



2019

05

08-10

北京新云南皇冠假日酒店

数据风云 十年变迁

DTCC

第十届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2019



+

○

○

○

VIPKID的数据库跨云架构设计

高级数据库专家

郝海民

目录

contents

1

历史与背景
云数据库演进之历史

2

MySQL Utilities
云数据库演进之现状

3

配置中心-VKDS
云数据库演进之现状

4

多云架构
云数据库演进之未来

DTCC 2019

第十届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2019

历史与背景



云计算的弹性 01

可根据业务的发展快速扩缩容
可使成本上最优化



弹性扩展

02 MySQL utilities

数据库的高可用，故障可隔离，可转移



异地多活

主从切换



多云容灾 04

可通过应用层、DB层在多云之间
实现容灾，流量的调度



动态配置

03 配置可动态加载

通过配置中心VKDS实现配置的动态加载，
流量可调度

故障无法隔离

多级从库

不易维护

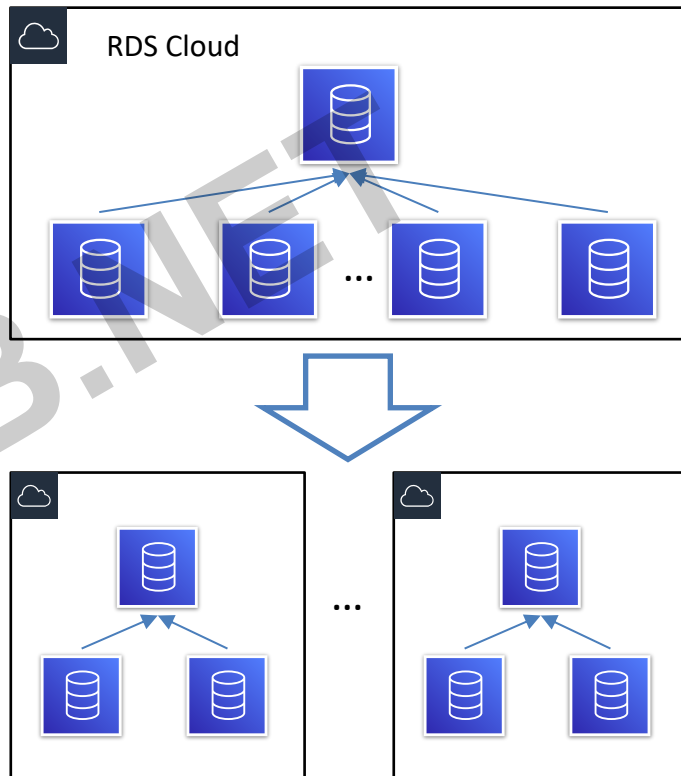
无法横向扩展

成本高

EC2自建DB

规范化

集群的元信息管理



MySQL Utilities

ITPUB.NET



MHA vs MySQL Utilities

| | MHA | MySQL Utilities |
|-------|----------|-----------------|
| 部署 | 复杂 | 简单 |
| 开发语言 | Perl | Python |
| 开发者 | 第三方 | 官方 |
| 生存周期 | 只切换一次 | 可循环 |
| 数据补齐 | 对比binlog | 对比GTID |
| 用户数量 | 多 | 少 |
| 节点数要求 | 3 | 2 |

环境要求

官方版本5.6+

开启GTID

master_info_repository = TABLE, relay_log_info_repository = TABLE

配置report_host, report_port变量

足够权限的数据库用户

建议开启半同步/增强半同步

mysqlfailover.py

daemon或者screen
主库故障

mysqlrpladmin.py

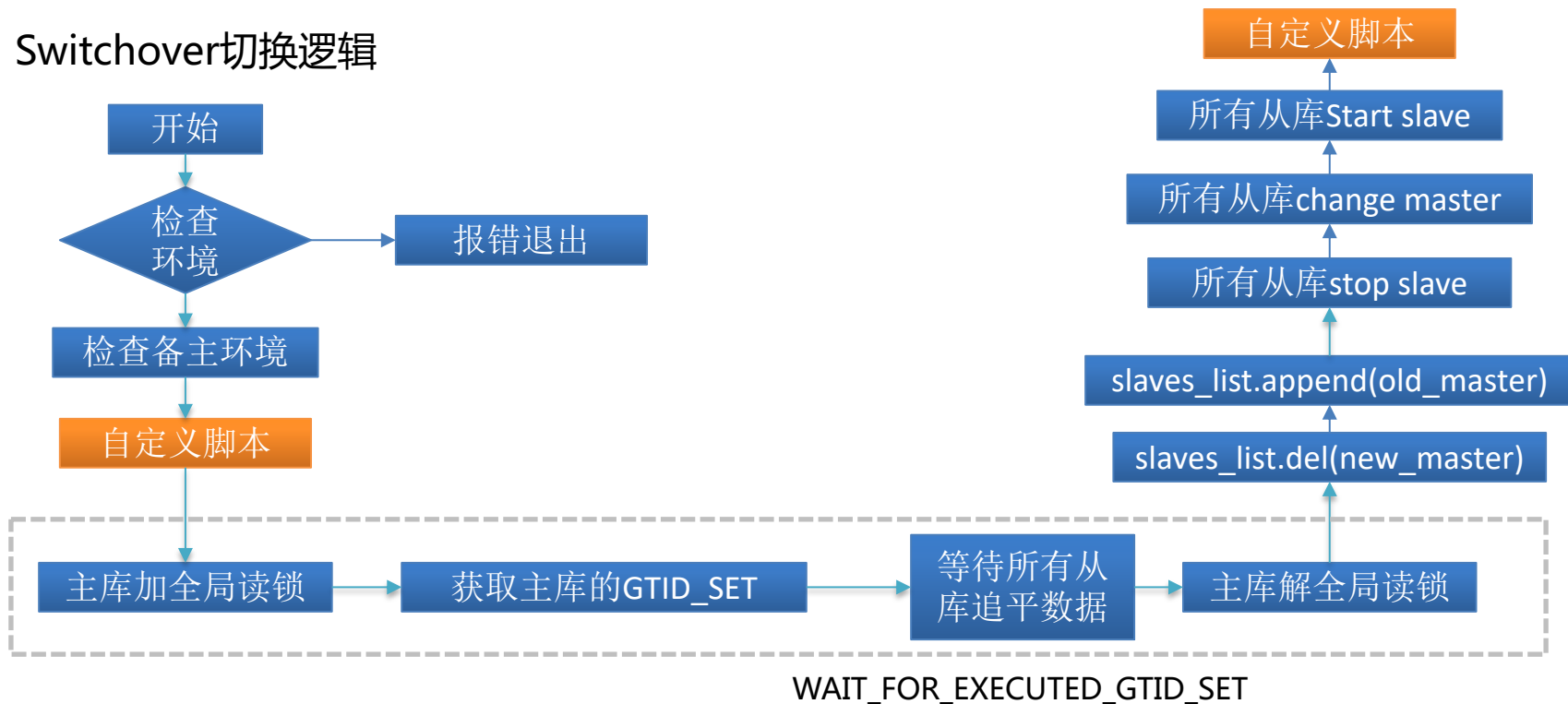
主库硬件升级
主库停机修改配置
主库故障需手动执行

...

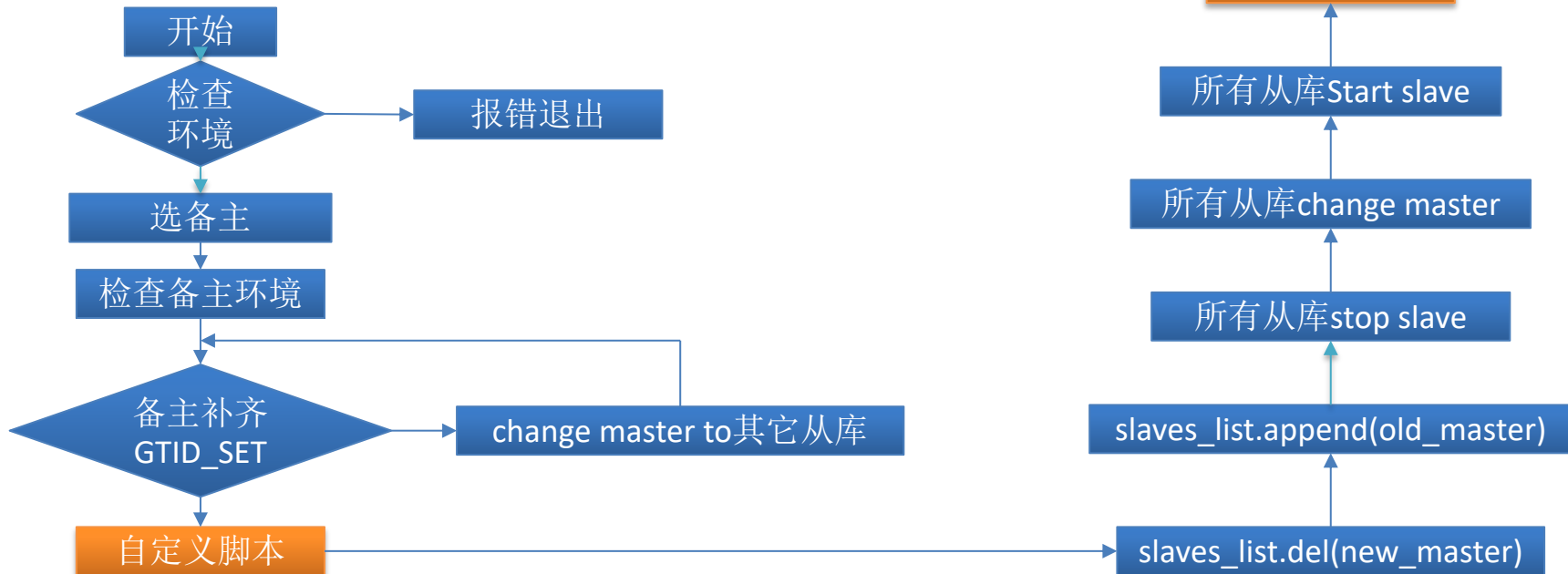
ITPUB.NET



Switchover切换逻辑



Failover切换逻辑



通过自定义脚本实现主从差异化配置以及与其它平台的通信

切换前调用

--exec-before

切换后调用

--exec-after

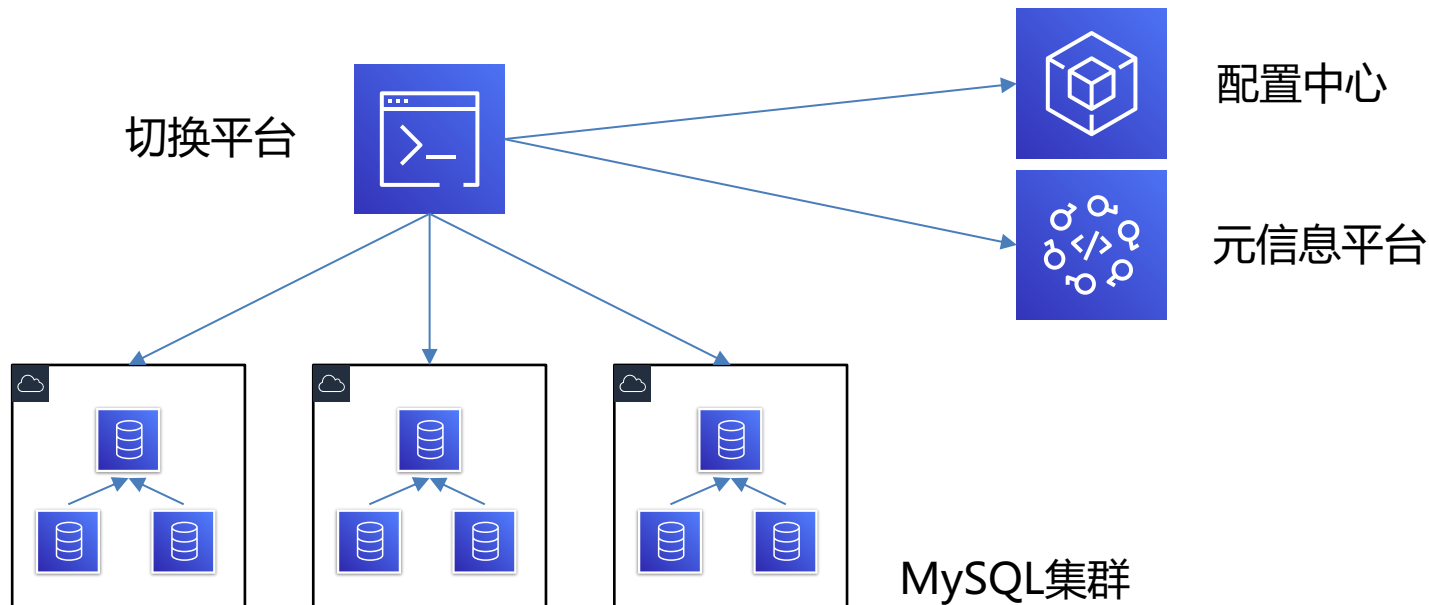


命令示例

```
/usr/bin/mysqlrpladmin --master=admin:admin@192.168.100.128:3306 \  
--slaves=admin:admin@192.168.100.129:3306,admin:admin@192.168.100.130:3306 \  
--log=/data/logs/mysqlfailover/switchover.log \  
--exec-before=/data/scripts/switchover_before.sh \  
--exec-after=/data/scripts/switchover_after.sh -vvv --demote-master switchover
```



MySQL utilites架构图

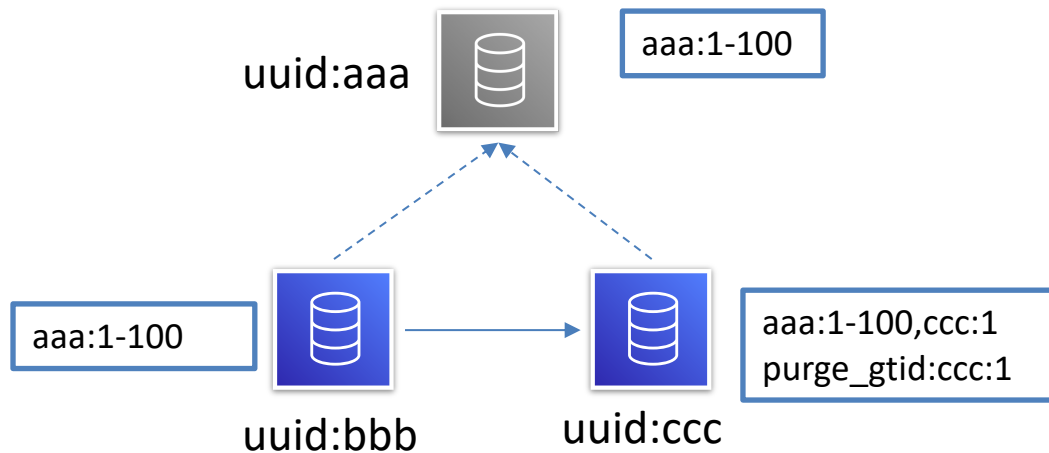


失败的案例

Change filter

Errant Transaction

数据库延迟大



如何发现Errant transaction

方式1 : `mysqlfailover -master=xxx@xxx -slaves=xxx@xxx,xxx@xxx -p`

方式2 : 使用`gtid_subtract`函数



如何解决Errant transaction

建议

在主库手动插入空事务

mysqlslavetrx在从库上插入空事务

不建议

手动删除从库的Errant Transaction

- 1.把mysql.gtid_executed里多余的删除了，然后kill -9
- 2.再使用innodb_force_recovery=6启动



配置中心-VKDS





VKDS大禹-数据库中间件

VKDS (VIPKID Database Service) —— VIPKID 分布式数据访问层中间件, 旨在为全公司提供一个通用数据访问层服务, 支持 MySQL 动态数据源、读写分离、分布式唯一主键生成器、分库分表、动态化配置等功能, 并且支持从客户端角度对数据源的各方面 (比如连接池、SQL等) 进行监控, 后续考虑实时合并计算、SQL智能分析诊断等。

管理后台





读写分离

自动分发SQL，可配置权重分流，可关闭读写分离



主库故障切换

可在线动态切换主库



从库高可用

支持从库故障自动切换、摘除、自动恢复、全部故障转主库



从库负载均衡

自动均衡负载，并支持按权重负载配置



对应用端透明

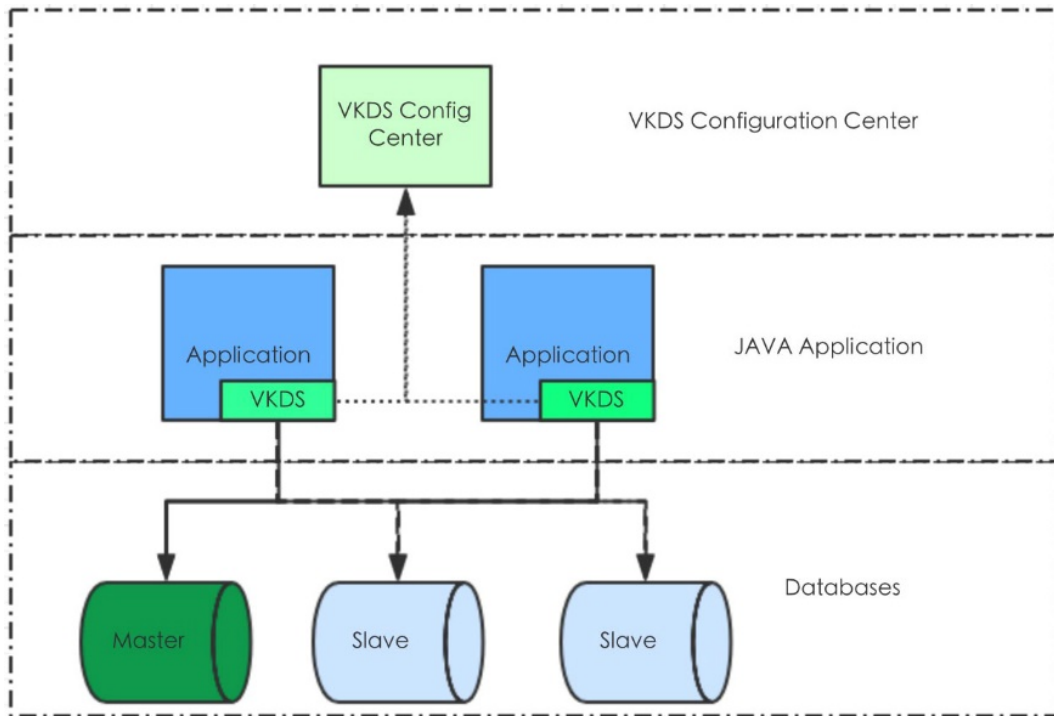
参数动态变更，主从在线动态切换，从库在线动态变更等



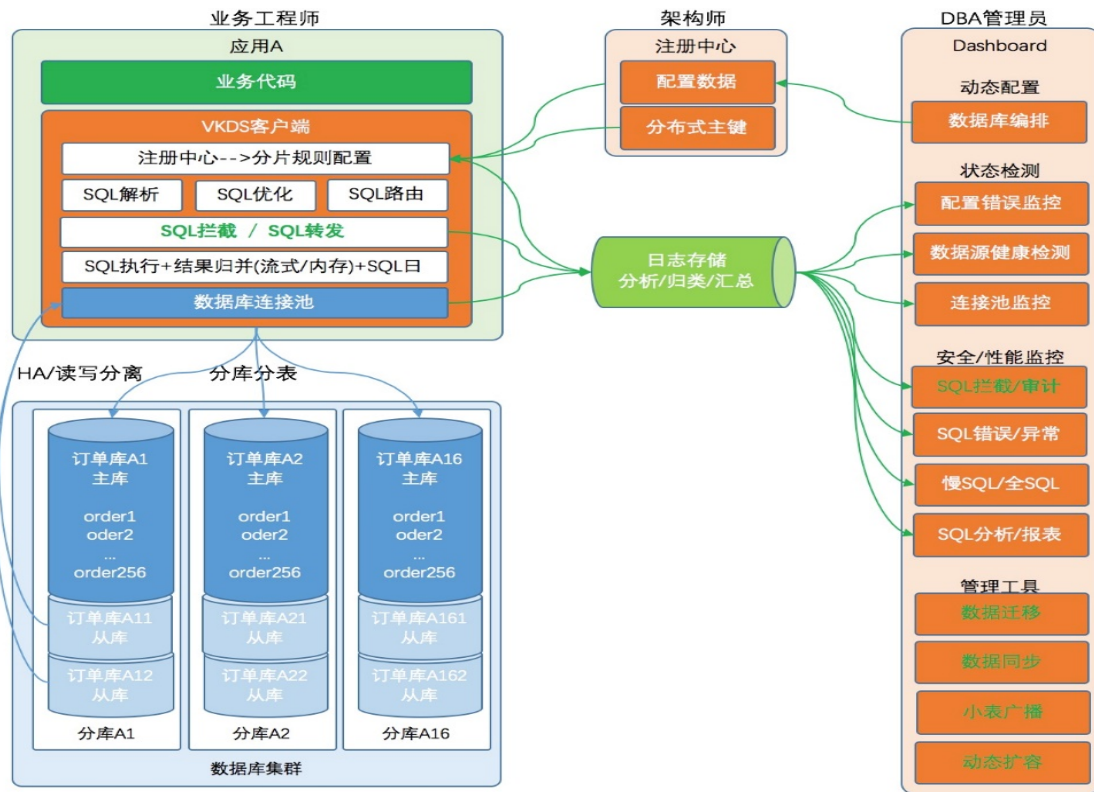
运行时监控

针对运行时慢查，错误SQL，DB连接异常实时预警

VKDS架构图1



VKDS架构图2



场景1：主库IO异常

场景2：流量的调度

场景3：从库延迟大

场景4：扩缩容数据库实例

场景5：应用连接池修改



多云架构



单云的风险

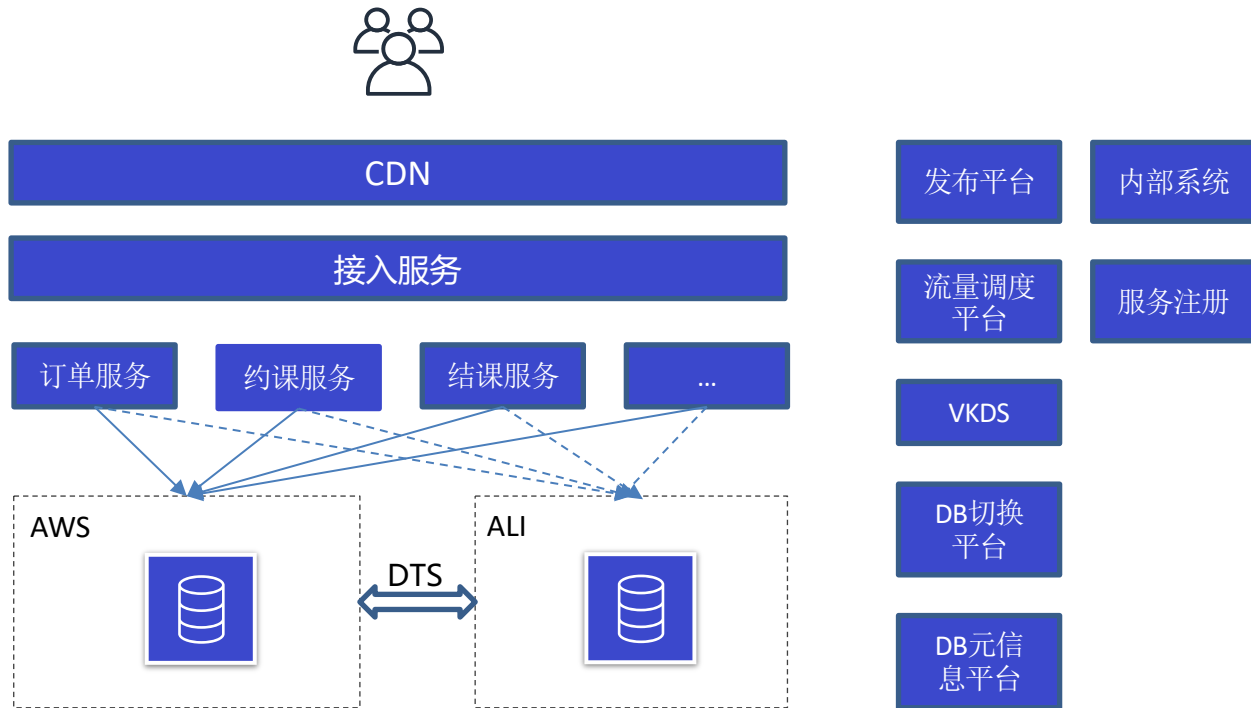
光缆被挖断

云厂商的故障

云厂商被攻击

...

保护光缆
人人有责
运维部宣



目标与展望

多云多活

容器化

动态扩缩容





THANKS