instance as close to source if multi-region
- Understand different engine specific concepts like binary reader or logminer
- Break out large tables over multiple tasks
- Break out LOB in separate tasks
- LOB considerations and different options
- Understand Batch Apply and when to use it
- Understand Extra Connection Attributes

Monitoring DMS performance

- AWS DMS automatically sends metrics to CloudWatch every minute for each replication instance

CloudWatch Metrics:

- CPUUtilization
- FreeableMemory
- SwapUtilization
- TargetLatency
- SourceLatency



- For list of available DMS metrics https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Monitoring.html

Troubleshooting DMS performance

Poor DMS performance symptoms:

- High CPU utilization
- High memory utilization
- High I/O usage
- High Target Latency
- High Source Latency
- Understanding logging options

Questions: performance efficiency

- How do you size your replication instance?
- How do you monitor performance of your replication instance?
- How do you handle LOB Data?
- How do you design your tasks?
- How do you handle full load vs CDC?
- Do you have secondary indexes and triggers enabled for full load?
- What Extra Connection Attributes have you set?
- How do you interpret service log messages? Do you have a process to turn on detailed debugging?

AWS Well-Architected Framework – Cost Optimization

Design considerations: cost optimization

- Size compute capacity for sustained load
- Use smaller replication instances in lower environments with less throughput
- Monitor metrics and set billing alarms
- Use cost allocation tags to track expense at a granular level
- Do burstable instances make sense for dev/test?
- Is your storage appropriately sized?

Data Transfer: cost optimization strategies

Monitoring

- CloudWatch Metrics:
  - Network In (Bytes)
  - Network Out (Bytes)

Additional Considerations

- Data Transfer into DMS is free.
- Data transferred between DMS and RDS or Amazon EC2 instances in the same Availability Zone is free.
- Data transferred between AZ’s, regions, or outside of AWS will incur standard charges.

AWS Well-Architected Framework - Security

Design considerations: security

- Configure security group to limit surface area of attack
- Isolate credentials from application by using AWS Secrets Manager or equivalent
- Configure AWS Secrets Manager to automatically rotate the secrets (credential management and rotation) for your AWS DMS endpoint credentials
- Use AWS IAM to control access to AWS DMS resources
- For SSL/TLS connections, rotate certificates before the expiration date

Encryption in DMS

Encryption of data at rest

- AES256-based storage encryption, incl. (CloudWatch Logs publishing, Enhanced Monitoring)
- Key management using AWS KMS
- No performance impact on workloads

Encryption of data in transit (TLS)

- Certificates are signed by an authority
- Endpoint connection is CN of certificate (protects against spoofing attacks)
- Custom endpoints not included in certificate (cannot verify identity)
- Supports TLS 1.2

More: https://aws.amazon.com/blogs/database/selecting-the-right-encryption-options-for-amazon-rds-and-amazon-aurora-database-engines/

Compliance

- Compliance and assurance programs for finance, healthcare, government, and more
- Third-party auditors assess the security and compliance of Amazon Aurora as part of multiple AWS compliance programs. These include SOC, PCI, FedRAMP, HIPAA, and others
- You can download third-party audit reports using AWS Artifact
- For a list of specific compliance programs, see AWS services in scope by compliance program

DMS is compliant with:

- SOC 1, 2, 3
- PCI
- ISMAP
- IRAP Protected
- OSPAR
- C5
- FedRAMP
- DoD SRG IL 4,5,6

DMS Resources - Getting Started

- Visit our website: https://aws.amazon.com/dms
- Technical User Guide: https://docs.aws.amazon.com/dms
- Pricing: https://aws.amazon.com/dms/pricing
- Support: https://forums.aws.amazon.com/forum.jspa?forumID=216
- Get started with Free Tier: https://aws.amazon.com/free

Thank you!

Please complete the session survey in the mobile app

Rotem Agmon

Solutions Architect

linkedin.com/in/rotemagmon/