



AWS 如何实现数据跨区域同步

宋烨
AWS





宋焯

现任亚马逊AWS解决方案架构师高级经理，帮助中国区客户利用云计算技术进行业务创新和数字化转型，基于AWS先进的云服务平台构建企业核心业务系统，并持续不断地推进架构演进，满足业务成长的需求。

宋焯有近二十年从事研发、技术培训、业务拓展和技术团队管理的经验。他在云计算、大数据、企业架构等方面拥有丰富的经验，参与过金融、运营商、制造业和教育等行业的云平台建设，可以利用这些经验帮助客户采用创新技术，创造业务价值。

宋焯于2017年加入AWS担任解决方案架构师高级经理；曾任思杰系统大中华区技术总监；之前在浙江大学、黎明网络和微软中国等单位任职。

宋焯于浙江大学获学士学位。

目录

- 1 全球架构及状态数据的全球区域复制
- 2 Amazon DynamoDB 及全局表概述
- 3 全局表 – 跨区域多主架构解析
- 4 更多的跨区域数据同步服务展望
- 5 总结

AWS 十大经验教训之一



“Failures are a given and everything will eventually fail over time.”

“故障总是在意料之外，情理之中发生”

Werner Vogels
CTO – Amazon.com

<https://www.allthingsdistributed.com/2016/03/10-lessons-from-10-years-of-aws.html>

业务真的需要多区域部署?

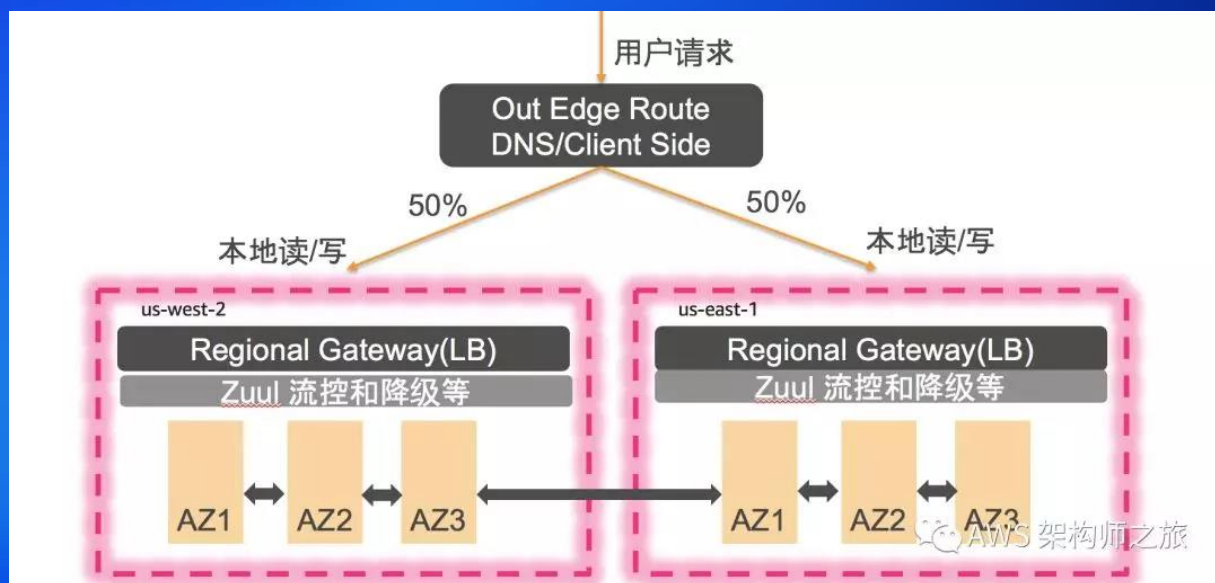


Netflix 的多区域多活架构经验

1 服务必须无状态

2 区域内服务必须访问本区域内的资源

3 用户的访问路径中不能有跨区域的服务调用



状态数据分类分级

交易等事务类

- 订单
- 支付

商品信息

- 商品目录
- 详情信息
- 报价
- 卖家等

事件/消息

- 点击流
- 对象状态变化事件流
- 服务消息

日志类

- 用户行为日志
- 后台服务器日志

低容量，但高一致性



大容量，相对低完整性要求

数据复制的主要方式

最难实现

同步复制，保证数据
跨区域强一致性

事务类

异步实时，连续复制

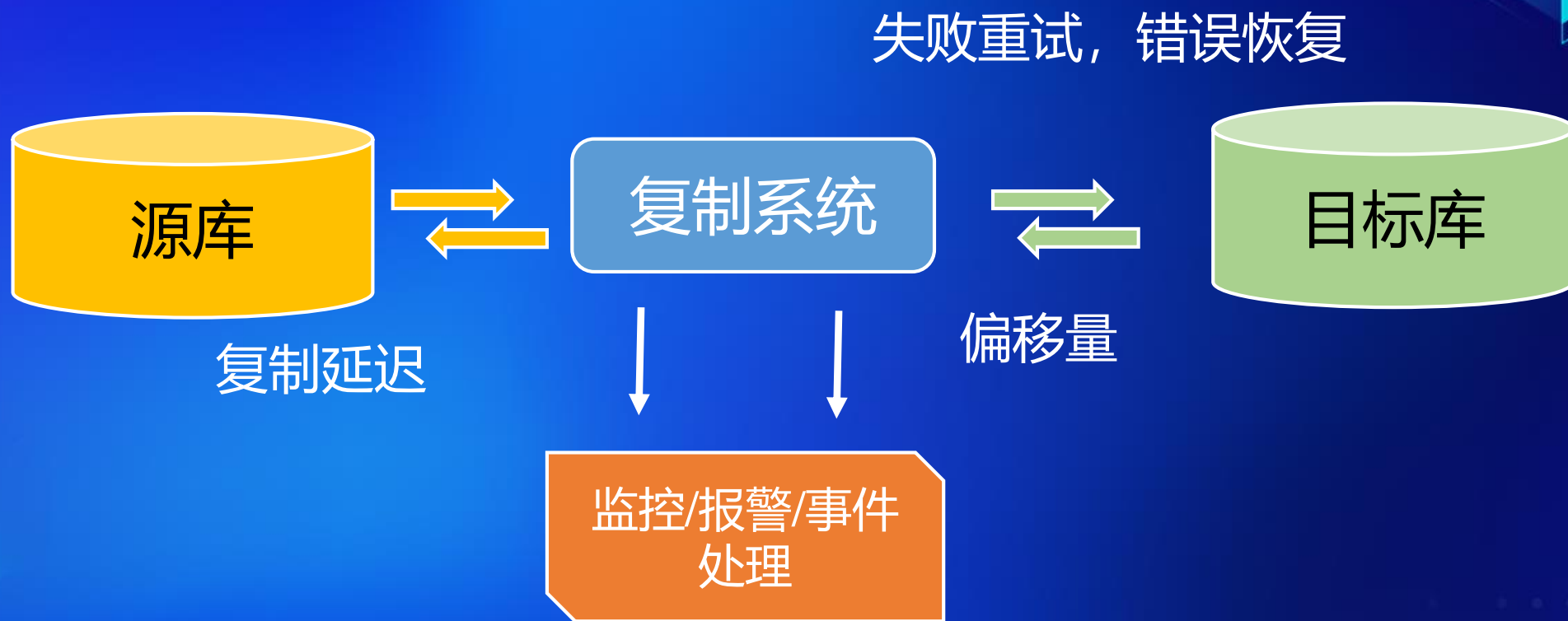
商品信息

异步批量复制

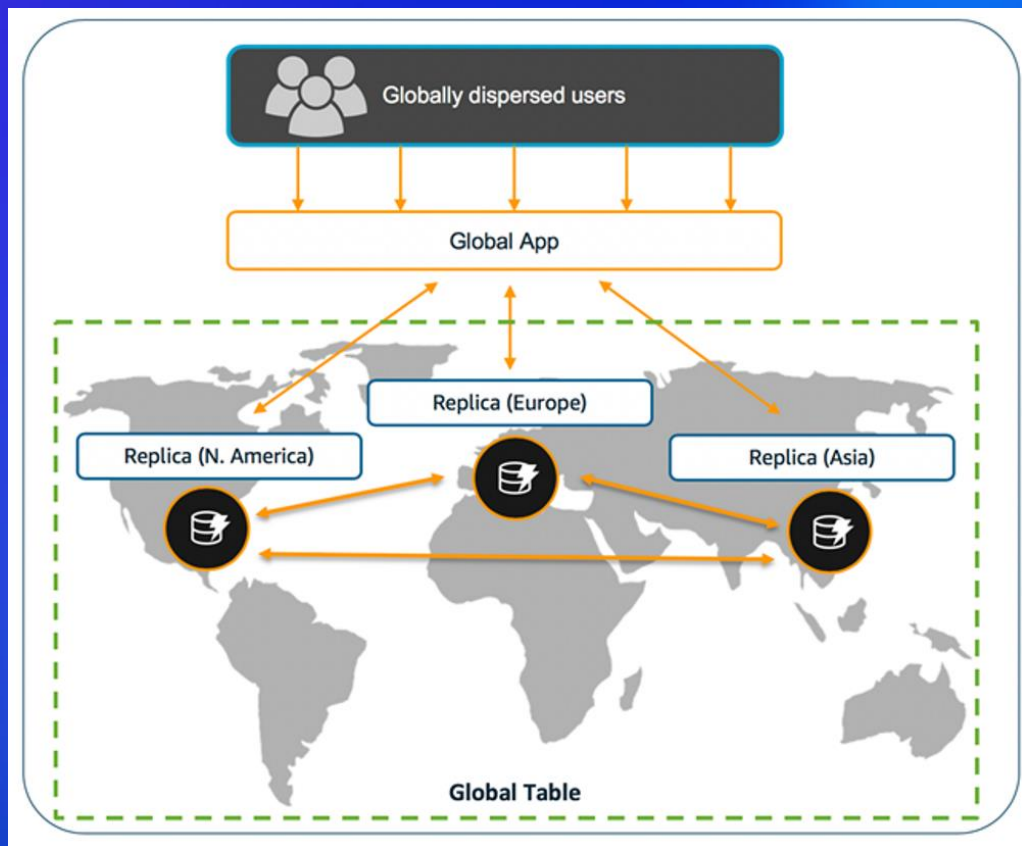
事件/消息
日志

容易管理

理想的数据复制架构



Amazon DynamoDB 及其全局表



- 多区域多活
- 没有额外的成本
- 简单有效的冲突处理
- 基于流的异步复制
- 目前单区域读写支持强一致性
- 目前多区域支持最终一致性
- 增加三个系统字段

QuickTime Player

File Edit View Window Help

DynamoDB · AWS Console

DynamoDB · AWS Console

DynamoDB · AWS Console

+

https://console.aws.amazon.com/dynamodb/home?region=us-east-1#tables:

...

☆

↓

≡

New

🔍

👤

☰

aws

Services ▾

Resource Groups ▾

★

🔔

Admin/zhetang-Isengard @ sh...

N. Virginia ▾

Support ▾

DynamoDB

Dashboard

Tables

Backups

Reserved capacity

Preferences

DAX

Dashboard

Clusters

Subnet groups

Parameter groups

Events

Create table

Delete table

Filter by table name

×

Choose a table group

▾

Actions ▾

?

Viewing 5 of 5 Tables

	Name	Status	Partition key	Sort key	Indexes	Total read capacity	Total write capacity	Auto Scaling	Encryption
<input type="radio"/>	Edu	Active	name (String)	-	0	5	5	-	DEFAULT
<input type="radio"/>	GlobalTableEx	Active	ID (String)	-	0	5	5	-	DEFAULT
<input type="radio"/>	GomokuPlayerInfo	Active	PlayerName (String)	-	0	5	5	-	DEFAULT
<input type="radio"/>	testpython	Active	ID (String)	-	0	5	5	-	DEFAULT
<input type="radio"/>	voteeditor	Active	name (String)	-	0	5	5	-	DEFAULT

Feedback

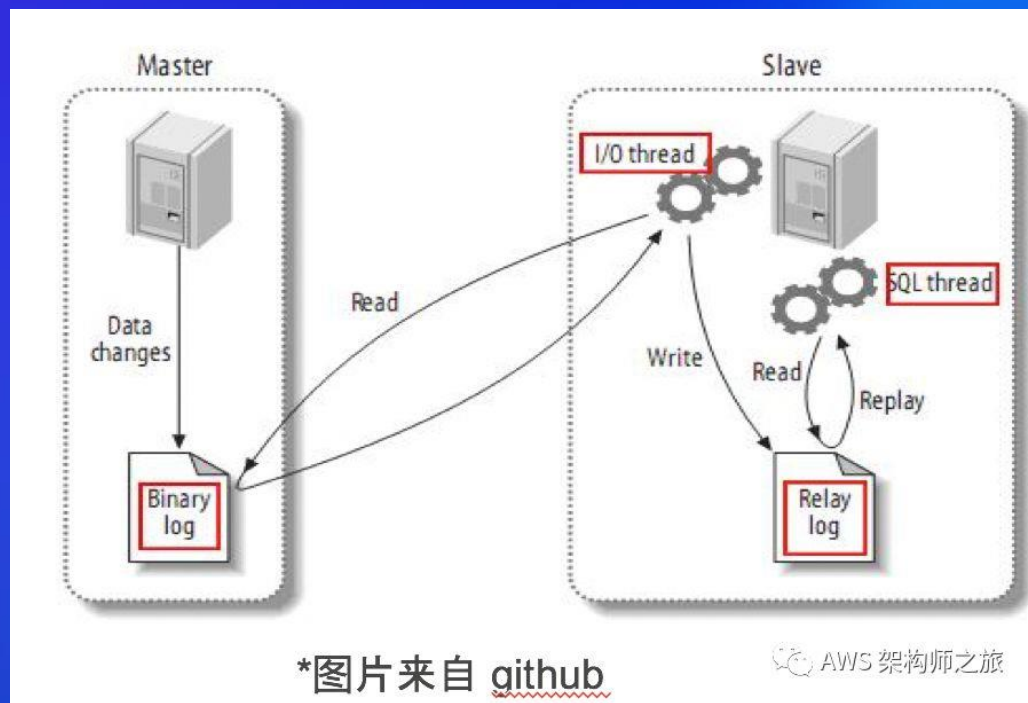
English (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

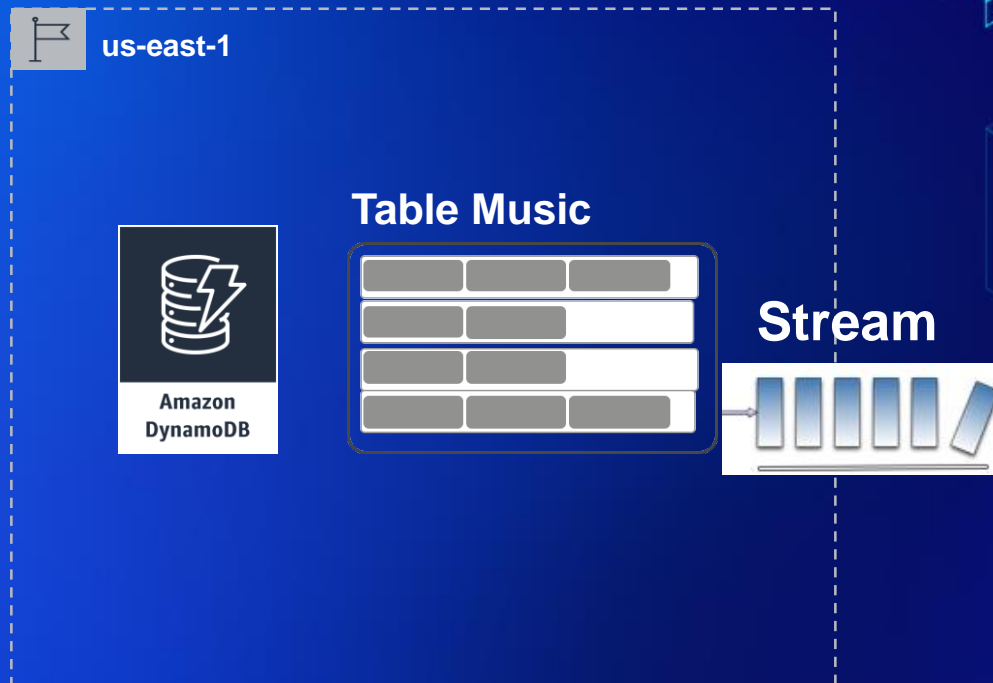
Privacy Policy

Terms of Use

数据更新：流 vs 日志



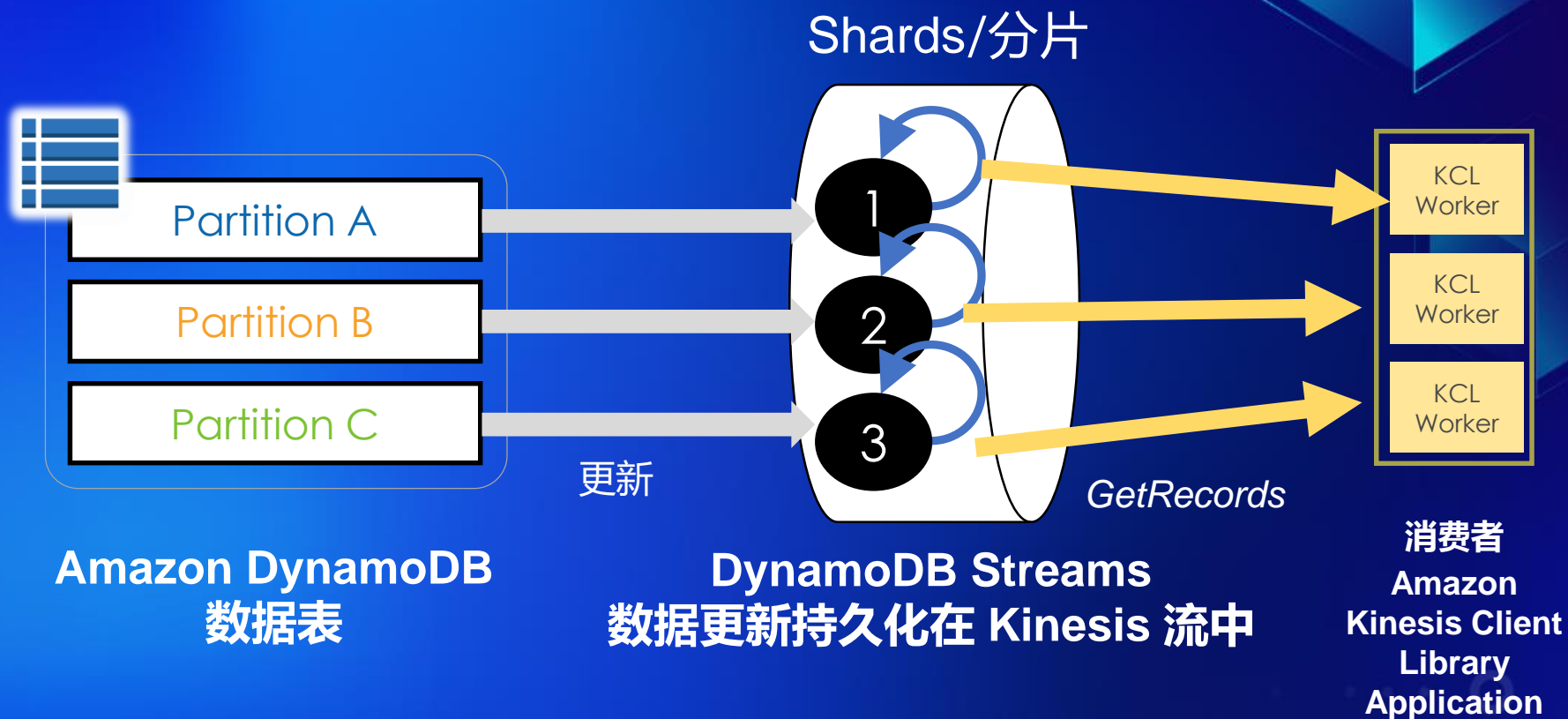
数据库: Binlog /redo Log, etc



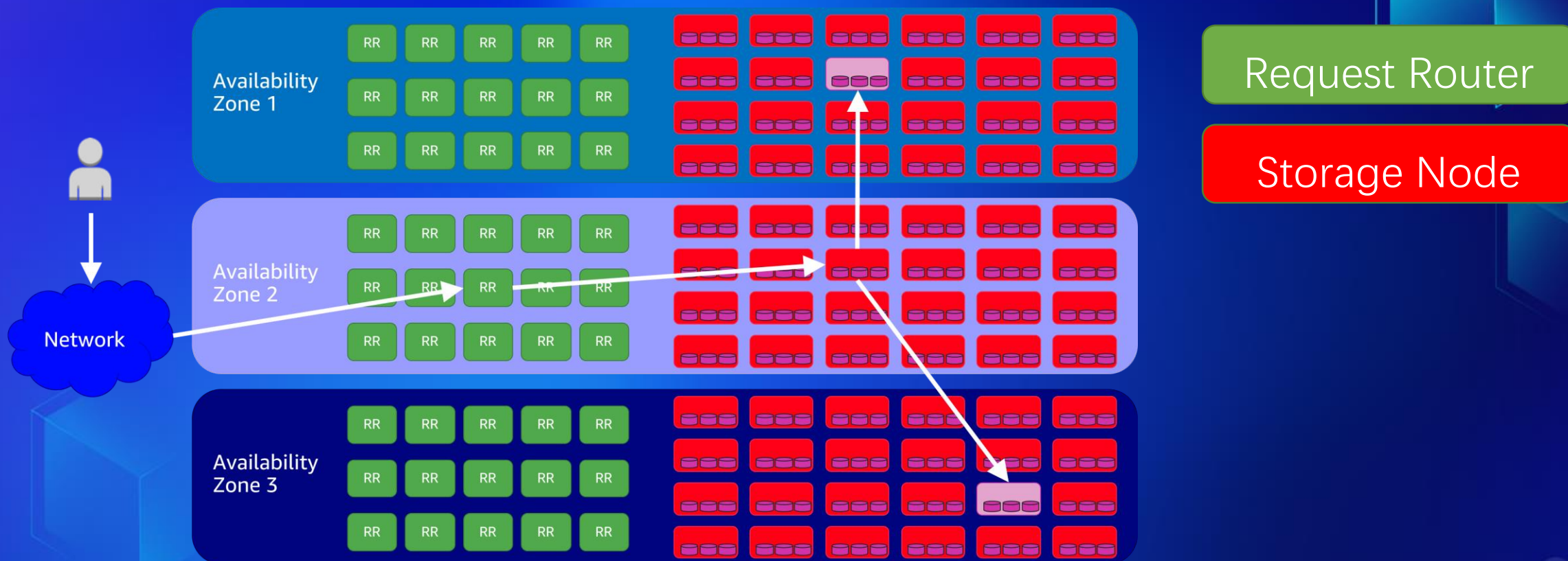
DynamoDB 流

DynamoDB 流

- ✓ 流记录将按照对该项目进行的实际修改的顺序显示
- ✓ 每个流记录仅在流中显示一次（基于主键）
- ✓ 可扩展、高可用、零运维
- ✓ 24 小时保存时间
- ✓ 亚秒级延迟
- ✓ 和 Kinesis Client Library 兼容

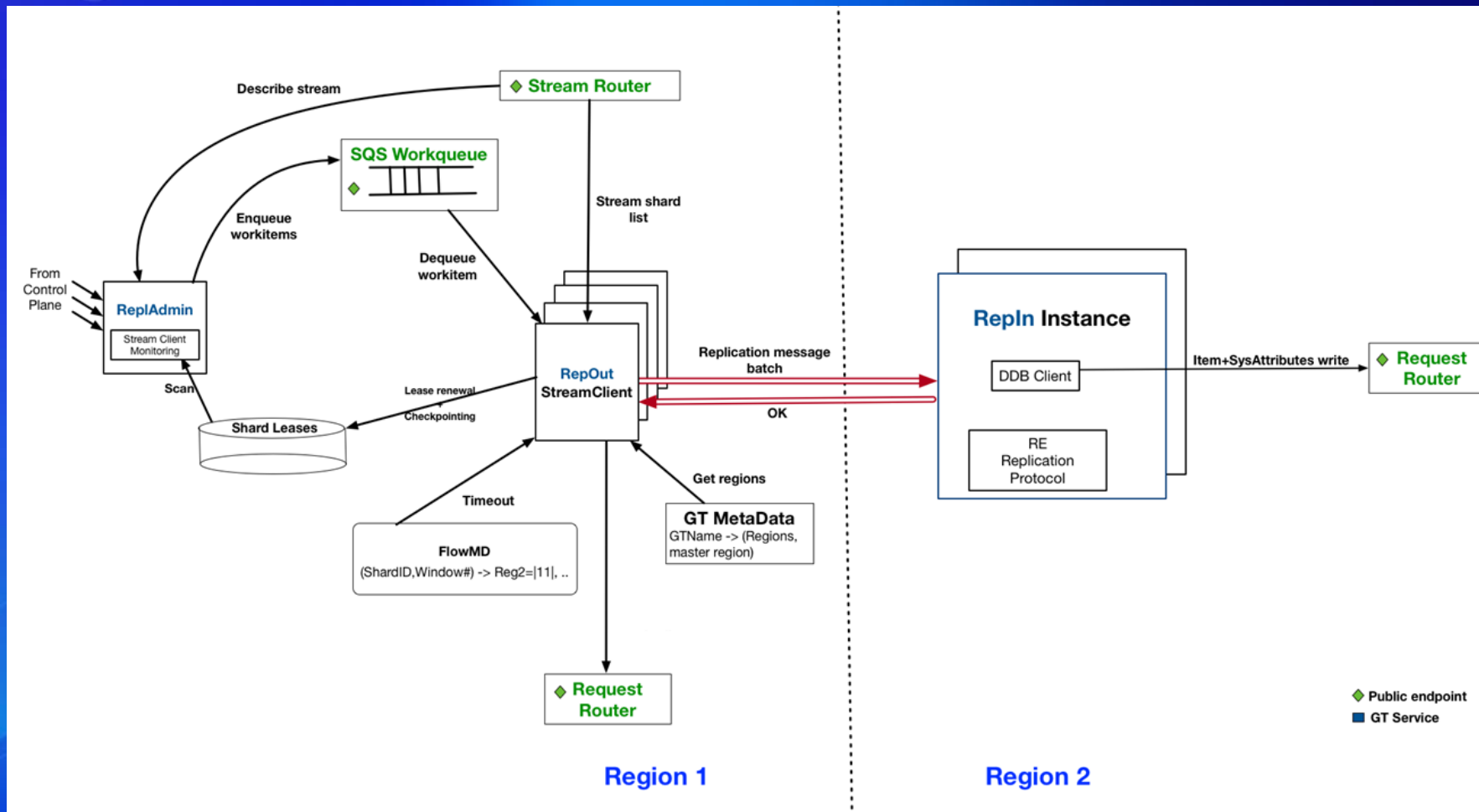


Amazon DynamoDB 核心架构



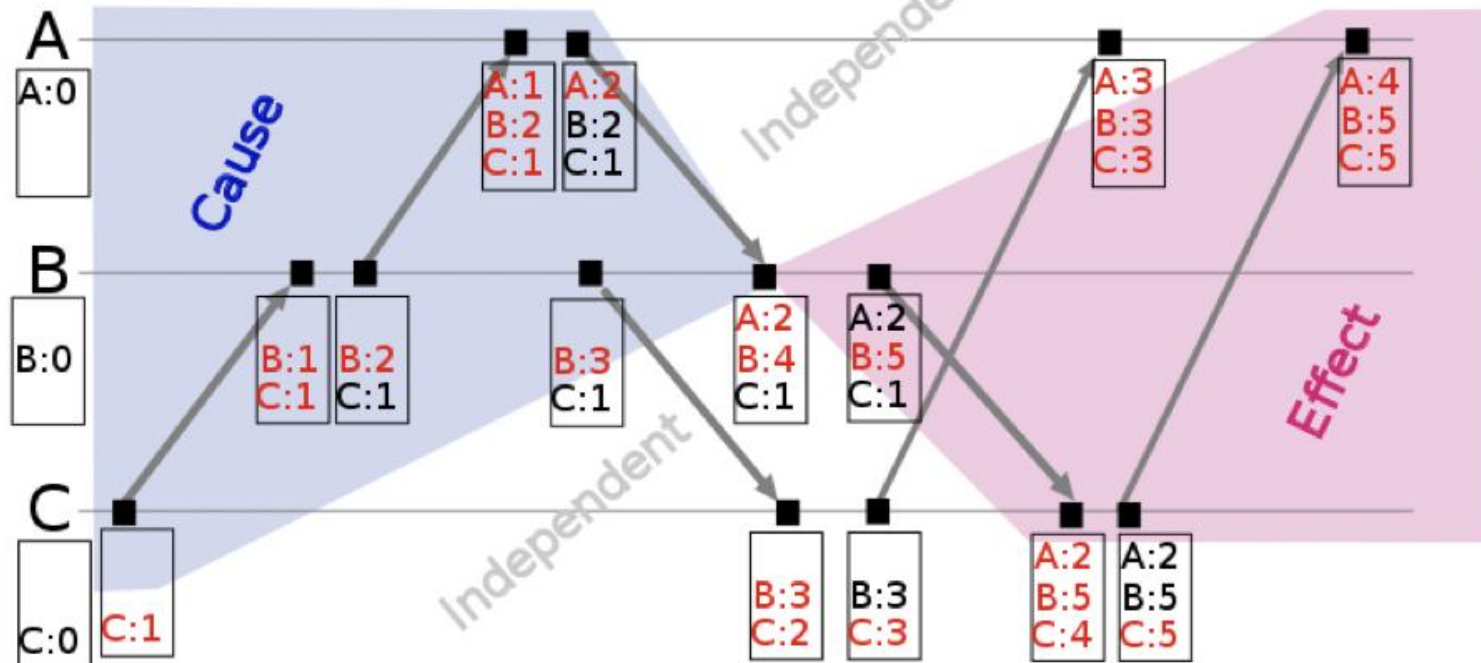
参考经典论文: 《Amazon's Dynamo》 https://www.allthingsdistributed.com/2007/10/amazons_dynamo.html

全局表：基于流的跨区域复制架构



时钟向量和冲突处理

Time →

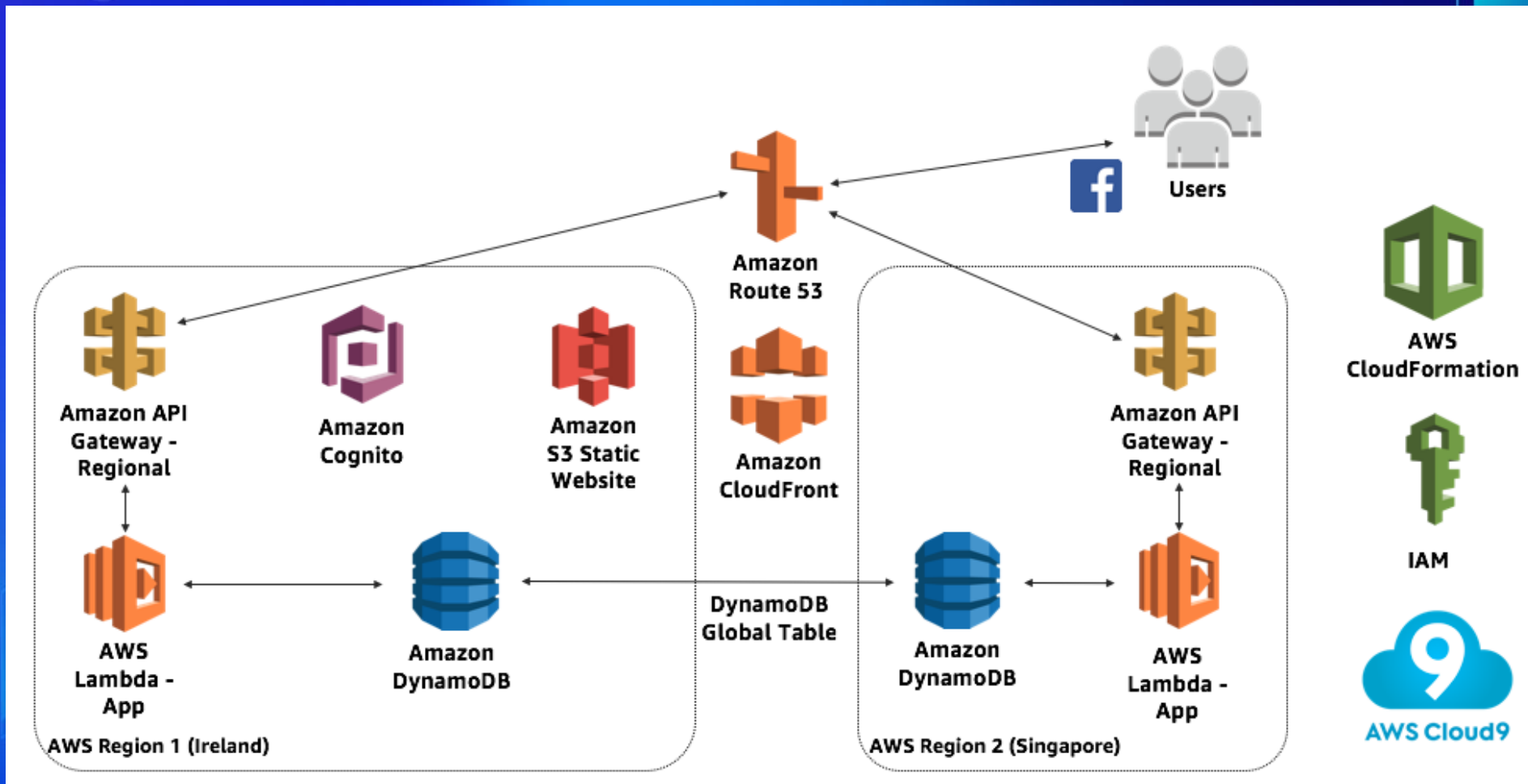


*图片来自: wikipedia

Last Writer wins.

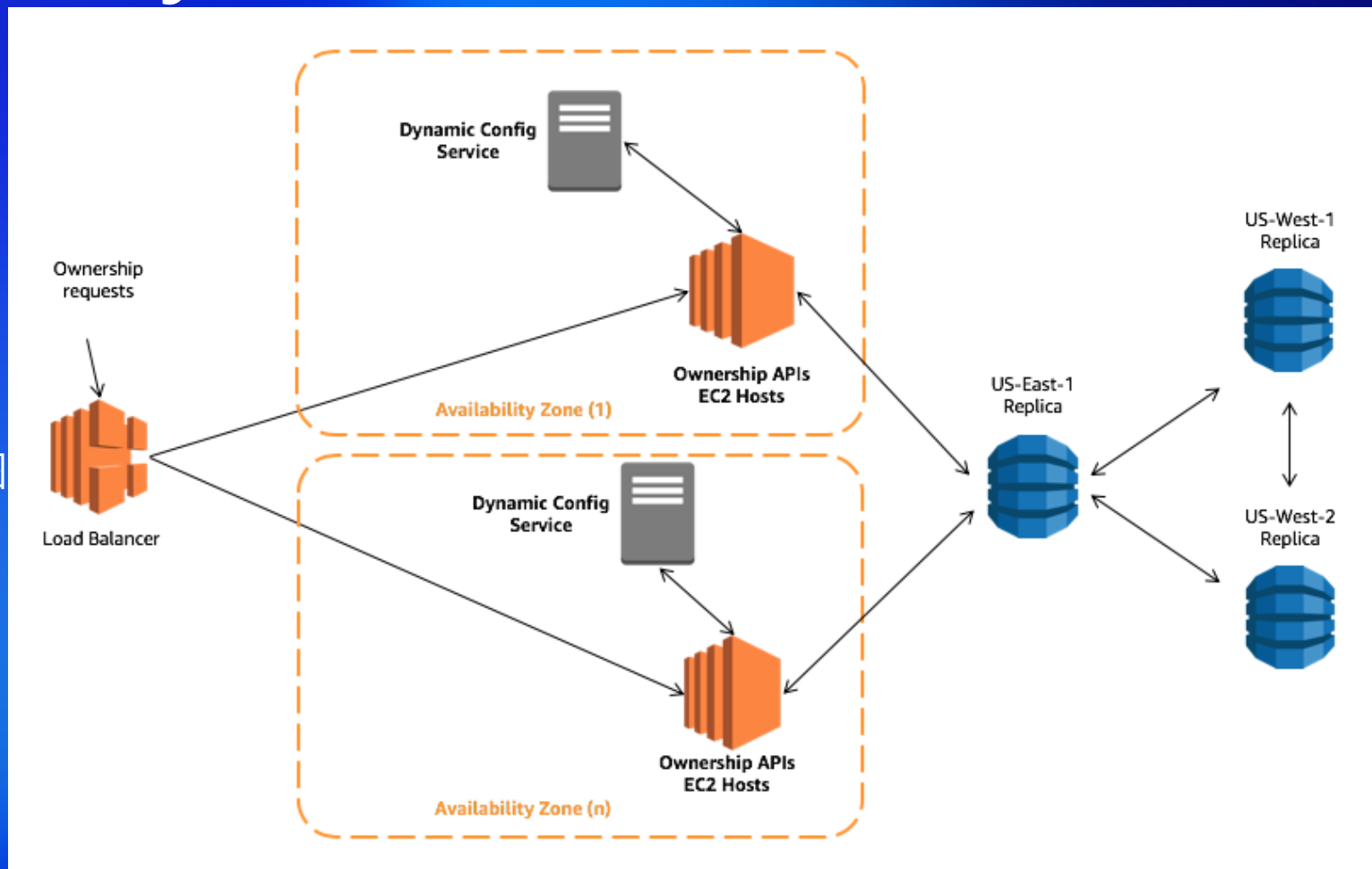
- aws:rep:deleting
- aws:rep:updatetime
- aws:rep:updateregion

基于DynamoDB 的多区架构



DynamoDB 案例 Amazon Video

Amazon Video 架构图



展望：数据铁三角 – CAP 定理

数据模型:

- Relational 关系
- Wide Column
- Document 文档
- Key/Value 键值对

可用性:

所有客户端任何时候都可以读写

A

CA

MSSQL
Oracle
DB2
MySQL

Greenplum
Postgres

同时满足两项

AP

DynamoDB
Cassandra
CouchDB
Riak
MongoDB

一致性:

所有的客户端任何时候都有一致性的数据视图

C

CP

Big Table
Hypertable
Hbase

MongoDB
Couchbase
Scalaris

DynamoDB
Memcache
Redis

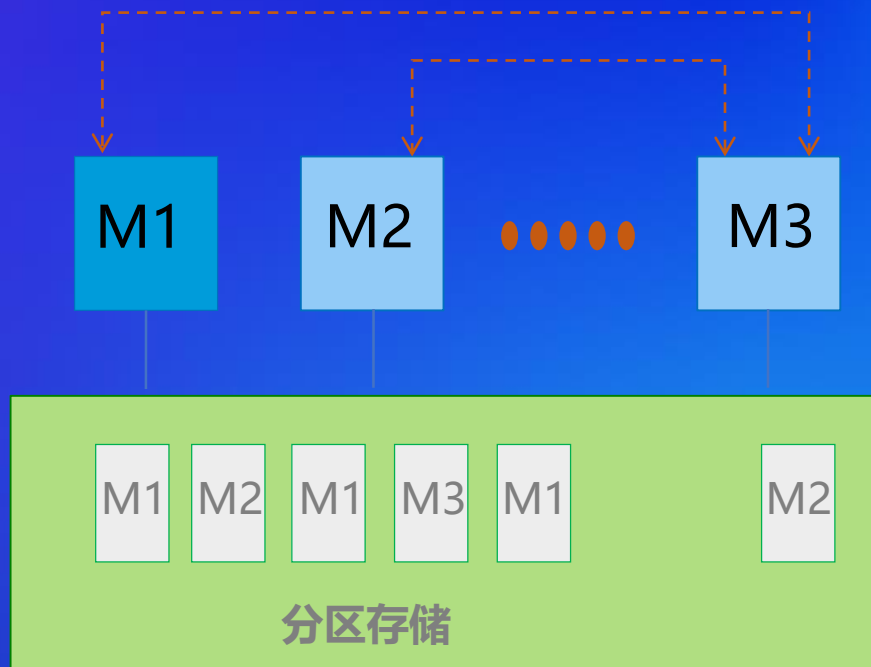
P

分区容错性:

哪怕物理网络发生隔离也不影响系统运行

现有的多主读写数据库实现方式

分布式锁管理器



重量级同步：悲观消极扩展

如 Oracle RAC, DB2 Purescale, Sybase

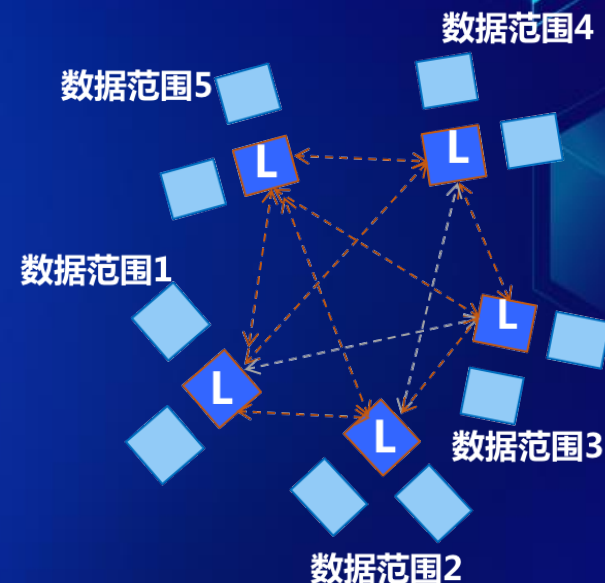
使用读写集进行全局排序



全局实体：扩展瓶颈

如 Galera, TangoDB, FaunaDB

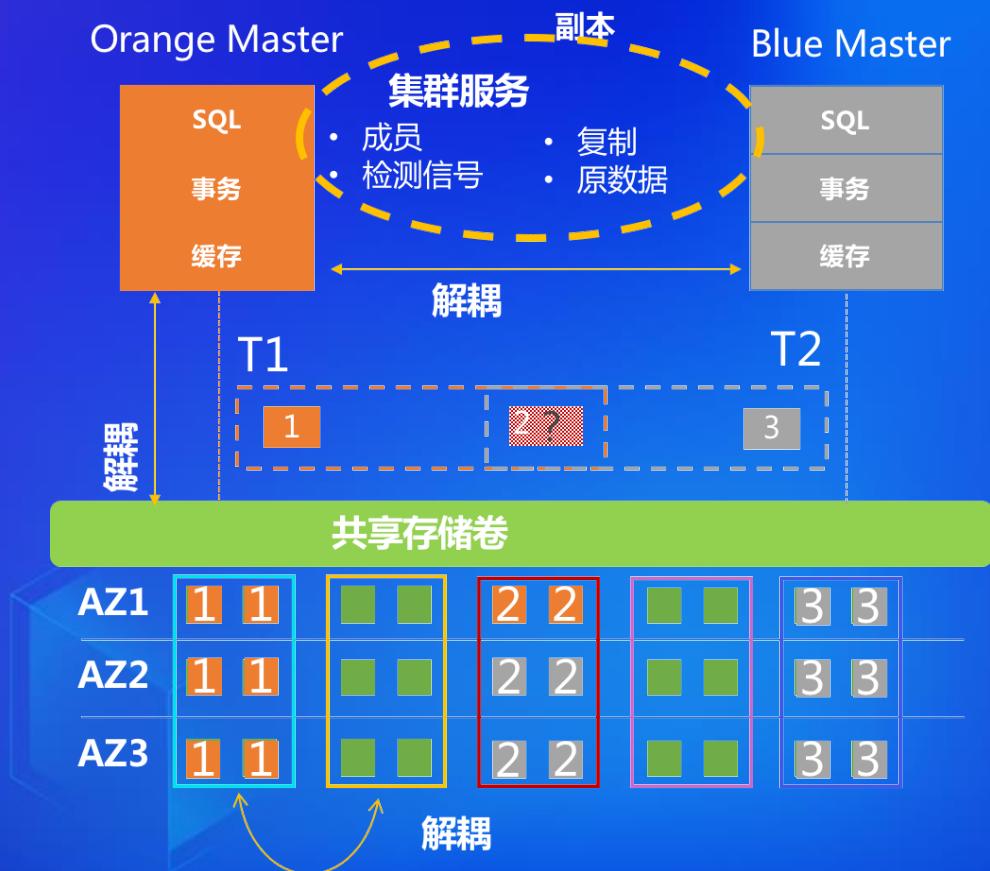
带有2PC的Paxos leader



重量级共识协议：热分区及努力实现跨分区查询

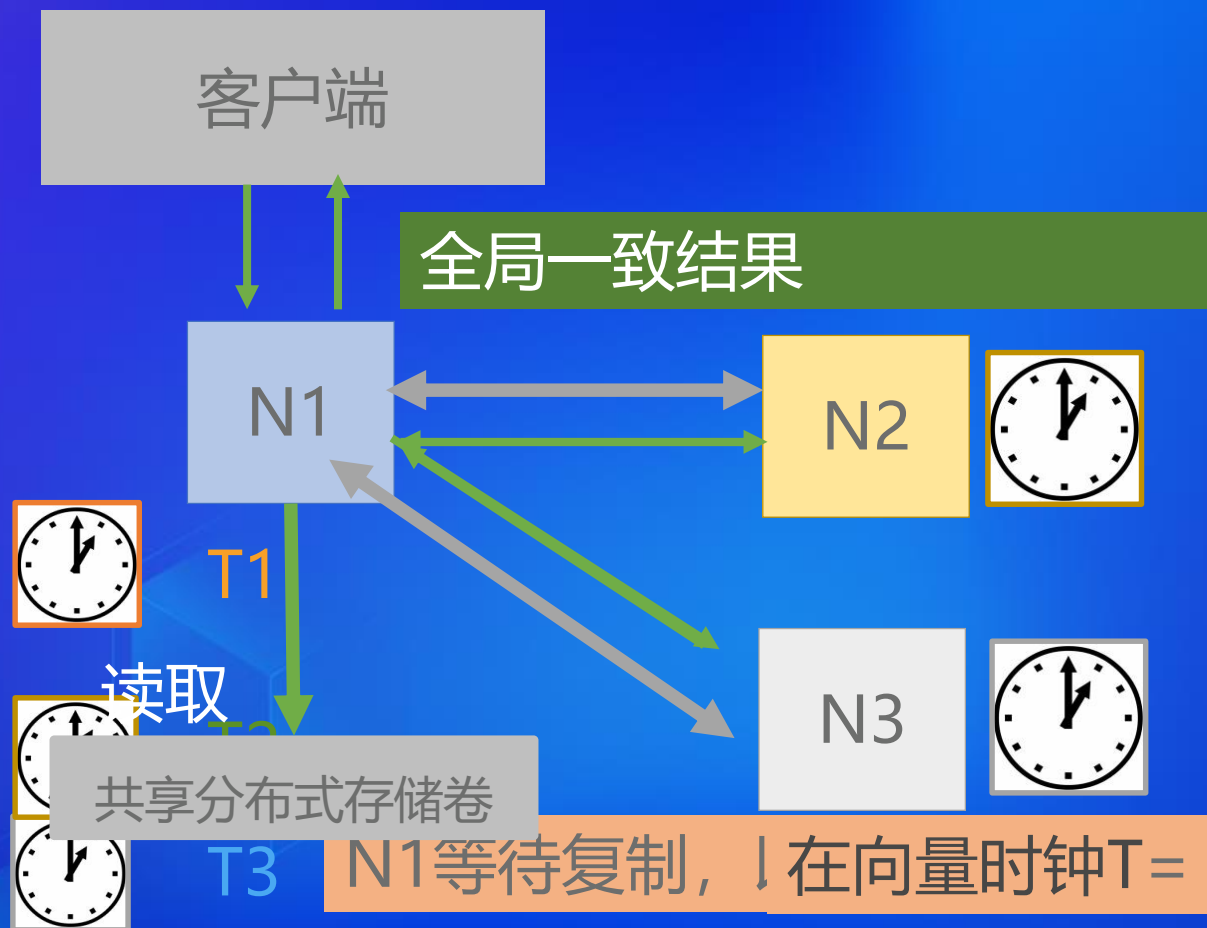
如 Spanner, CockroachDB, Ignite

Amazon Aurora 单区域多主关系数据库



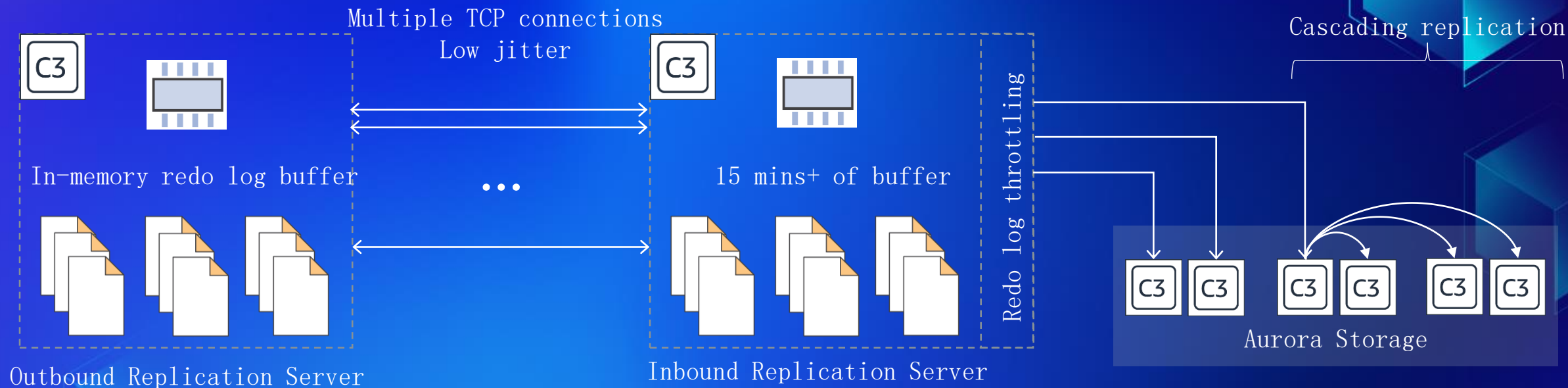
- 无悲观锁定
- 无全局排序
- 无全局提交协调
- 乐观冲突管理
- 解耦系统
- 微服务架构

全局一致性读用户体验?



- 写入路径无等待
- 仅增加了全局一致读取的延迟
- 每个会话均可配置

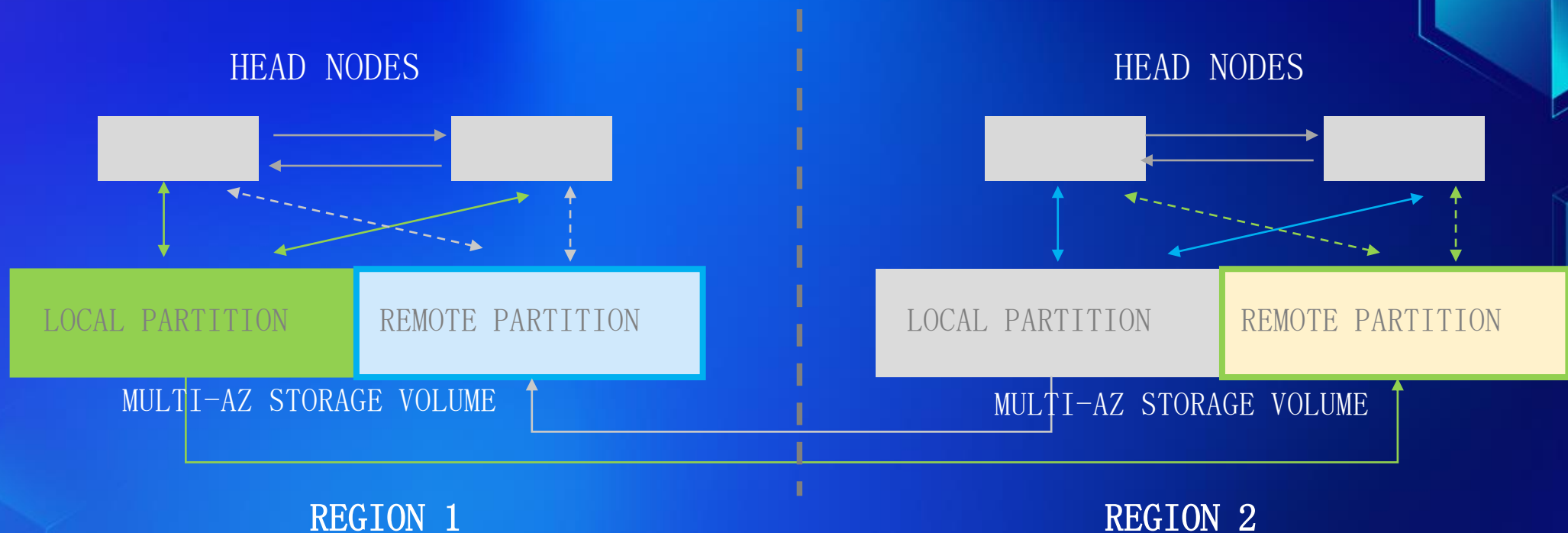
Amazon Aurora 全局数据库 – 亚秒级物理复制



- 内存缓存 Redo 日志
- 多个 TCP 链接以降低网络抖动影响

- 智能流控：安全将redo logs 回放到存储节点
- 级联复制：利用 Aurora 存储节点快速本区域复制

展望：多区域多主 Aurora 数据库



本地写
乐观锁并发控制

多层冲突感知和处理 – head nodes, 存储层, 可用区和区域级仲裁
如果有没存储层或没有冲突, 写性能是线性增长

总结

“故障总是在意料之外，情理之中发生”

数据复制范式：网络复制延迟、偏移量、故障恢复

客户全球架构：业务诉求与状态数据复制关系

客户全球架构：AWS 支持 NoSQL 和关系数据库跨区域复制

AWS 实践：DynamoDB 全局表和 Aurora 全局数据库

Thanks For Watching



本PPT来自2019携程技术峰会
更多信息请关注“携程技术中心”微信公众号~