



Postgres中国技术大会 2017 (PG大象会)

# PPT分享

Postgre中国用户会

## ◆ QQ交流群

PostgreSQL专业2群: 100910388

PostgreSQL专业3群: 150657323

PostgreSQL专业4群: 461170054

◆ 文档翻译群: 309292849

◆ 欢迎投稿: [press@postgres.cn](mailto:press@postgres.cn)

## ◆ 微信公众号



官方微信公众号

## ◆ 新浪微博



官方微博

## ◆ 官方网站

[www.postgres.cn](http://www.postgres.cn)

## ◆ FaceBook

China PostgreSQL User Group

## ◆ Twitter

China PostgreSQL User Group



# TBase 可靠性探究



嘉宾：赵海明

公司：腾讯

邮箱：ylhzhao@tencent.com

手机：15110093390

腾讯·互联网+



**1**

**TBase 简介**

**2**

**可靠性设计背景**

**3**

**可靠性解决方案**

**4**

**可靠性运营效果**



1

**TBase 简介**

2

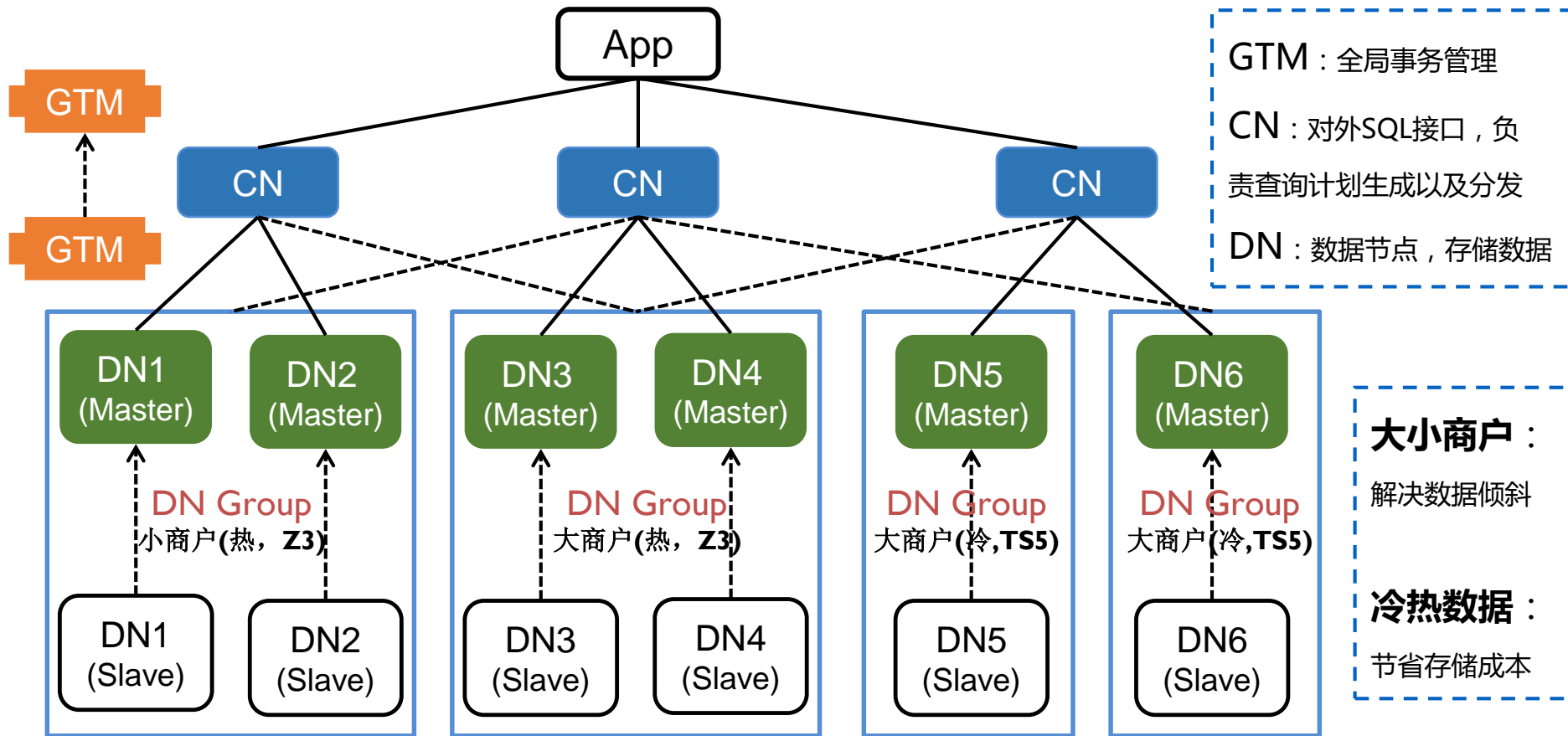
**可靠性设计背景**

3

**可靠性解决方案**

4

**可靠性运营效果**





**1.统一视图**：用户无感知，像使用单机版一样使用集群系统。

**2.接口丰富**：SQL99标准，兼容市场主流商用数据库语法。

**3.海量存储**：DN Group支持存储海量数据，双Key 分布灵活应对数据倾斜。

**4.市场需求**：微信，金融，保险，政企、公安等



1

**TBase 简介**

2

**可靠性设计背景**

3

**可靠性解决方案**

4

**可靠性运营效果**



### ➤ 易用性

- ✓ 组件多，部署成本高昂？
- ✓ 无监控，无法时时掌控系统运行状态？
- ✓ 无告警，怎么知道出了问题？
- ✓ 容量瓶颈，如何扩容和搬迁？

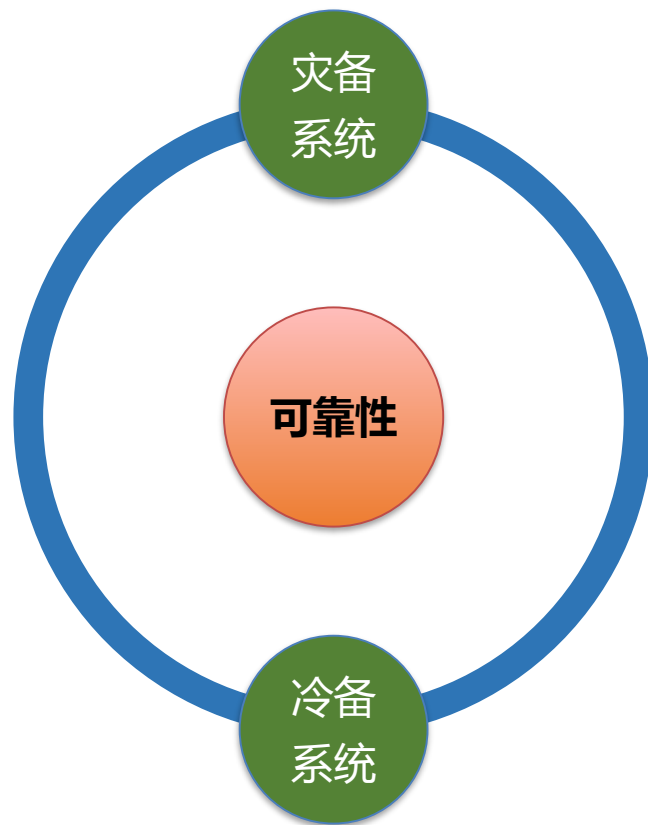
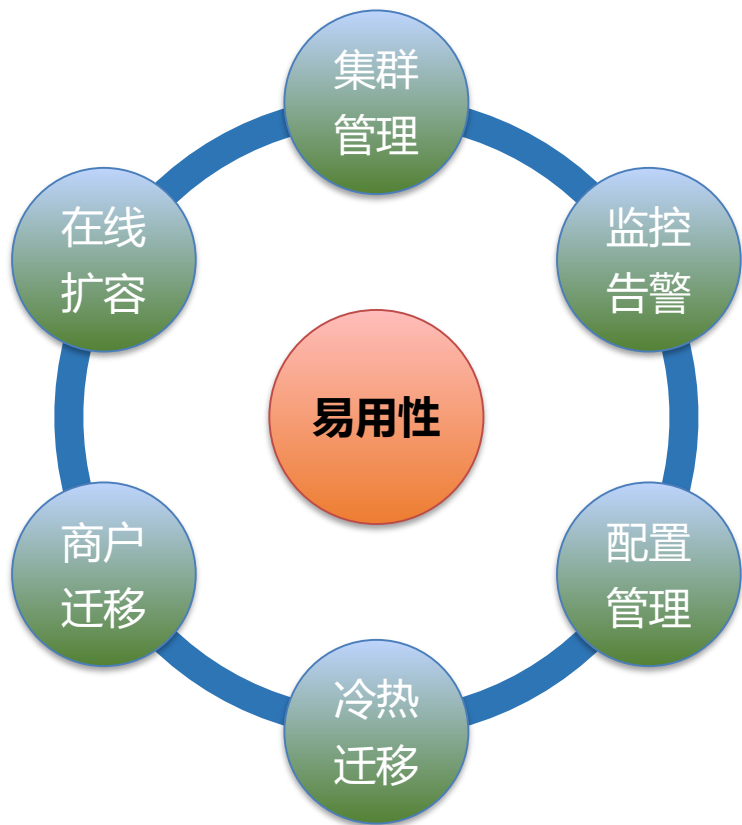
### ➤ 可靠性

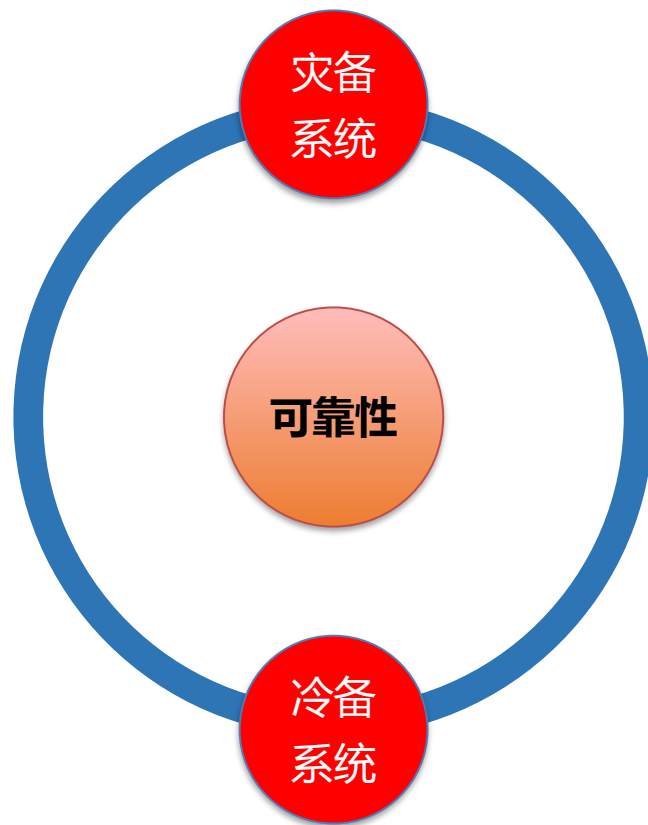
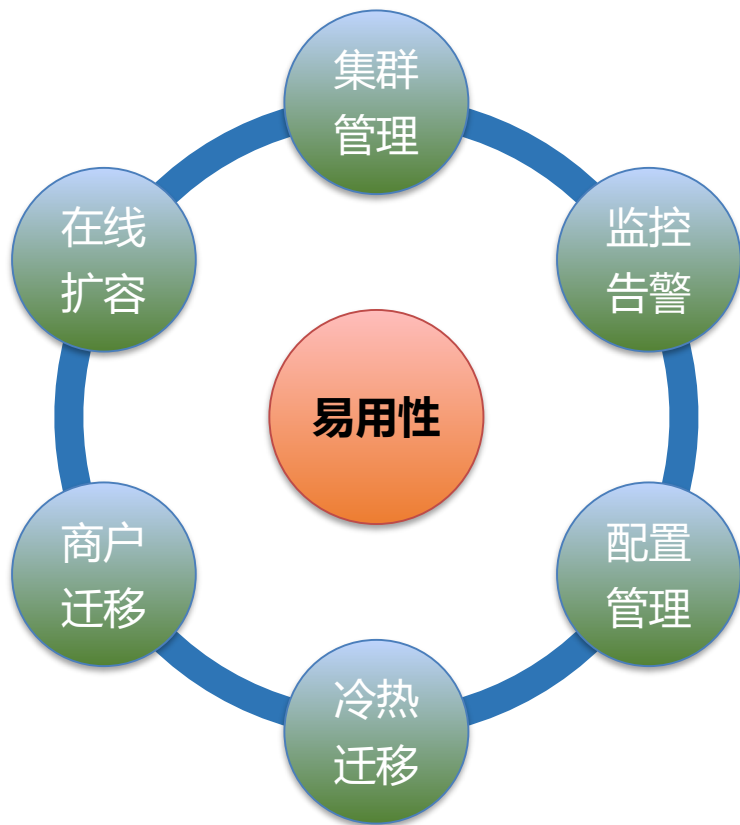
- ✓ 主机故障，数据库宕机怎么办？
- ✓ 人为因素，可怕的 `rm -rf *` ？



提升易用性

保证可靠性



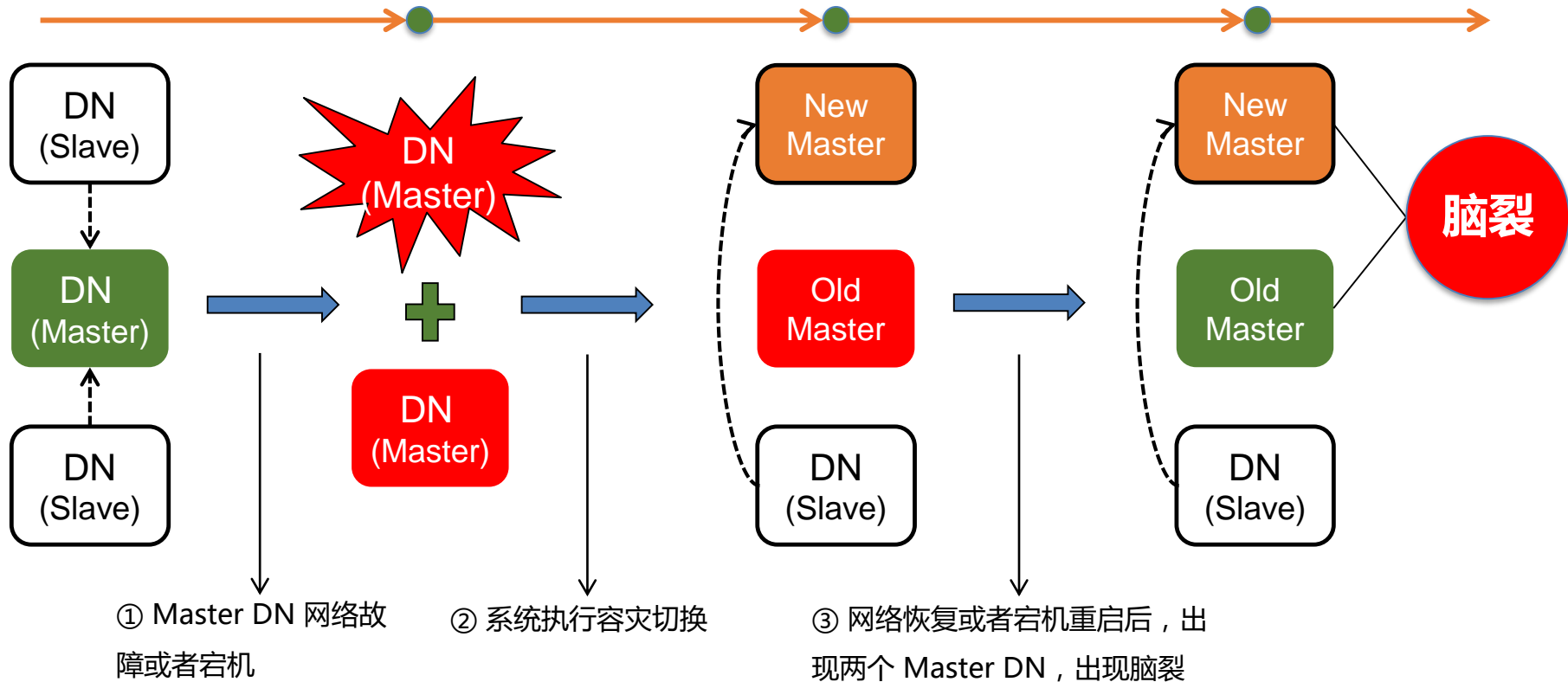




## ① 主 DN 网络故障 or 宕机

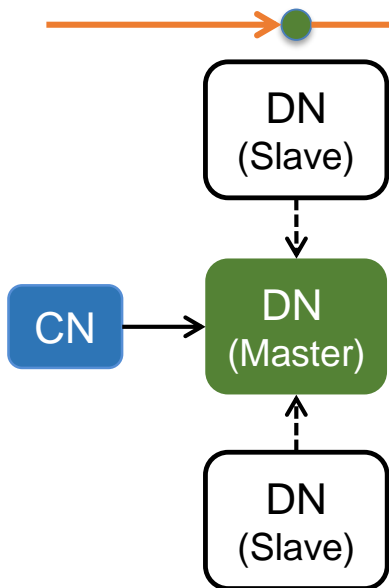
## ② 执行容灾切换

## ③ 网络恢复 or 宕机重启

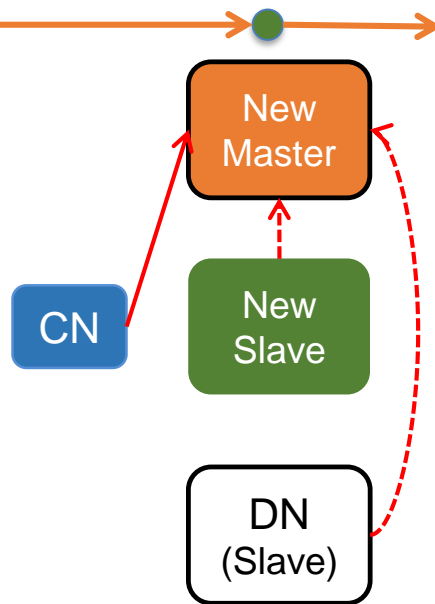


**① 切换流程**

- ① 锁 MasterDN
- ② 停止 MasterDN
- ③ 更新 Center Map
- ④ 切换 CN 路由
- ⑤ 备机升主
- ⑥ 其它备机 WAL 重连
- ⑦ 原主降备
- ⑧ 解锁原主

**② 切换前**

- 切换前，一主两备

**③ 切换后**

- 切换后，一个备机升主
- 另一个备机重新指向新主
- 原来的主机，成为新的备机



1

TBase 简介

2

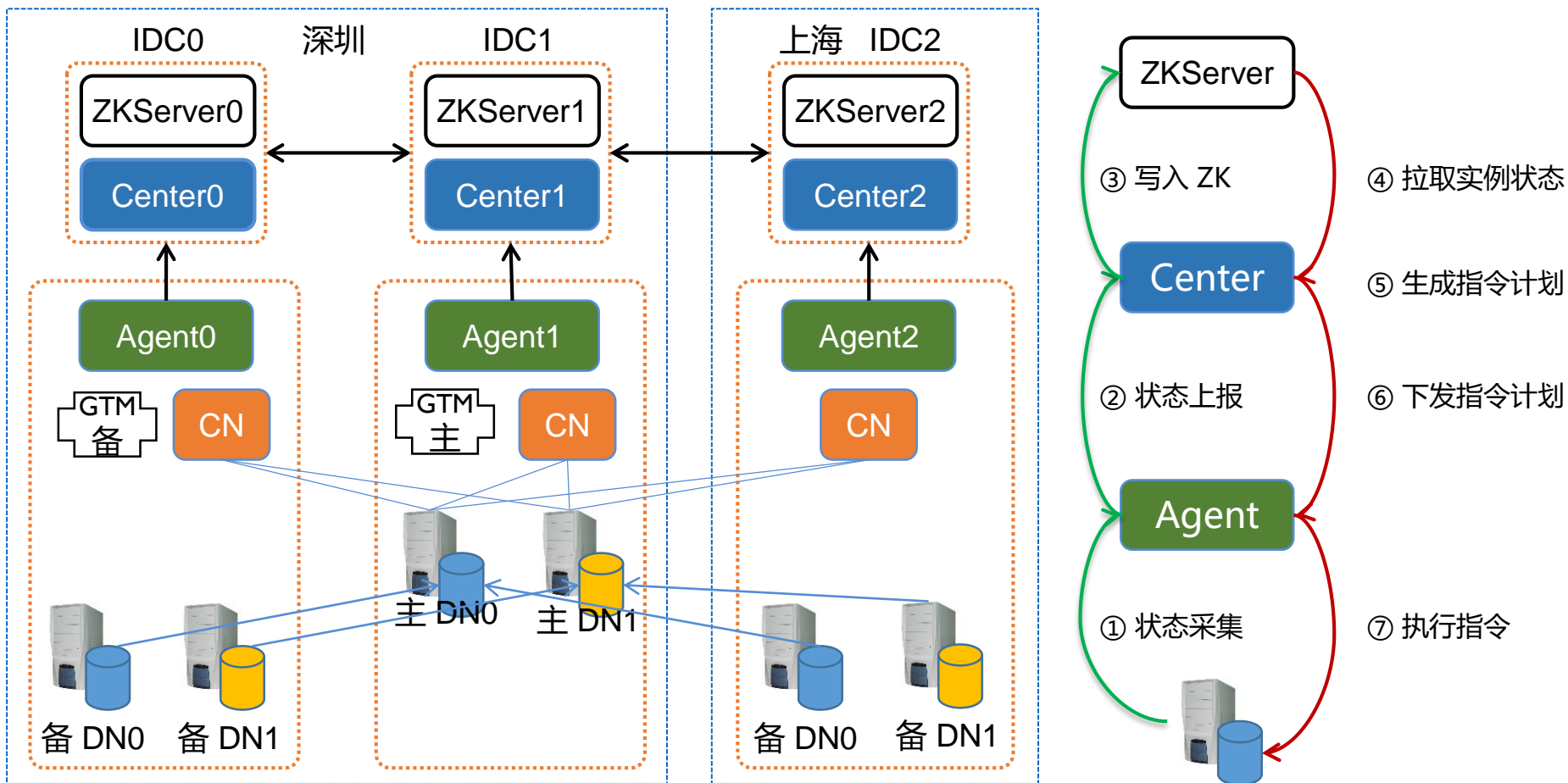
可靠性设计背景

3

可靠性解决方案

4

可靠性运营效果

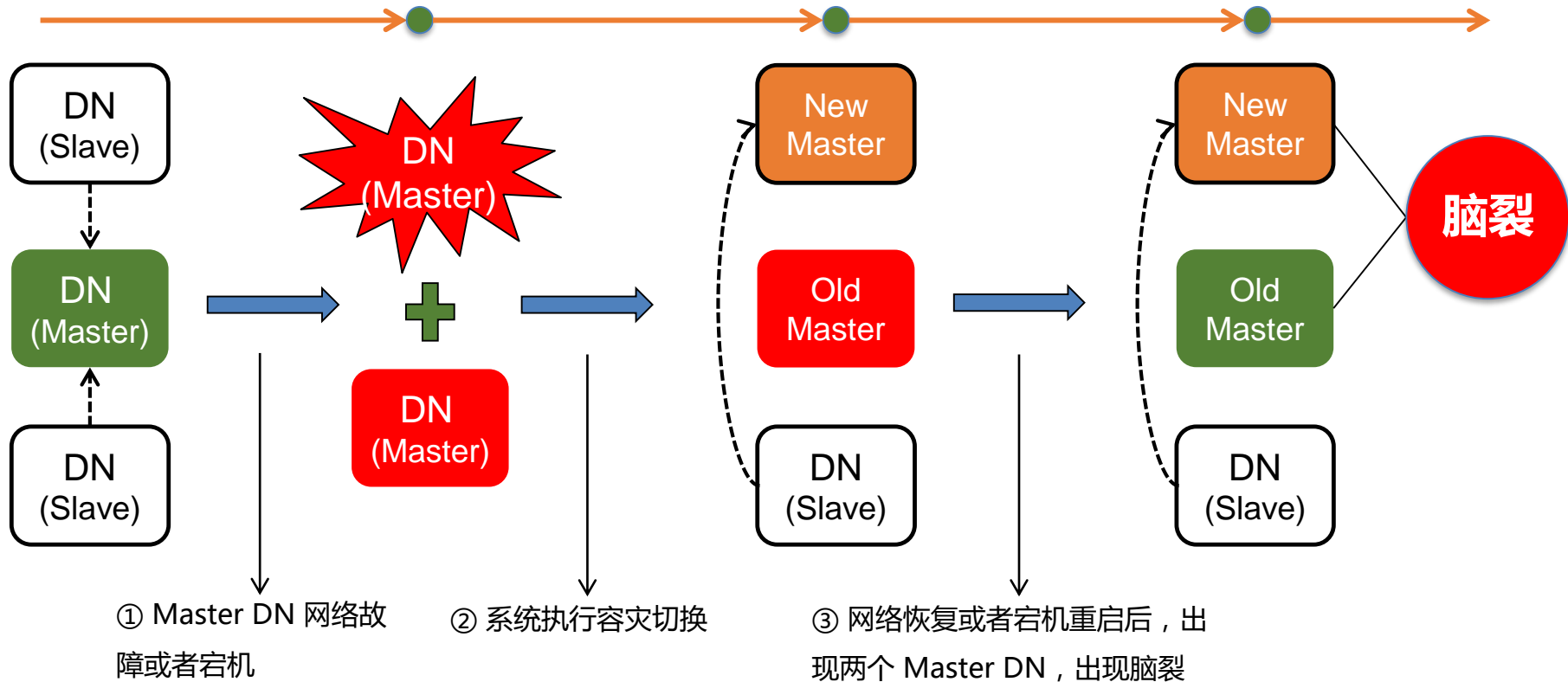


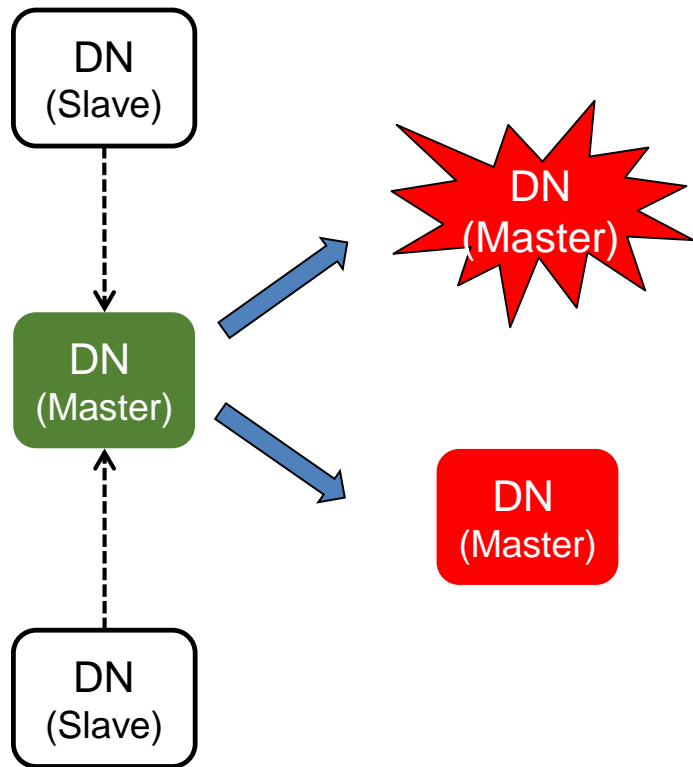


## ① 主 DN 网络故障 or 宕机

## ② 执行容灾切换

## ③ 网络恢复 or 宕机重启





- ① Master DN 网络故障  
导致容灾后脑裂的问题？

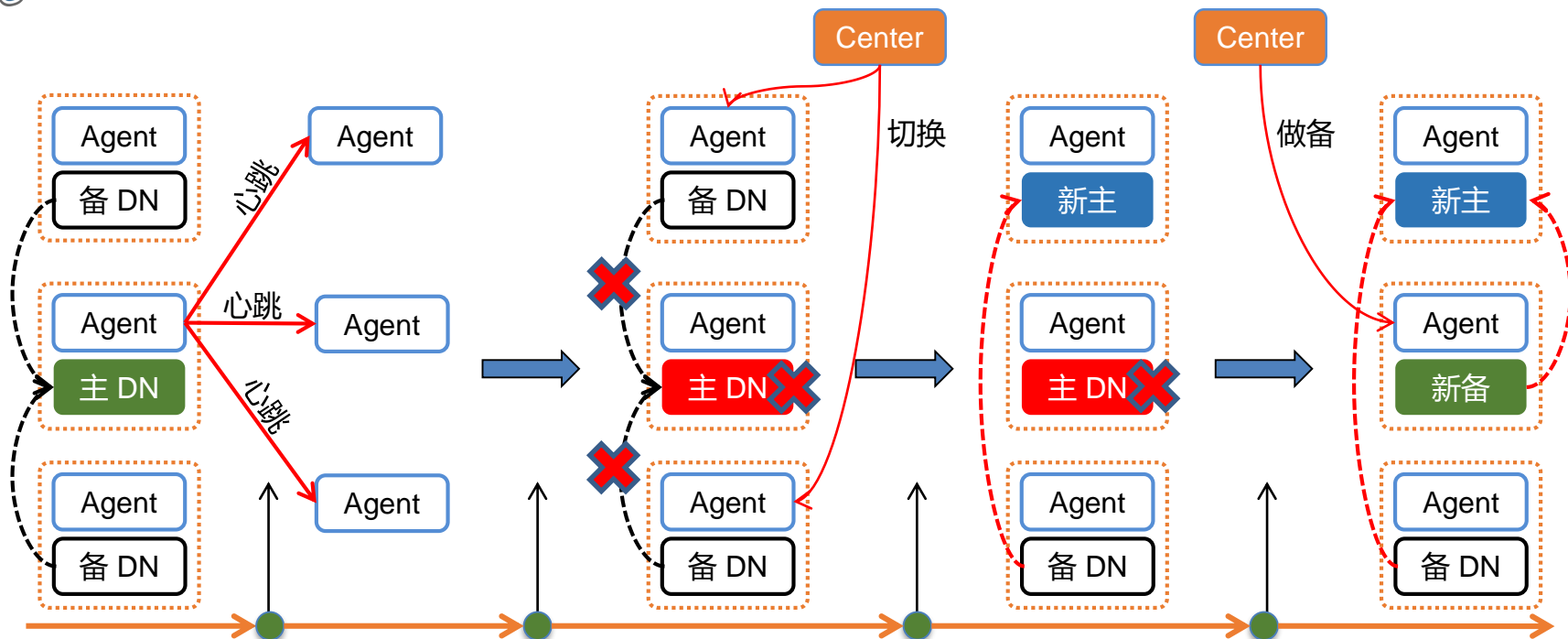


**孤岛检测**

- ② Master DN 主机宕机  
导致容灾后脑裂的问题？



**角色校验**

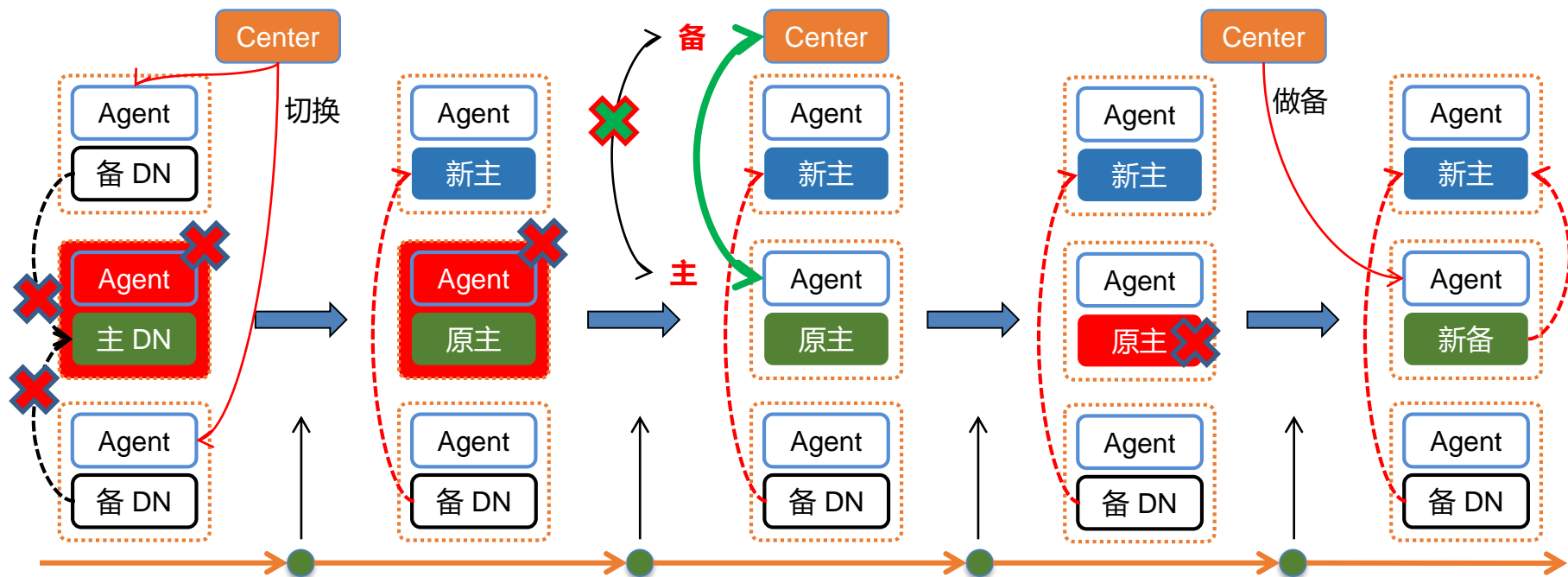


① **孤岛检测**：Agent 向其它主机发送**网络心跳**，探测本机是否成为**网络孤岛**

② **杀死实例**：Agent 发现自己成为网络孤岛，阻塞本机 Master DN

③ **容灾切换**：Center 发起容灾流程，进行主备切换，恢复读写服务

④ **恢复主备**：Agent 网络恢复后，Center 发起做备流程，恢复主备

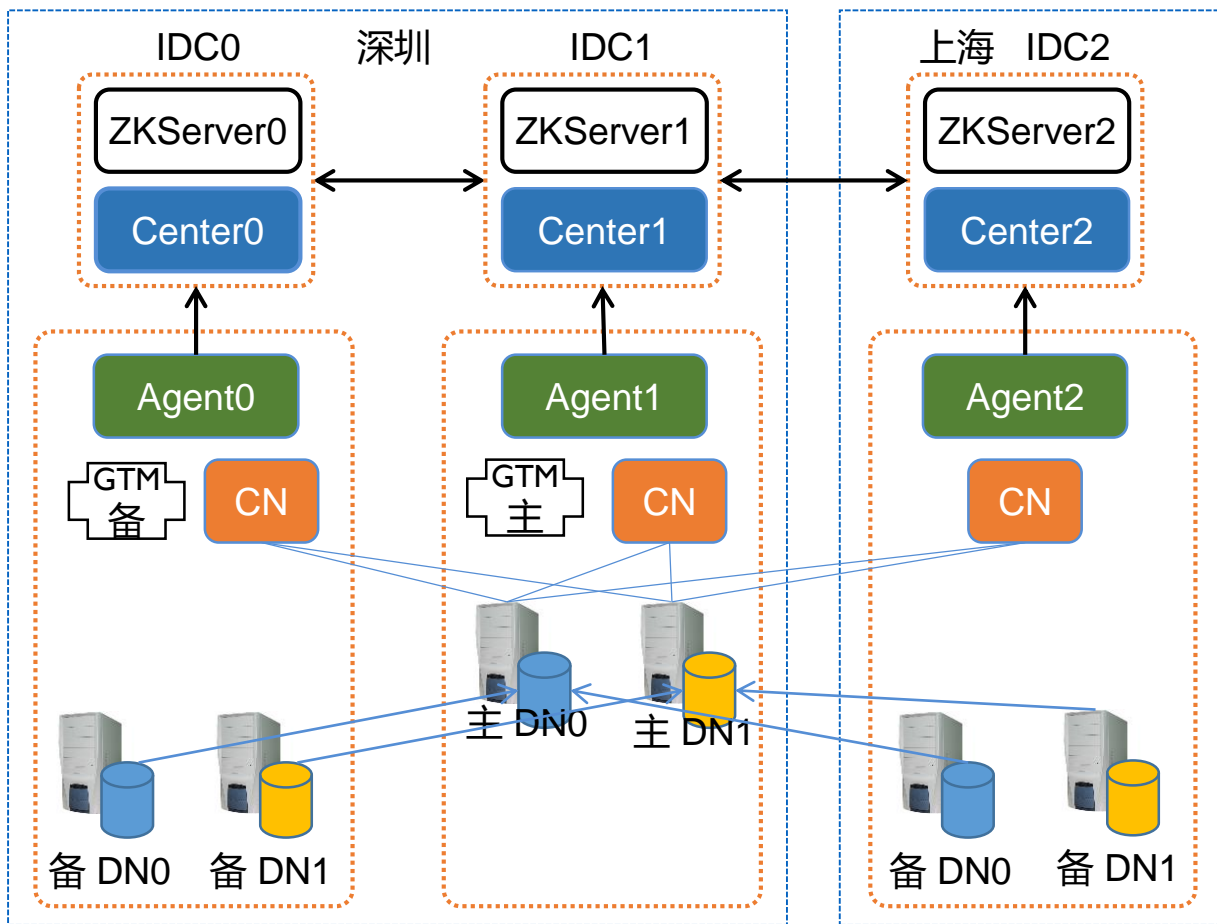


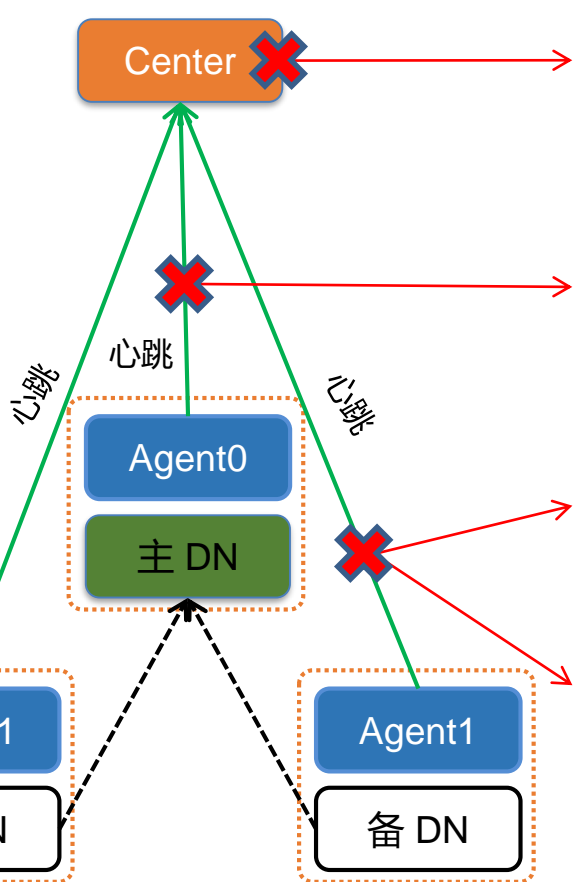
① **宕机切换**：Agent 主机宕机，Center 发起主备切换

② **角色校验**：宕机重启后，Agent 会和 Center 对 DN 实例**主备角色**进行一次校验

③ **杀死实例**：Agent 发现自己存储的角色与 Center 下发的角色冲突，阻塞原主 DN

④ **恢复主备**：Agent 杀死原主 DN 后，Center 发起做备流程，恢复主备





① **Center 宕机**：Center 在容灾过程中，出现宕机，导致容灾失败？

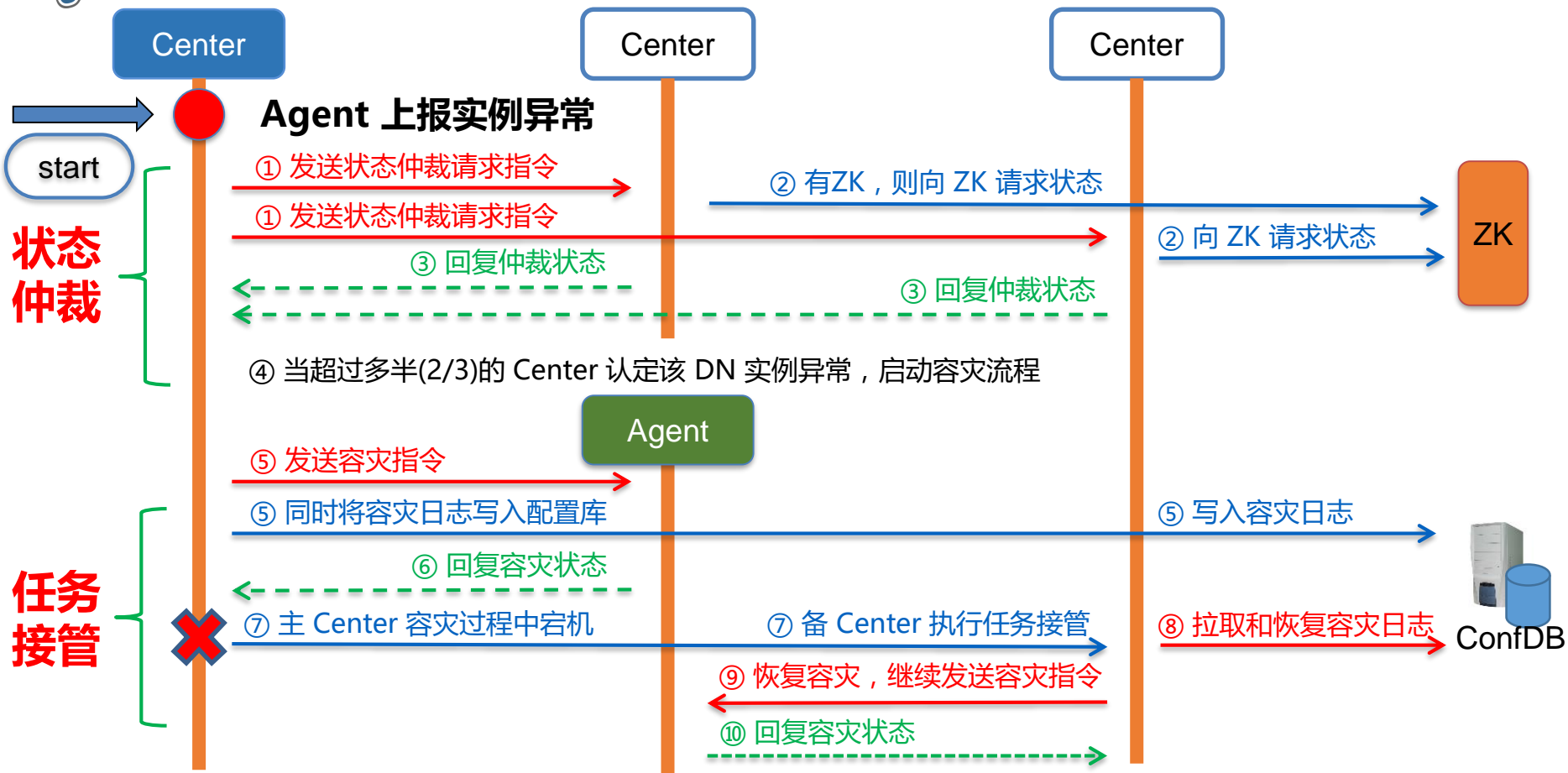
② **状态误判**：网络问题，使得 Center 对主 DN 状态误判，导致发出错误的倒换指令？

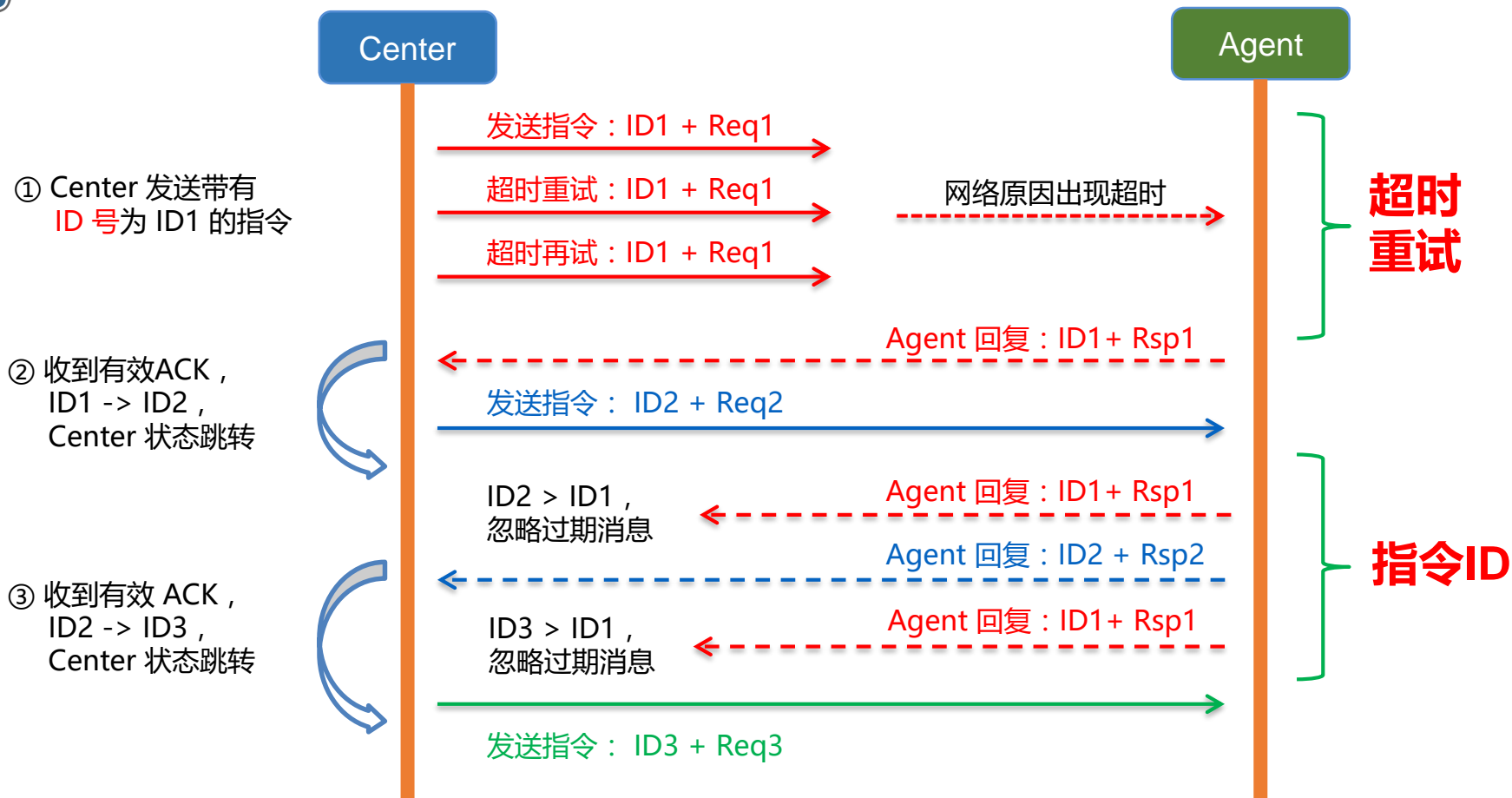
③ **指令超时**：Center 向 Agent 发送指令时，Agent 出现超时？

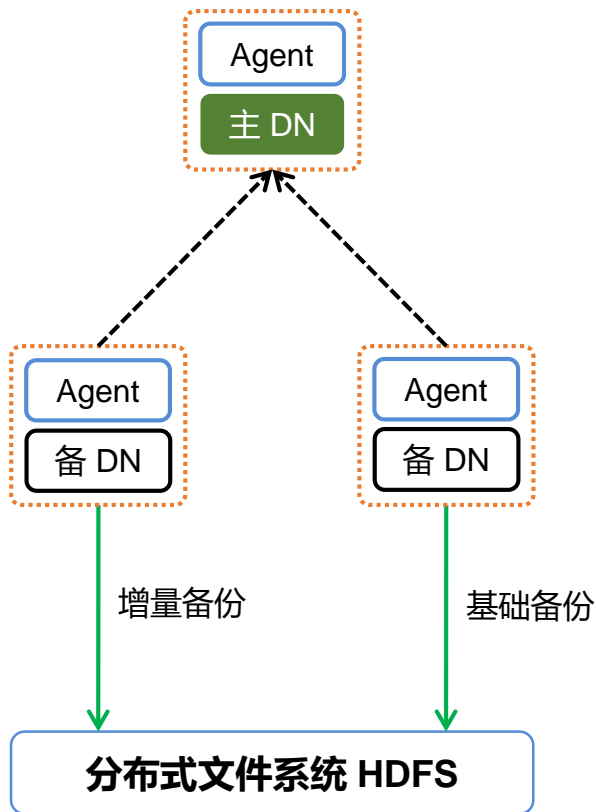
④ **指令乱序**：Agent 回复 Center 指令时，出现指令乱序？

**状态仲裁 + 任务接管**

**指令 ID + 超时重试**







## 备机做冷备

- ① **资源负载低**：备机冷备，主机无负载，业务无感知。
- ② **备份效率高**：并发冷备 + 不落盘透写 HDFS，零 I/O。
- ③ **可控性高**：网络等资源配额、随起随停。



1

**TBase 简介**

2

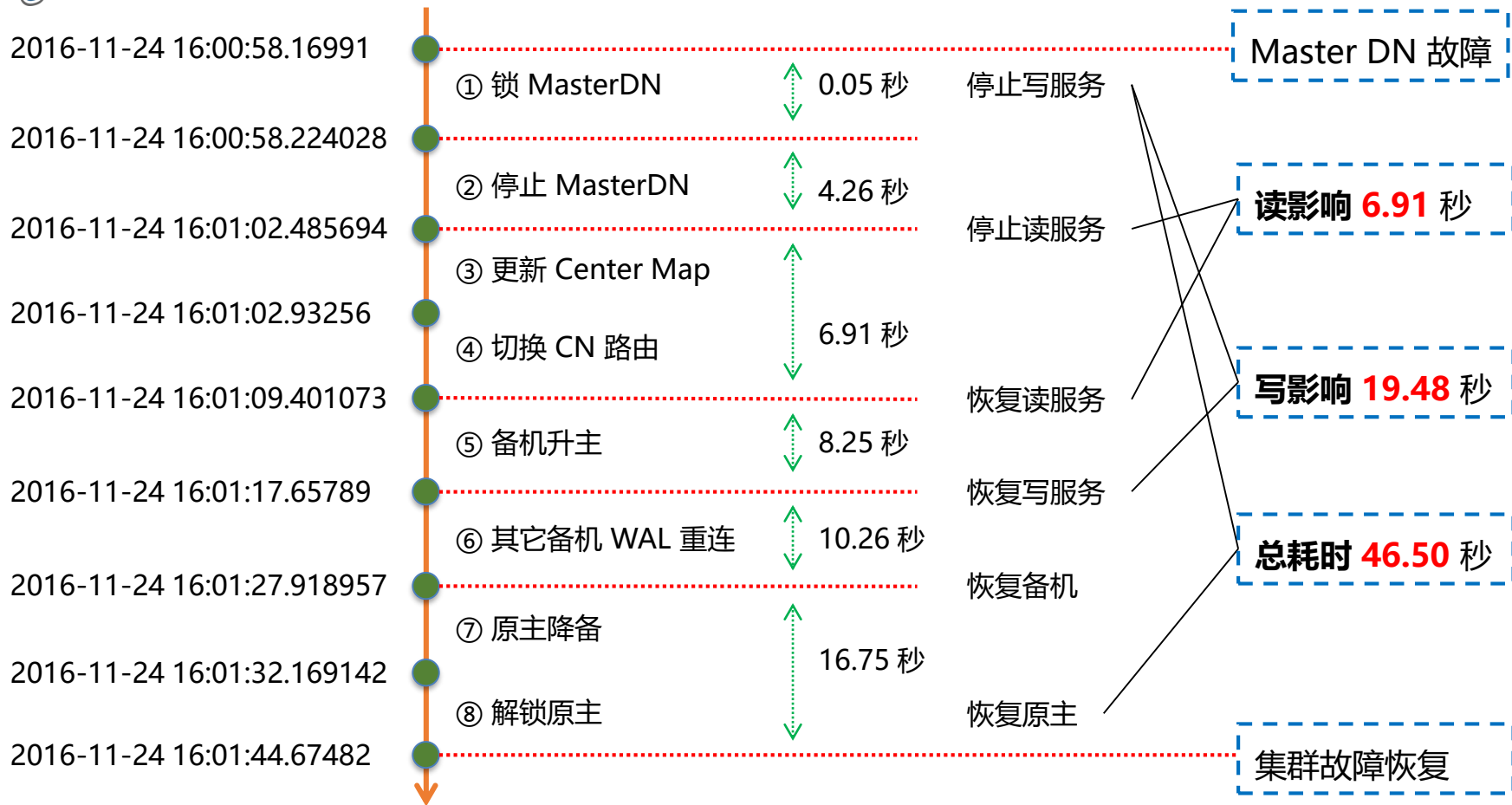
**可靠性设计背景**

3

**可靠性解决方案**

4

**可靠性运营效果**





Thanks!

