



# INNOVATE

ONLINE CONFERENCE

分会场二：数据库

# 数据库和数据仓库上云的最佳实践

韩思捷，AWS 解决方案架构师

# 议题

AWS Data Migration Service (AWS DMS)以及 AWS Schema Conversion Tool (AWS SCT) 介绍

AWS Snowball 与 AWS DMS 集成使用

AWS DMS 和 AWS SCT 的产品亮点

迁移手册

# AWS Data Migration Service 和 AWS Schema Conversion Tool

# 数据库的迁移是上云的一个重要部分

如何快速并方便的把我数据中心里的数据迁移到云上？

如何在数据库的迁移过程中尽量缩短应用系统的宕机时间？

是否存在方法能够进行 schema 的自动转换和冲突检查？

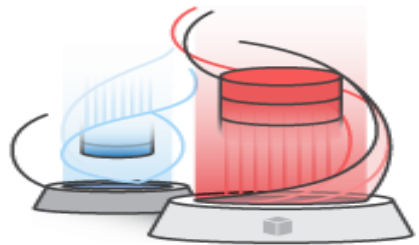
在迁移过程中如何把我的数据库里的数据进行整合？

我能在迁移过程中更换掉我的商业数据库吗？

# AWS DMS 和 AWS SCT 是什么？

我们的目标：所有客户可以自由的选择适合他们业务需求的数据平台  
#DBFreedom

AWS Database Migration Service (AWS DMS) 能够简单并安全的迁移或者复制你的数据库以及数据仓库到 AWS 上



AWS Schema Conversion Tool (AWS SCT) 能够把你的商业数据库或者数据仓库的 schema 转换为开源数据库引擎或者 AWS 原生的服务，比如 Amazon Aurora 和 Amazon Redshift

# 何时使用 AWS DMS 和 AWS SCT?

转换



迁移



复制





# AWS Schema Conversion Tool

- 转换你的数据库

把你的数据库，包括 Oracle, SQL Server, 或者 DB2 LUW，转换为 PostgreSQL, MySQL, 或 Amazon Aurora



- 转换你的数据仓库

把你的数据仓库，包括 Oracle, SQL Server, Netezza, Greenplum, Vertica 或 Teradata，转换为 Amazon Redshift





# AWS Data Migration Service

迁移



- 迁移关键业务系统
- 迁移数据仓库到 Amazon Redshift
- 归档老数据
- 升级小版本
- 合并多个数据分片到 Amazon Aurora
- 复制数据从而在云端分析数据
- 从 NoSQL 迁移到 SQL，或者从 SQL 迁移到 NoSQL，或者从 NoSQL 迁移到 NoSQL



Amazon Redshift



Amazon RDS



Amazon S3



Amazon DynamoDB



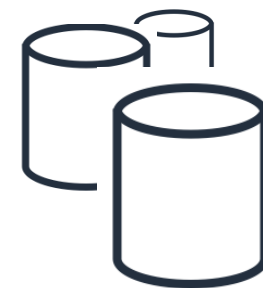
Amazon Aurora

# AWS Data Migration Service

复制



- 创建跨 region 的只读副本
- 在云上运行你的分析操作
- 为你的数据湖导入数据



# AWS DMS – 支持的源和目标类型

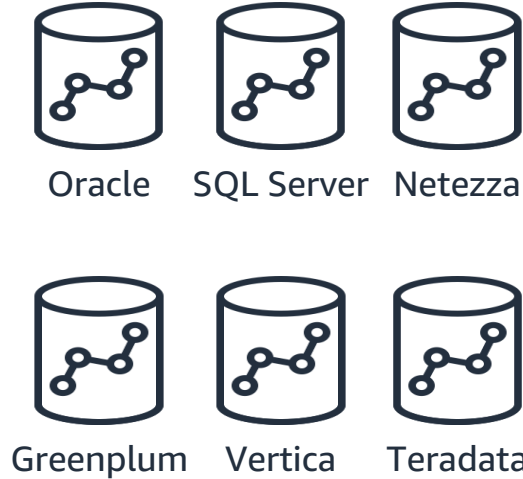
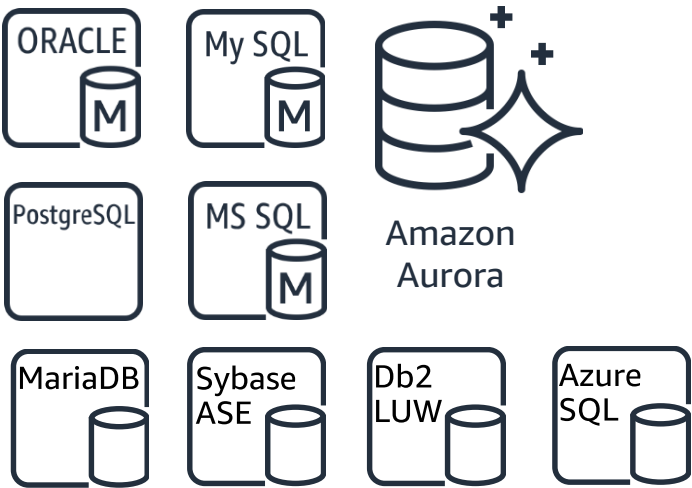
关系型数据库

NoSQL

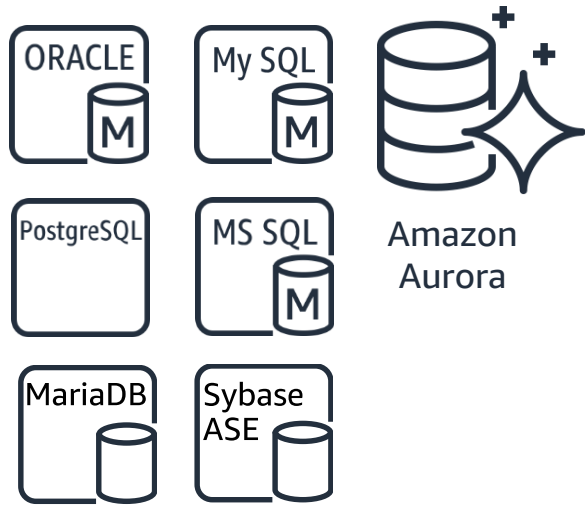
数据仓库

数据湖

源端



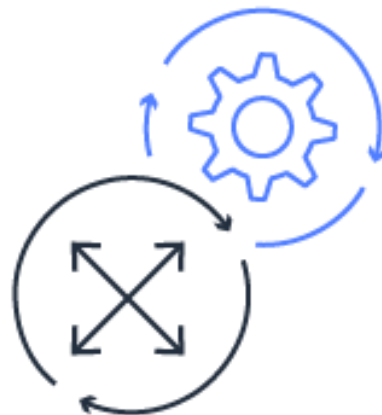
目标端



# AWS DMS / AWS SCT 产品特点



同构或异构



现有环境



转换数据结构和代码



验证



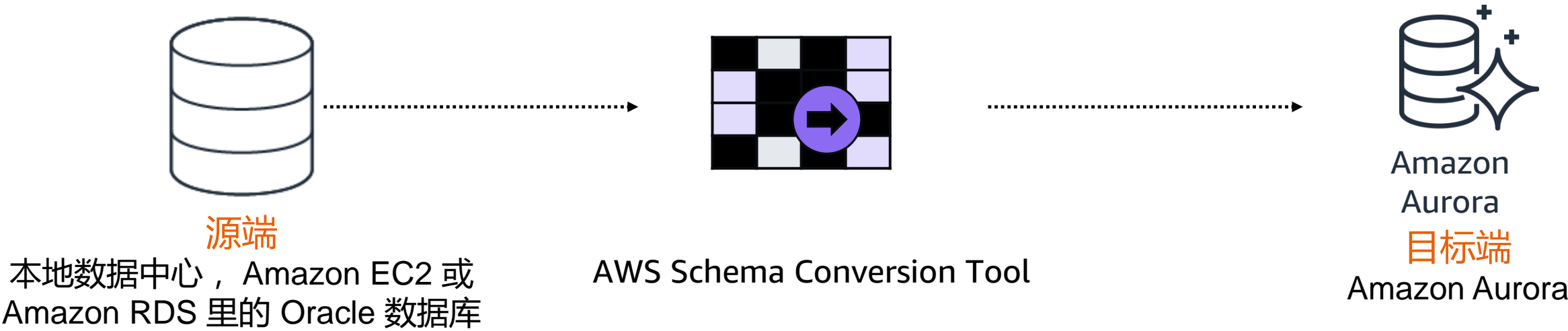
安全



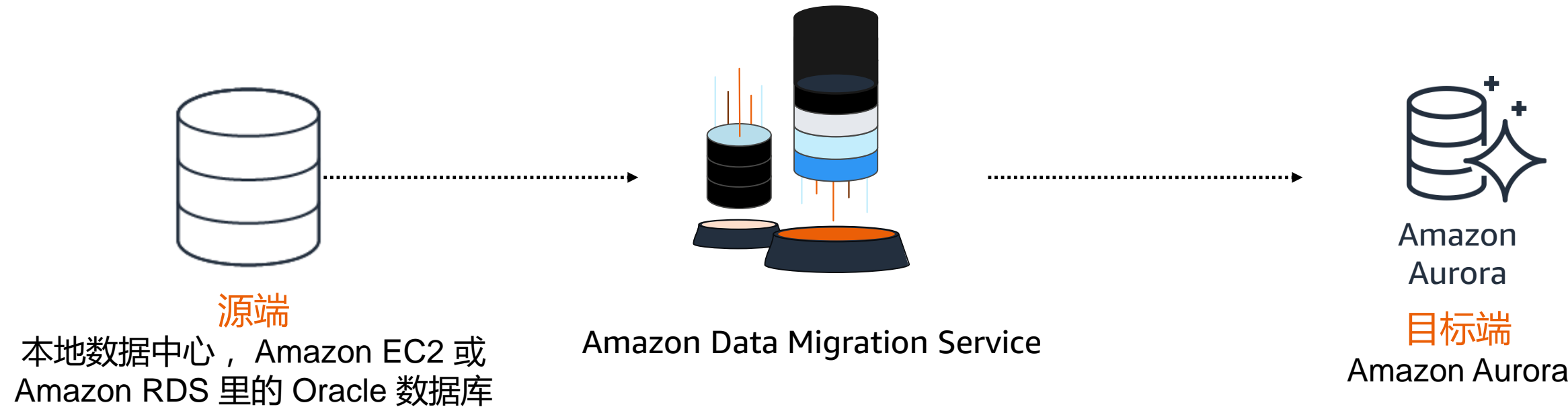
检查

# 数据库迁移过程

## 步骤1



## 步骤2



# AWS Snowball

# AWS Snowball

AWS Snowball 是一个 PB 级别的数据传输解决方案，该方案使用了安全的设备在 AWS 云上传入或者传出数据

## 特性

简单，快速以及安全的传输数据

成本只有基于网络传输的1/5

可以传输若干 PB 的数据



**AWS Snowball**



# AWS Snowball



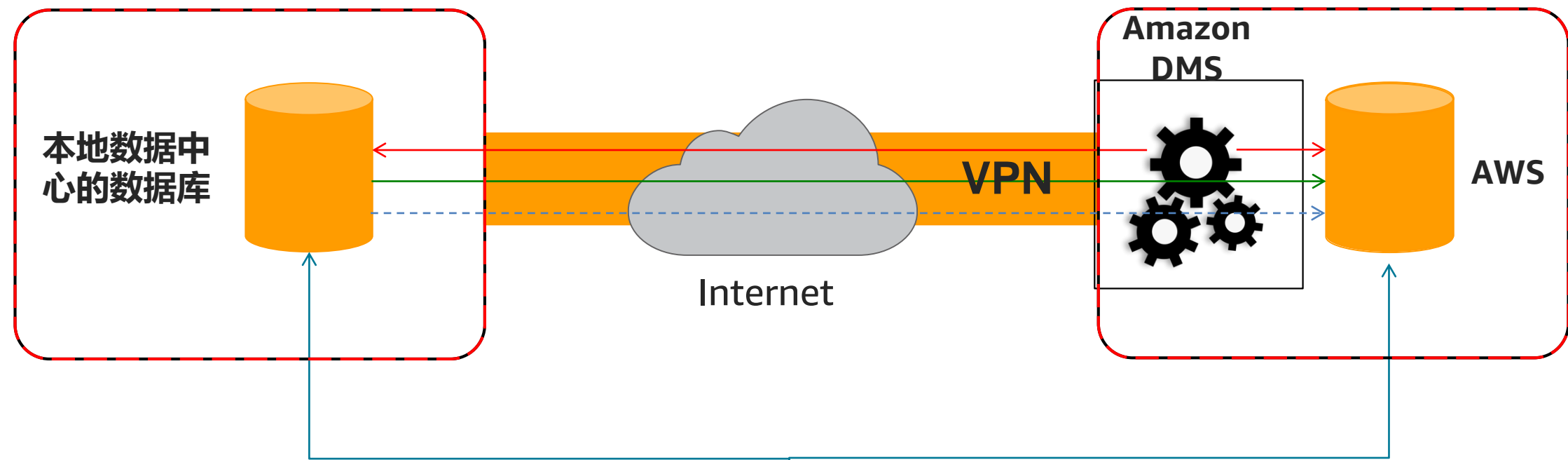
# 同时使用 AWS Snowball 和 AWS DMS

## 通常的用户案例 –

- 迁移大型数据库（超过5TB）
- 一次性的迁移多个数据库
- 通过低速网络迁移数据库

# AWS DMS 的架构，组件以及处理流程

# 在迁移过程中保持数据库可用



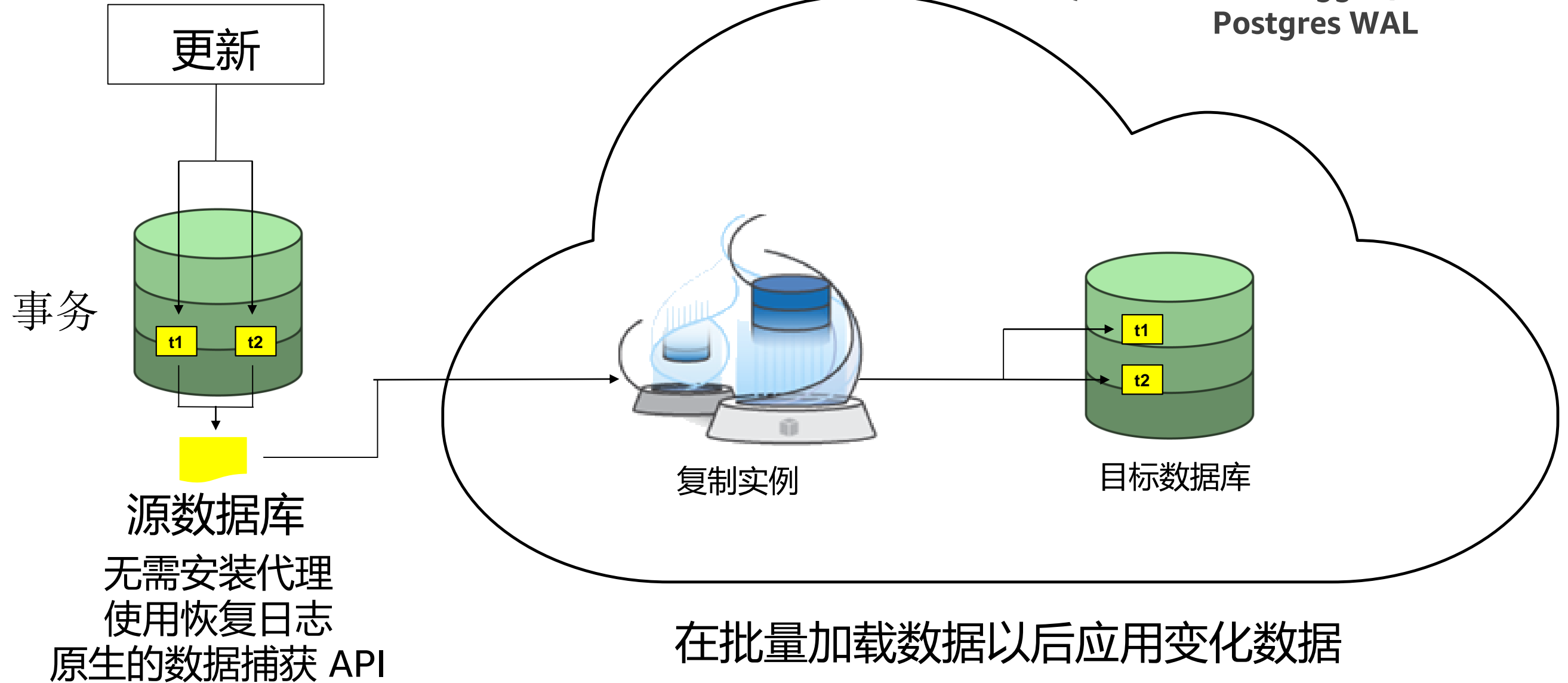
- ◆ 启动一个复制实例
- ◆ 连接到源和目标数据库
- ◆ 选择表，schema 或者数据库



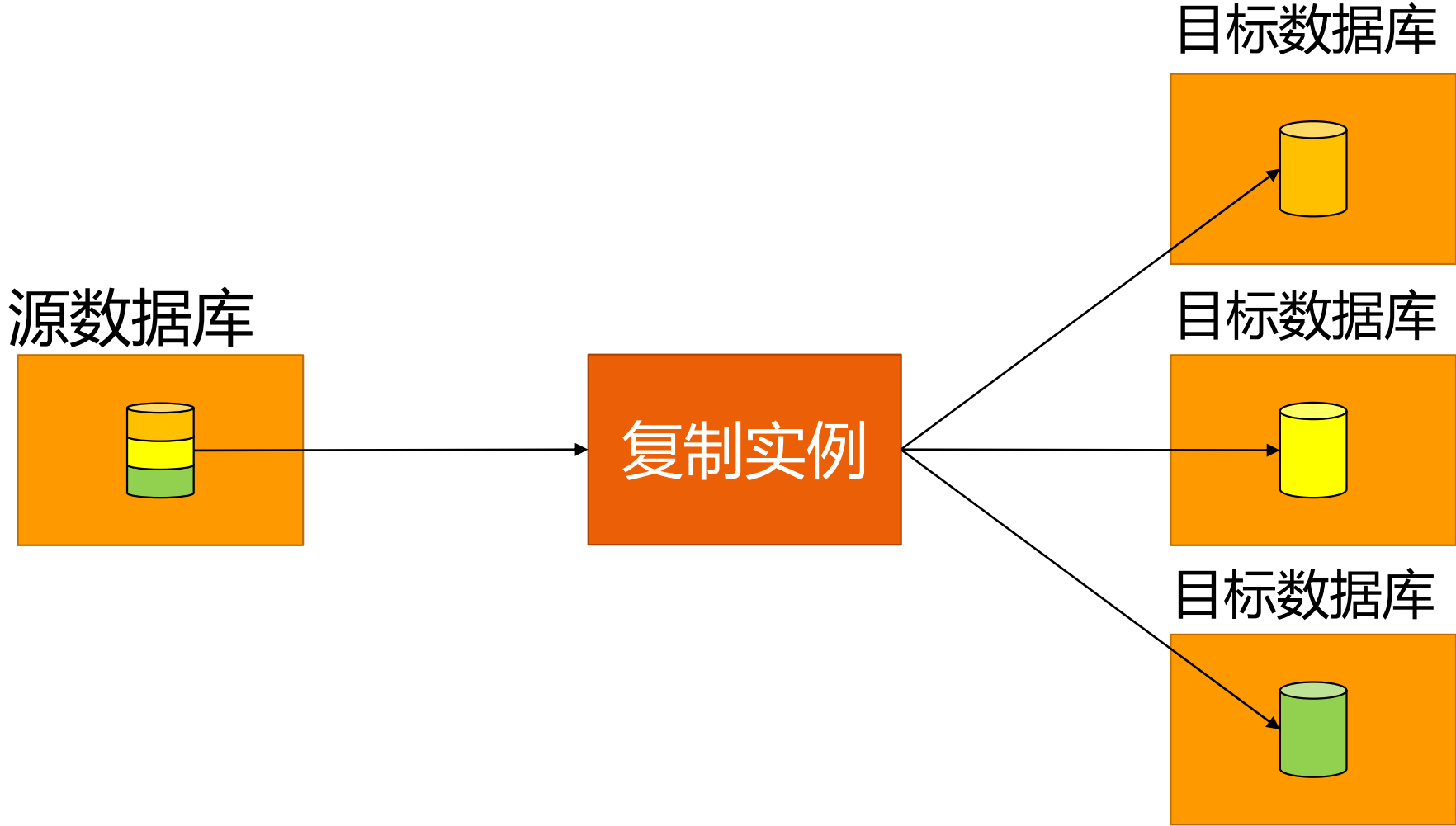
- ◆ 让 AWS DMS 创建表，加载数据以及保持数据同步
- ◆ 在合适的时候把你的应用系统切换到目标数据库上

# 捕获并应用变化数据 (CDC)

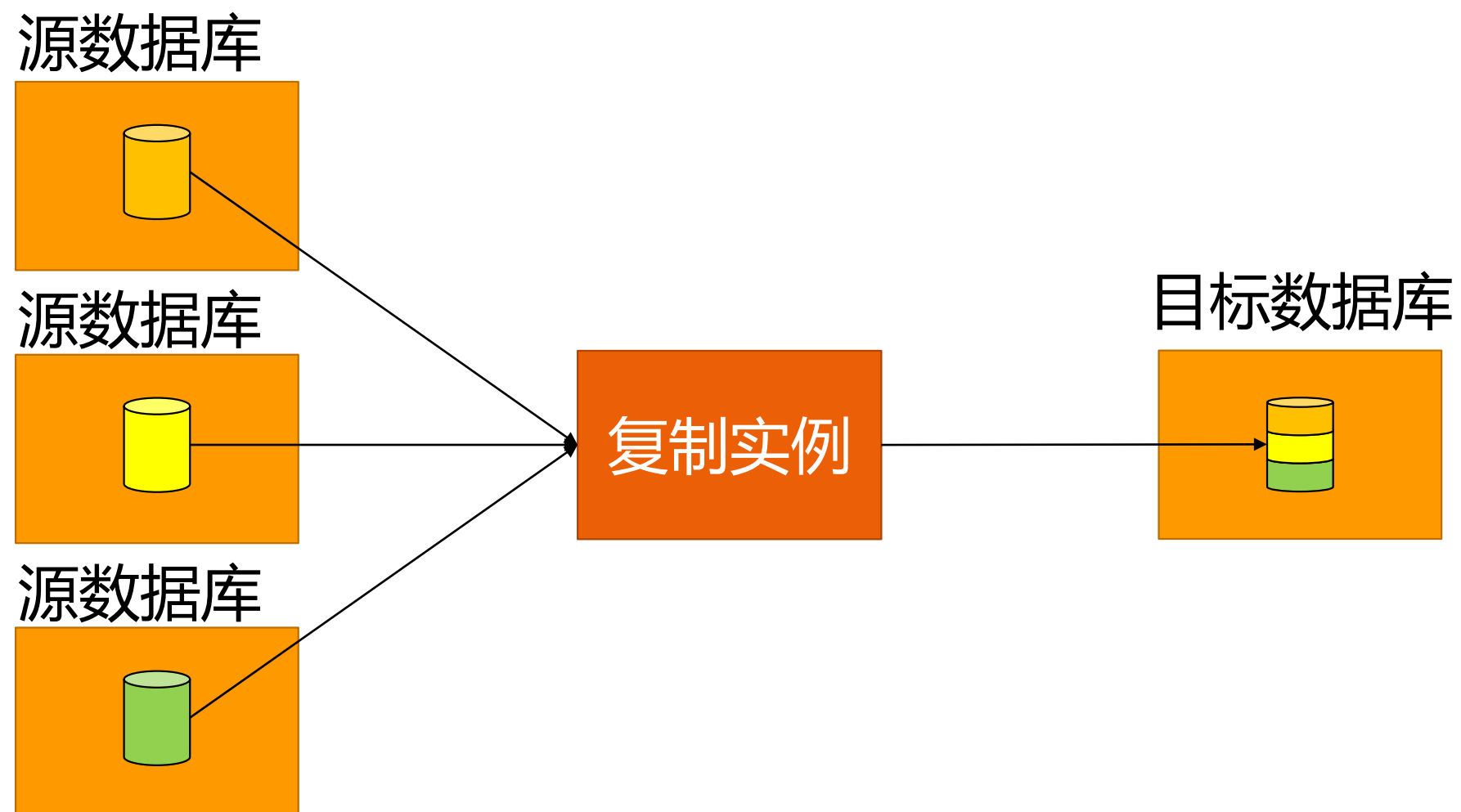
Oracle supplemental logging  
MySQL row-level bin logging  
SQL Server bulk logged/full recovery  
Postgres WAL



# 多个目标数据库

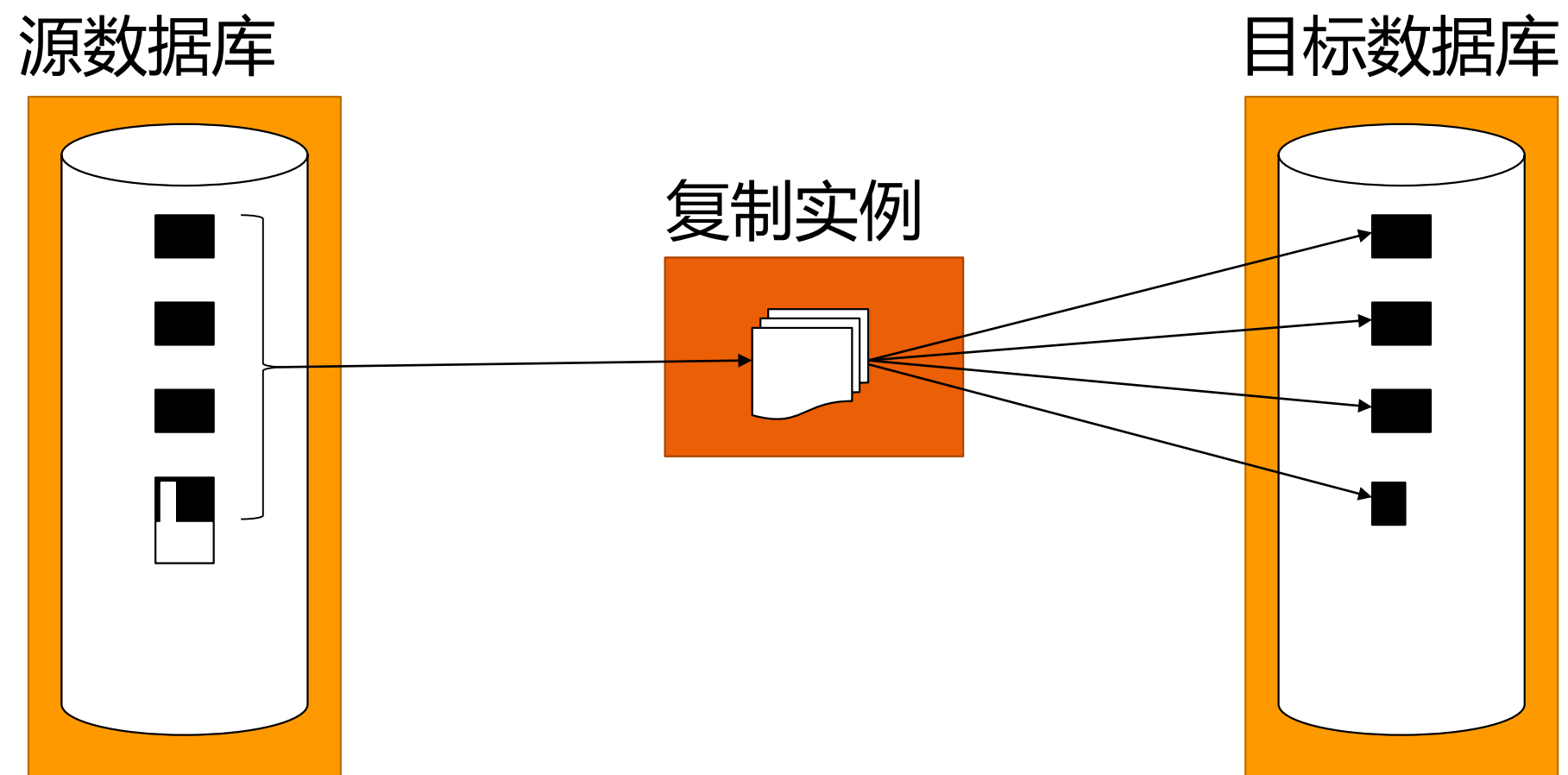


# 多个源数据库

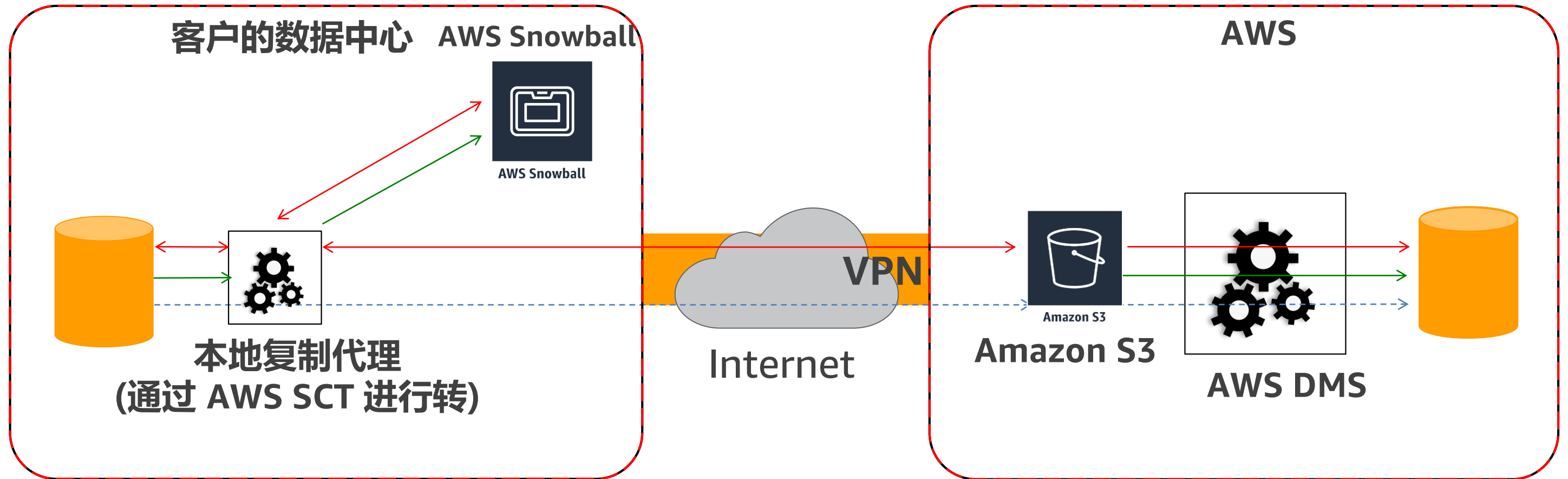




# 有选择的复制表



# 使用 AWS Snowball 与 AWS DMS 迁移数据



# 演示

## AWS DMS

## Dashboard

### ▼ Conversion & migration

## Database migration tasks

### ▼ Resource management

## Replication instances

## Endpoints

## Certificates

## Subnet groups

# AWS Database Migration Service

Migrate your databases to AWS with minimal downtime

## The quickest and easiest way to migrate databases to AWS with low cost and

# AWS DMS / AWS SCT 产品的主要特性



# 从已有环境迁移到 Amazon Aurora 相关的手册

- 按照不同的主题描述了如何把数据库和数据仓库迁移到 AWS 上
- 介绍了相关的特性和不同的数据库对象
- 迁移最佳实践
  - Oracle 到 Aurora PostgreSQL
  - SQL Server 到 Aurora MySQL
  - SQL Server 到 Aurora PostgreSQL

从下面的地址下载手册：

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

	Oracle Feature	PostgreSQL Feature	Compatibility
<a href="#">Link</a>	Index Organized Tables (IOTs)	PostgreSQL "Cluster" Tables	Yes*
<a href="#">Link</a>	Common Data Types	Common Data Types	Yes
<a href="#">Link</a>	Table Constraints	Table Constraints	Yes
<a href="#">Link</a>	Table Partitioning including: RANGE, LIST, HASH, COMPOSITE, Automatic LIST	Table Partitioning including: RANGE, LIST	Yes*
<a href="#">Link</a>	Exchange & Split Partitions	N/A	None
<a href="#">Link</a>	Temporary Tables	Temporary Tables	Yes*
<a href="#">Link</a>	Unused Columns	ALTER TABLE DROP COLUMN	Yes
<a href="#">Link</a>	Virtual Columns	Views and/or Function as a Column	Yes*
<a href="#">Link</a>	User Defined Types (UDTs)	User Defined Types (UDTs)	Yes
<a href="#">Link</a>	Read Only Tables & Table Partitions	Read Only Roles and/or Triggers	Yes*
<a href="#">Link</a>	<a href="#">Link</a> Recovery Manager (RMAN)	AWS Aurora Snapshots	Yes
<a href="#">Link</a>	<a href="#">Link</a> Flashback Database	AWS Aurora Snapshots	Yes
<a href="#">Link</a>	<a href="#">Link</a> 12c Multi-tenant architecture: PDBs and CDB	Databases	Yes*
<a href="#">Link</a>	<a href="#">Link</a> Tablespaces & DataFiles	Tablespaces	Yes*
<a href="#">Link</a>	<a href="#">Link</a> Data Pump	pg_dump & pg_restore	Yes
<a href="#">Link</a>	<a href="#">Link</a> Resource Manager	Separate AWS Aurora Clusters	Yes
<a href="#">Link</a>	<a href="#">Link</a> Database Users	Database Roles	Yes
<a href="#">Link</a>	<a href="#">Link</a> Database Roles	Database Roles	Yes
<a href="#">Link</a>	<a href="#">Link</a> SGA & PGA Memory	Memory Buffers	Yes
<a href="#">Link</a>	<a href="#">Link</a> V\$ Views & Data Dictionary	System Catalog Tables, Statistics Collector, AWS Aurora Performance Insights	Yes*
<a href="#">Link</a>	<a href="#">Link</a> Log Miner	Logging Options	Yes
<a href="#">Link</a>	<a href="#">Link</a> Instance & Database Parameters (SPFILE)	AWS Aurora Parameter Groups	Yes
<a href="#">Link</a>	<a href="#">Link</a> Session Parameters	Session Parameters	Yes
<a href="#">Link</a>	<a href="#">Link</a> Alert.log (error log)	Error Log via AWS Console	Yes
<a href="#">Link</a>	<a href="#">Link</a> Automatic and Manual Statistics Collection	Automatic and Manual Statistics Collection	Yes
<a href="#">Link</a>	<a href="#">Link</a> Viewing Execution Plans	Viewing Execution Plans	Yes

# AWS SCT 产品亮点

## 评估报告

- 对使用了开源数据库引擎的源数据库的迁移进行评估 – Amazon RDS for MySQL, Amazon RDS for PostgreSQL, 以及 Amazon Aurora
- 推荐目标数据库引擎
- 为完成迁移提供了详细的步骤

## 转换数据结构和代码

- 尝试转换数据库结构以及代码到目标数据库，包括存储过程和函数
- 扫描并转换应用程序代码里内嵌的 SQL 语句
- 产生建议报告

## 迁移数据仓库到 Amazon Redshift

- 通过本地迁移代理抽取数据
- 文件会被加载到一个 Amazon Simple Storage Service (Amazon S3) 桶里，然后导入 Amazon Redshift
- Netezza
- Vertica
- Greenplum
- Teradata
- Oracle
- SQL Server



# AWS DMS 产品亮点

## 迁移前的评估

- 在进行数据迁移之前检查迁移的任务配置
- 对潜在的问题进行告警，比如不支持的数据类型等
- 防止不必要的，消耗时间的运行时迁移故障

## 数据检查

- 对所有被选择要进行迁移的数据进行检查
- 包括 Full load 和 CDC 阶段
- 可以针对同构以及异构的数据库迁移

## 集成 Snowball

- 使用 AWS Snowball & AWS DMS 来迁移数据到 AWS
- 迁移特别大的数据库 (超过 5TB)
  - 一次迁移多个数据库
  - 基于较慢的网络进行迁移

# Amazon Data Migration Service

## 迁移数据库到 AWS

**120,000+**  
已经迁移的数据库



在本地数据中心和 AWS 之间进行迁移



在不同的数据库之间进行迁移



自动转换数据结构



持续数据复制实现零宕机迁移

请参考我们的博客：

<https://aws.amazon.com/blogs/database>

大类: database migration service

# 感谢参加 AWS INNOVATE 2019 在线技术大会

我们希望您在这里找到感兴趣的内容！

也请帮助我们完成**投票打分**和**反馈问卷**。

欲获取关于 AWS 的更多信息和技术内容，可以通过以下方式找到我们：



微信公众号：AWSChina



新浪微博：<https://www.weibo.com/amazonaws/>



领英：<https://www.linkedin.com/company/aws-china/>



知乎：<https://www.zhihu.com/org/aws-54/activities/>



视频中心：<http://aws.amazon.bokecc.com/>



更多线上活动：<https://aws.amazon.com/cn/about-aws/events/webinar/>