



QCon

QCon全球软件开发大会（杭州站）2012

讲师资料下载

支付宝弹性计算架构

阿玺

支付宝-技术部

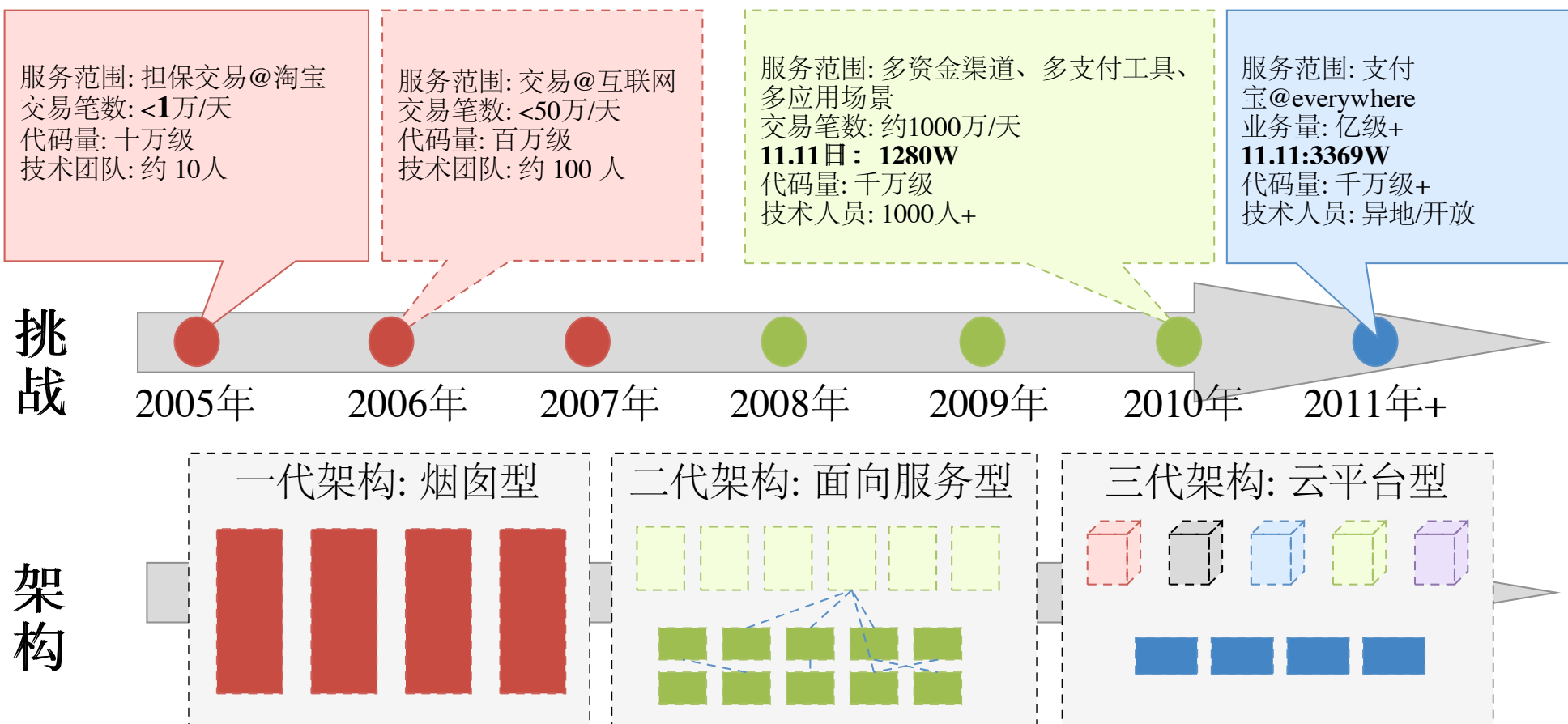
Mail: xi.hux@alipay.com

新浪微博: 支付宝_阿玺

个人介绍

- 胡喜，花名阿玺，2007年加入支付宝，主持支付平台基础技术的架构设计与研发工作，并且参与支付宝核心支付平台的架构设计和系统升级。

支付宝系统发展历程



2012.11.11系统必须具备交易处理能力：

1亿+

80亿+数据库事务

500亿+的SQL执行

1000亿+服务调用

500+个应用协同完成

我们需要什么样的架构

底层计算资源（IAAS）做到弹性是否满足？

应用层面如何做到可伸缩性？

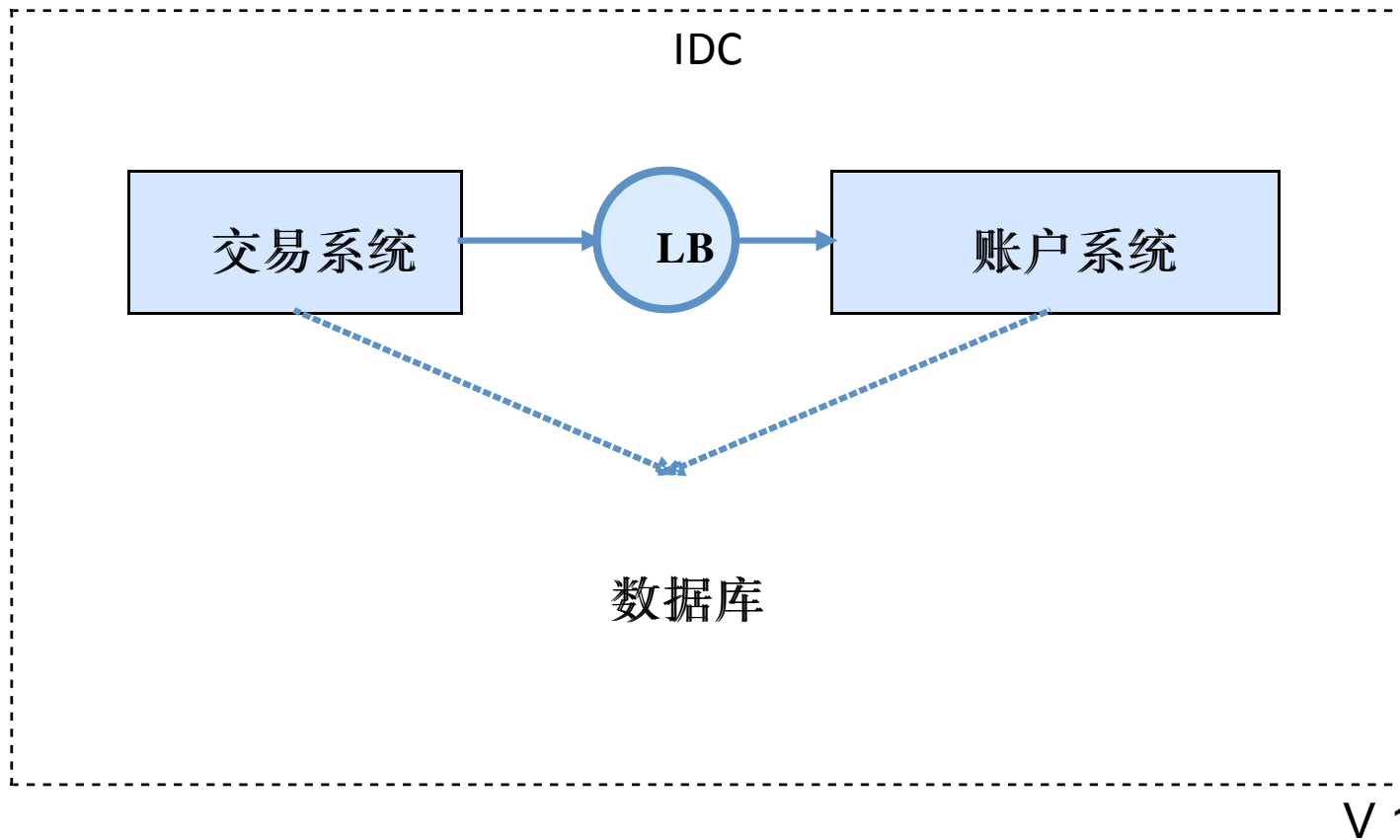
出现故障后是否能够做到快速恢复？

这一切是否能够做到自动化控制？

Agenda

- 可伸缩性：提升容量百万级到亿级
- 故障容忍：99.9%到99.99%+
- 弹性控制：人工控制到秒级自动调度

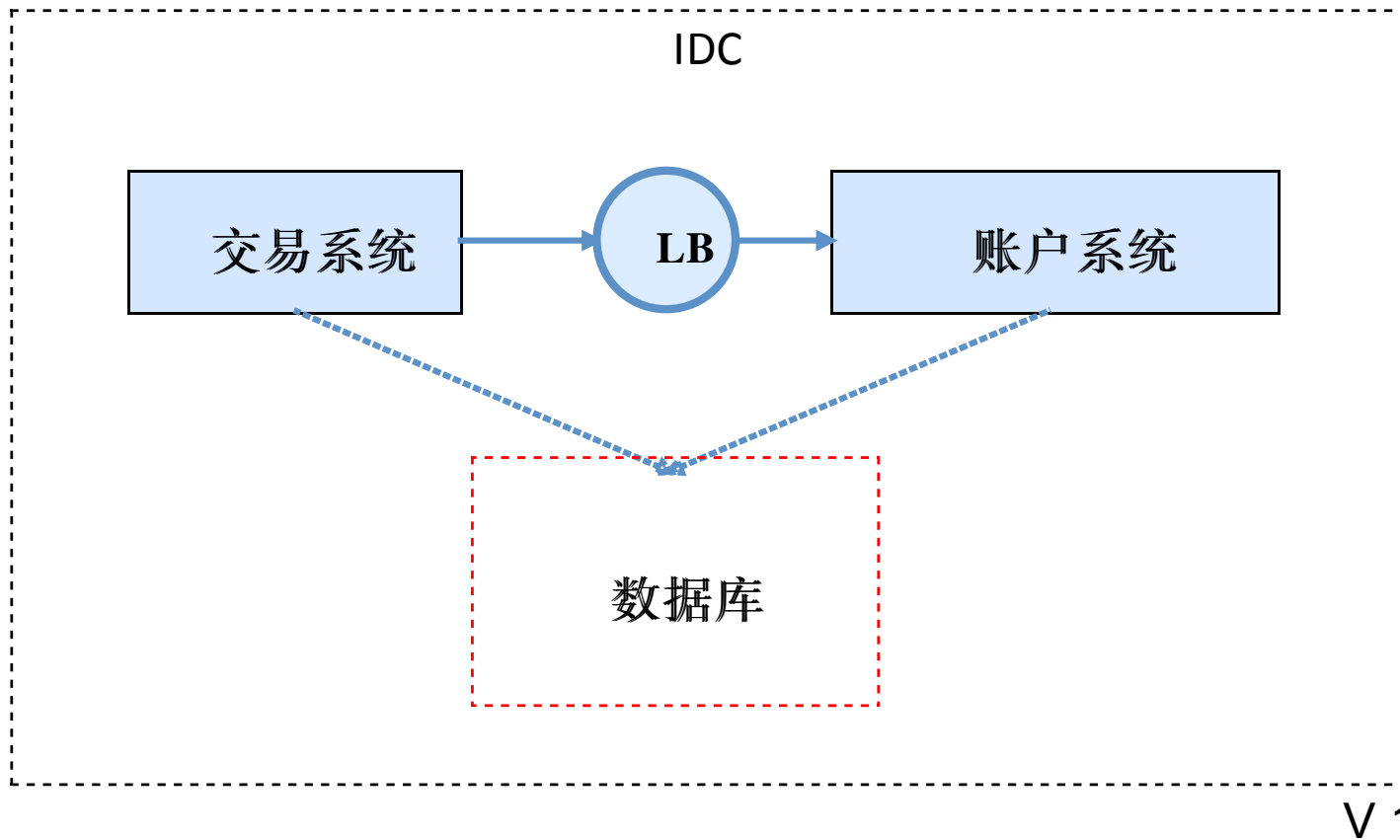
一个简化的支付宝系统模型



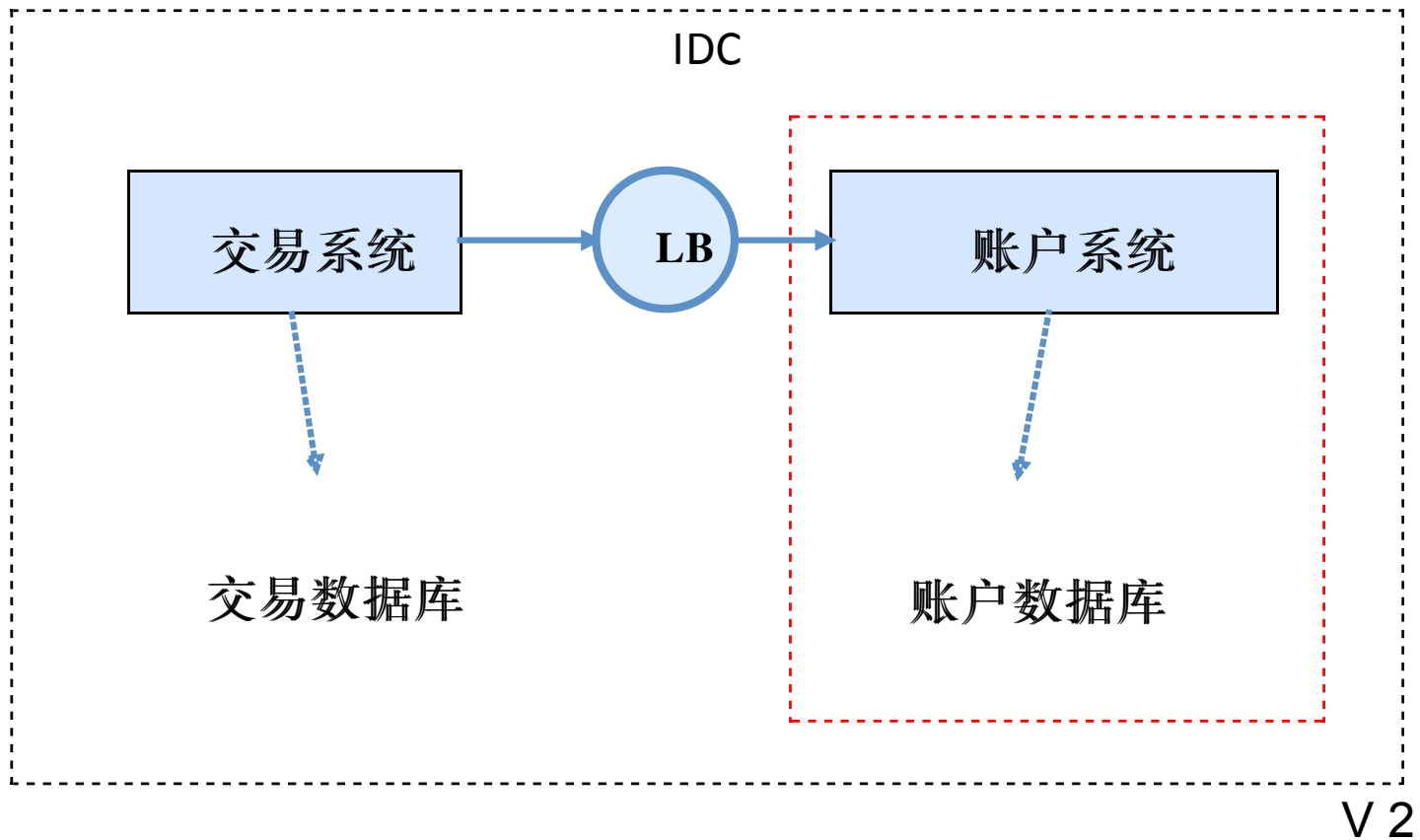
提升容量百万级到亿级

可伸缩

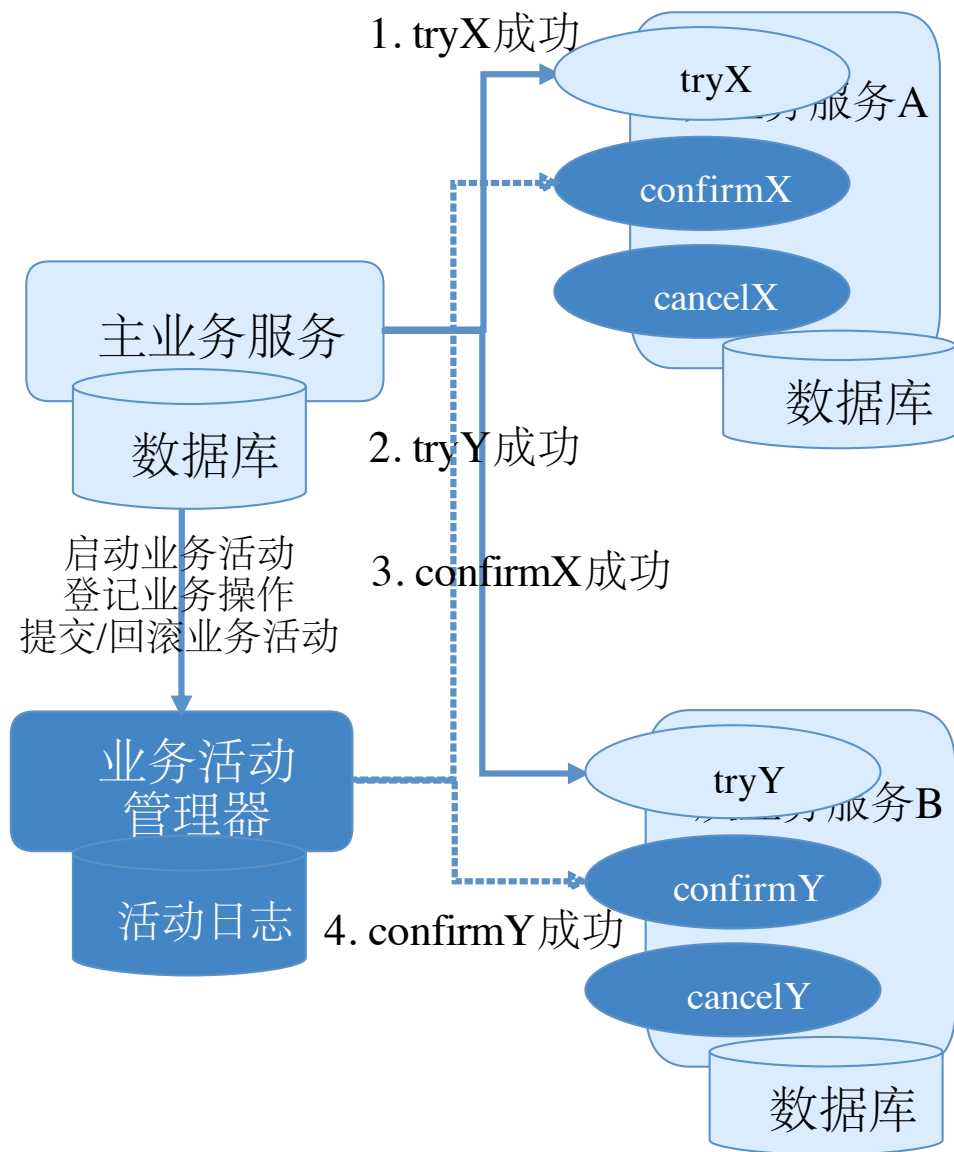
数据库的瓶颈



一致性瓶颈



业务一致性：service层的分布事务



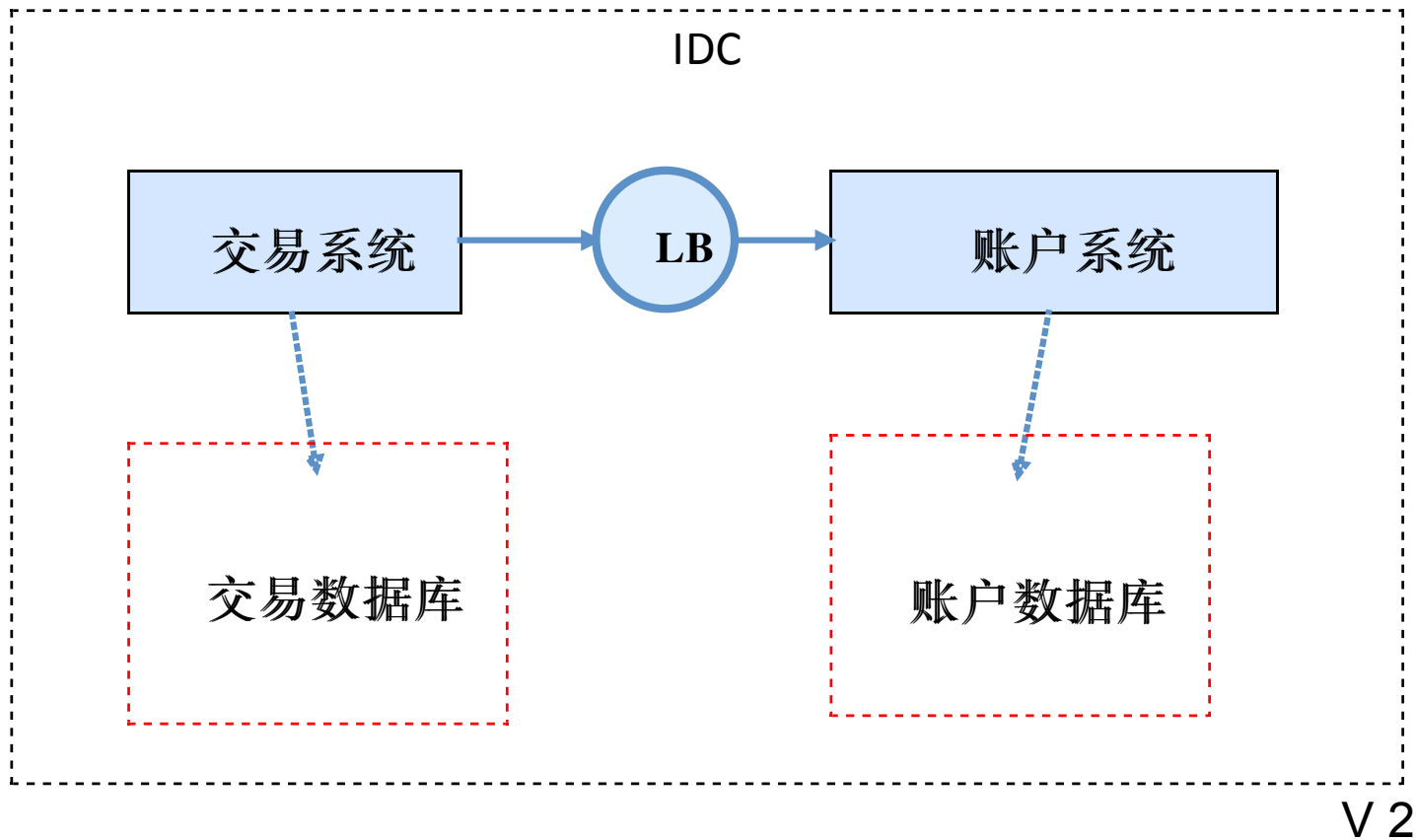
实现

- 一个完整的业务活动由一个主业务服务与若干从业务服务组成
- 主业务服务负责发起并完成整个业务活动
- 从业务服务提供TCC型业务操作
- 业务活动管理器控制业务活动的一致性，它登记业务活动中的操作，并在业务活动提交时确认所有的TCC型操作的confirm操作，在业务活动取消时调用所有TCC型操作的cancel操作

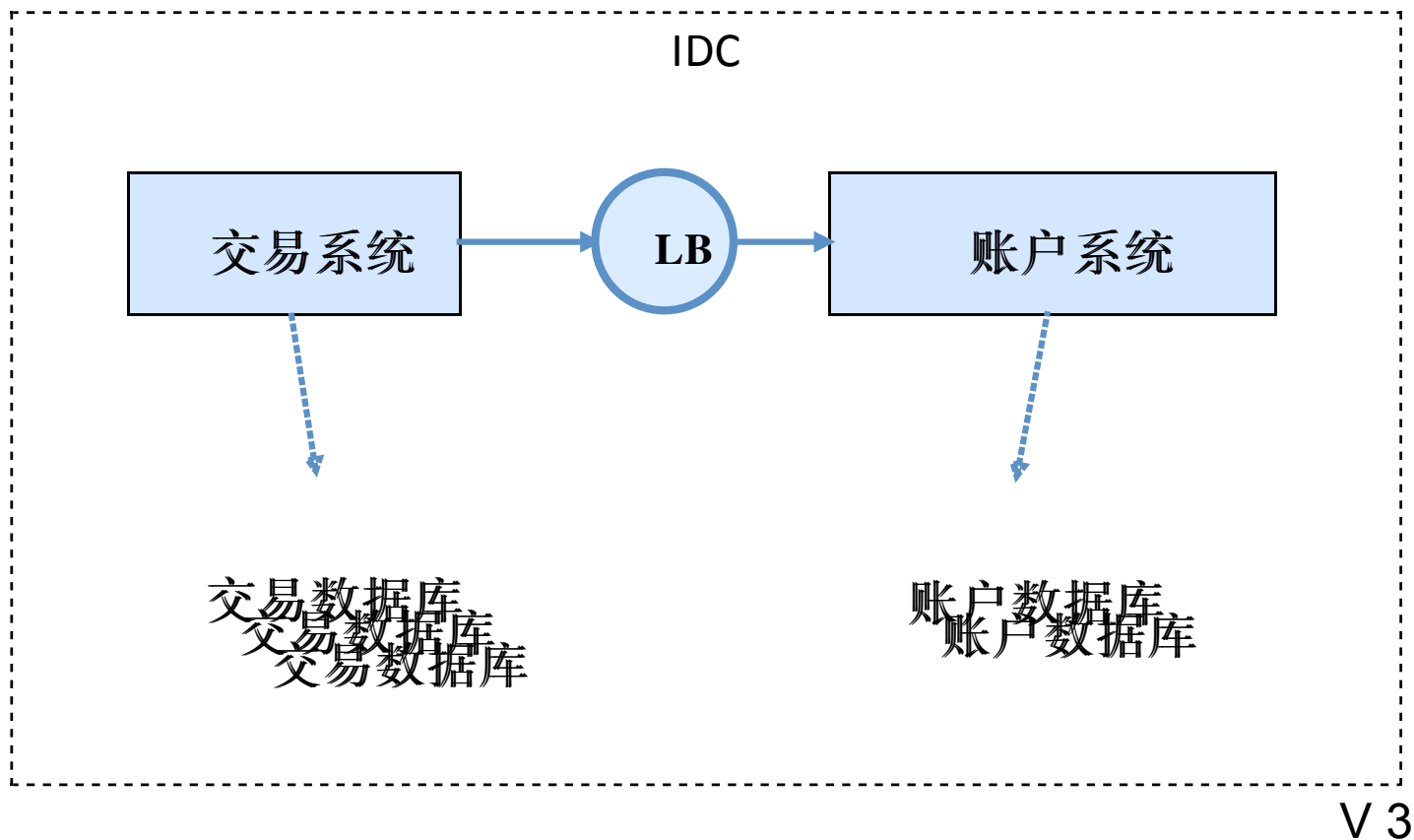
与2PC协议比较

- 没有单独的Prepare阶段，降低协议成本
- 系统故障容忍度高，恢复简单

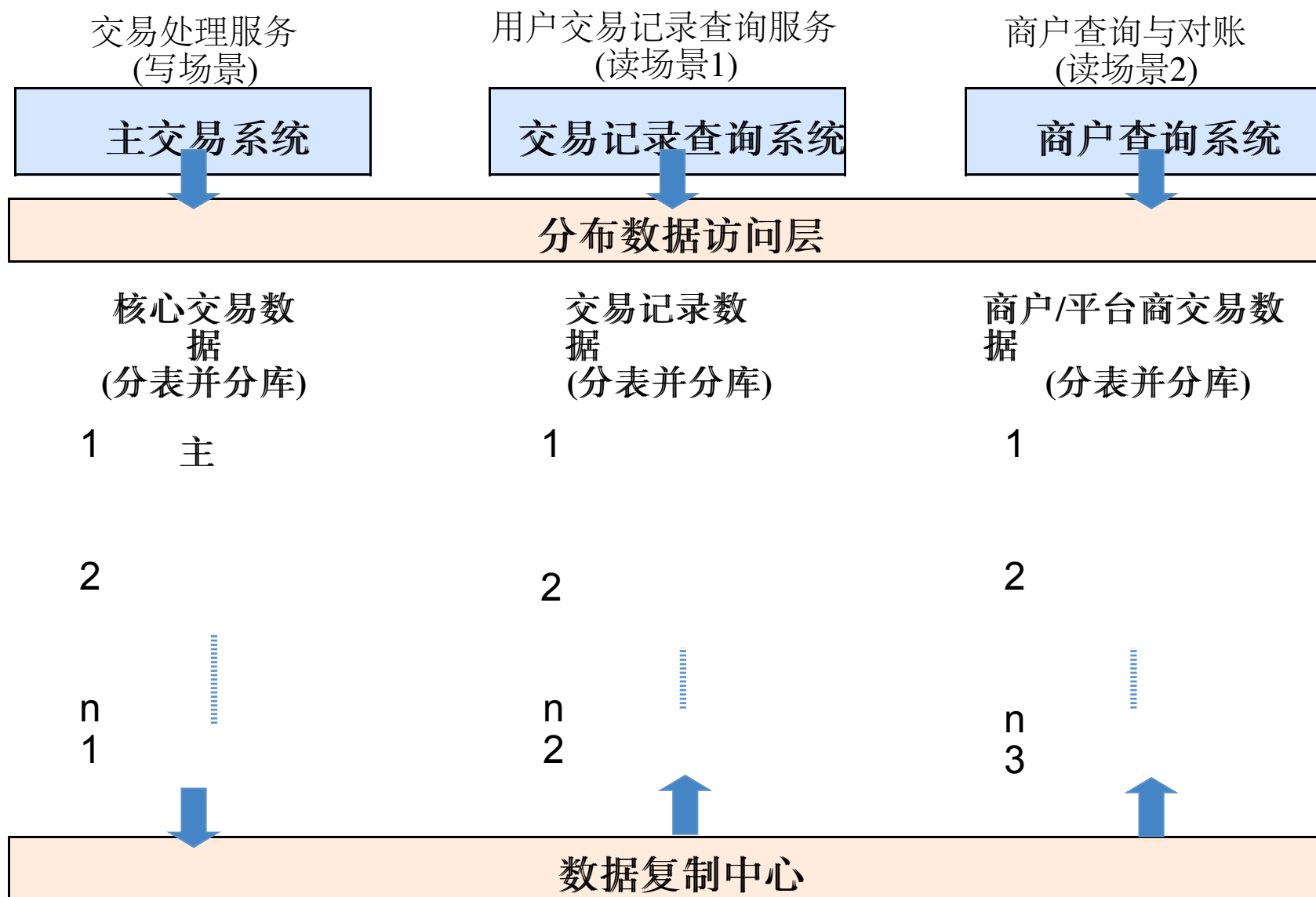
单个库的瓶颈



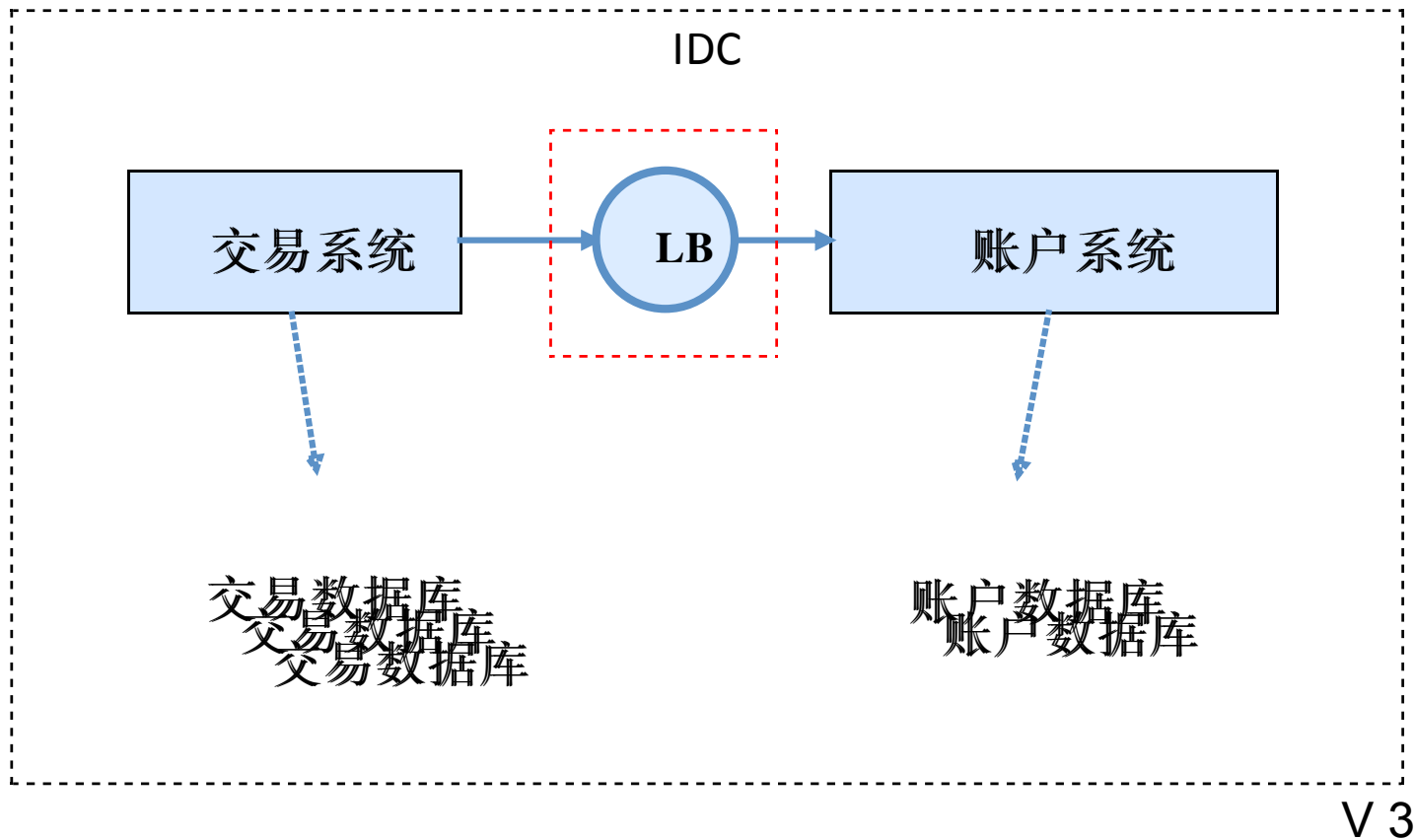
数据可伸缩性：数据水平拆分与复制



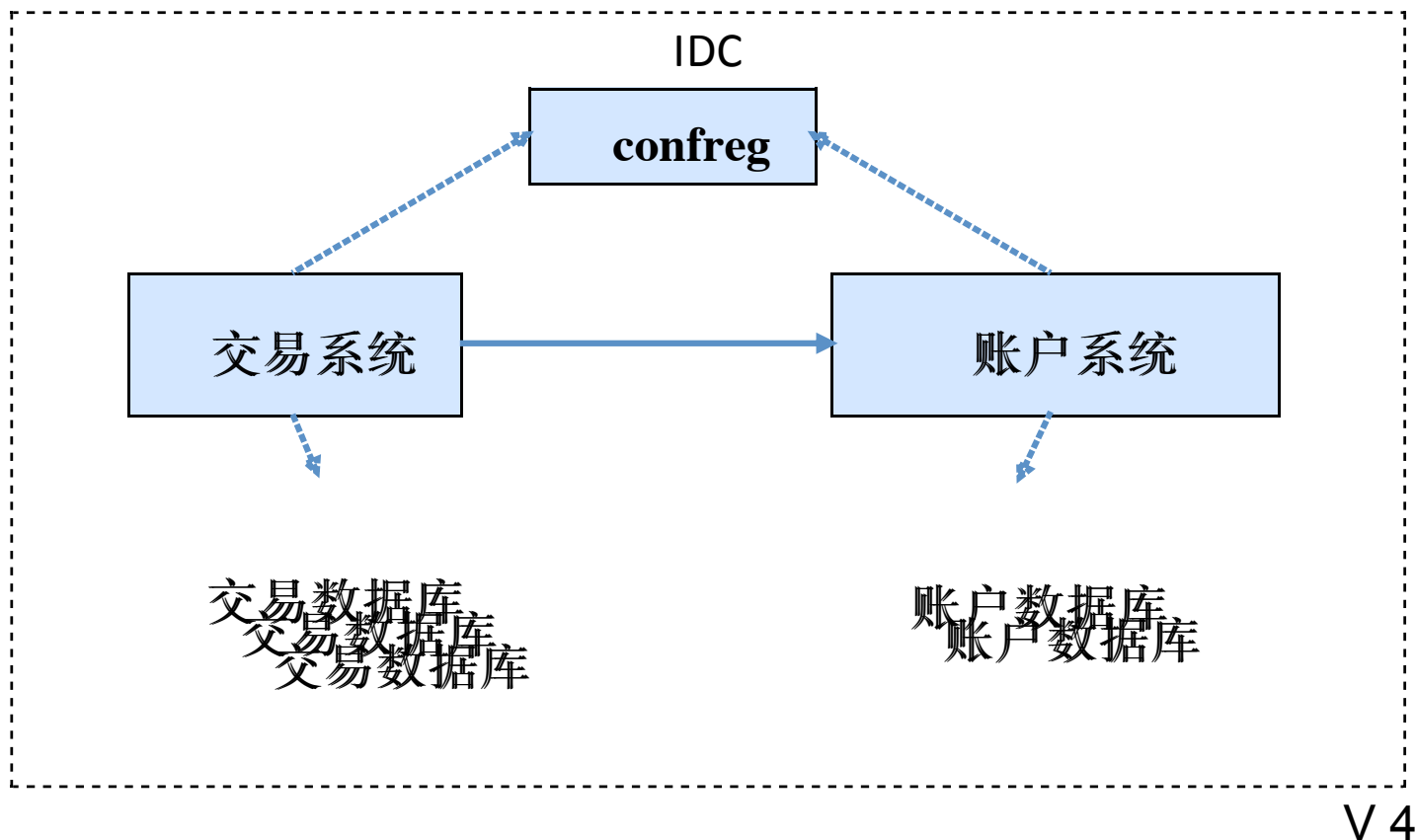
数据可伸缩性：交易数据拆分



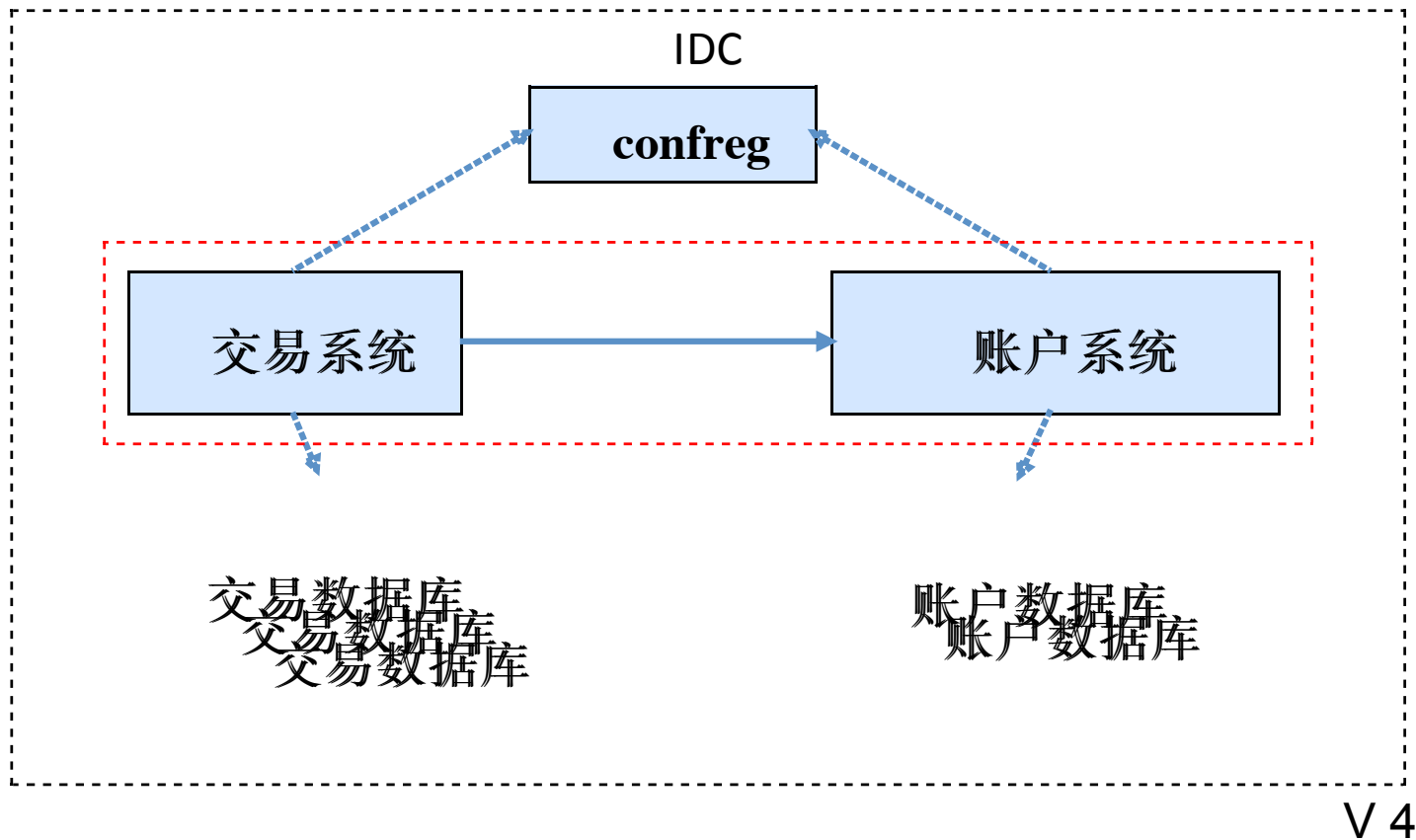
网络伸缩瓶颈



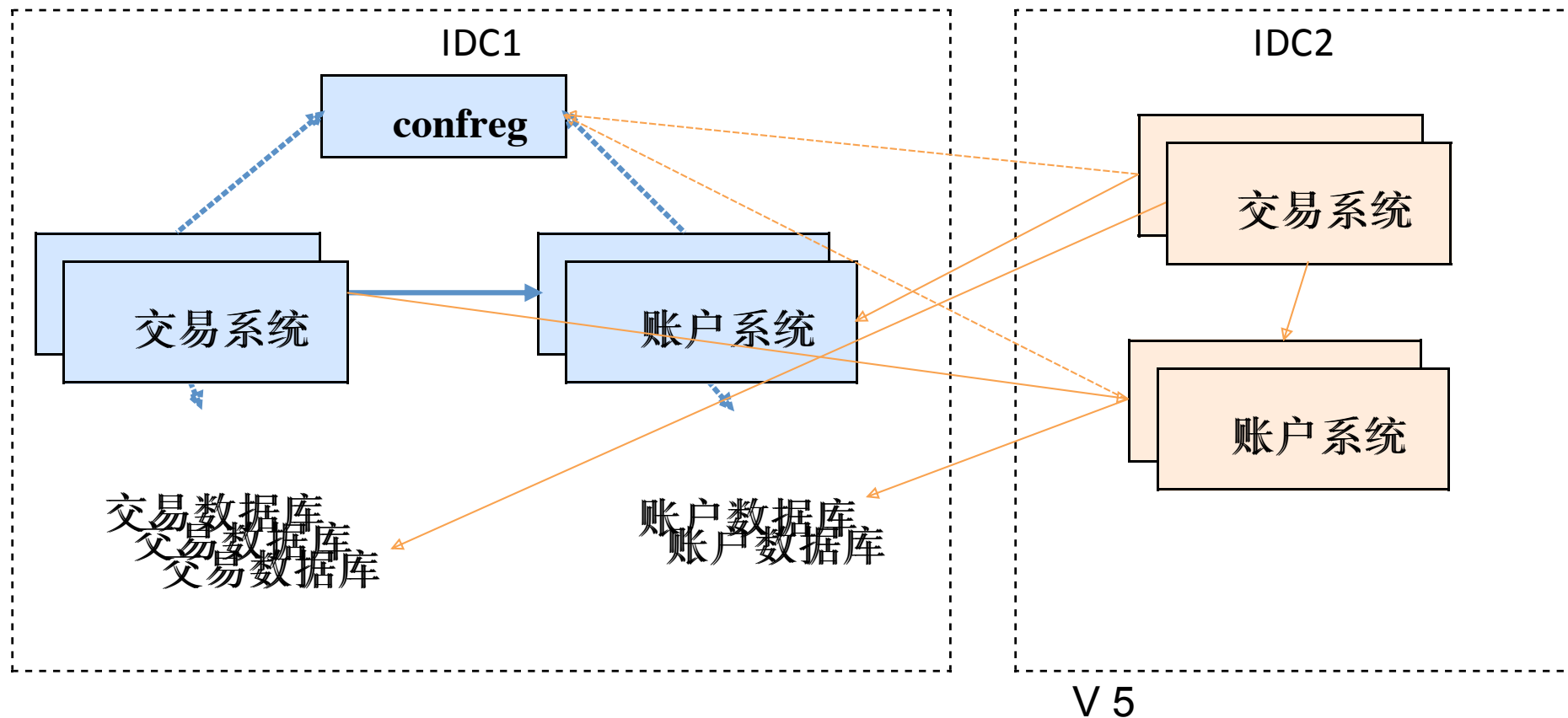
网络可伸缩性：消除网络设备瓶颈



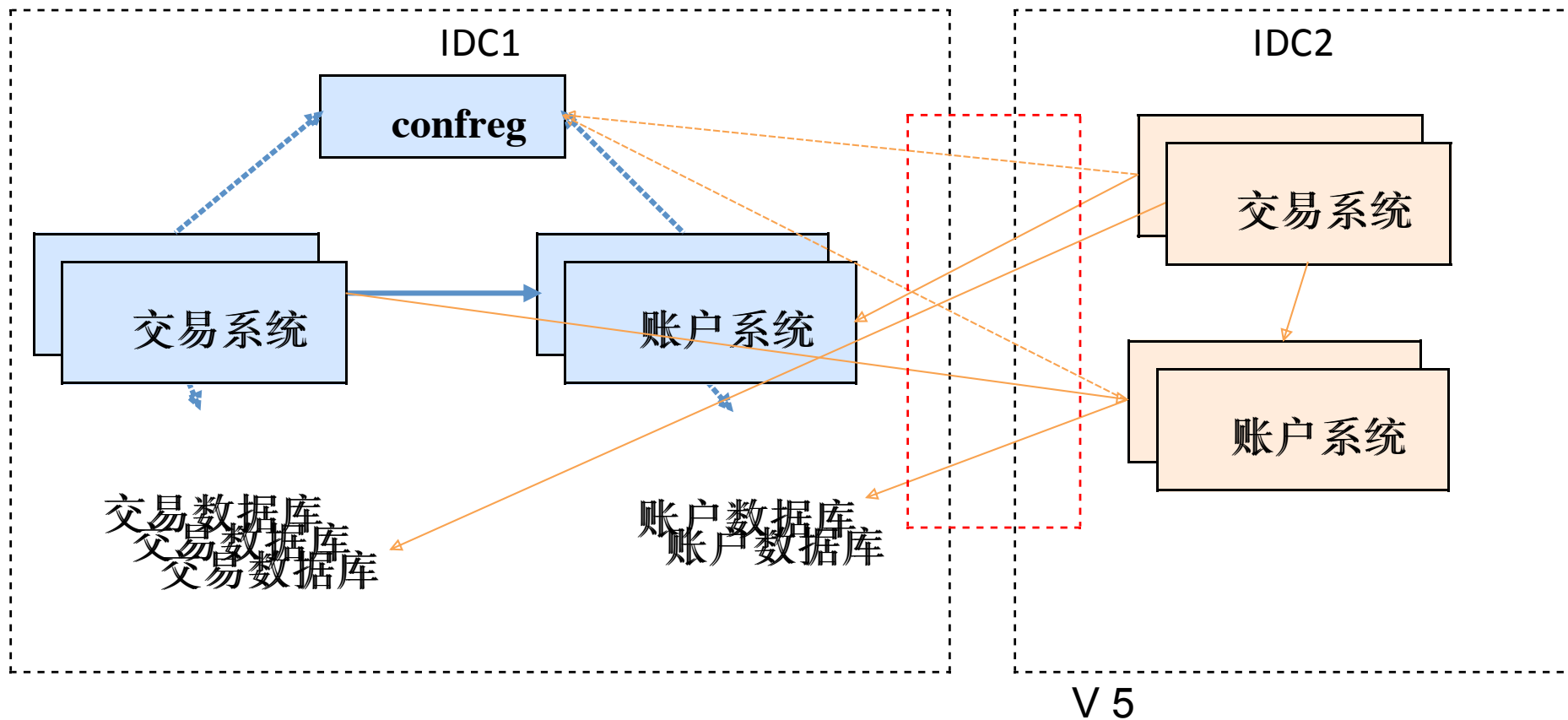
☐ 服务器伸缩瓶颈



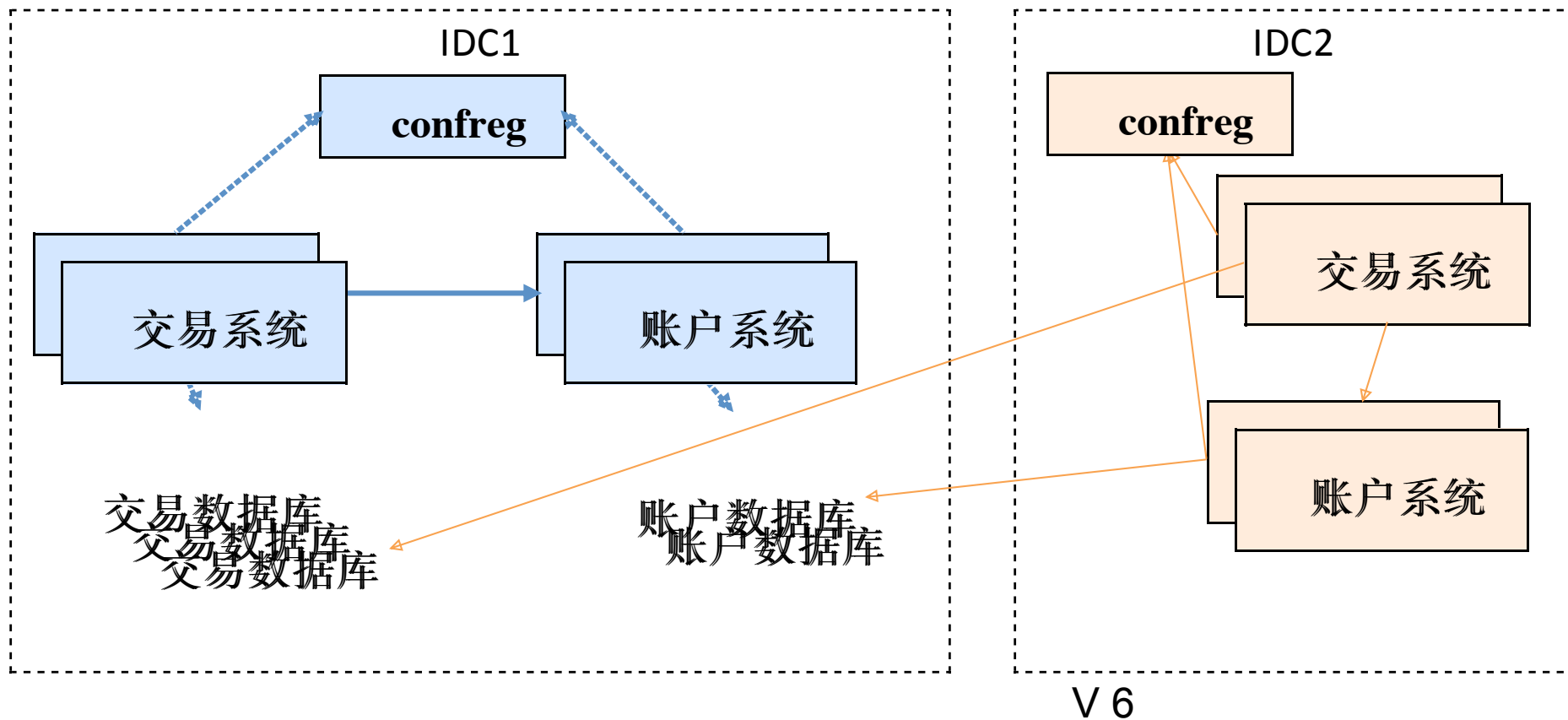
服务器伸缩：服务器扩展到多个IDC



跨机房通讯的瓶颈



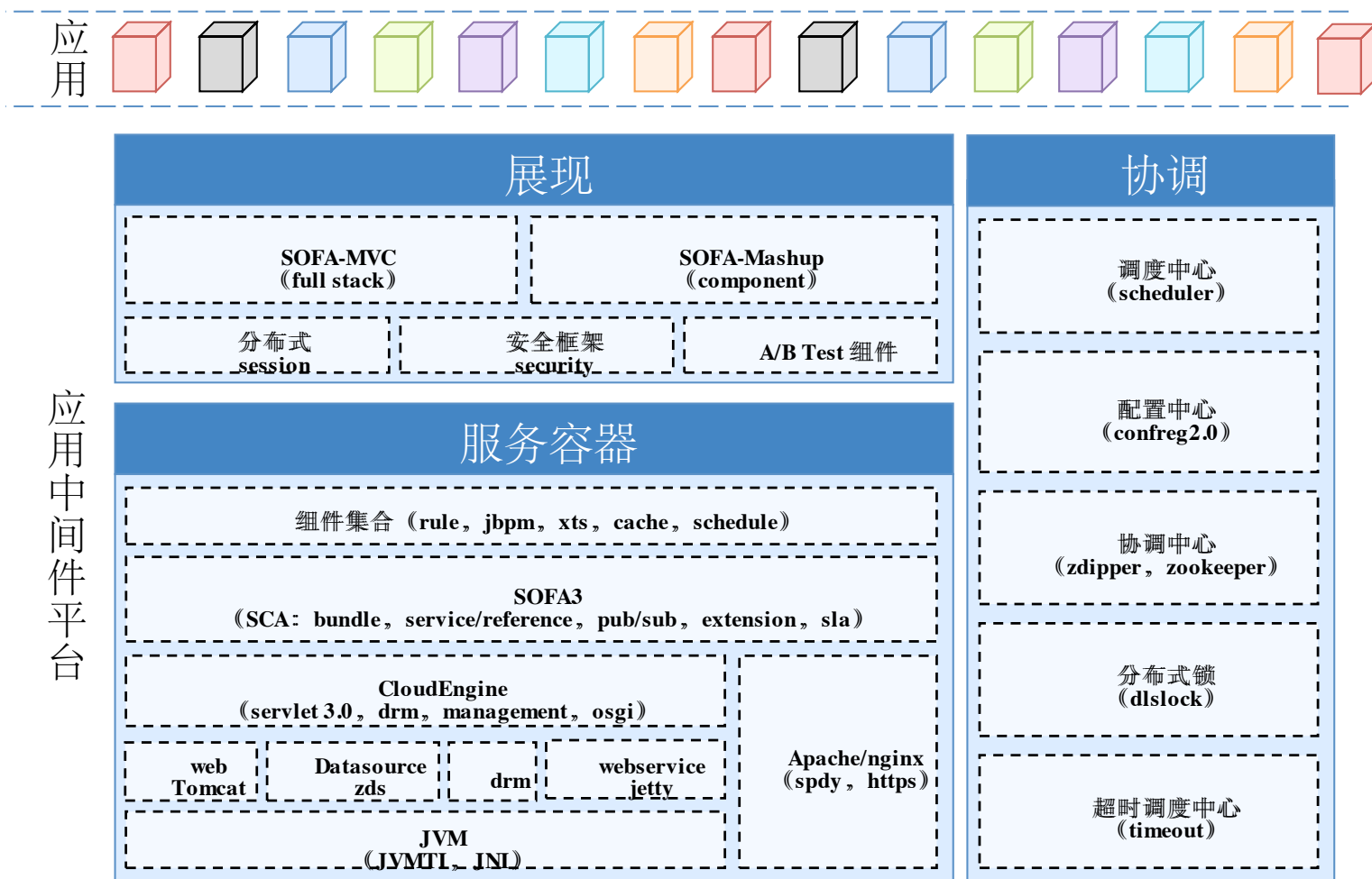
IDC伸缩：部分独立IDC



📌 小结：提升容量百万级到亿级

- ❑ 数据的可伸缩性
 - ✓ 垂直，水平拆分，复制，分布式事务
- ❑ 网络可伸缩性
- ❑ IDC可伸缩性

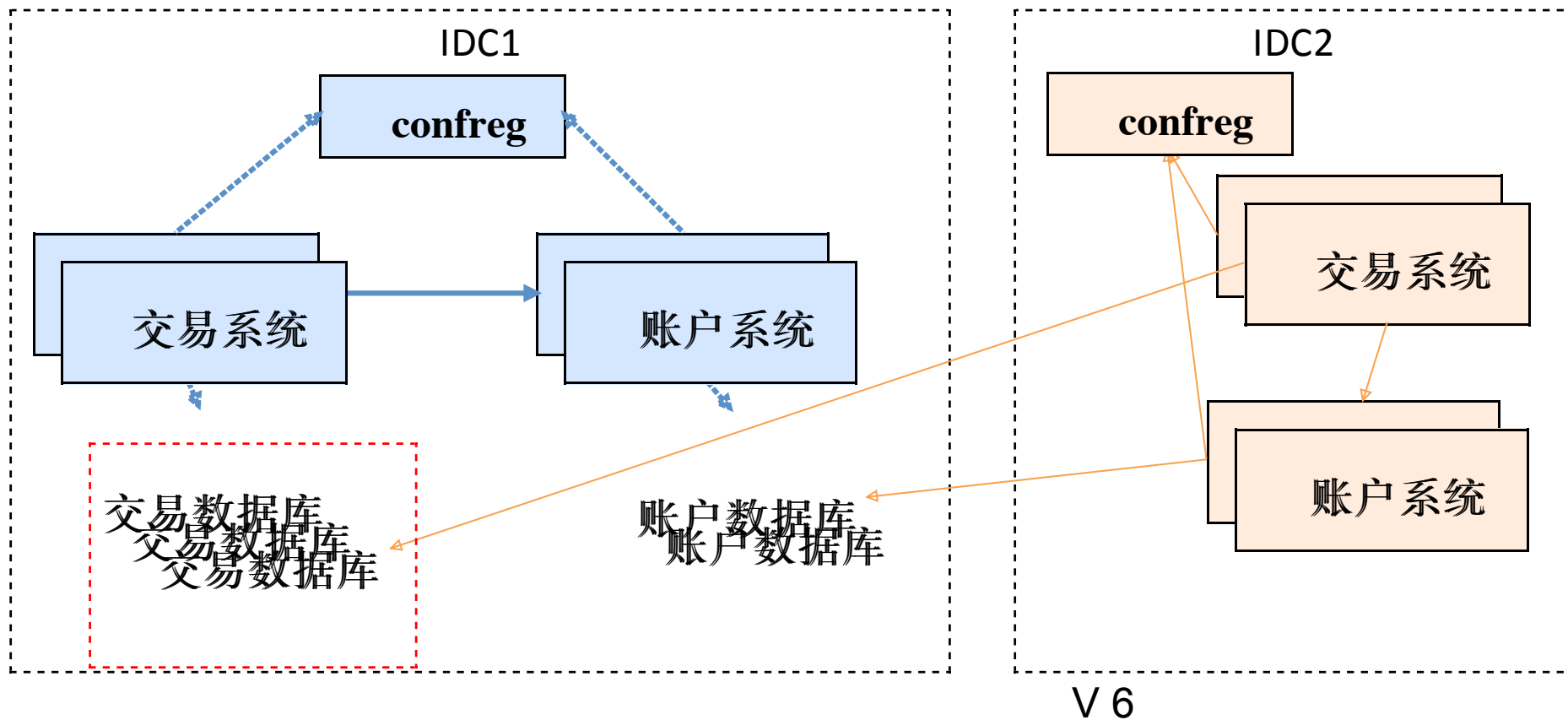
应用中间件技术架构



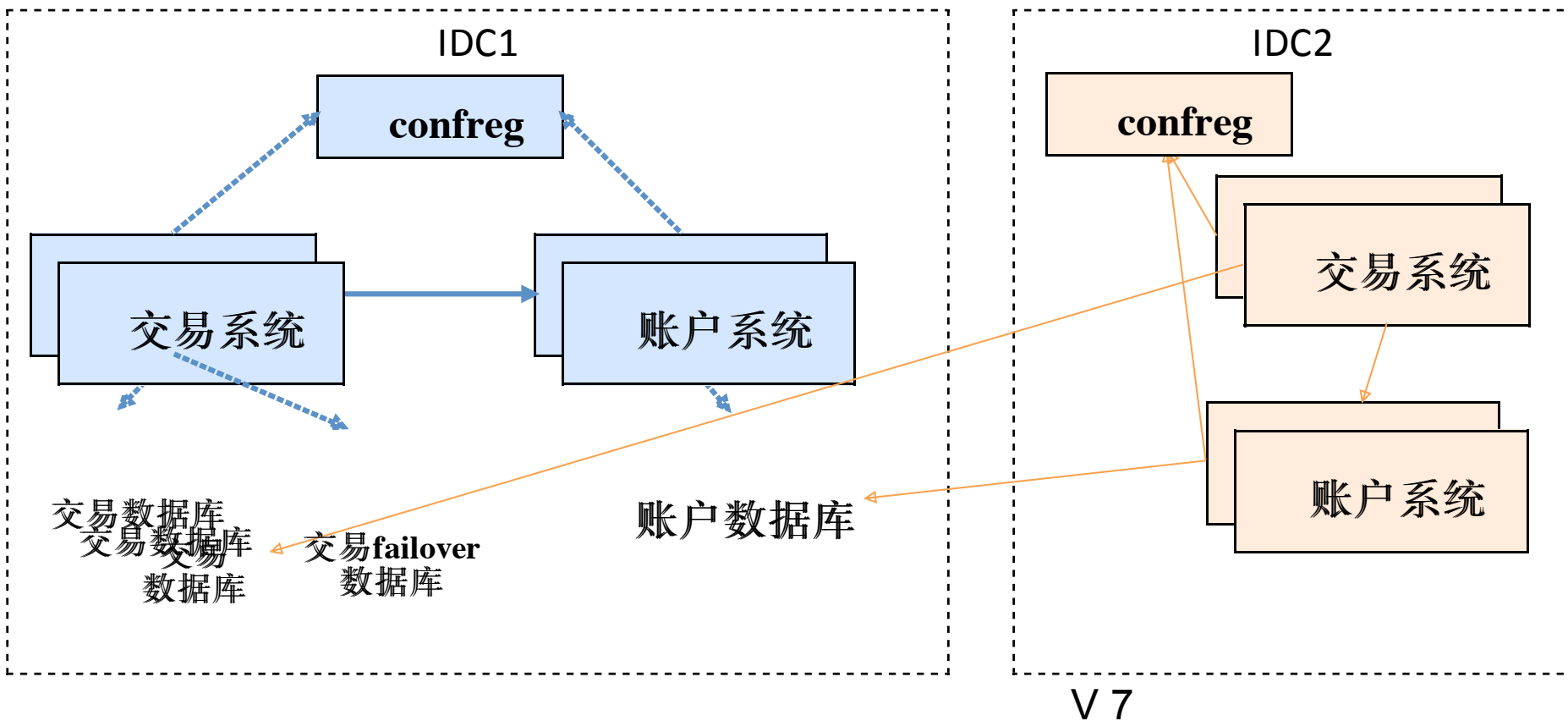
99.9% 到 99.99%+

故障容忍

数据库单点故障

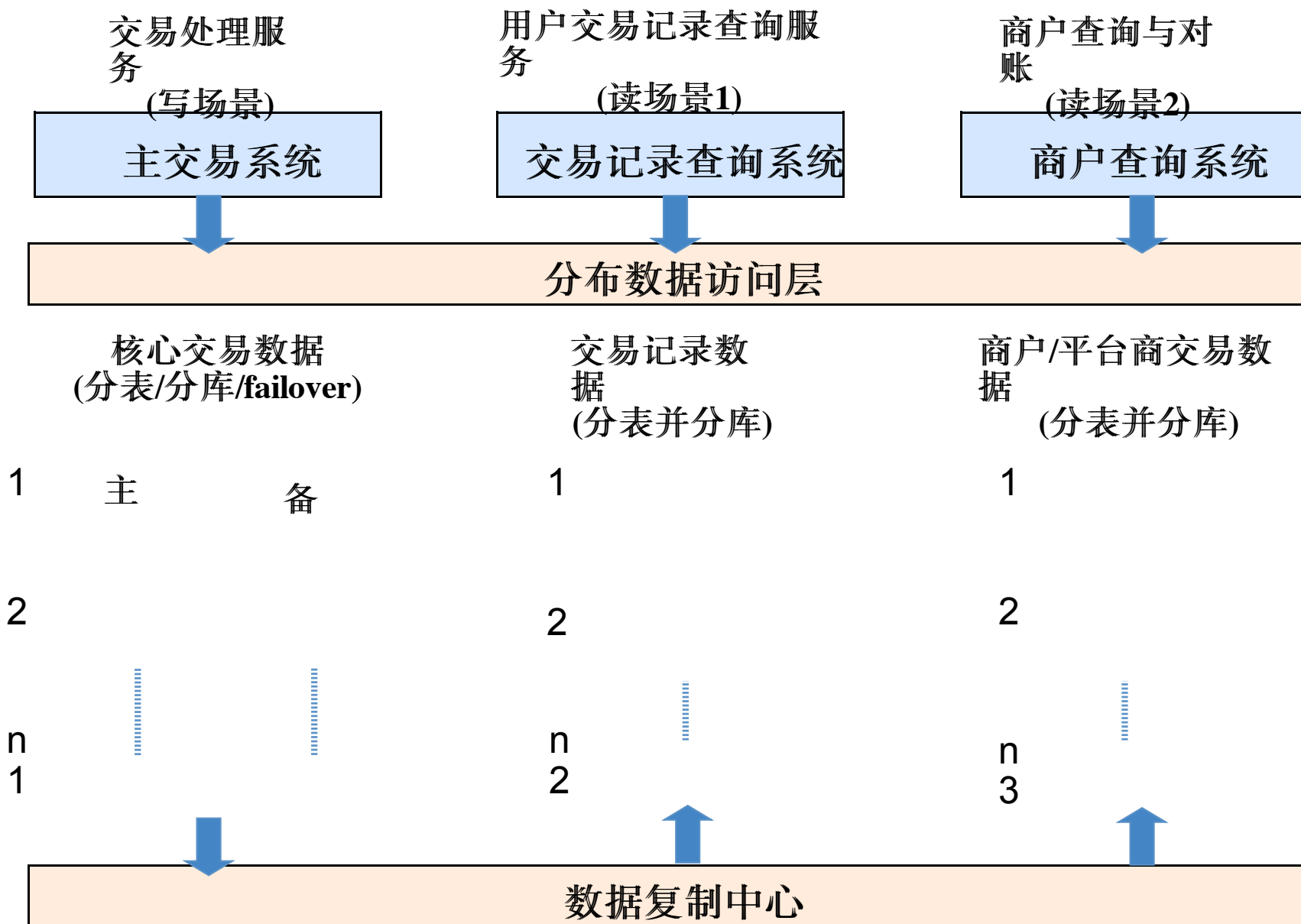


故障容忍-消除数据库单点

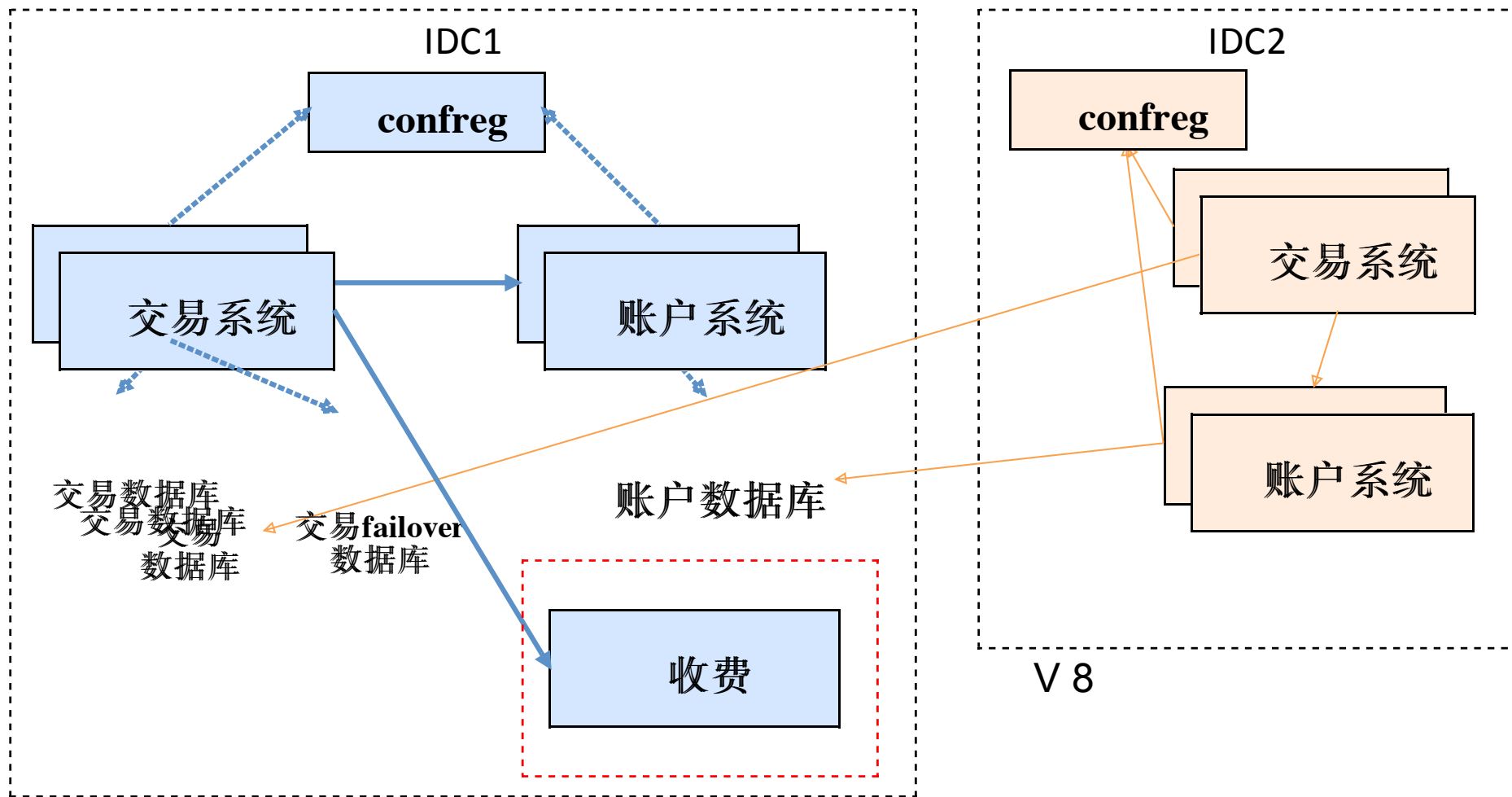




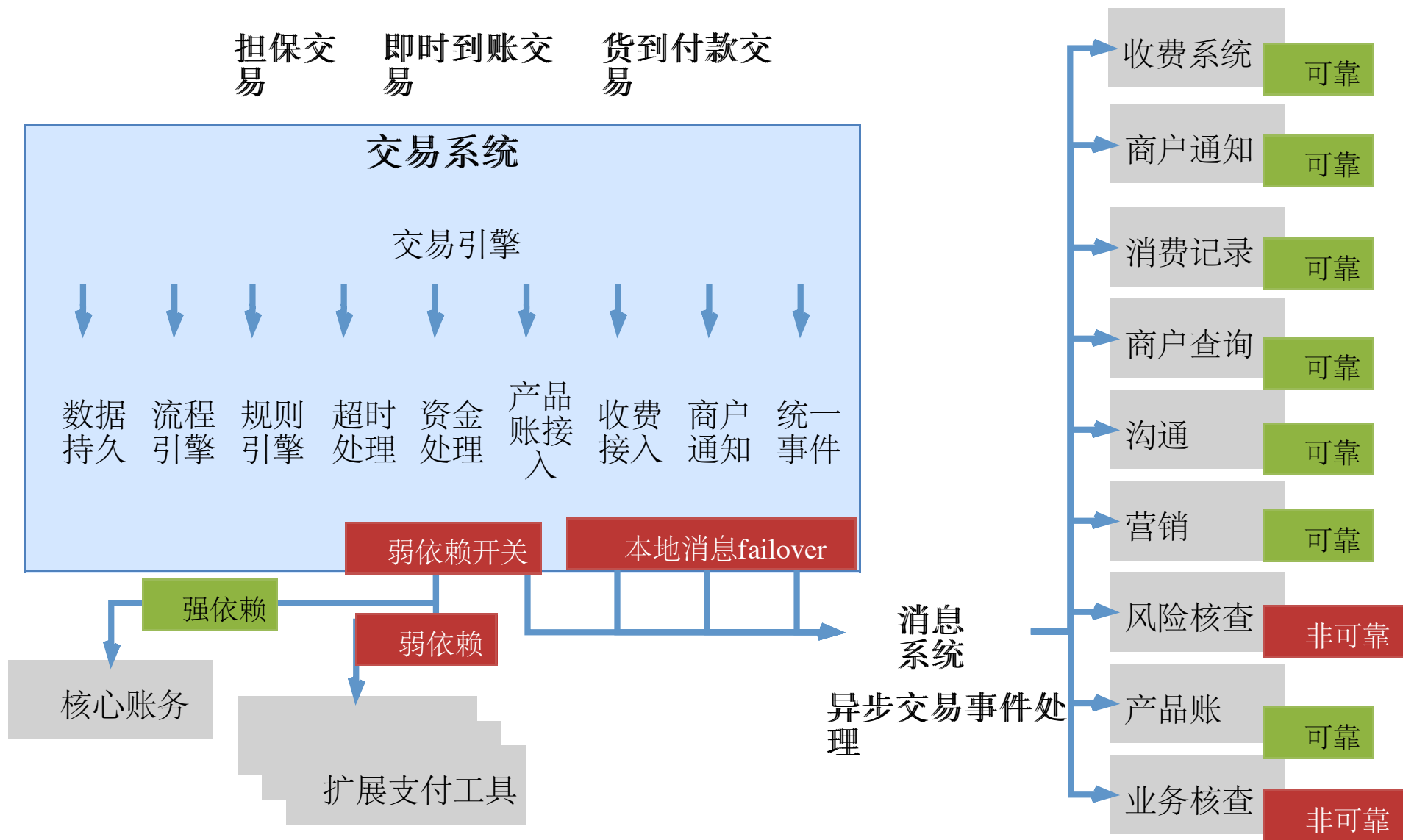
故障容忍-数据库的failover



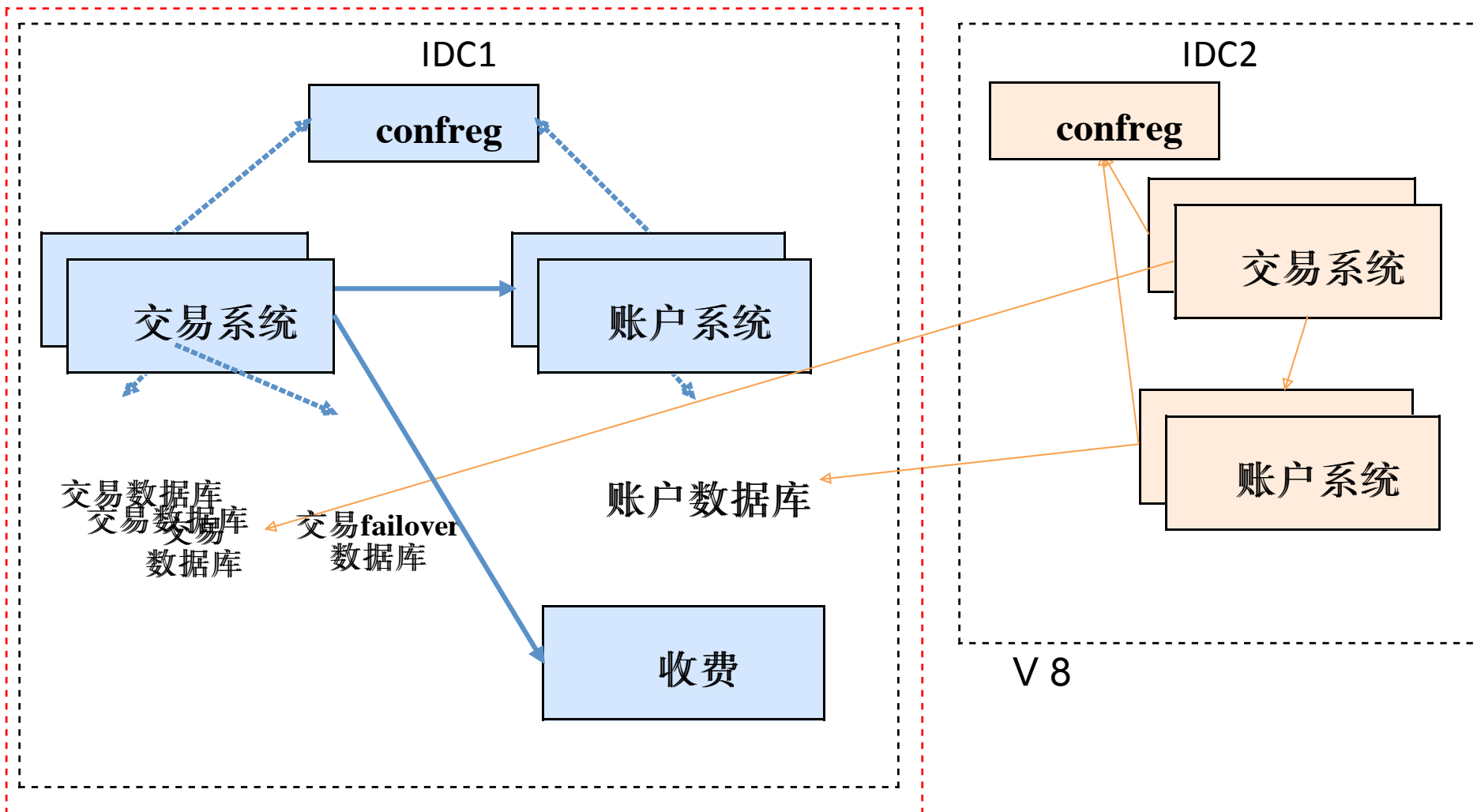
服务依赖故障



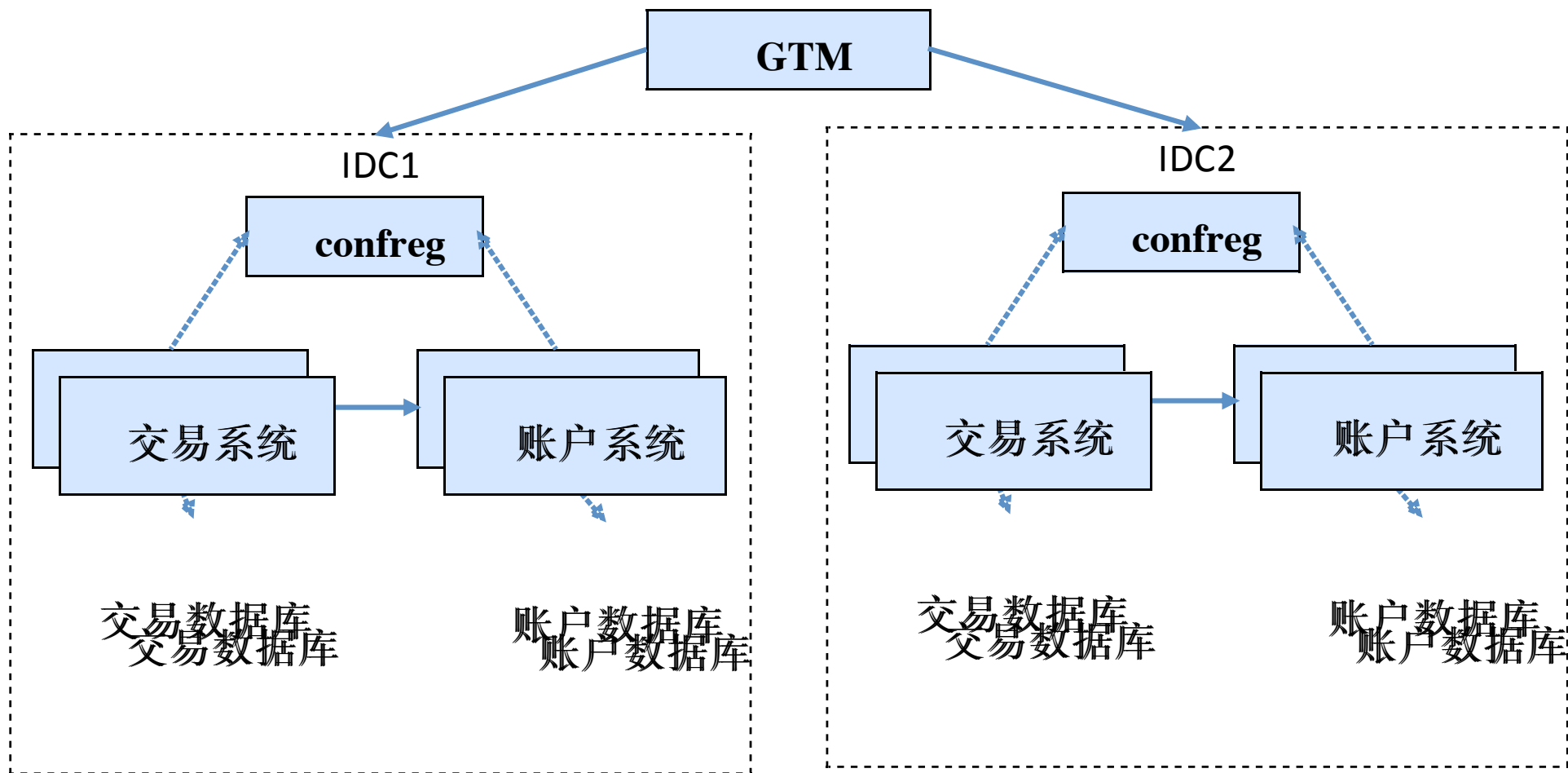
故障容忍-控制服务依赖



IDC故障



故障容忍-完全独立IDC



小结： **99.9%到99.99%+**

- ❑ 消除任何数据库单点
- ❑ 控制服务依赖
- ❑ 完全独立的IDC

人工控制到秒级自动调度

弹性控制

弹性能力-监控平台

维度配置

❗ 维度列一旦提交以后无法删除，只能做增加/修改(序号与类型无法改变)，修改后需要提交表单!

接口 方法 结果 来源 结果码 app server 交易来源 原始交易来源

维度名: 接口

维度序号: 0

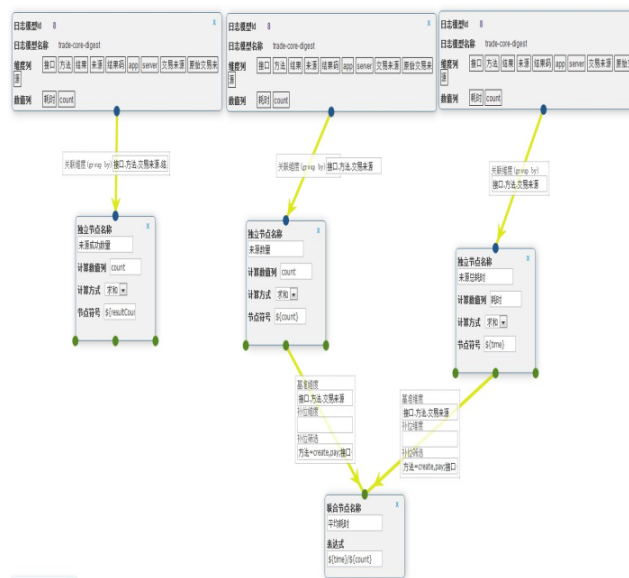
维度类型: 真实维度

分词序号: 3

数值配置

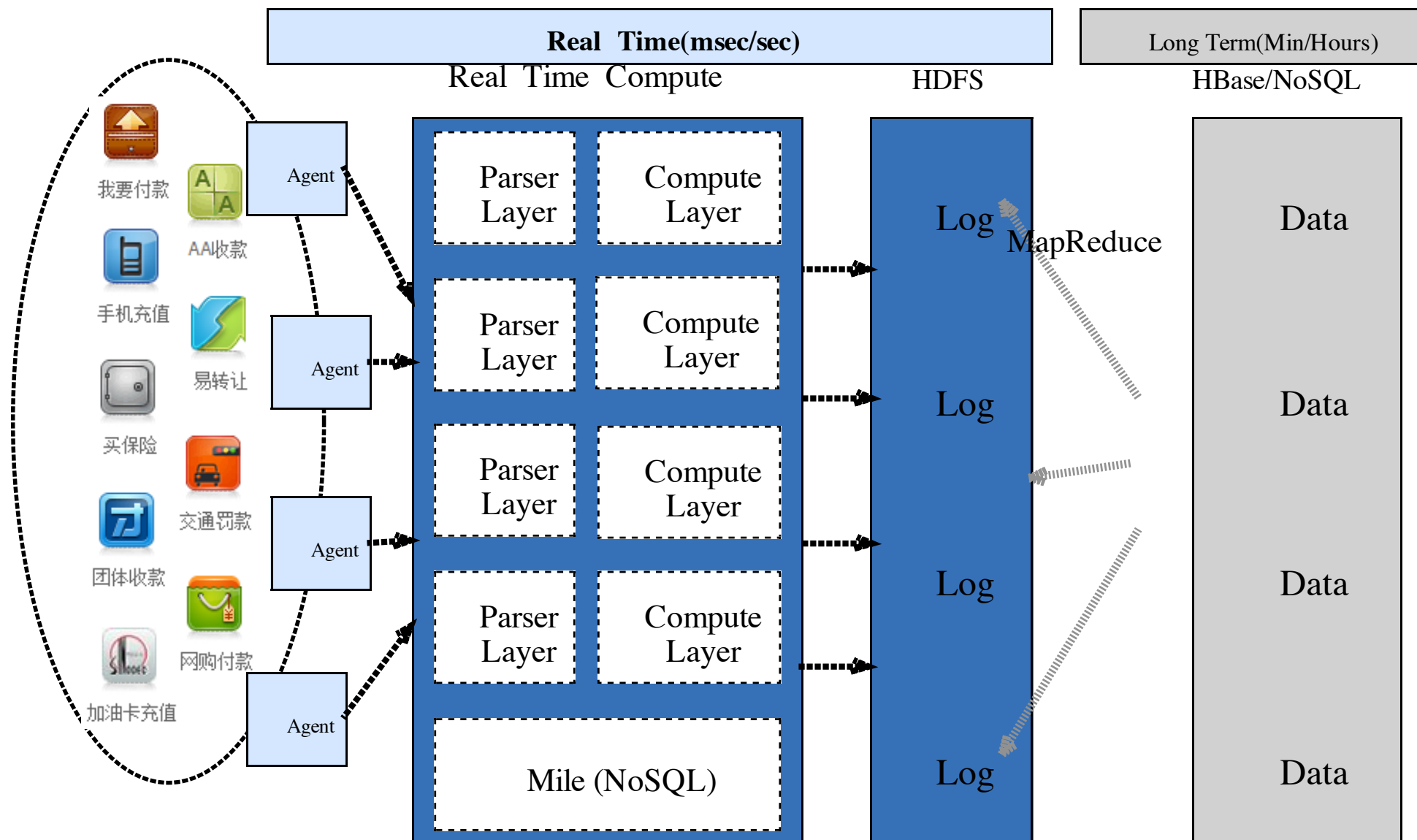
❗ 数值列一旦提交以后无法删除，只能做增加/修改(序号与类型无法改变)，修改后需要提交表单!

耗时 count



时间	来源成功数量	基线
09:42	12510	12348
09:41	12528	12290
09:40	12277	12229
09:39	12240	12166
09:38	12327	12101
09:37	12217	12035
09:36	12195	11968
09:35	11963	11900
09:34	11979	11832
09:33	11954	11764

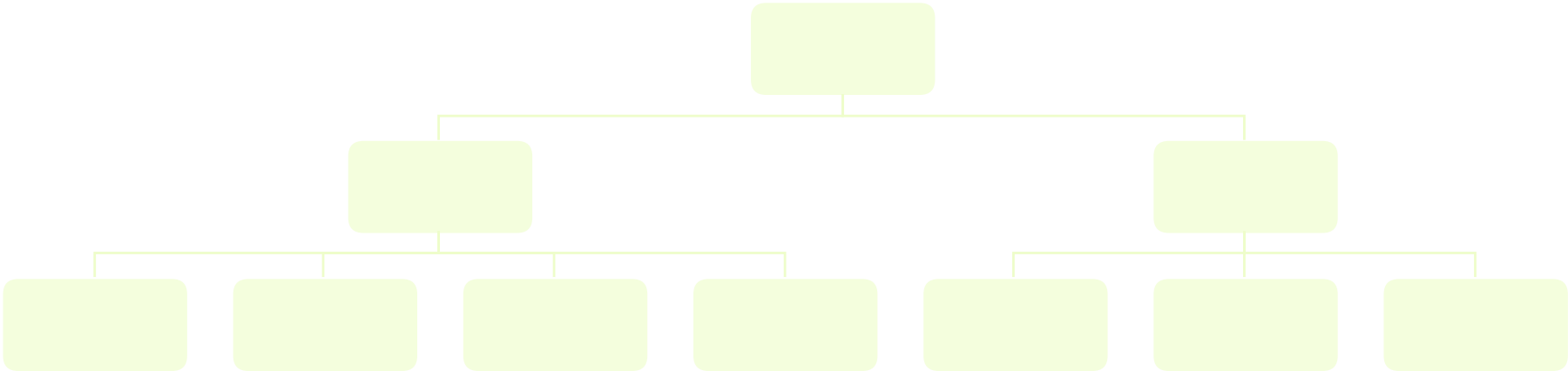
弹性能力-秒级监控系统



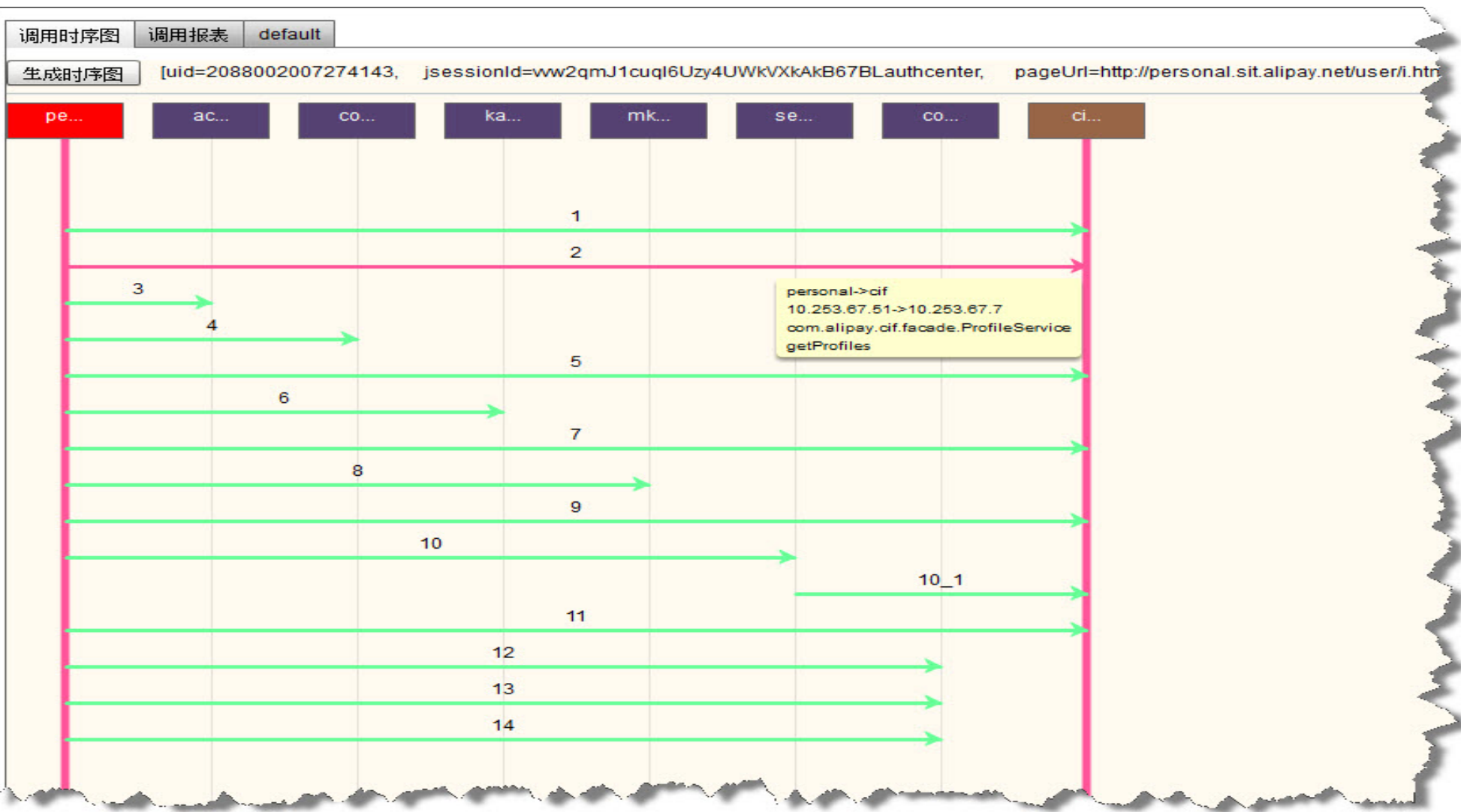
弹性能力 – 容器级的精细化控制

- ❑ **help** 显示所有的指令；
- ❑ **showall** 以树型显示所有的资源属性。
- ❑ **setapp** <appname> 设置app的上下文，如果只有一个app默认选择这个app，这样可以简化整个命令行的操作；
- ❑ Display Osgi Bundles Status
 - ✓ **bundle** <appname> <id> 显示对应应用的ace中的bundle资源，如果只有一个app显示默认的app中的bundle信息；
- ❑ Display Base Config
 - ✓ **config** <appname> 显示某个app；
 - ✓ **app** [appname] 显示所有app的信息，包含这个app的基本属性（如：有多少个服务，多少个应用，多少个datasouce，多少个drm资源）；
 - ✓ **drm** <appname> 显示所有drm的信息；
- ❑ Display Service Component Runtime （服务组件模型）
 - ✓ **service** <appname> [servicename 注意：可以通配符][**-b** [jvmltrlwshttp 注意：可以多选] <id> 显示所有service信息，如果只有一个app显示默认的app中的service信息；
 - ✓ **reference** <appname> [servicename 注意：可以通配符][**-b** [jvmltrlwshttp 注意：可以多选] <id> 显示所有reference信息，如果只有一个app显示默认的app中的reference信息；
 - ✓ **context** <appname> <id> 显示所有的上线问信息，如果只有一个app显示默认的app中的context信息；
 - ✓ **consumer** <appname> <id> 显示所有的上线问信息，如果只有一个app显示默认的app中的consumer信息；
 - ✓ **publisher** <appname> <id> 显示所有的上线问信息，如果只有一个app显示默认的app中的publisher信息；
- ❑ Display Web Status
 - ✓ **tomcat** <appname> 显示tomcat的运行关键属性，如果只有app显示默认的app中的bundle信息；
 - ✓ **mvc** mvc框架属性；
- ❑ Display DataLevel Status
 - ✓ **datasource** (缩写 ds)
 - ✓ **tddl**
- ❑ Display Transports Status
 - ✓ **ws**
 - ✓ **tr**
- ❑ Display Performance Status
 - ✓ **ipstats** (ip stats)
 - ✓ **qps** (query per second)
 - ✓ **pvs** (web)
 - ✓ **tpr** (time per request)
 - ✓ **thread** 线程运行状态
- ❑ Display Misc Status
 - ✓ **cron** (schedule)

弹性能力-自动化的调度

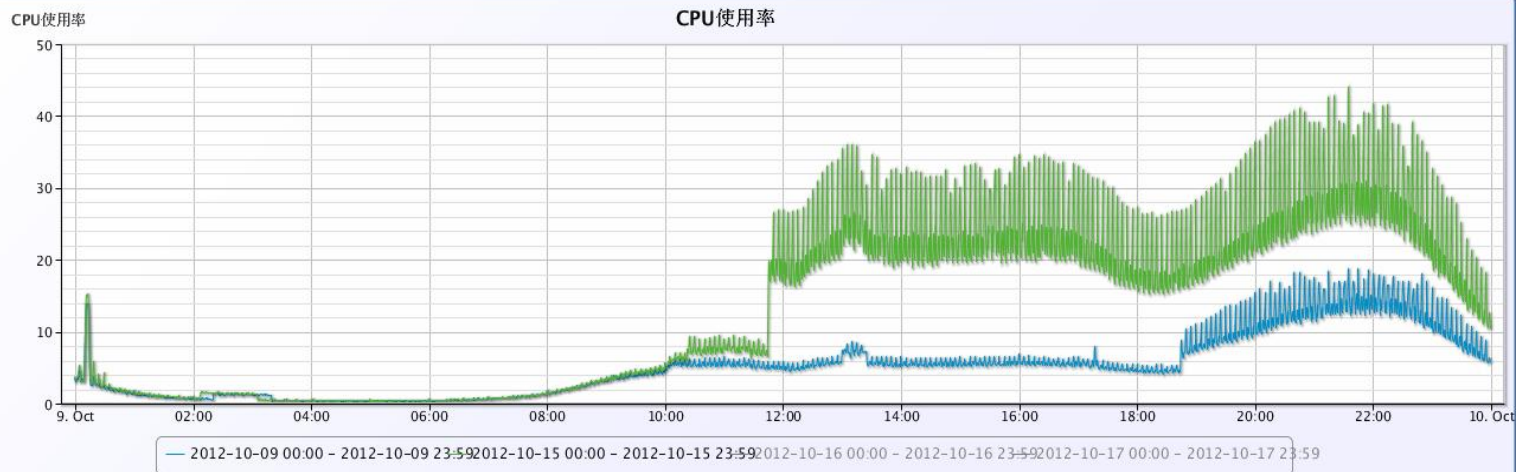


治理 - SOA调用

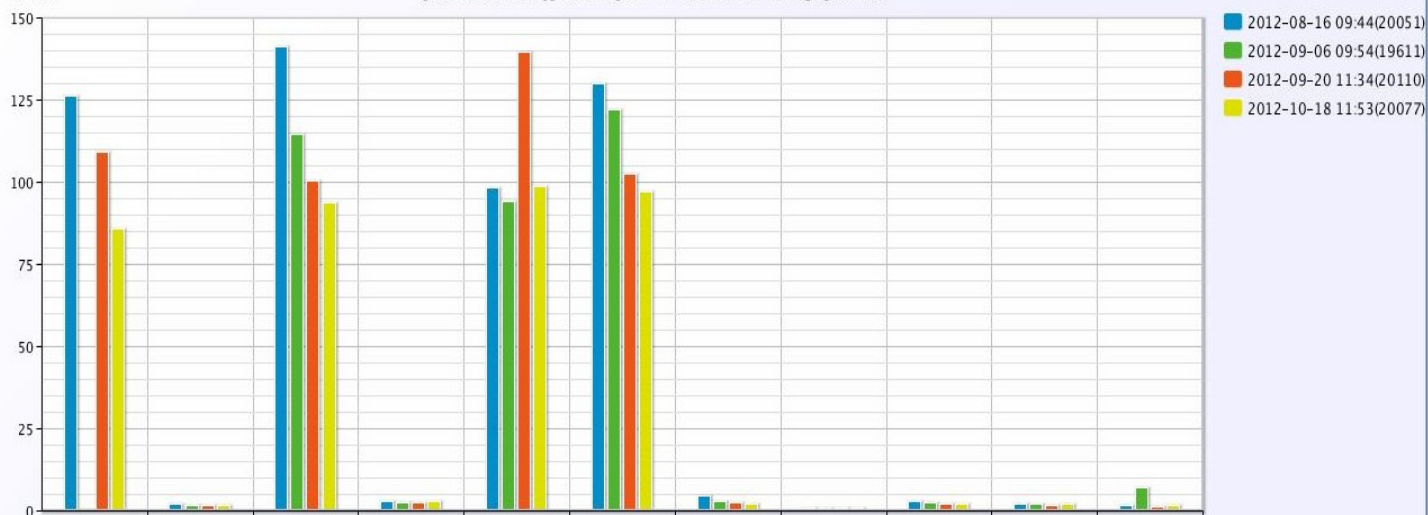


治理 - 性能分析

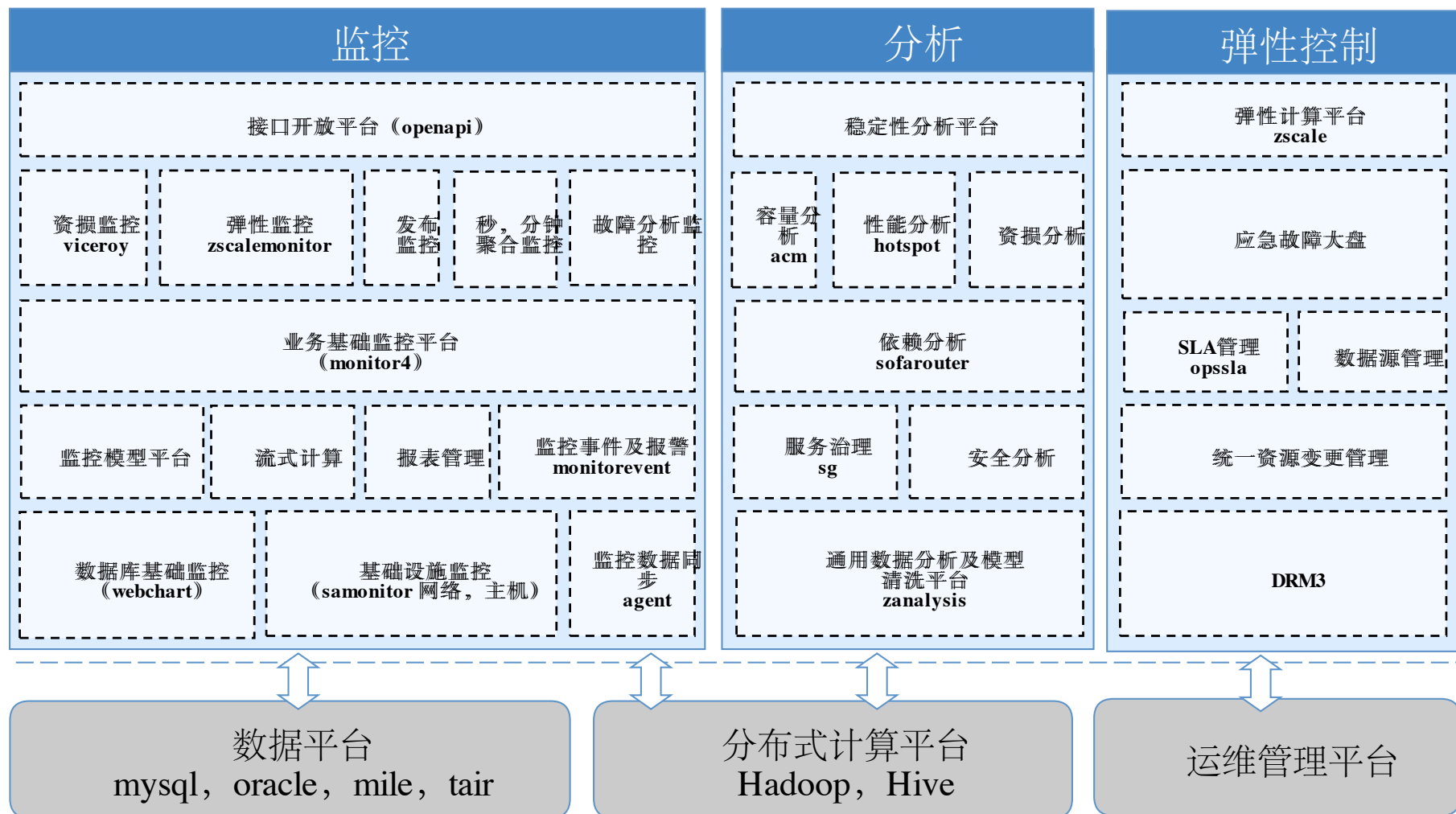
- 在线上提供性能历史对比图，可以对比每周发布之后的关键链路应用系统的性能变化和增长趋势



响应时间 [淘宝交易创建][20000]笔-关键系统响应时间[4]次对比



弹性控制平台



小结：人工控制到秒级自动调度

- 实时的系统监控能力
- 快速自动化的系统调度能力
- 精细化的系统治理能力

总结

谢谢



- 阿玺
- 支付宝-技术部
- Mail: xi.hux@alipay.com
- 新浪微博: 支付宝_阿玺



QCon全球软件开发大会（北京站）2013

即将启动敬请关注

www.qconbeijing.com

全球企业开发大会

INTERNATIONAL
SOFTWARE DEVELOPMENT
CONFERENCE