



TRANSFORMATION DAY



宁夏区域
由西云数据运营

北京区域
由光环新网运营

巧用 AWS 数据库迁移服务 加速企业数字化转型

蒋华

AWS APN合作伙伴解决方案架构师

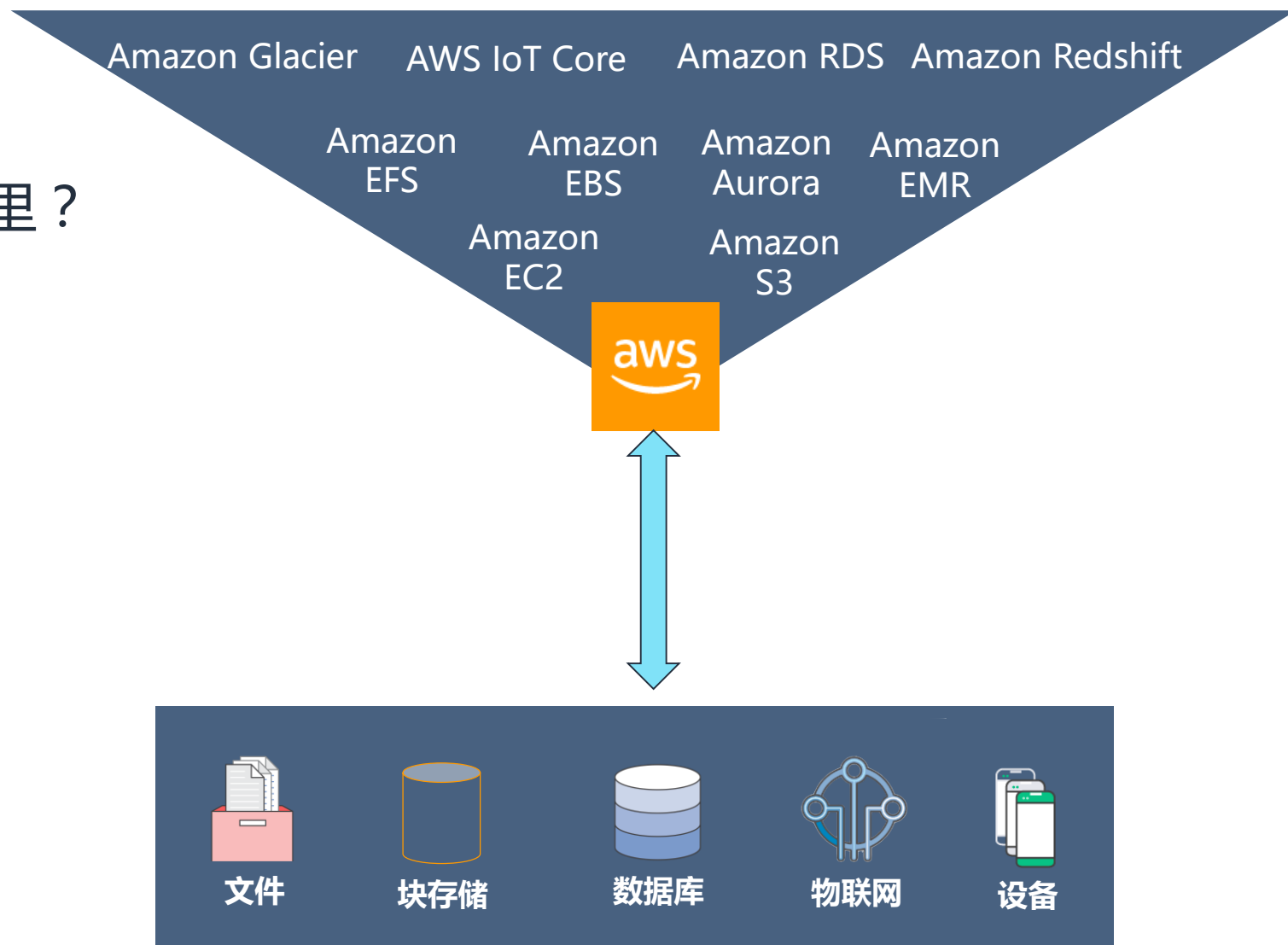
议程

- 关于数据库迁移
- 数据库迁移的工具
- AWS 迁移框架介绍
- 数据库迁移过程
- AWS DMS 迁移案例
- 参考资源

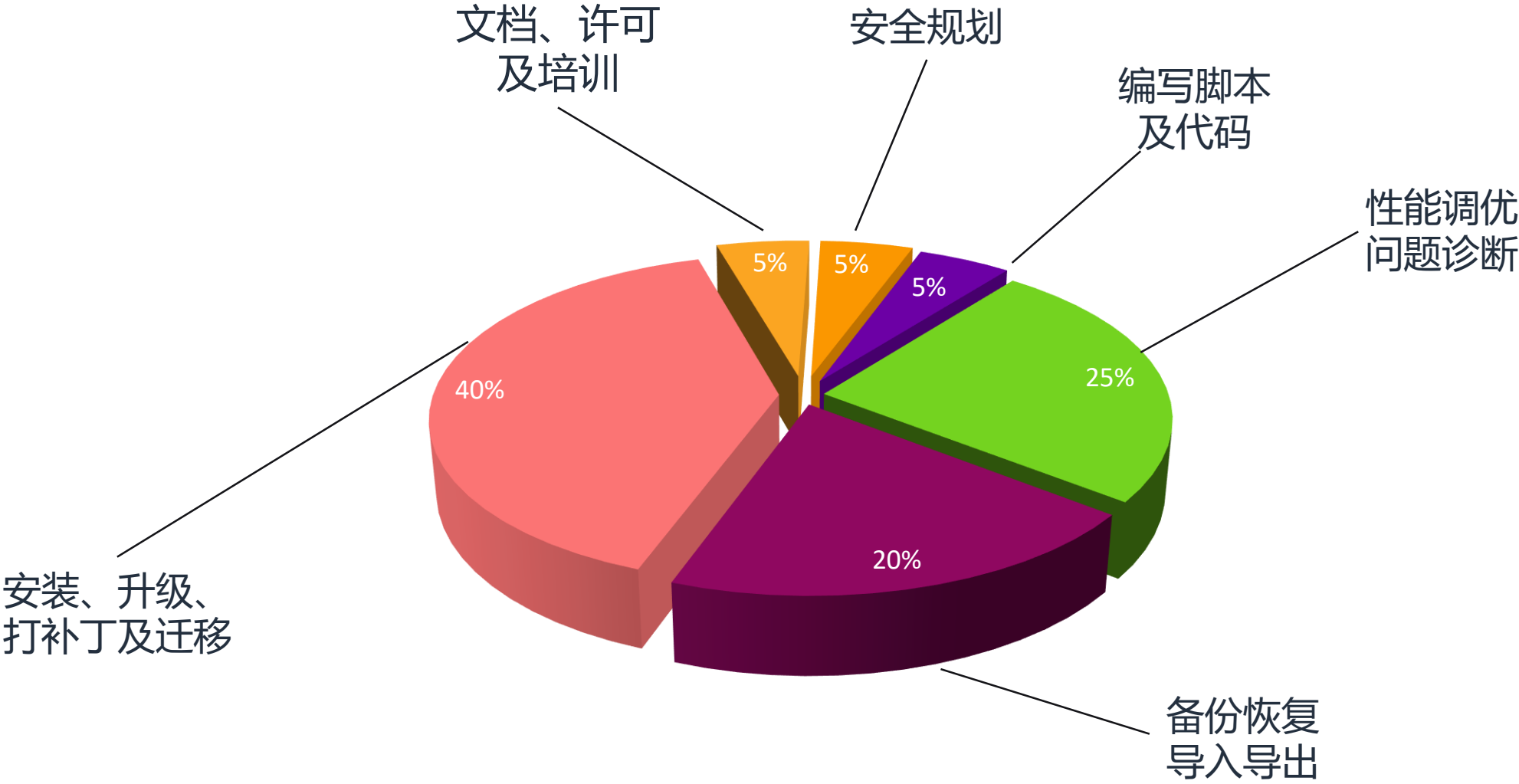
关于数据库迁移

数据迁移的五个关键问题

- ① 需要迁移的数据类型？迁往哪里？
- ② 一次性或持续数据迁移？
- ③ 单向或双向访问？
- ④ 迁移的数据量和时间？
- ⑤ 到云端网络的连通性及带宽？



传统数据库管理工作分类



Amazon RDS 关系型数据库服务

全面支持开源及商业数据库引擎

Amazon
Aurora

MySQL

PostgreSQL

MariaDB

Microsoft
SQL Server

ORACLE

轻松管理



利用 AWS 管理控制台、命令行界面或者简单的 API 调用，即可在几分钟之内就绪。无需预置基础设施、也不需要安装和维护数据库软件。

高可靠性和持久性



Amazon RDS 具有众多其他功能可以增强关键生产数据库的可靠性，包括多AZ、自动备份、快照以及自动主机更换。

高度可扩展



可在线扩展数据库的计算和存储资源，多数引擎都允许启动一个或多个只读副本以从主数据库分流读取流量。

快速且安全



针对高性能 OLTP 及通用型存储优化的SSD选项。轻松控制对数据库的网络访问，并提供了静态加密和传输时加密。

Amazon Aurora 数据库

兼容 MySQL 和 PostgreSQL 的关系数据库，为云打造。
性能和可用性与商用数据库相当，成本只有 1/10。

高性能和高可扩展性



5 倍于标准 MySQL 的吞吐量
3 倍于 PostgreSQL 的吞吐量
性能相当而成本仅为商用DB的1/10
可以跨3个AZ,至多 15 个可读副本
存储自增长，单实例可达 64TB

高可用性和持久性



可用性高于 99.99%
具有容错及自我修复能力
跨3个AZ复制6个数据副本
数据持续备份到 S3
实例故障转移小于30 秒

高度安全



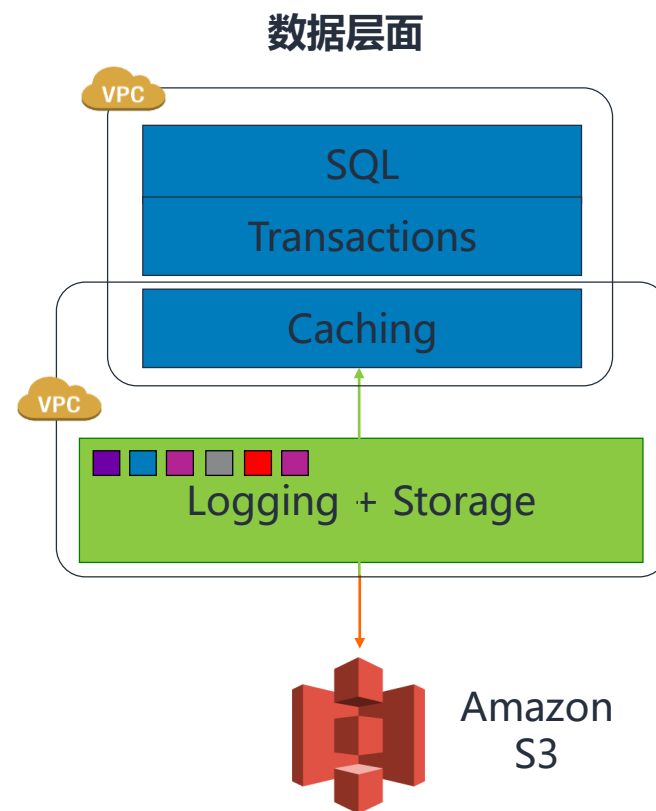
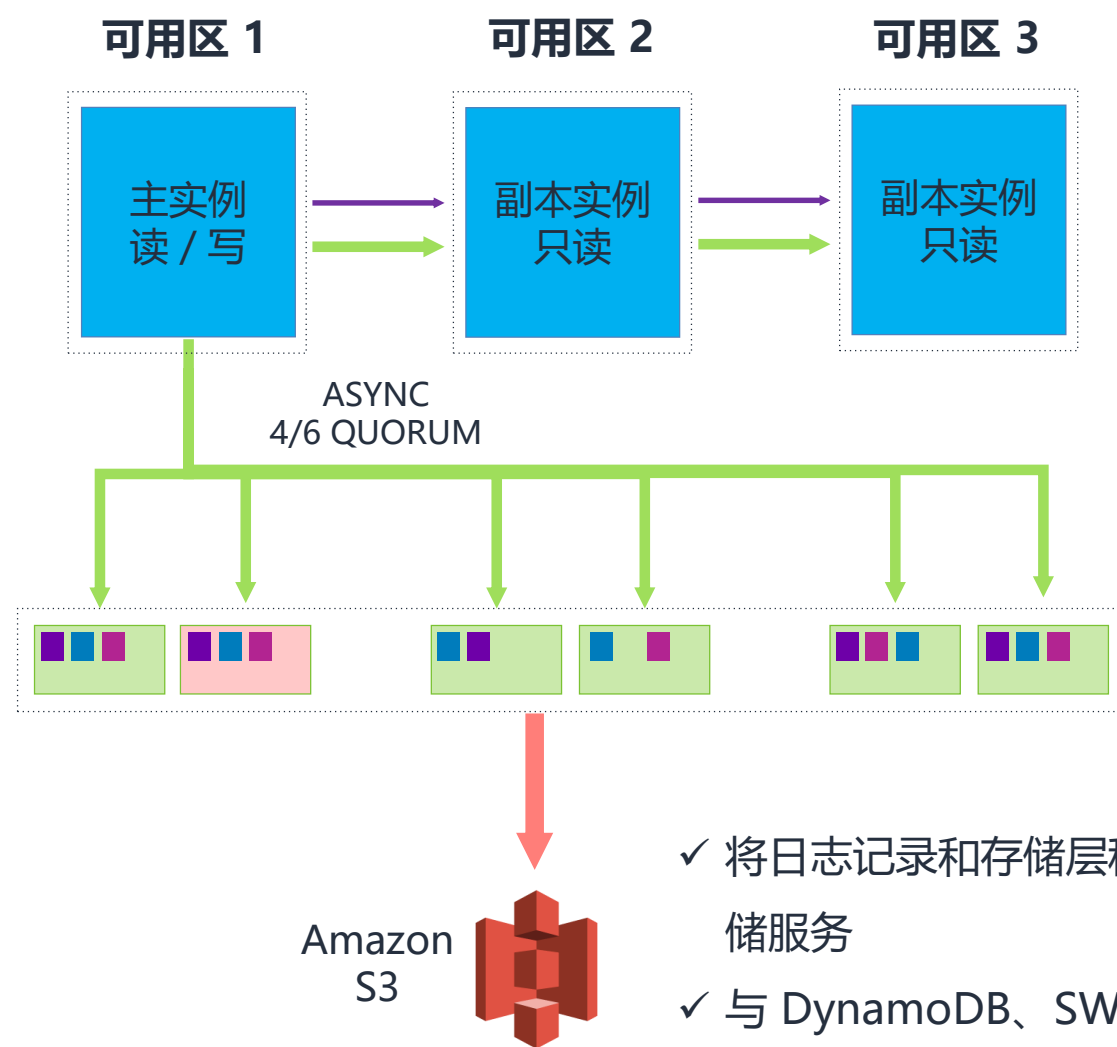
通过VPC 进行网络级
隔离，支持静态存储
及传输时加密，集群
中的备份、快照和副
本自动加密

完全托管



无需担心硬件、软件补
丁、设置、配置或备份等
数据库管理任务。会自动
持续监控并将其备份到
S3，可以实现精细的时间
点恢复。

Amazon Aurora 数据库的高可用及高扩展能力



控制层面



Amazon DynamoDB



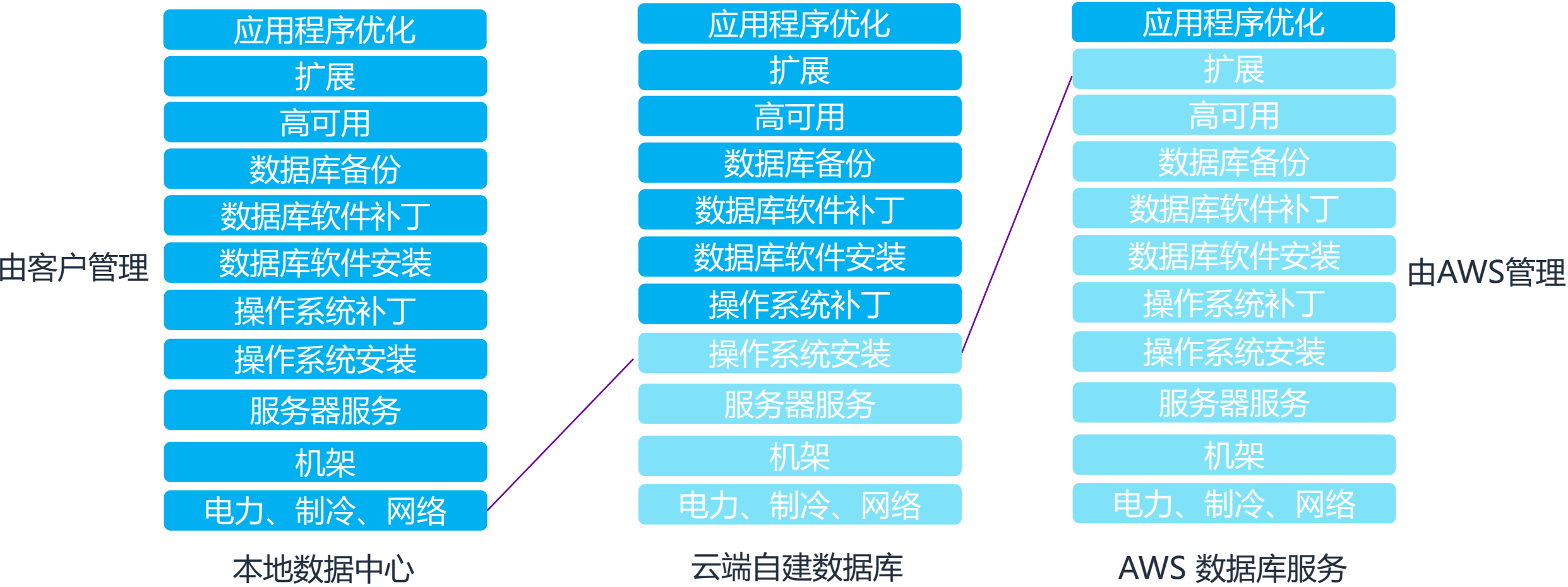
Amazon SWF



Amazon Route 53

- ✓ 将日志记录和存储层移入多租户、横向扩展、为数据库优化的存储服务
- ✓ 与 DynamoDB、SWF、Route 53 等其他 AWS 服务集成，用于控制层面的操作
- ✓ 持续备份与S3集成，并具有 99.999999999% 的持久性

完全托管服务 – 自动执行管理任务



数据库迁移面临的挑战

✖ 需要商业的复制软件及迁移工具 <

✖ 数据库迁移导致的应用停机时间长 6

✖ 新环境下复杂的配置与管理 ⌚

✖ 数据库迁移导致的应用程序代码修改 1001001
0101001
1000100

数据库迁移服务及工具

AWS Database Migration Service 数据库迁移服务

1. 非关系型数据库



在停机时间尽可能短的状态下将您的数据库迁移至 AWS。

已有 90000 多个数据库使用 AWS DMS 完成迁移。

2. 关系型数据库

ORACLE

Microsoft®
SQL Server®



PostgreSQL

MySQL®

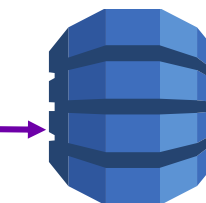
3. 其他数据源



Amazon S3



AWS DMS

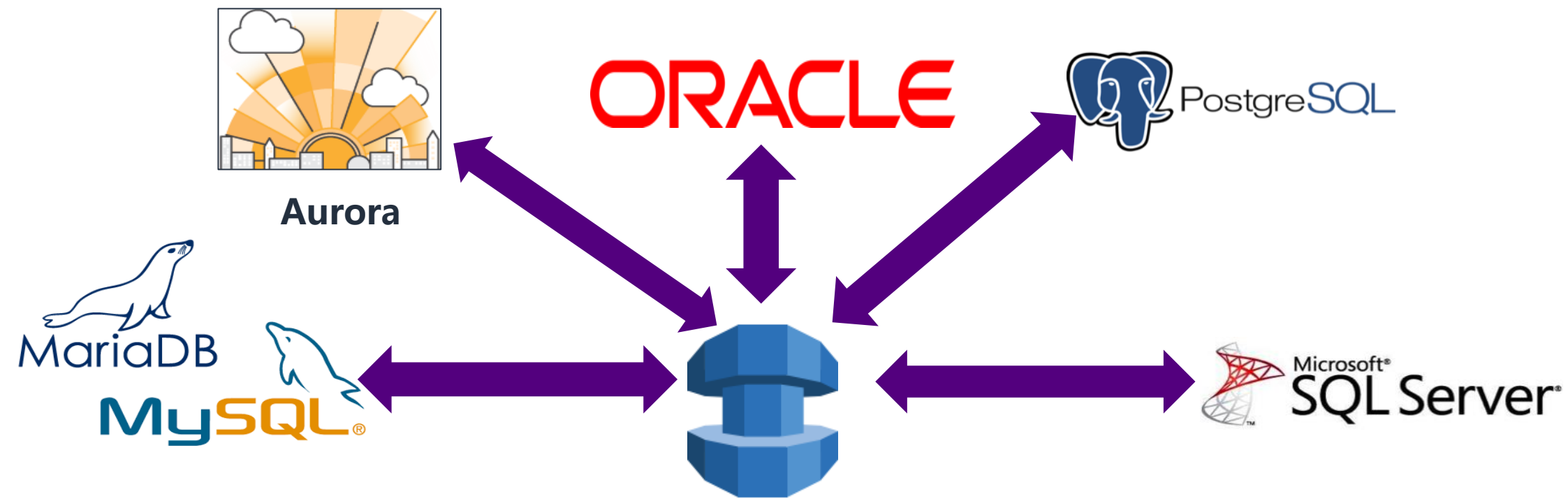


DynamoDB
RedShift
Aurora

- ① 简单易用
- ② 最少停机时间
- ③ 持续数据复制
- ④ 多数据源支持
- ⑤ 运行可靠

<https://aws.amazon.com/dms/free-dms/>

AWS DMS 迁移与复制支持的数据库引擎



支持的终端节点	本地数据中心	Amazon RDS	Amazon EC2
本地数据中心	不支持	支持	支持
Amazon RDS	支持	支持	支持
Amazon EC2	支持	支持	支持

AWS Schema Conversion Tool 数据模型转换工具

在从源数据库迁移到目标数据库的过程中，AWS SCT 能够对很多数据库 Schema 和代码的转换任务进行自动化处理。



ORACLE

Microsoft®
SQL Server®

TERADATA®

ORACLE

PostgreSQL

MariaDB

MySQL®

NETEZZA

Greenplum

Microsoft®
SQL Server®

VERTICA



MariaDB



Amazon
Aurora

MySQL®

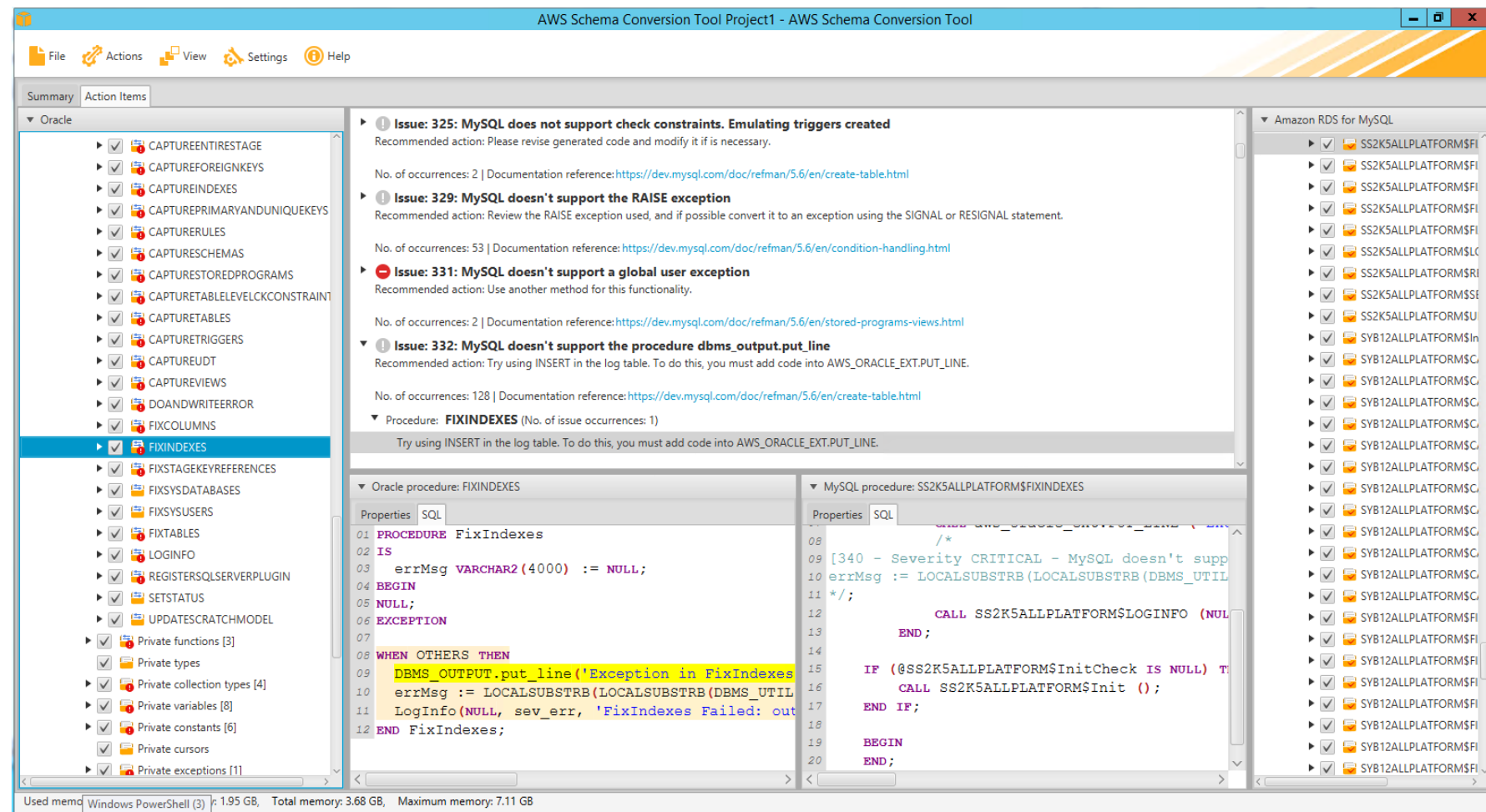
PostgreSQL



转换表结构、视图以及存储过程



Amazon Aurora



- ① 序列号
- ② 用户定义数据类型
- ③ 同名词
- ④ 包
- ⑤ 存储过程
- ⑥ 函数
- ⑦ 触发器
- ⑧ Schemas
- ⑨ 表
- ⑩ 索引
- ⑪ 视图

对数据库迁移进行评估

Database Migration Assessment Report

Source Database: RDS_ADMINISTRATION rds_administration@ec2-54-172-36-60.compute-1.amazonaws.com:81
92.0RCL
Oracle Database 12c Enterprise Edition 12.1.0.1.0 (64bit Production)



Executive Summary

We completed the analysis of your Oracle source database and estimate that 91% of the database storage objects and 100% of database code objects can be converted automatically or with minimal changes if you select Amazon Aurora as your migration target. Database storage objects include schemas, tables, columns, constraints, indexes, sequences, synonyms, user define types and types. Database code objects include functions, procedures, packages, triggers, views, materialized views, events, SQL scalar functions, SQL inline functions, SQL table functions, attributes, variables, constants, table types, public types, private types, cursors, exceptions, parameters and other objects. Based on our analysis of SQL syntax elements of your source database schema, we estimate that 99.9% of your entire database schema can be converted automatically to Amazon Aurora. To complete the migration, we recommend 597 conversion action(s) ranging from simple tasks to medium-complexity actions to significant conversion actions.

Database Objects with Conversion Actions for Amazon Aurora

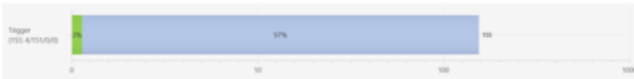
Of the total 1,576 database storage object(s) and 155 database code object(s) in the source database, we were able to identify 1,427 (91%) database storage object(s) and 155 (100%) database code objects that can be converted automatically or with minimal changes to Amazon Aurora.

149 (9%) database storage object(s) required 149 significant user action(s) to complete the conversion.

Figure: Conversion statistics for database storage objects



Figure: Conversion statistics for database code objects



Detailed Recommendations for Amazon Aurora Migrations

If you choose to migrate your Oracle database to Amazon Aurora, we recommend the following actions.

在 AWS SCT 里连接到源
数据库和目标数据库

运行评估报告

阅读简介部分

遵循后续详细的指导步骤

Database Migration Assessment Report

Source Database: RDS_ADMINISTRATION rds_administration@ec2-54-172-36-60.compute-1.amazonaws.com:81
92.0RCL
Oracle Database 12c Enterprise Edition 12.1.0.1.0 (64bit Production)



Storage Object Actions

Sequence Changes

Some changes are required to sequences that cannot be converted automatically. You'll need to address these issues manually.

Issue 341: MySQL doesn't support sequences

Recommended Action: Try developing a system for sequences in your application.

Issue Code: 341 | No. of Occurrences: 134 | Estimated Complexity: Significant

Schemas.RDS_ADMINISTRATION.Sequences.BACKUP_ID_SEQUENCE
Schemas.RDS_ADMINISTRATION.Sequences.CERTIFICATE_ID_SEQUENCE
Schemas.RDS_ADMINISTRATION.Sequences.CHARACTER_SET_ID_SEQ
Schemas.RDS_ADMINISTRATION.Sequences.CUSTOMER_SUBNET_GROUP_ID_SEQ
Schemas.RDS_ADMINISTRATION.Sequences.CUSTOMER_SUBNET_ID_SEQ
+129 more

Index Changes

Some changes are required to indexes that cannot be converted automatically. You'll need to address these issues manually.

Issue 207: MySQL doesn't support function indexes

Recommended Action: Revise your code and try to use simple index.

Issue Code: 207 | No. of Occurrences: 3 | Estimated Complexity: Significant

Documentation References: <https://dev.mysql.com/doc/refman/5.6/en/create-table.html>

Schemas.RDS_ADMINISTRATION.Tables.DBI_ENGINE_SEEDS.Indexes.I_DBI_ENG_SEED_DBI_ENG_CONF_ID
Schemas.RDS_ADMINISTRATION.Tables.RDS_SYSTEM_ACCOUNTS.Indexes.I_SYS_ACCOUNT_DEFAULT
Schemas.RDS_ADMINISTRATION.Tables.RUNNABLE_DBI_CONFIG.Indexes.U_RNBL_DBI_CFG_PREFERRED

Constraint Changes

Some changes are required to constraints that cannot be converted automatically. You'll need to address these issues manually.

Issue 210: MySQL doesn't support FUNCTION AS DEFAULT VALUE

Recommended Action: Try using a trigger.

Issue Code: 210 | No. of Occurrences: 2 | Estimated Complexity: Simple

Documentation References: <https://dev.mysql.com/doc/refman/5.6/en/create-table.html>

Schemas.RDS_ADMINISTRATION.Tables.CUSTOMERS.Constraints.CK_CUSTOMER_TRUST_LEVEL_STATE: 0:10
Schemas.RDS_ADMINISTRATION.Tables.STORAGE_VOLUMES.Constraints.CK_SV_LIFECYCLE: 0:8

Issue 325: MySQL does not support check constraints. Emulating triggers created

Recommended Action: Please revise generated code and modify it if necessary.

Issue Code: 325 | No. of Occurrences: 283 | Estimated Complexity: Simple

Documentation References: <https://dev.mysql.com/doc/refman/5.6/en/create-table.html>

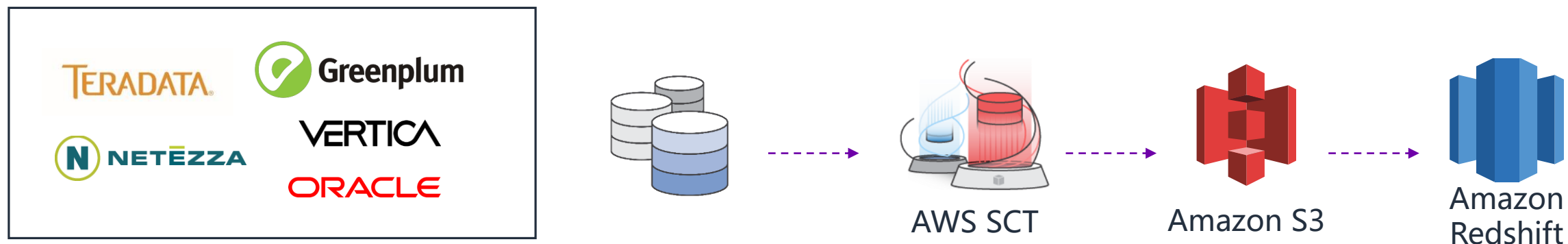
AWS SCT 的数据抽取代理

从你的数据仓库里抽取数据并迁移到 Amazon Redshift

通过本地代理进程抽取数据

抽取出来的数据已经针对 Amazon Redshift 进行了优化，并保存到本地文件

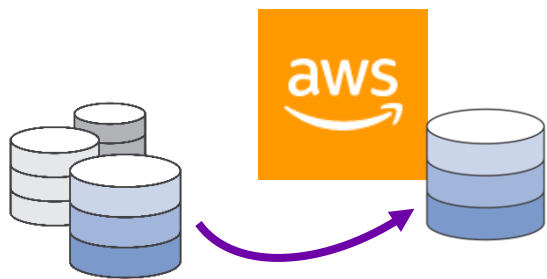
文件会通过网络或者 AWS Snowball 传输到 Amazon S3 的桶里，然后加载到 Amazon Redshift



何时使用数据库迁移服务?

何时使用 AWS DMS 数据库迁移服务?

迁移



- ① 迁移关键业务系统
- ② 从 Classic 网络迁移到 Amazon VPC
- ③ 迁移数据仓库到 Amazon Redshift
- ④ 归档旧数据
- ⑤ 升级小版本
- ⑥ 合并多个数据分片到 Amazon Aurora
- ⑦ 复制数据从而在云端分析数据
- ⑧ 从 NoSQL 迁移到 SQL，或者从 SQL 迁移到 NoSQL，或者从 NoSQL 迁移到 NoSQL

源端

ORACLE

Microsoft SQL Server

MySQL

mongoDB

PostgreSQL

Amazon S3

MariaDB

SQL

目标端

Amazon S3

Microsoft SQL Server

MariaDB

Amazon Redshift

MySQL

PostgreSQL

Amazon Aurora

Amazon Dynamo DB

ORACLE

项目迁移阶段的服务及工具

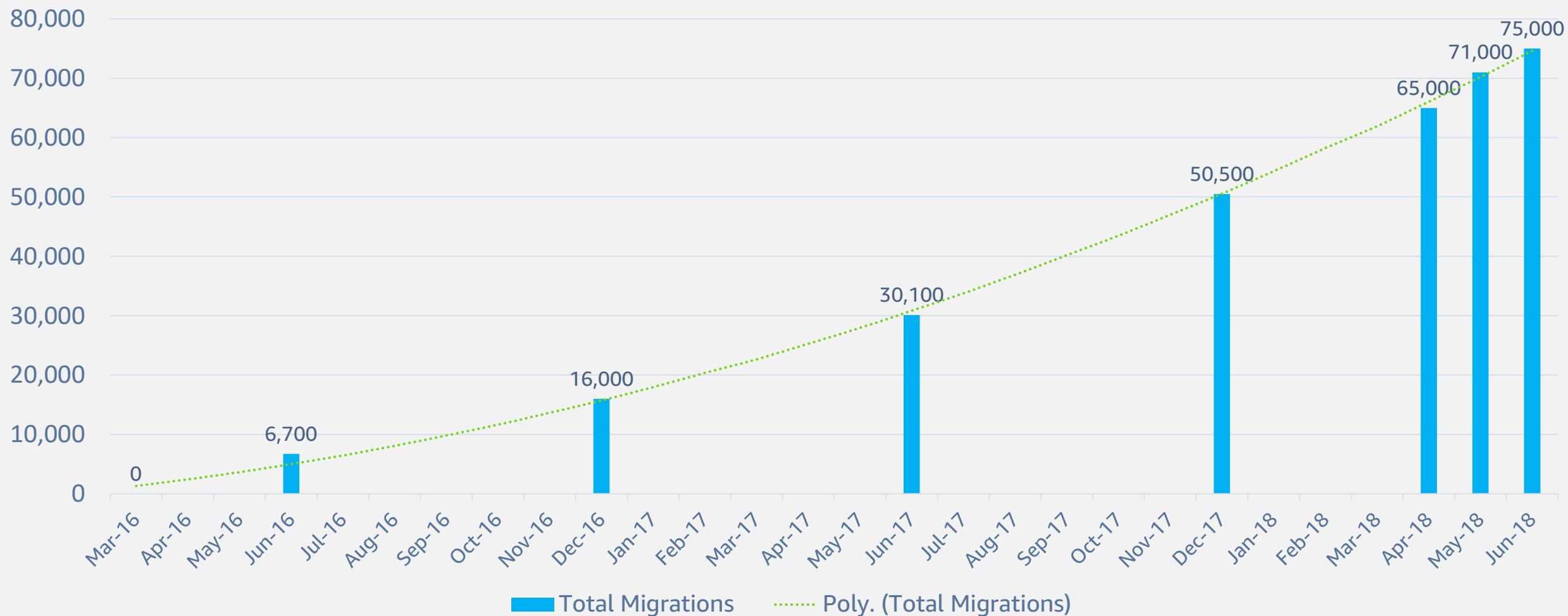
阶段	服务 / 工具	备注
评估	AWS SCT	关于数据库对象、复杂性和迁移问题的报告
模型迁移	AWS SCT	复制模型或迁移模型，具体取决于它是同构迁移，还是异构迁移
数据迁移	AWS DMS AWS SCT	批量加载和更改数据捕获(CDC)选项 大型数据仓库(包括 AWS Snowball)的提取和加载、集成
应用迁移	AWS SCT	应用程序代码中的 SQL 语句迁移
数据验证	AWS DMS	确保源和目标上的数据相同
功能测试	第三方工具	确保应用程序按预期运行
性能测试	第三方工具	确保应用程序性能按预期运行

迁移服务的使用场景

场景	例子	推荐
同构迁移到相同的数据库版本	Migration of Oracle Database 11gR2 Enterprise Edition from on-premise to EC2	使用数据库原生复制技术创建备用数据库，然后故障转移到备用数据库
同构迁移到不同的数据库版本	Migration of MySQL 5.5 to MySQL 5.7	AWS SCT AWS DMS
同构迁移到不同的数据库版本	Migration of SQL Server Enterprise Edition to Standard Edition	AWS SCT AWS DMS
异构迁移	Migration from Oracle Database to PostgreSQL	AWS SCT AWS DMS

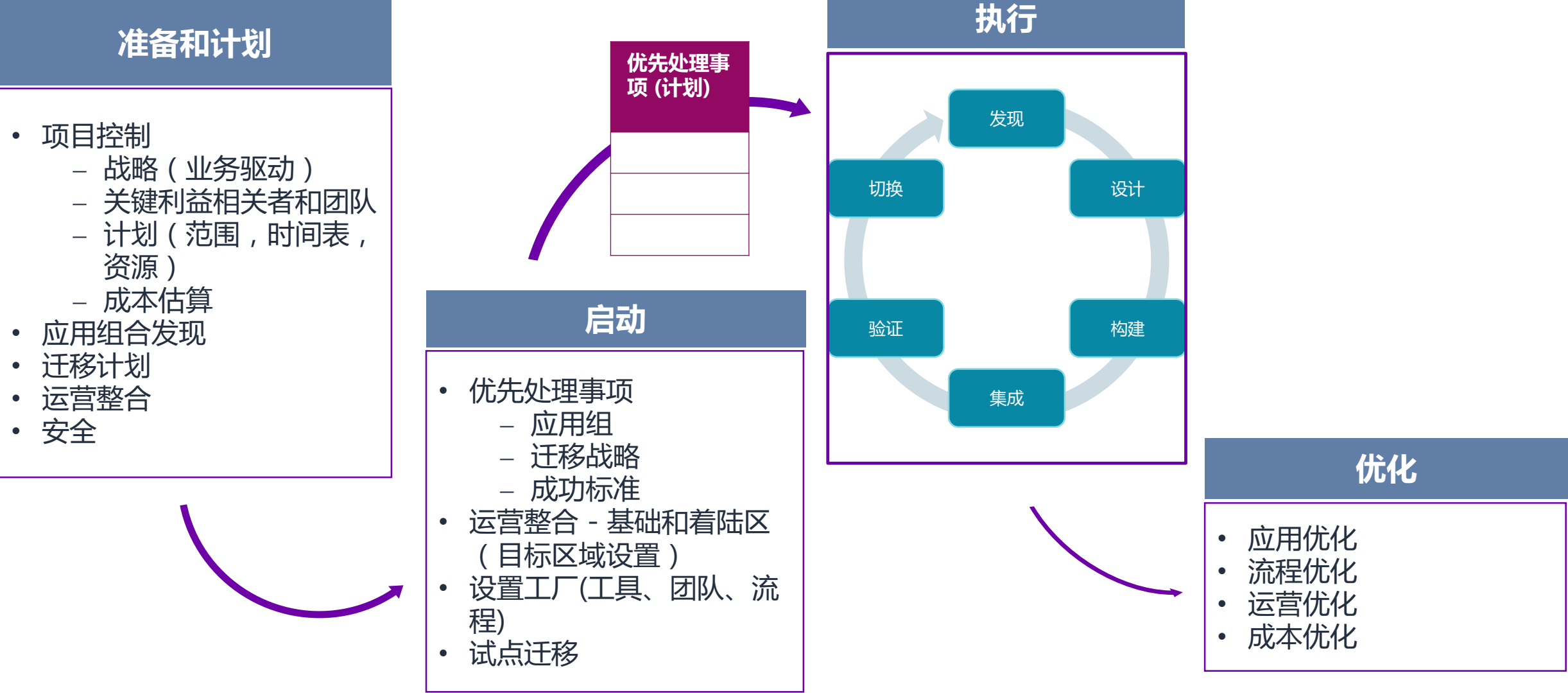
AWS DMS 数据库迁移服务的采用趋势

Total Database Migrations by DMS



AWS 迁移框架介绍

AWS 迁移框架



AWS 迁移框架 - 准备和计划

准备和计划

- 项目控制
 - 战略（业务驱动）
 - 关键利益相关者和团队
 - 计划（范围，时间表，资源）
 - 成本估算
- 应用组合发现
- 迁移计划
- 运营整合
- 安全

项目控制 侧重于确保有一个得到组织中 **关键利益相关方** 支持的迁移战略。此外，我们还研究如何界定将开展此项工作的团队，以及相关的 **时限** 和 **成本估算**。

决策点参考示例：

- ① 谁是执行发起人？
- ② 是否有任何引人注目的事件会影响迁移战略？
- ③ 我们有合适的资源吗？他们是如何组织的？
- ④ 我们正在使用的时间范围是什么？
- ⑤ 我们有必要的预算吗？

AWS 迁移框架 - 准备和计划

快速 了解哪些应用程序是符合 **适用云**、**云友好** 或 **云原生**，然后仅对这部分应用程序子集执行深入分析。

准备和计划

- 项目控制
 - 战略（业务驱动）
 - 关键利益相关者和团队
 - 计划（范围，时间表，资源）
- 成本估算
- **应用组合发现**
- 迁移计划
- 运营整合
- 安全



专有硬件设施
非 x86 环境的工作负载
不适合云的软件许可方式
特定位置(例如, 工厂车间)或本地低延迟的限制



已虚拟化的 x86
明确定义应用程序边界
适合云的软件许可方式
了解应用程序和数据的相互依赖关系



可水平扩展
利用服务 (例如 SOA、API等)
应用程序设计的“12 要素”
软件供应商提供的云映像 (AMI、部署模板)



微服务架构
API 优先的设计
内置容错功能
捆绑的监控指标和主动的故障测试

* 云友好 包含 适用云 的特征
** 云原生/SaaS 包含 云友好 与 适用云 的特征

应用评估

- 业务驱动力和预期的投资回报率？
- 迁移的发起者 (业务线, C-level) ？
- ISV 应用程序？ ISV 是否支持目前环境？
- 迁移的维护时间窗口？
- 设计文档？
- 原开发者、DBA 仍然可找得到？

数据库评估

- 多少数据库对象（表、视图、触发器、存储过程等）？
- 多大的数据量？
- 存储过程与触发器的复杂性？
- 专有的数据库特性？
- 非标准或用户自定义的数据类型？
- 数据库字符集转换（UTF8、GBK）？
- 时区或 UTC？
- 用户的认证方式？
- 软件许可方式（CPU核、用户数、ULA等）？

应用程序技术评估

- 数据库的访问方式:
 - 程序代码中的 SQL 语句？
 - 调用数据抽象层？
 - API 调用？
- 尽可能使用标准 SQL 语句？
- SQL 语句的复杂性, 例如：很多表关联的分析或简单的增、删、改、查询？
- SQL 代码行数？
- 应用访问方式, 例如：LDAP、数据库用户等

AWS 迁移框架 – 启动

启动

- 优先处理事项
 - 应用组
 - 迁移战略
 - 成功标准
- 运营整合 - 基础和着陆区
(目标区域设置)
- 设置工厂(工具、团队、流程)
- 试点迁移

- 确定应用程序的优先级和集成应用程序分组；
- 概述每个应用程序迁移的成功标准；
- 创建 AWS 着陆区（账户、VPC、子网、IAM 角色、VPN / 专线等）；
- 配置 AWS DMS、AWS SCT 和其他迁移工具；
- 成立迁移团队；
- PoC / 试点应用迁移；

建立迁移团队

- **应用程序架构师/开发人员**：应用程序专家,可以识别哪些组件重要、复杂、冗余等；
- **源数据库 DBA**：了解数据库设计、架构、使用的功能以及哪些必须迁移到目标；
- **目标数据库 DBA**：目标数据库专家, 帮助完成源到目标库的功能映射；
- **AWS 解决方案架构师**：确定AWS中的正确目标架构, 并熟悉 DMS/SCT；
- **应用程序/数据库开发人员**：用于迁移存储过程、触发器和应用程序代码的客户和/或APN合作伙伴资源；

招募和培养人才

- 如果从本地数据中心迁移到云端，目标库与 AWS 服务通常需要 **新技能**；
- 为现有员工制定 **培训计划**；
- 如有需要，**招募** 相应的技术人员；
- 对技能不再相关人员进行**再培训**、**转岗** 或者 **不再雇佣**；

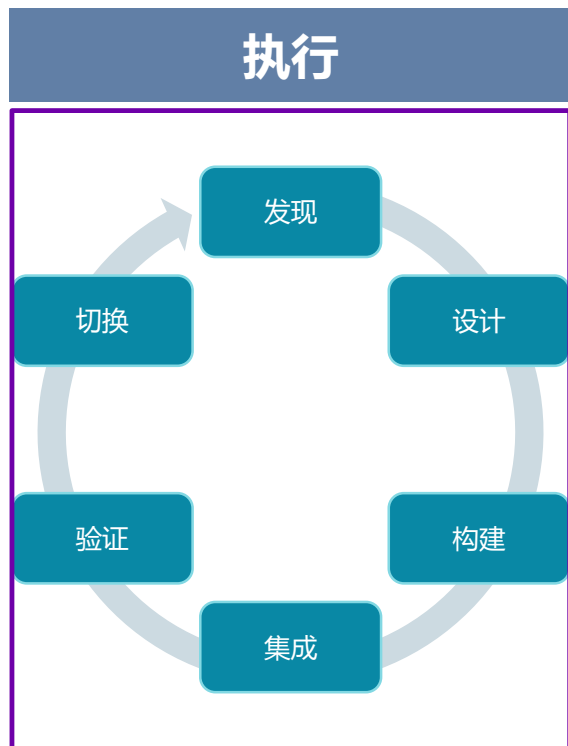
PoC/试点应用迁移

在迁移 **启动阶段**，选择 **复杂度合理** 的模块 / 组件，通过迁移来 **验证** 你的设想。

你能够：

- ①获得更精准的迁移评估报告；
- ②确定哪些可以自动化；
- ③了解迁移工具的行为方式（限制、错误、需要改进的地方）；
- ④了解当前的团队缺乏的技能；

AWS 迁移框架 – 执行



- 总是有备用的计划；
- 根据试点应用与 PoC 的经验来执行；
- 通常迁移数据库的时间与迁移应用程序的时间相同
- 确定如何切换：
 - ① 并行运行：昂贵且困难；
 - ② 尽可能短的停机时间：AWS DMS + CDC(增量)；
 - ③ 大型维护窗口：上线前的应用和数据验证；

AWS 迁移框架 – 优化

优化

- 应用优化
- 流程优化
- 运营优化
- 成本优化

- AWS DMS 实例与复制任务优化；
- 数据库性能优化；
- 合适的数据库实例类型；
- 应用程序优化；
- 合适的应用程序实例类型；
- 通过停 / 起 Amazon EC2 、 Amazon RDS 实例优化成本；
- 购买预留实例或竞价实例；
- 定位性能瓶颈，评估使用缓存、NoSQL等方法优化性能的可能性，采用微服务框架；
- 优化 AWS 托管服务的使用以帮助执行高可用 / 容灾方案，例如，Amazon RDS 多可用区部署，自动扩展；

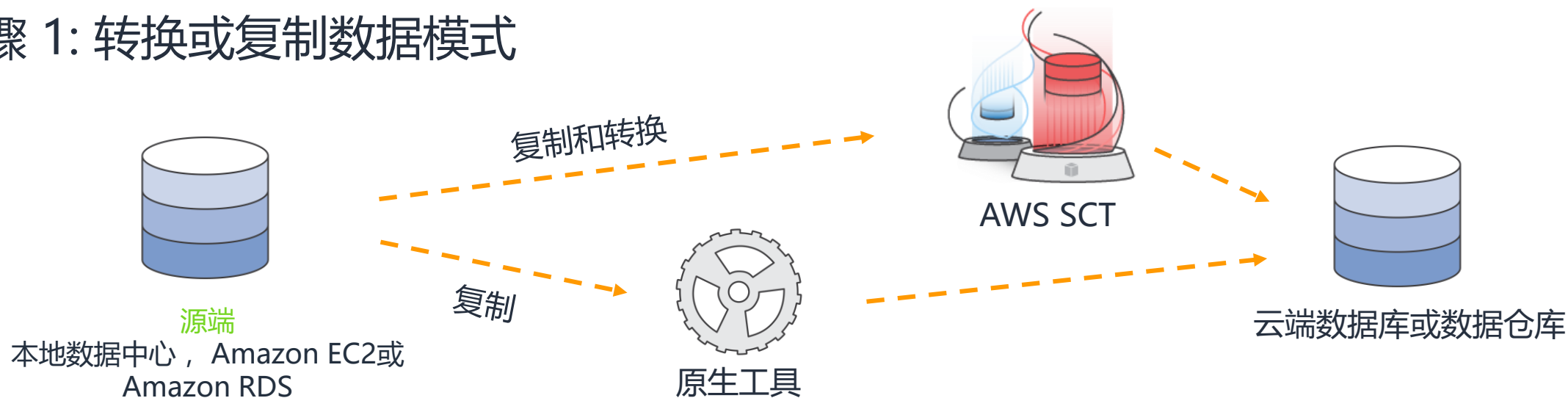
数据库迁移的过程

数据库迁移各阶段工作量评估

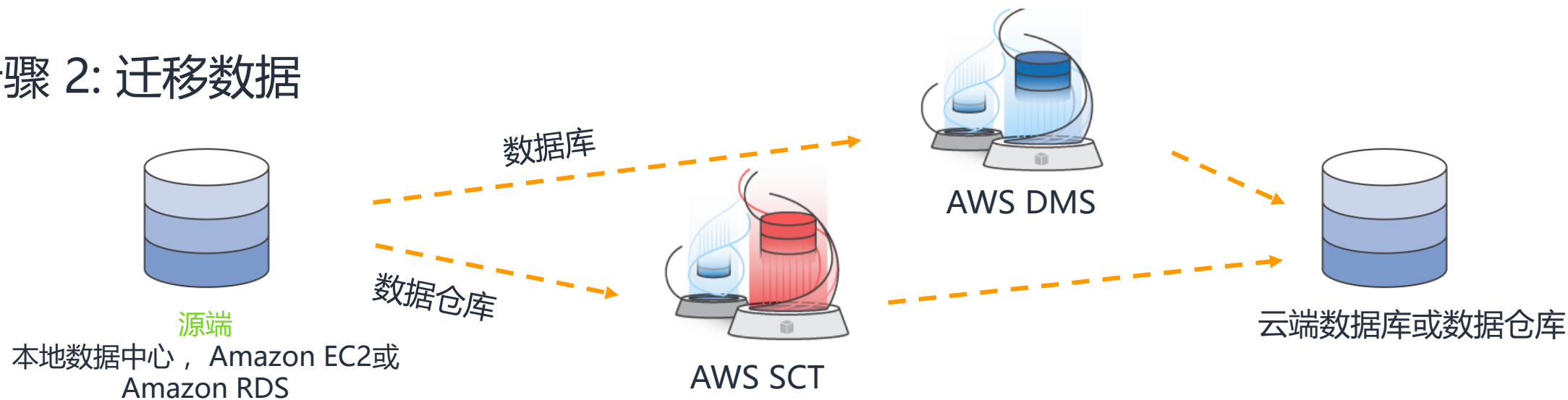
阶段	工作描述	自动化工具	工作量 (%)
1	评估	AWS SCT	2
2	数据模式转型	AWS SCT & AWS DMS	14
3	应用程序转换 / 修复	AWS SCT	25
4	脚本转换	AWS SCT	7
5	与第三方应用程序集成		3
6	数据迁移	AWS SCT & AWS DMS	4
7	全系统的功能测试		29
8	性能调整	AWS SCT	2
9	集成和部署		7
10	培训、知识		2
11	文档和版本控制		2
12	后期生产支持		3

数据库迁移过程

步骤 1: 转换或复制数据模式

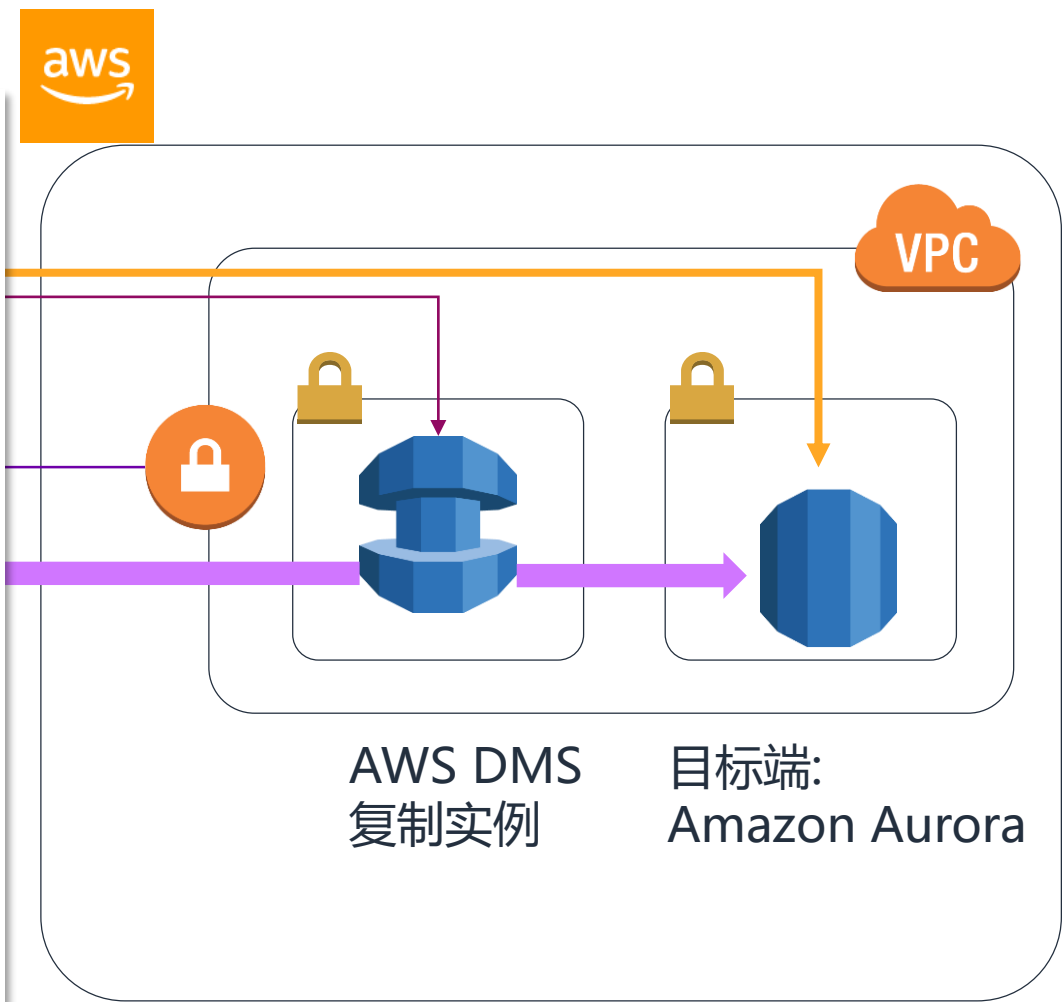
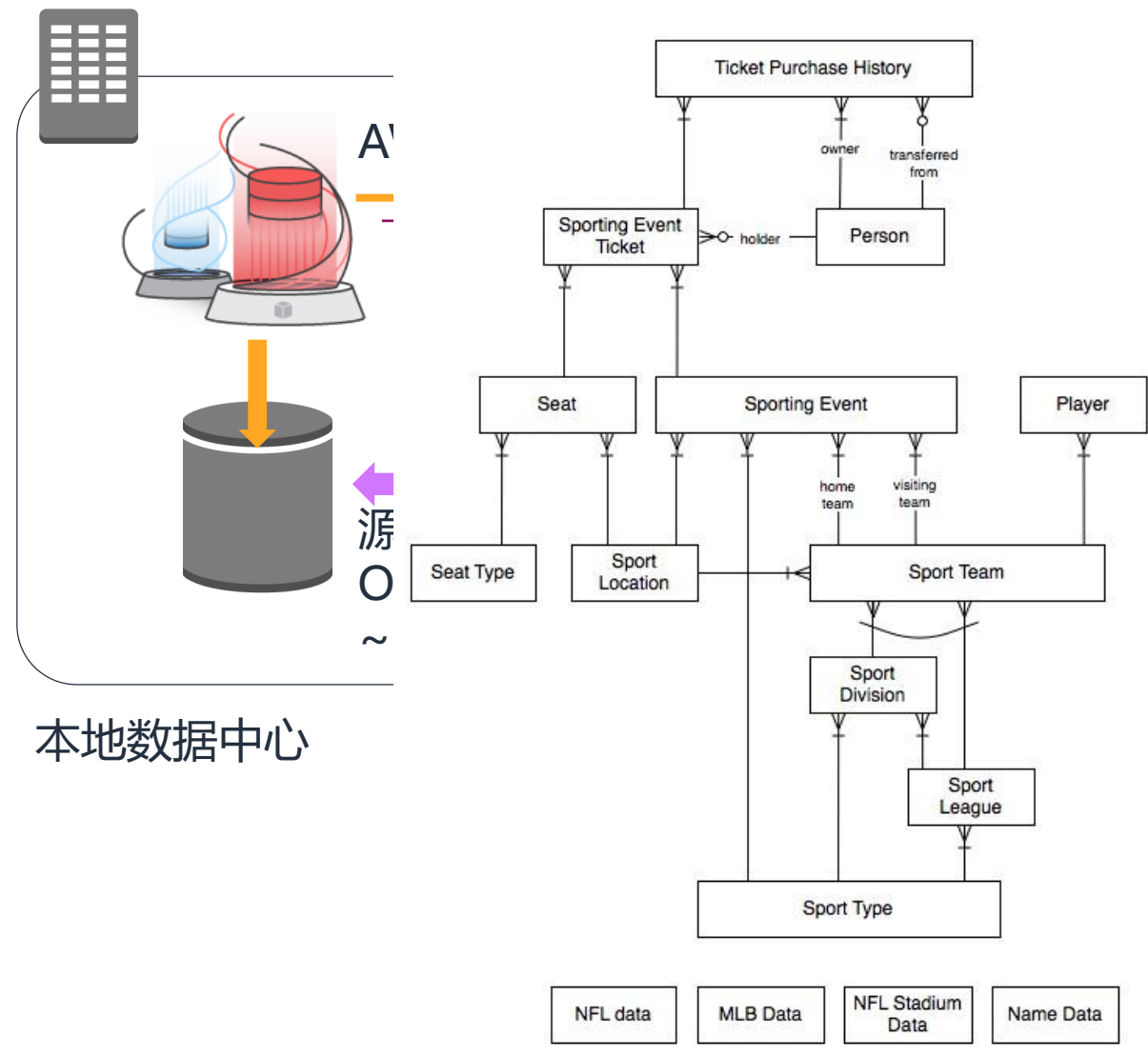


步骤 2: 迁移数据

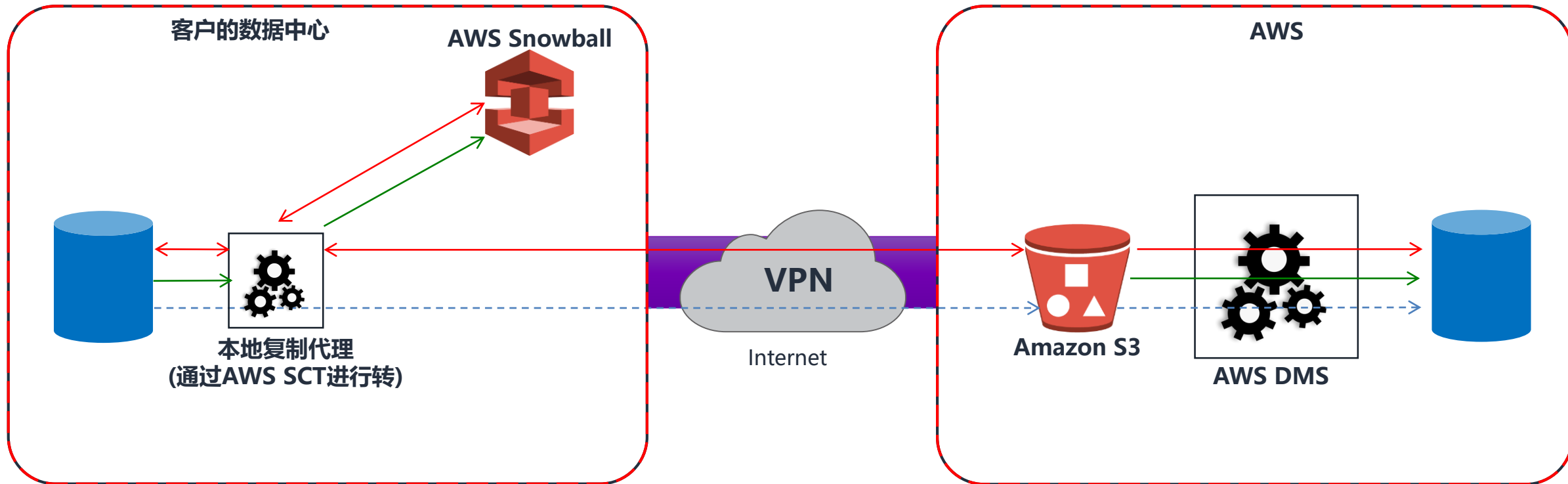


从本地数据中心 Oracle 到 Amazon Aurora 迁移

数据模型



使用 AWS Snowball 迁移数据



AWS DMS 迁移了 90,000+ 数据库到 AWS



技术资源

AWS SCT 详细文档：

http://docs.aws.amazon.com/zh_cn/SchemaConversionTool/latest/userguide/Welcome.html

AWS DMS 服务介绍：

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

AWS DMS 详细文档：

http://docs.aws.amazon.com/zh_cn/dms/latest/userguide/Welcome.html

AWS DMS 分步迁移指南：

http://docs.aws.amazon.com/zh_cn/dms/latest/sbs/DMS-SBS-Welcome.html

常见问题

<https://aws.amazon.com/cn/dms/faqs/>



TRANSFORMATION DAY



宁夏区域
由西云数据运营

北京区域
由光环新网运营

谢谢