

架构实战营模块7

第4课：异地多活架构的三种模式

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

李运华

前阿里资深技术专家（P9）

教学目标

1. 掌握异地多活架构三种模式和案例



没有银弹！

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

目录

1. 业务定制型异地多活
2. 业务通用型异地多活
3. 存储通用型异地多活

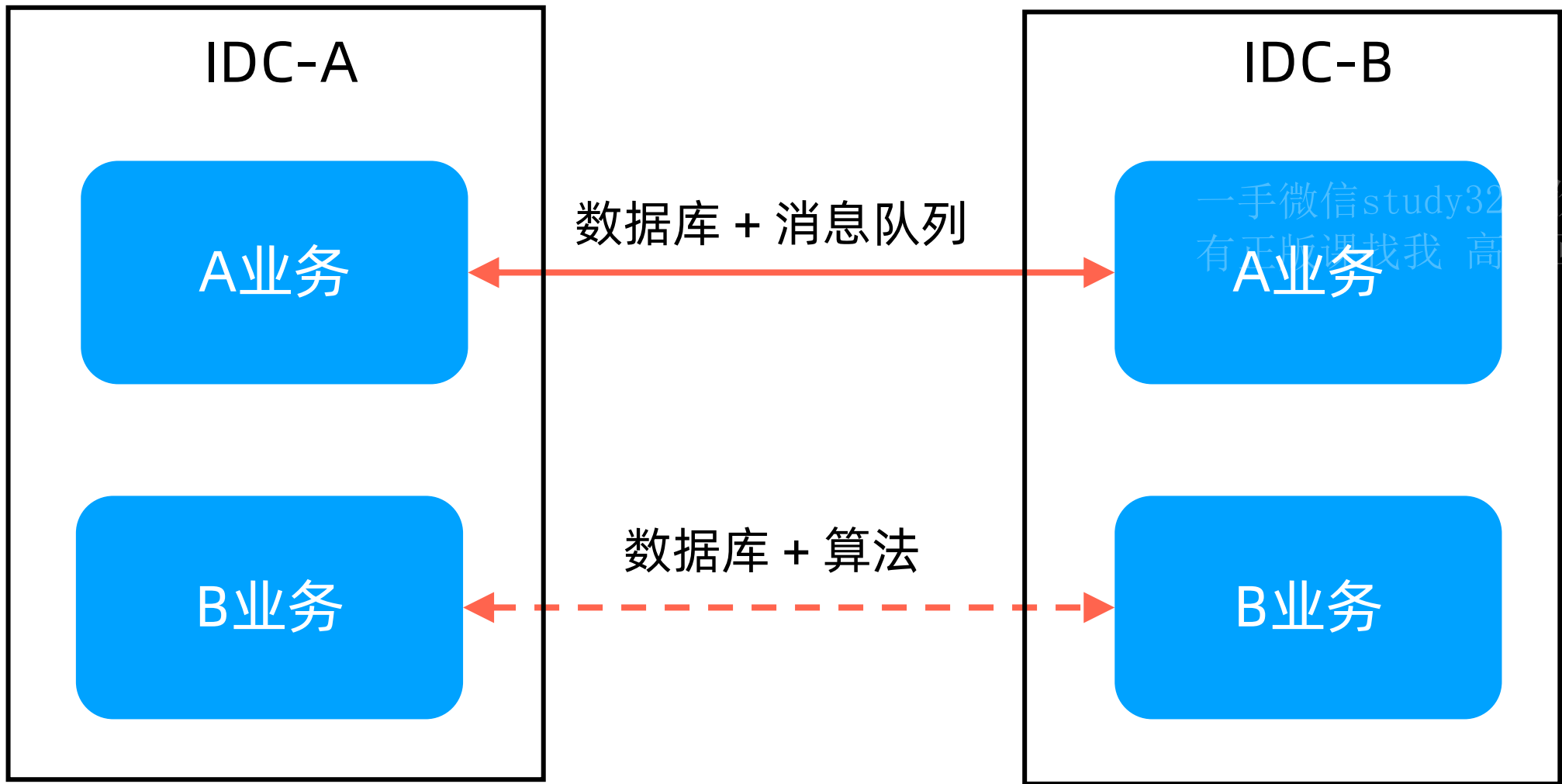
一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

1. 业务定制型异地多活

一手微信study822 价格更优惠
有正版课找我 高价回收帮回血

业务定制型异地多活

按照业务的**优先级**进行排序，优先保证核心业务异地多活；基于核心业务的流程和数据，设计**定制化**的异地多活架构。



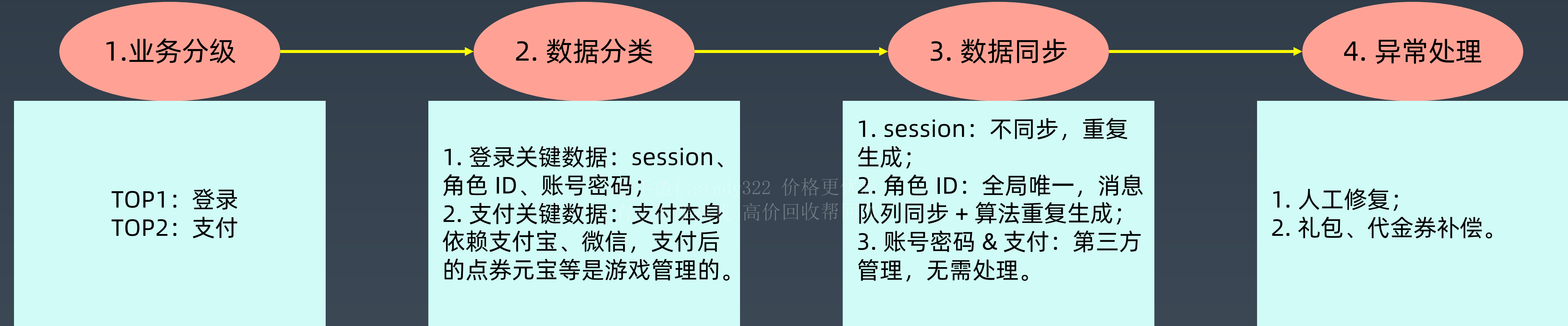
【优点】

对基础设施无强要求，例如机房部署、存储系统、时延等，一般部署在远距离的两个城市，可以支持区域级别故障处理。

【缺点】

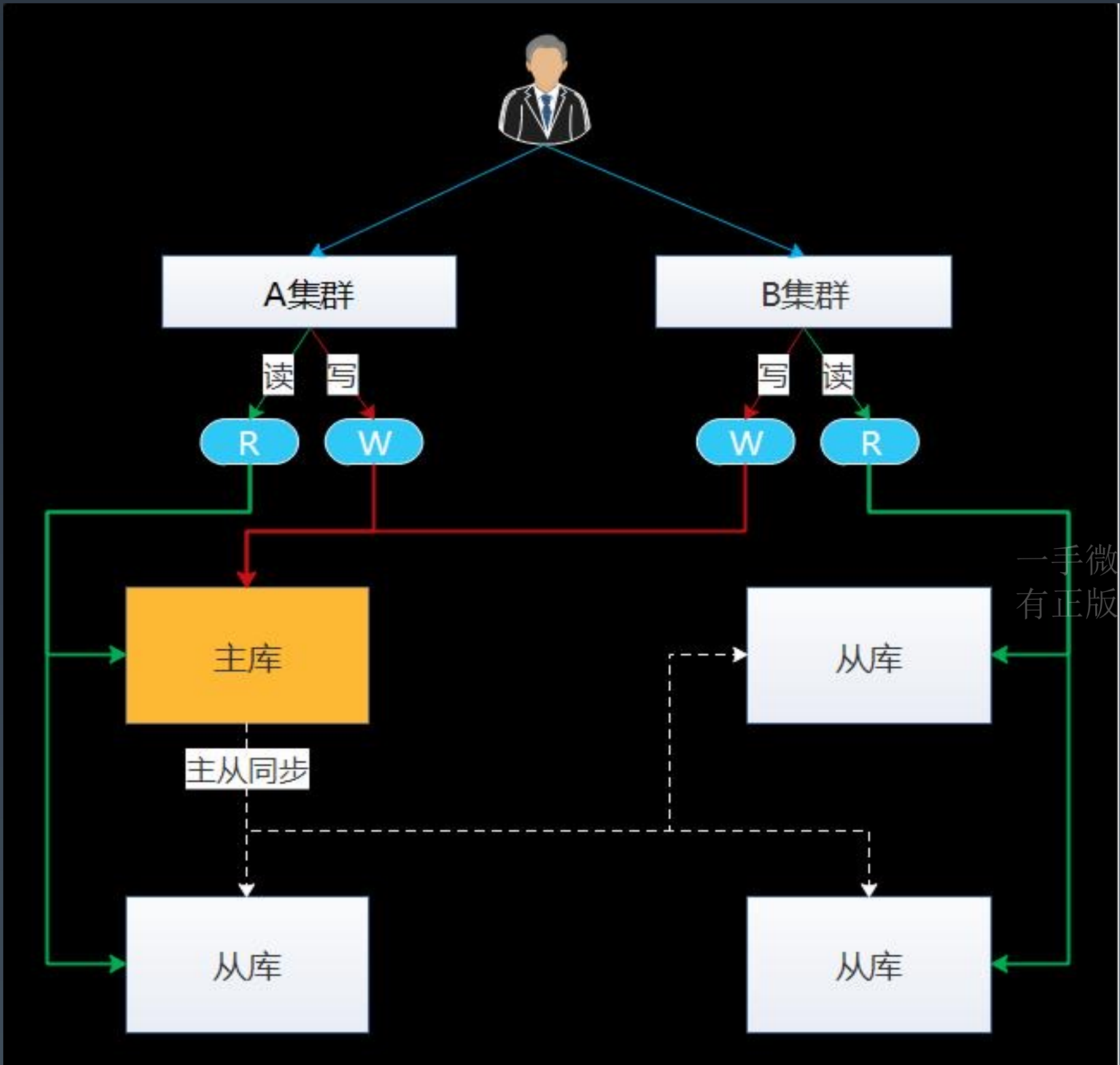
1. 不通用，每个业务都需要单独来做异地多活，业务需要改造；
2. 难扩展，核心业务如果有重大变化，异地多活方案要重新设计。

阿里游戏异地多活设计

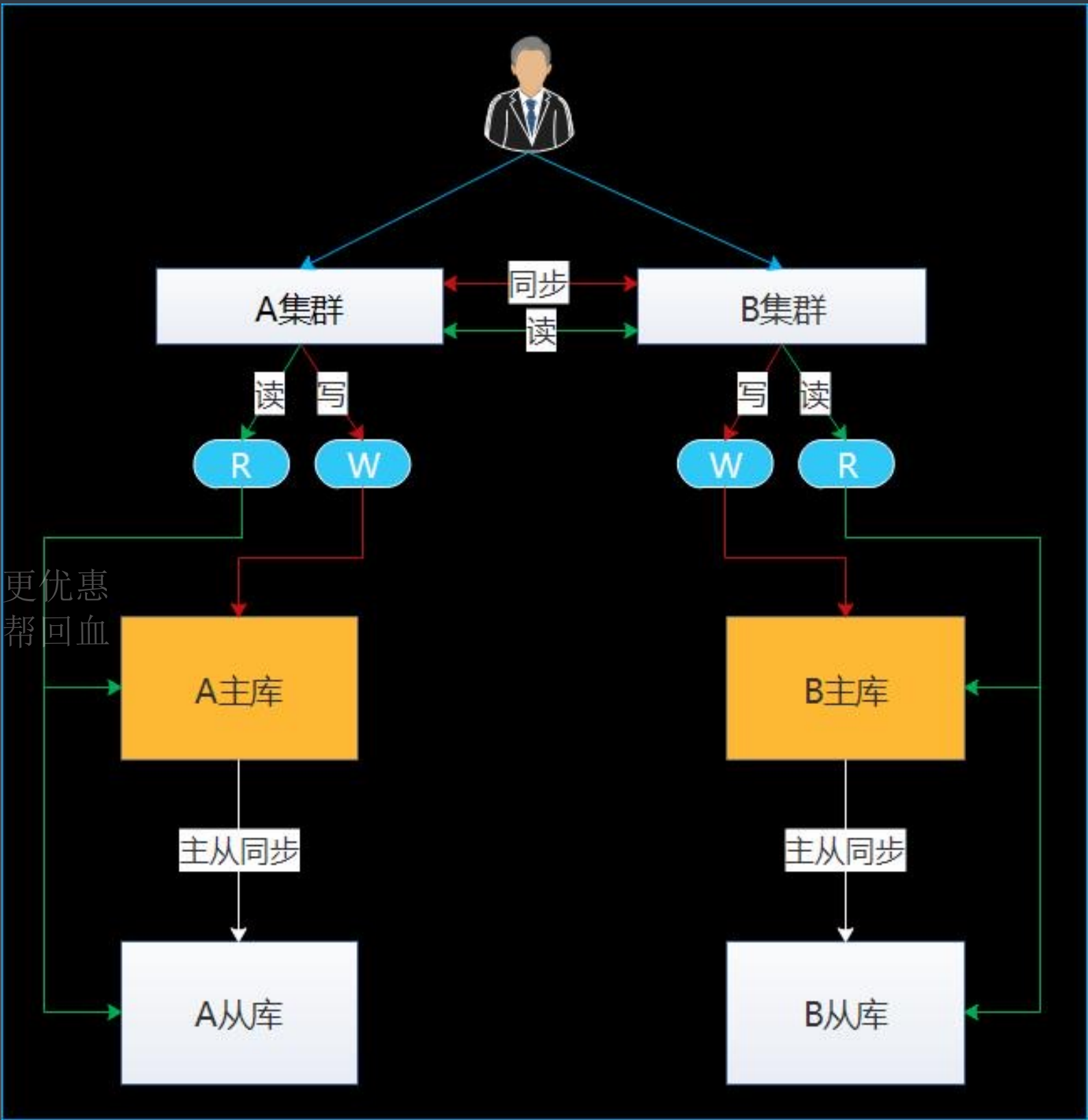


估计一下礼包和代金券补偿一个月大概多少钱?

异地多活架构



存在全局单点、跨机房同步时延问题。



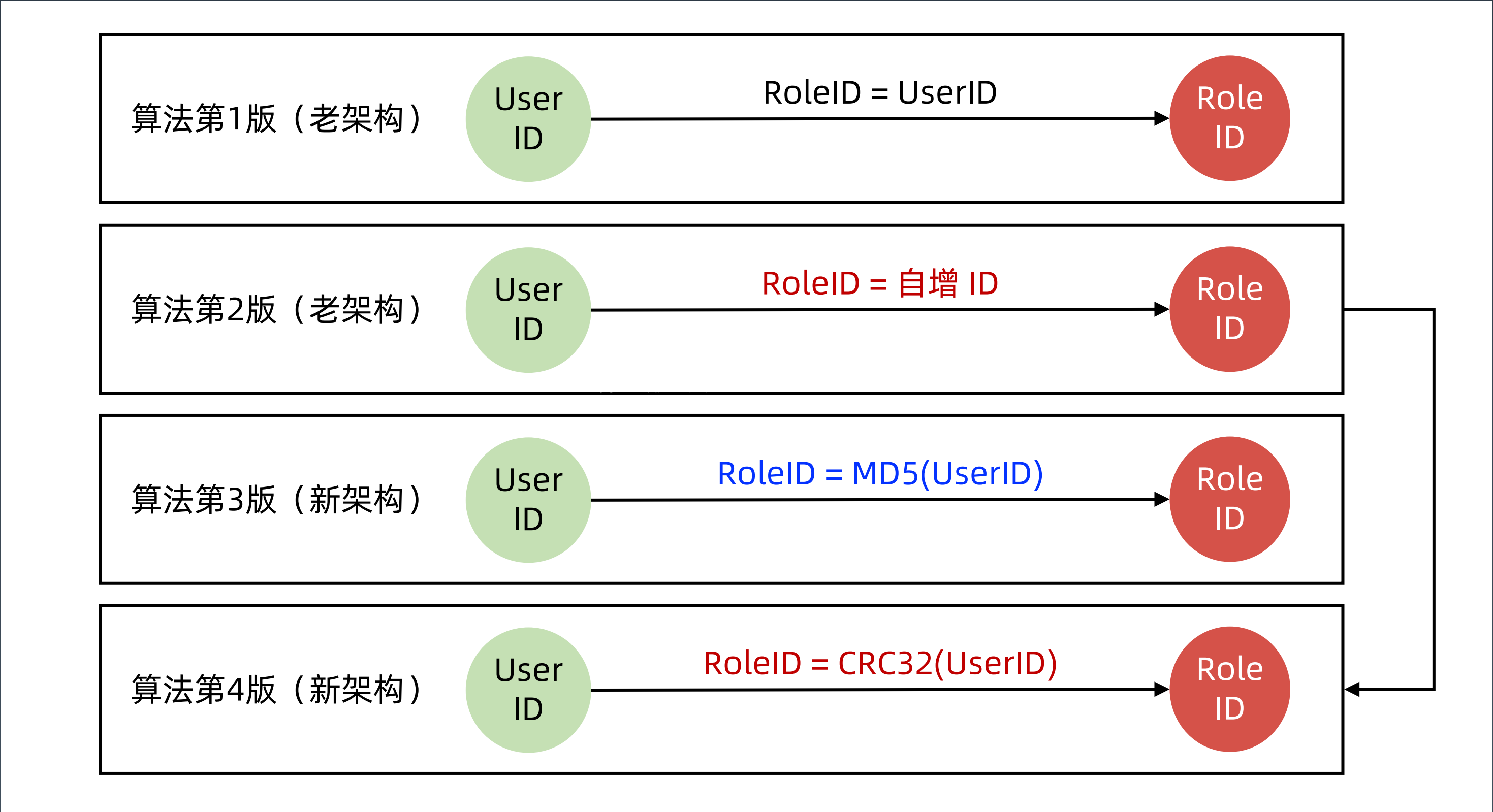
1. 业务层数据同步；
2. 二次读取；
3. 可重复生成全局唯一数据。

异地多活设计核心 - RoleID



为什么不直接返回用户名或者用户 ID?

异地多活设计核心 - 算法生成 RoleID



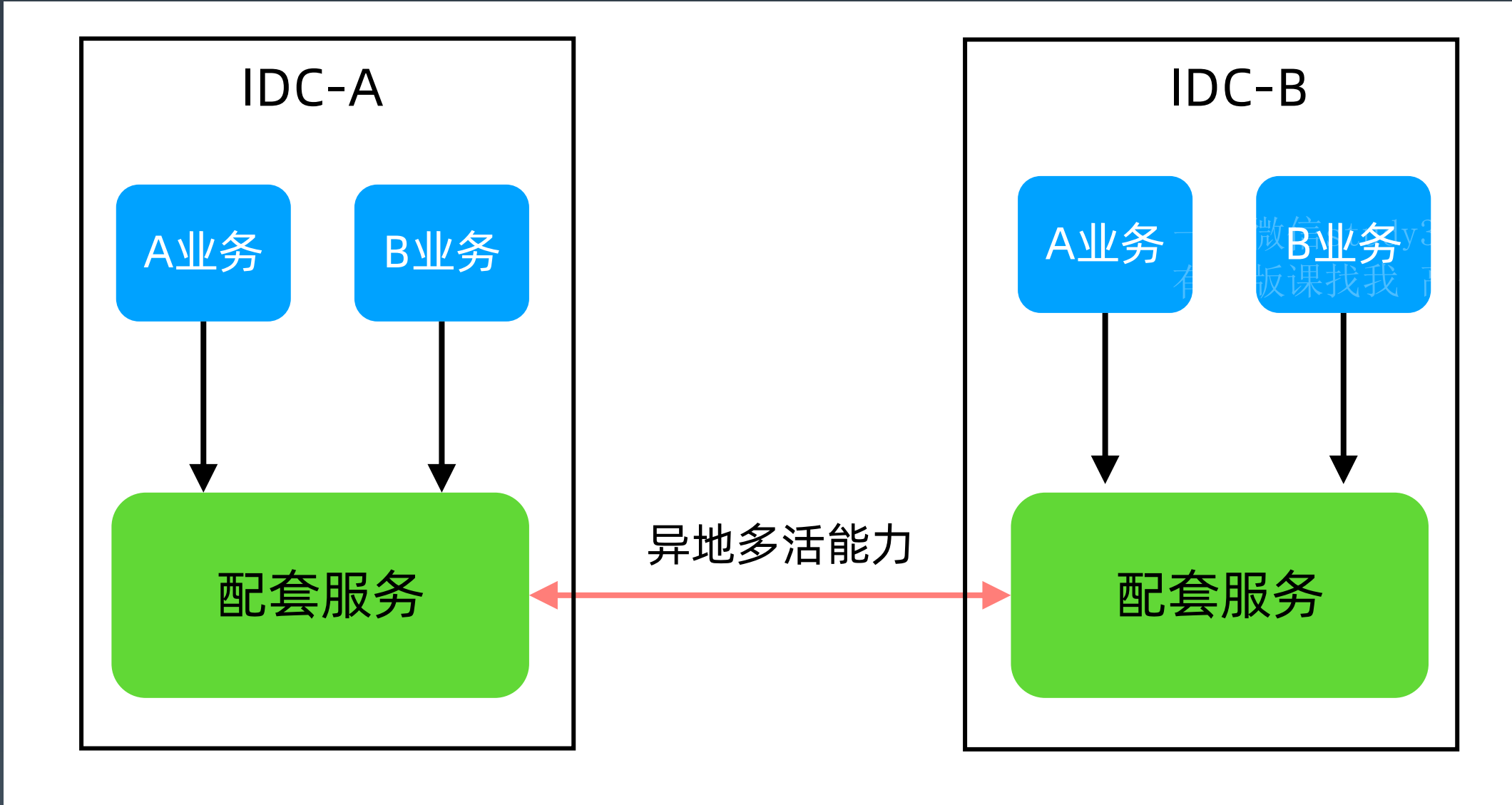
不要死记
硬背算法

2. 业务通用型异地多活

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

业务通用型异地多活

通过配套服务来支持异地多活，无需按照业务优先级排序来挑选某些业务实现异地多活，只需要判断业务是否能异地多活。



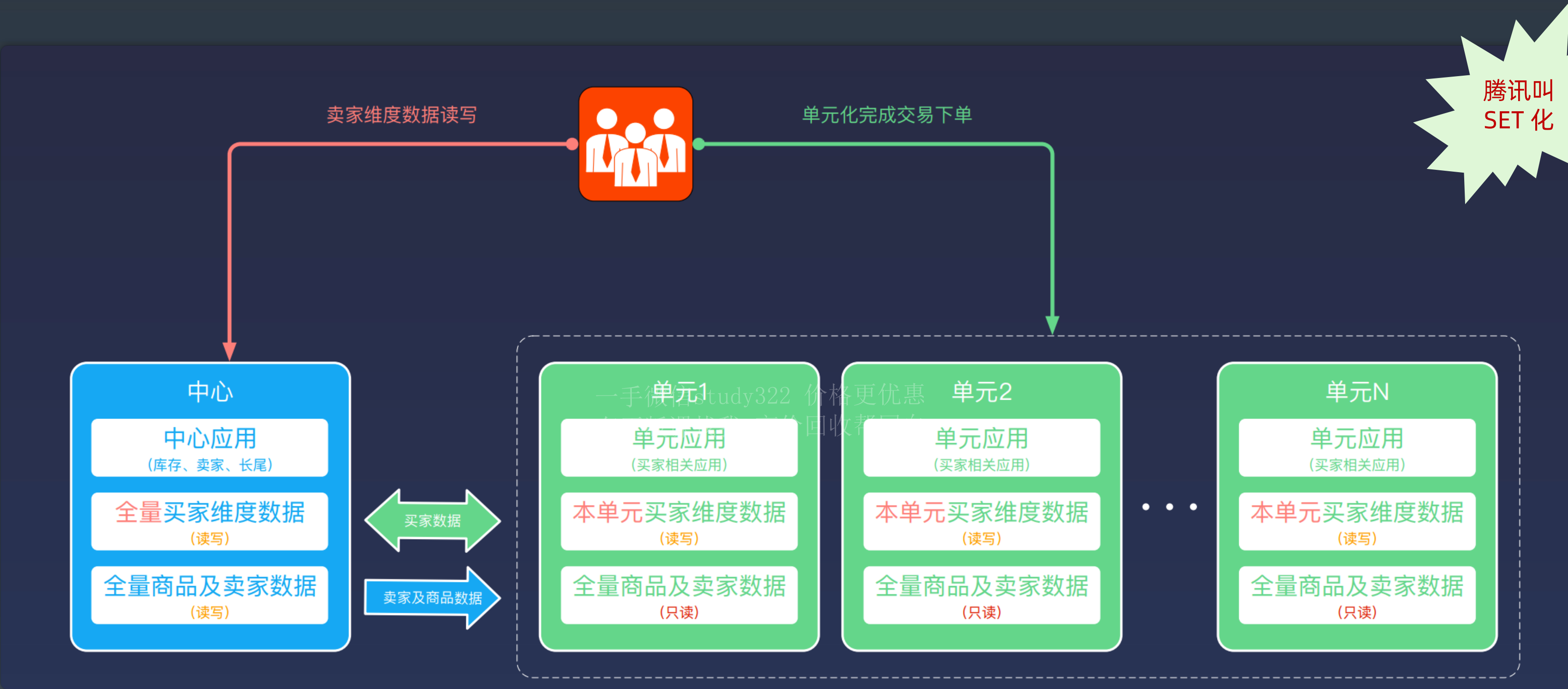
【优点】

- 1. 对硬件基础设施无强要求，例如机房部署、存储系统、时延等，一般部署在远距离的两个城市，可以支持区域级别故障处理。
- 2. 业务基本不用改造，只需判断业务是否支持 **BASE**，支持就可以异地多活，不支持就单点部署。

【缺点】

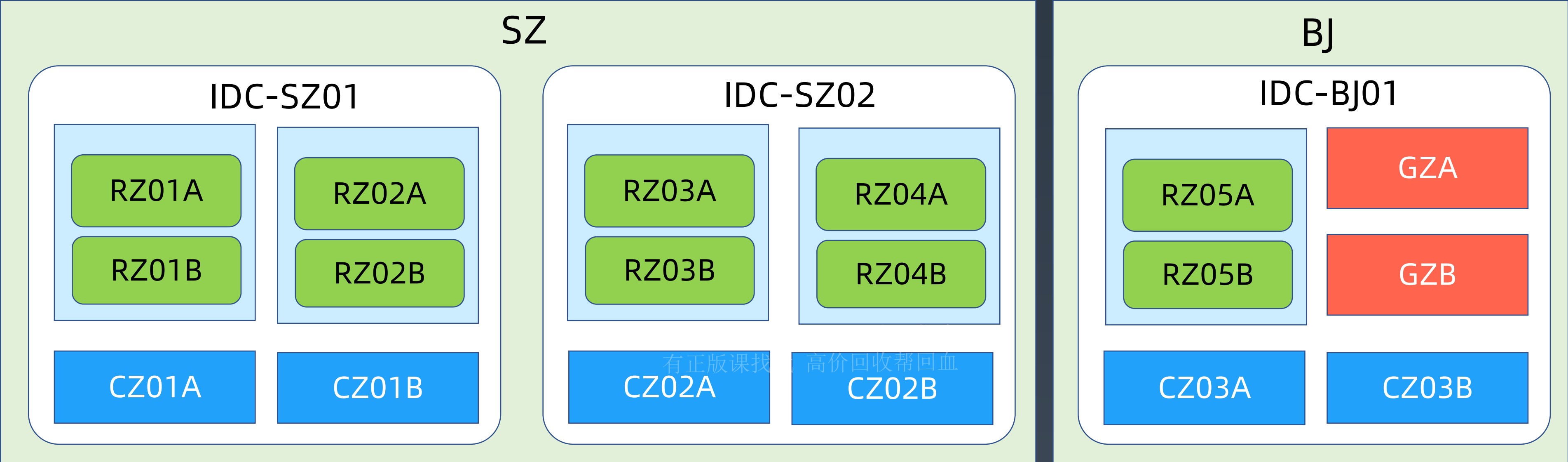
- 1. 配套服务复杂，包括流量调度、容灾切换、建站平台、配置管理等。
- 2. 存在全局单点业务，例如库存、卖家等相关业务。
- 3. 机房距离较远的时候，RTO 比较大，可能达到分钟级。

业务通用型案例 - 淘宝的单元化架构



LDC (Logic Data Center) 架构，又称为单元化架构，是指一个能完成所有业务操作的自包含集合，在这个集合中包含了所有业务所需的所有服务，以及分配给这个单元的数据。[参考链接](#)

业务通用型案例 - 蚂蚁的 LDC 架构



RZone (Region Zone) : 部署按用户维度拆分的关键业务系统。核心业务和数据单元化拆分，拆分后分片均衡，单元内尽量自包含（调用封闭），拥有自己的数据，能完成所有业务。一个可用区可以有多个 RZone。[参考链接](#)

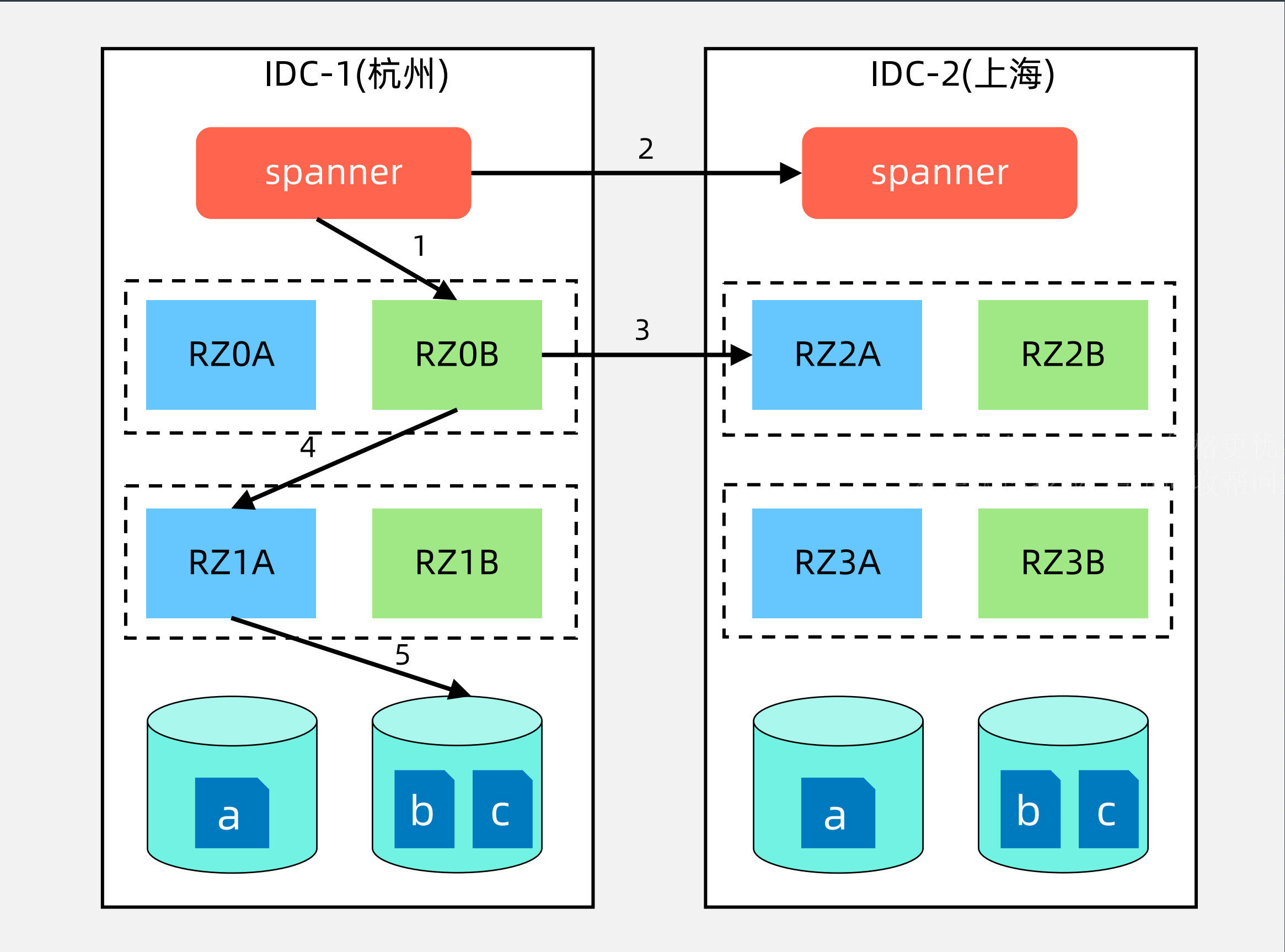
CZone (City Zone) : 部署未按用户维度拆分的系统，被 RZone 高频访问，解决跨域通信延时问题。为了解决异地延迟问题而特别设计，适合读多写少且不可拆分的业务。一般每个城市一套应用和数据，是 GZone 的快照。

GZone (Global Zone) : 部署未按用户维度拆分的系统，被 RZone 依赖，提供不可拆分的数据和服务，如配置型的业务。数据库可以和 RZone 共享，多租户隔离，全局只有一组，可以配置流量权重。



能看出来淘宝单元化架构与蚂蚁 LDC 架构的区别么？

LDC 路由



Spanner 是蚂蚁基于 Nginx 自研的反向代理网关。

【一次路由】

箭头1：对于应该在本 IDC 处理的请求，就直接映射到对应的 RZ 即可

箭头2：不在本 IDC 处理的请求，Spanner 可以转发至其他 IDC 的 Spanner

【二次路由】

有些场景来说，A 用户的一个请求可能关联了对 B 用户数据的访问，比如 A 转账给 B，A 扣完钱后要调用账务系统去增加 B 的余额，这时候就涉及到：

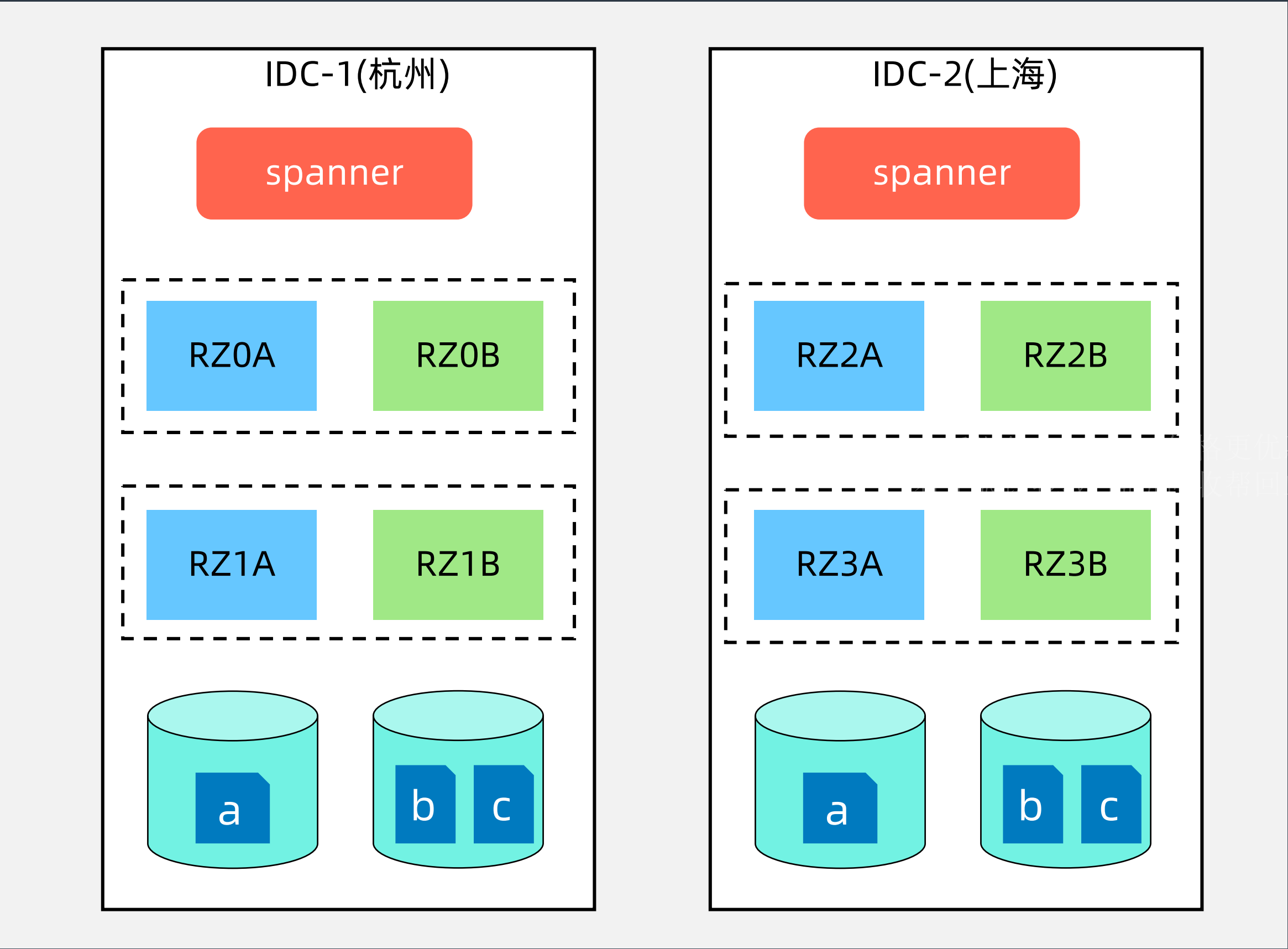
箭头3：跳转到其他 IDC 的 Rzone；

箭头4：跳转到本 IDC 的其他 Rzone。

【数据路由】

RZ 访问哪个数据库，是可以配置的，对应图中箭头5。

LDC 容灾



惠血

【一级容灾】

Rzone 自容灾，例如 RZ0A 挂掉，spanner 会将所有流量会分给到 RZ0B。

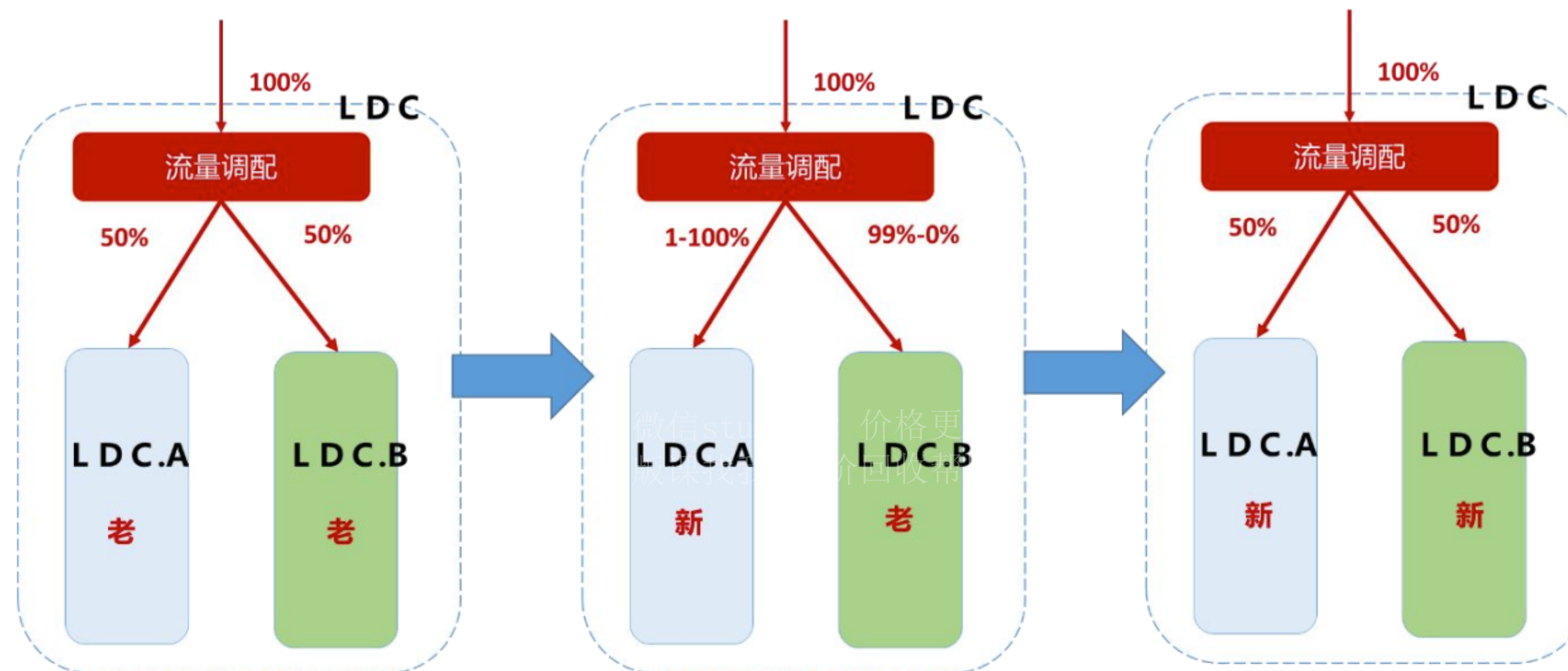
【二级容灾】

如果 RZ0A 和 RZ0B 都挂掉，但数据库没挂掉，直接 **新建** Rzone 代替。
需要建站平台快速新建整个 Rzone 。

【三级容灾 - 异地多活】

如果 IDC-1 挂掉，则在 IDC-2 新建 RZone，然后读写 IDC-2 的数据库。
这种情况可能丢数据，RTO 是分钟级，所以官方要求，不能接受这个 RTO 就不要做切换，也就不支持异地多活。

LDC 蓝绿发布



Step1.发布前，将“蓝”流量调至0%，对“蓝”的所有应用整体无序分2组发布。

Step2. “蓝”引流1%观察，如无异常，逐步上调分流比例至100%。

Step3. “绿”流量为0%，对“绿”所有应用整体无序分2组发布

Step4. 恢复日常运行状态，蓝、绿单元各承担线上50%的业务流量

单元化架构的配套服务

流量调度

负责将用户请求流量分配到对应的 RZone，例如蚂蚁的 Spanner。

配置中心

负责 Zone 的配置和容灾切换，例如 Rzone 的业务范围，Zone 访问哪些数据库等。

建站平台

快速新建一个完整的 Zone，目前基本都是基于容器技术来实现。

发布平台

基于流量调度，完成 Zone 的蓝绿发布、灰度发布等。

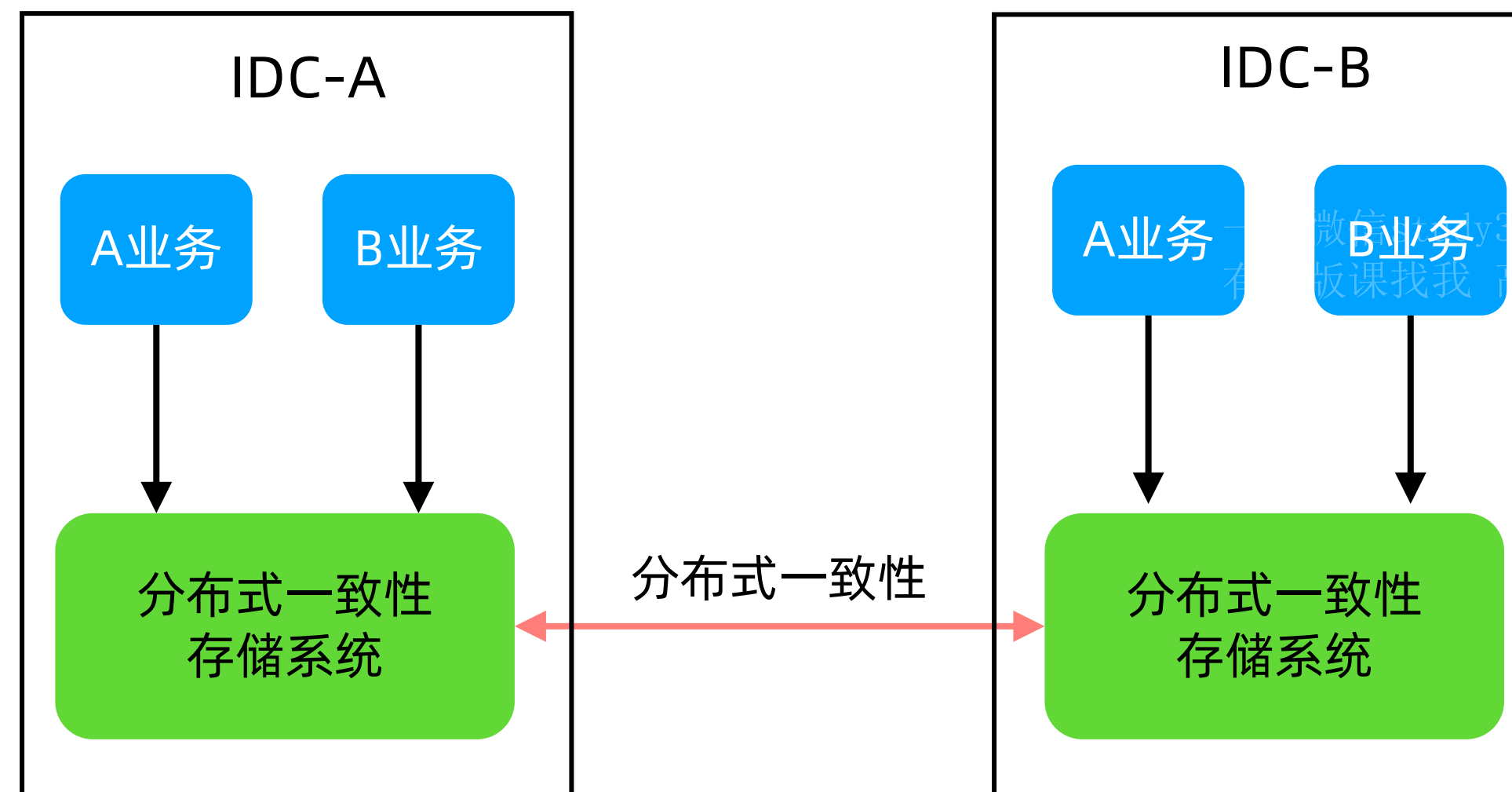
手微信study522 价格更优惠
有正版课找我 高价回收帮回血

3. 存储通用型异地多活

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

存储通用型异地多活

采用本身已经支持分布式一致性的存储系统，架构天然支持异地多活。



【优点】

1. 架构天然就支持异地多活，业务除了切换存储系统外，其它基本不用改造。

【缺点】

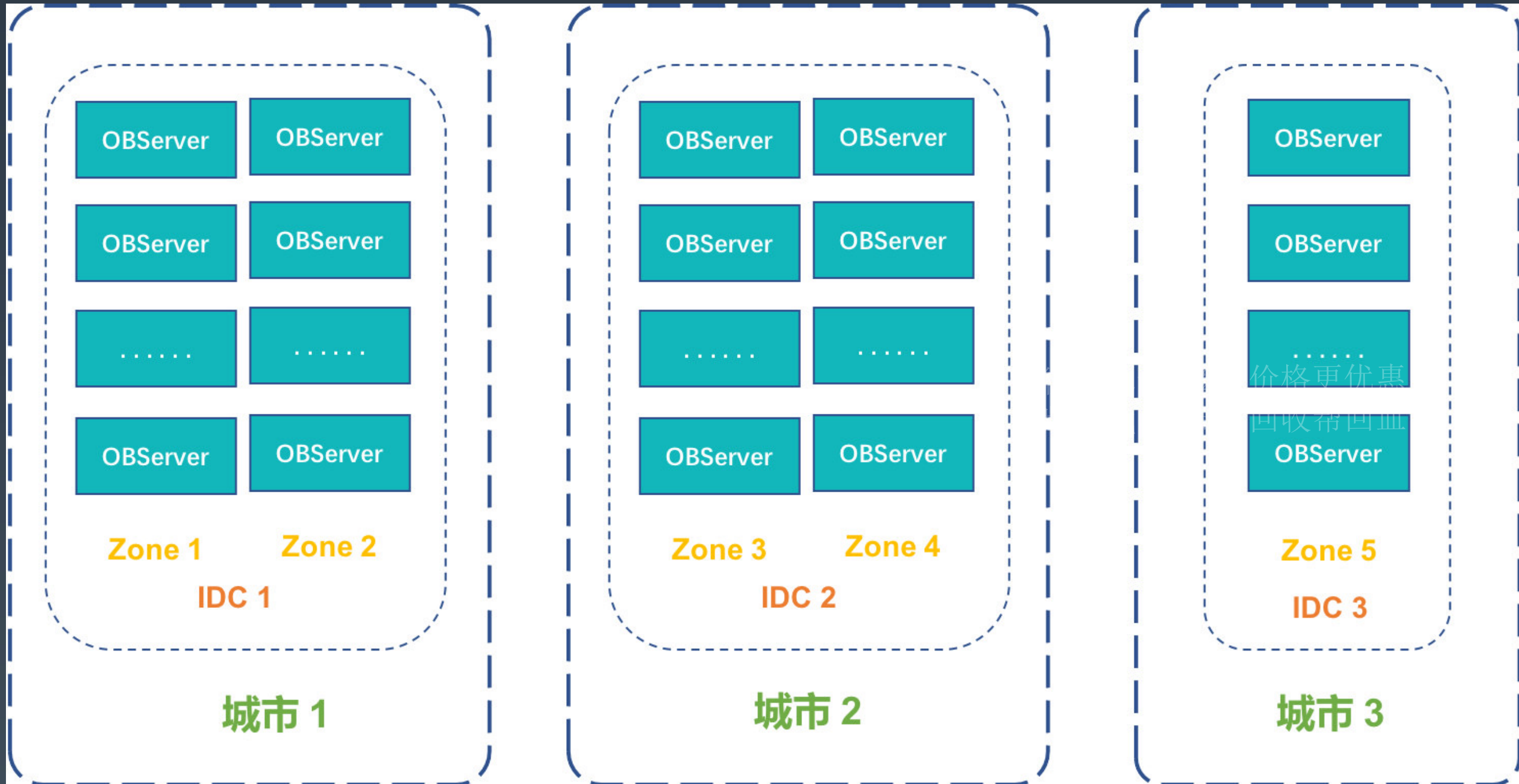
1. 需要分布式一致性的存储系统支持，目前这样的存储系统可选不多，例如 ZooKeeper、etcd、OceanBase。
2. 对机房部署有强要求，如果要想实现异地多活，只能采用近邻部署（参考模块7第3课）。

OceanBase 数据库是阿里巴巴和蚂蚁集团不基于任何开源产品，完全自研的原生分布式关系数据库软件，在普通硬件上实现金融级高可用，首创“三地五中心”城市级故障自动无损容灾新标准，具备卓越的水平扩展能力，全球首家通过 TPC-C 标准测试的分布式数据库，单集群规模超过 1500 节点。产品具有云原生、强一致性、高度兼容 Oracle/MySQL 等特性，承担支付宝 100% 核心链路，在国内几十家银行、保险公司等金融客户的核心系统中稳定运行。[参考链接](#)



- 1. Zone 是一个机房内的一组服务器，包含多台 OceanBase 数据库服务器（OBServer），一般会把数据的多个副本分布在不同的 Zone 上，可以实现单个 Zone 故障不影响数据库服务。
- 2. 每台 OBServer 包含 SQL 引擎、事务引擎和存储引擎，并服务多个数据分区，其中，每个 Zone 有一台 OBServer 会同时使能总控服务（RootService），用于执行集群管理、服务器管理、自动负载均衡等操作。
- 3. OBServer 上会运行 SQL 引擎、事务引擎和存储引擎，用户的 SQL 查询经过 SQL 引擎解析优化后转化为事务引擎和存储引擎的内部调用。
- 4. OceanBase 数据库还会执行**强一致的分布式事务**，从而在分布式集群上实现数据库事务 ACID。[参考链接](#)

OceanBase 三地五中心异地多活架构



【部署要求】

- A. 2近 (10ms) 1远 (30ms) ;
- B. 2 + 2 + 1 = 5 Zone。

【优缺点】

- 1. 可以应对城市级别的故障；
- 2. 无法应对区域级别故障，例如杭州和上海同时停电；
- 3. 成本最高，至少三个机房，如果采用分区架构，每个分区都要按照三地五中心来部署。

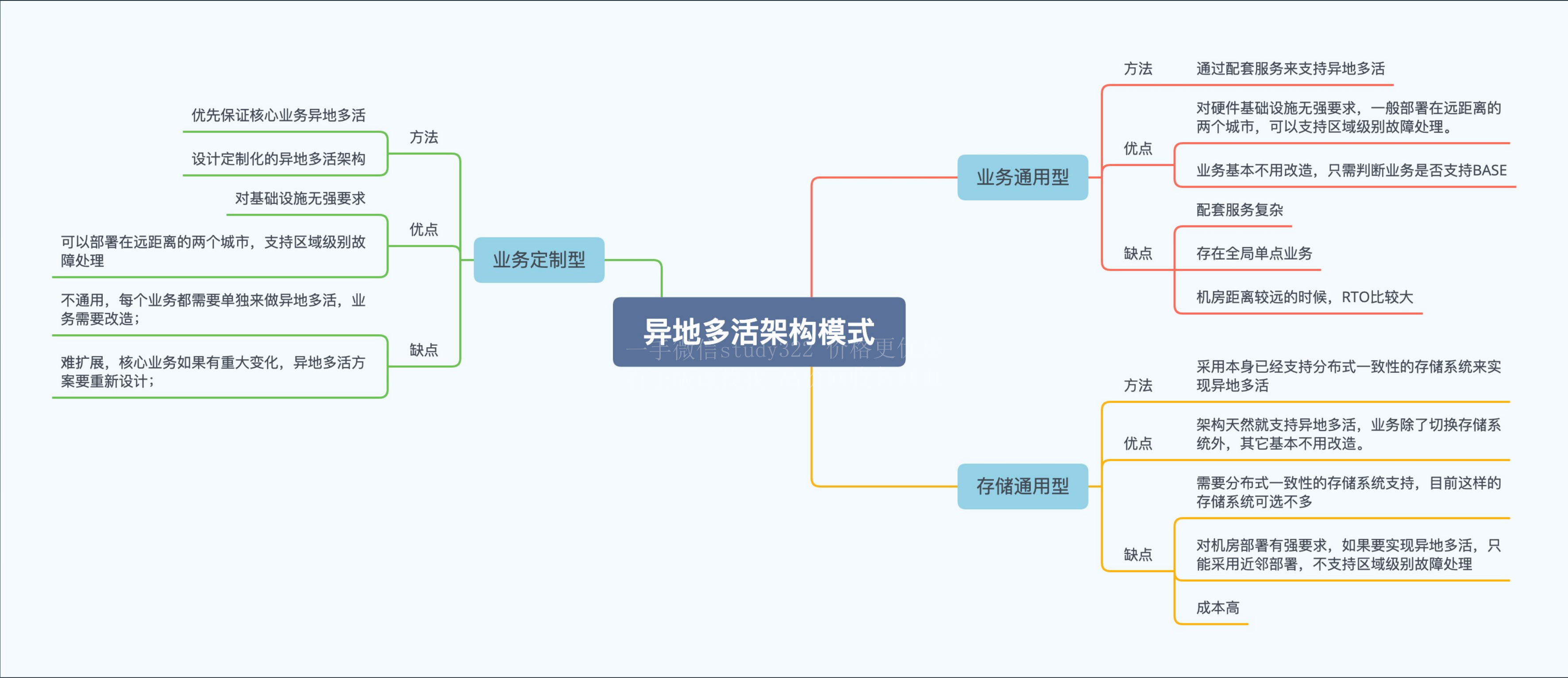


为什么每个分区都要求三地五中心？

异地多活架构模式对比

	应对机房级别灾难	应对城市级别灾难	应对区域级别灾难	是否通用	配套服务	硬件成本	应用建议
业务定制型	是	是	是	否	低，只需很少配套服务	低，无特别要求	中小公司 中小业务
业务通用型	是	是	是	是	高，强大的配套服务	低，无特别要求	大公司，业务无强一致性，例如视频、资讯、微博、社交类业务
存储通用型	是	是	否，要求近邻部署	是	低，只需很少配套服务	高，对机房数量和时延有强要求	大公司，业务有强一致性要求，例如交易、支付、金融类业务

本节思维导图



随堂测验

【判断题】

1. 业务定制型异地多活架构优先保证核心业务异地多活。
2. 业务通用型异地多活架构可以支持所有业务异地多活。
3. 存储通用型异地多活架构技术水平最高，应该优先采用。
4. 分布式强一致性存储系统如果要兼顾高性能和高可用的话，只能采取近邻部署或者同城部署。
5. 微信朋友圈采用业务通用型异地多活架构是比较合适的。

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血

【思考题】

三种异地多活架构模式可以排列组合么？



茶歇时间



八卦，趣闻，内幕.....

THANKS

一手微信study322 价格更优惠
有正版课找我 高价回收帮回血