



百度数据库中间层技术介绍

尹博学
dba@baidu.com



DTCC2012



MySQL数据库集群面临的问题

数据库中间层整体设计

数据库中间层主要功能模块和策略

数据库中间层应用

数据库中间层发展

- ✓ 随着数据库流量和服务器的数量增长，数据库集群面临很多问题：
- 如何实现应用程序与集群的解耦，降低运维成本？
- 如何实现数据库服务的高可用？
- 如何将并发控制前移，保护数据库系统？
- 如何降低数据库开发成本？
-

MySQL数据库集群面临的问题



数据中间层整体设计

数据库中间层主要功能模块和策略

数据库中间层应用

数据库中间层发展

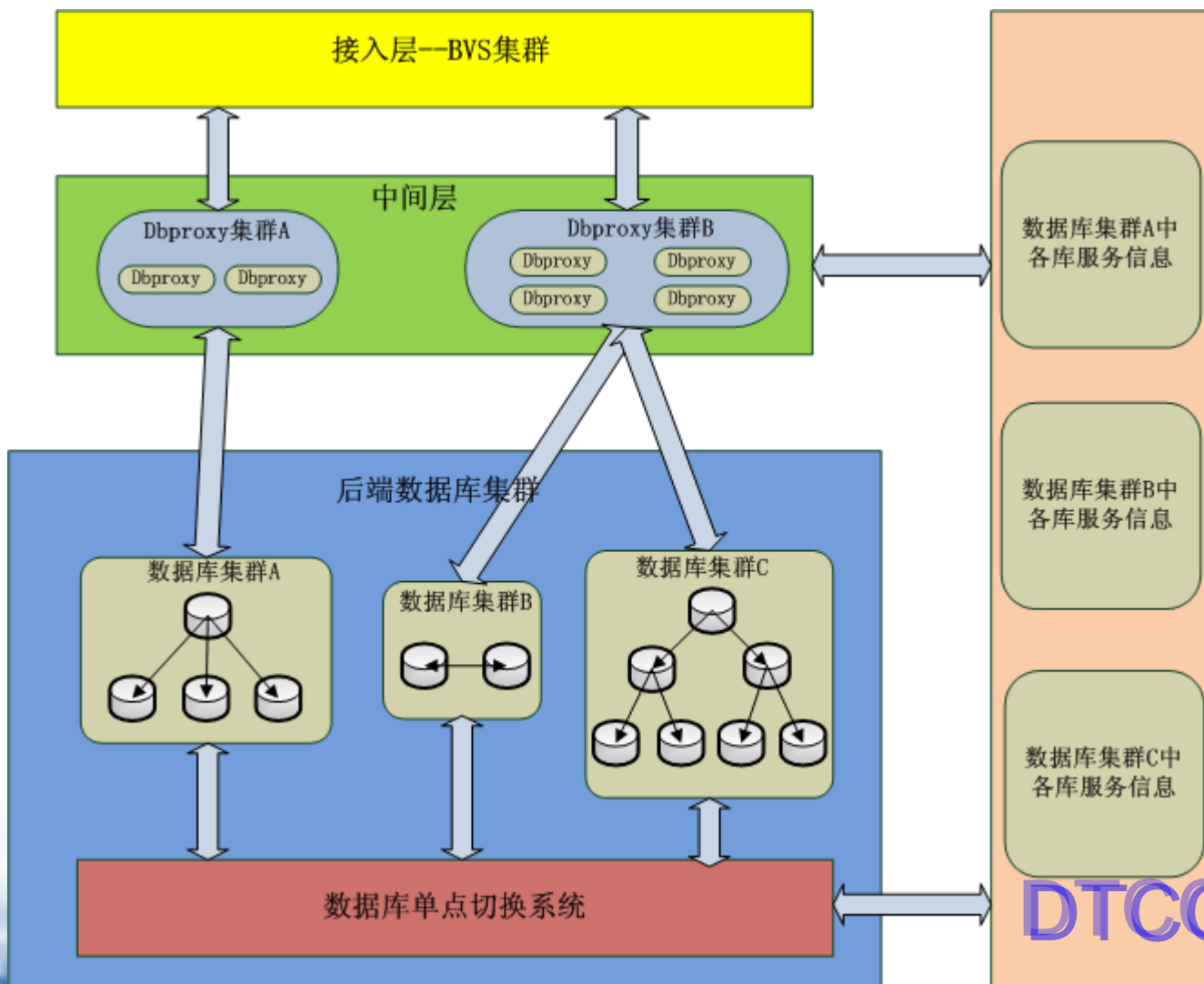
✓ 数据库中间层（dbproxy）的定义：

- 屏蔽集群内部细节
- MySQL集群读写分离
- 连接池
- 负载均衡
- 访问控制
- 可集群化部署，运维

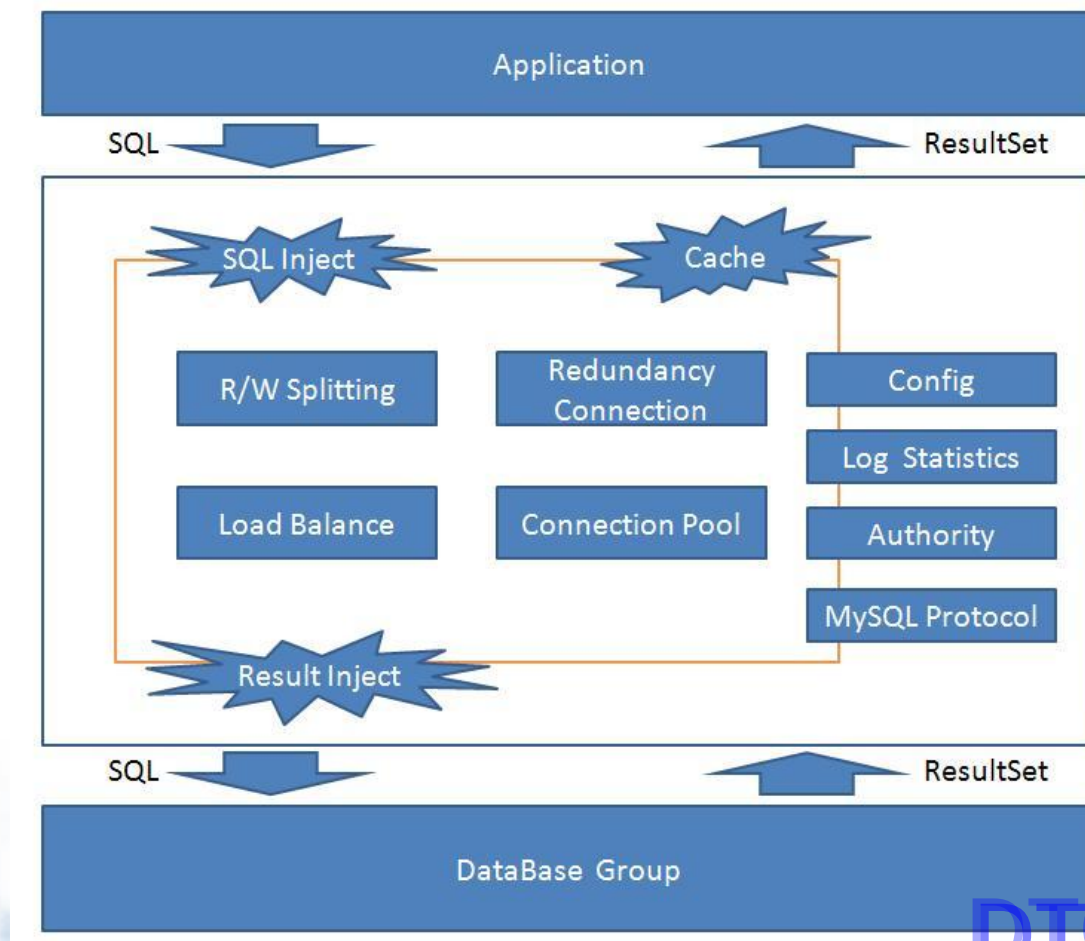
✓ 关键要求：

- 对数据库应用透明
- 性能要好：高并发，低延迟



✓ 百度数据库集群三层架构



✓ Dbproxy逻辑框架设计：

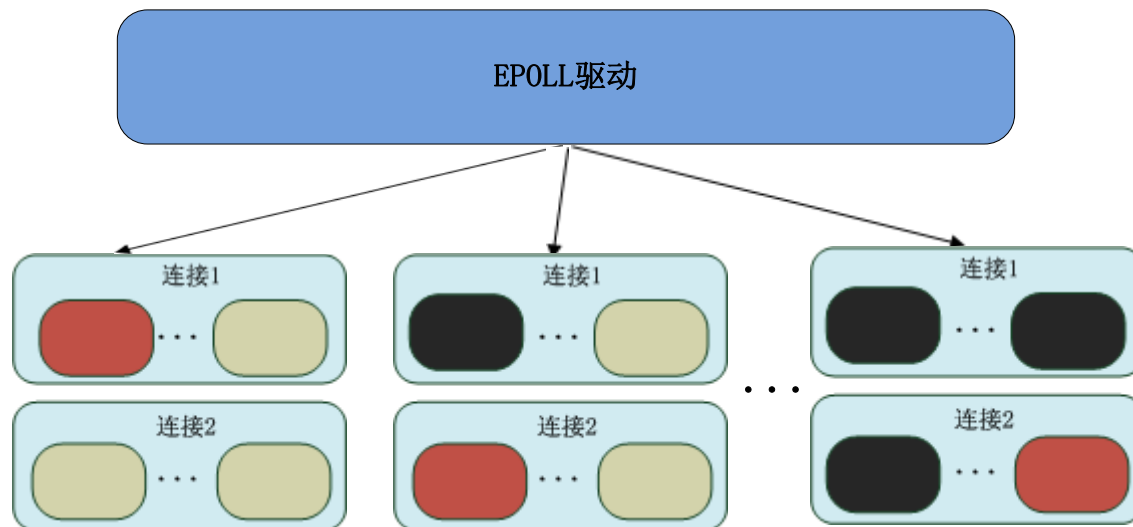


✓ Dbproxy的程序框架设计：

- 进程 ?  or 线程 ?
- MySQL API ? or 按照协议与机制自己实现 ? 

✓ Dbproxy支持并发设计：

- 进程并发设计：



- 应用程序 <-> dbproxy <-> MySQL 阶段划分
handshake_send, auth_read, auth_result_send, query_read,
.....

MySQL数据库集群面临的问题

数据中间层整体设计



数据库中间层主要功能模块和策略

数据库中间层应用

数据库中间层发展

✓ 读写分离策略—基本原则

- 读请求发往从库；写请求发往主库
- 读写请求如何划分？

Select * from student where class=XXX 发往slave

Select * from student where class=XXX for update ?

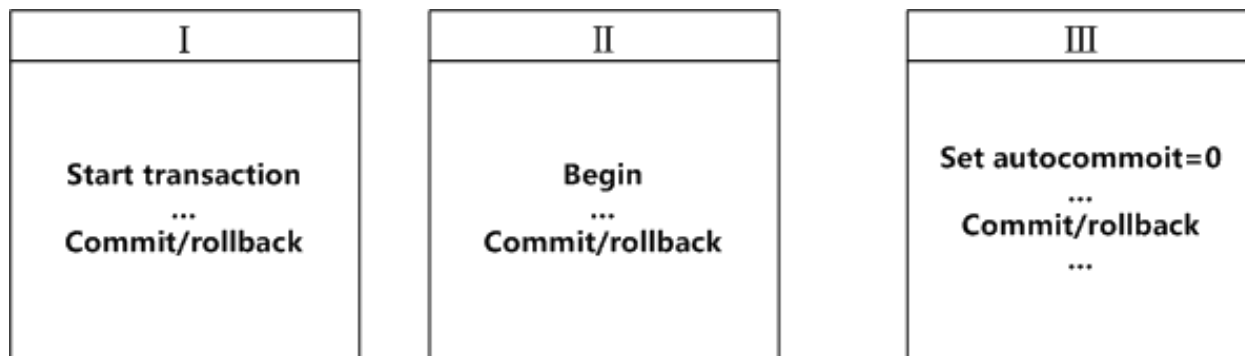


划分原则

枚举确定为读的SQL请求，将其发往从库
其余均发往主库

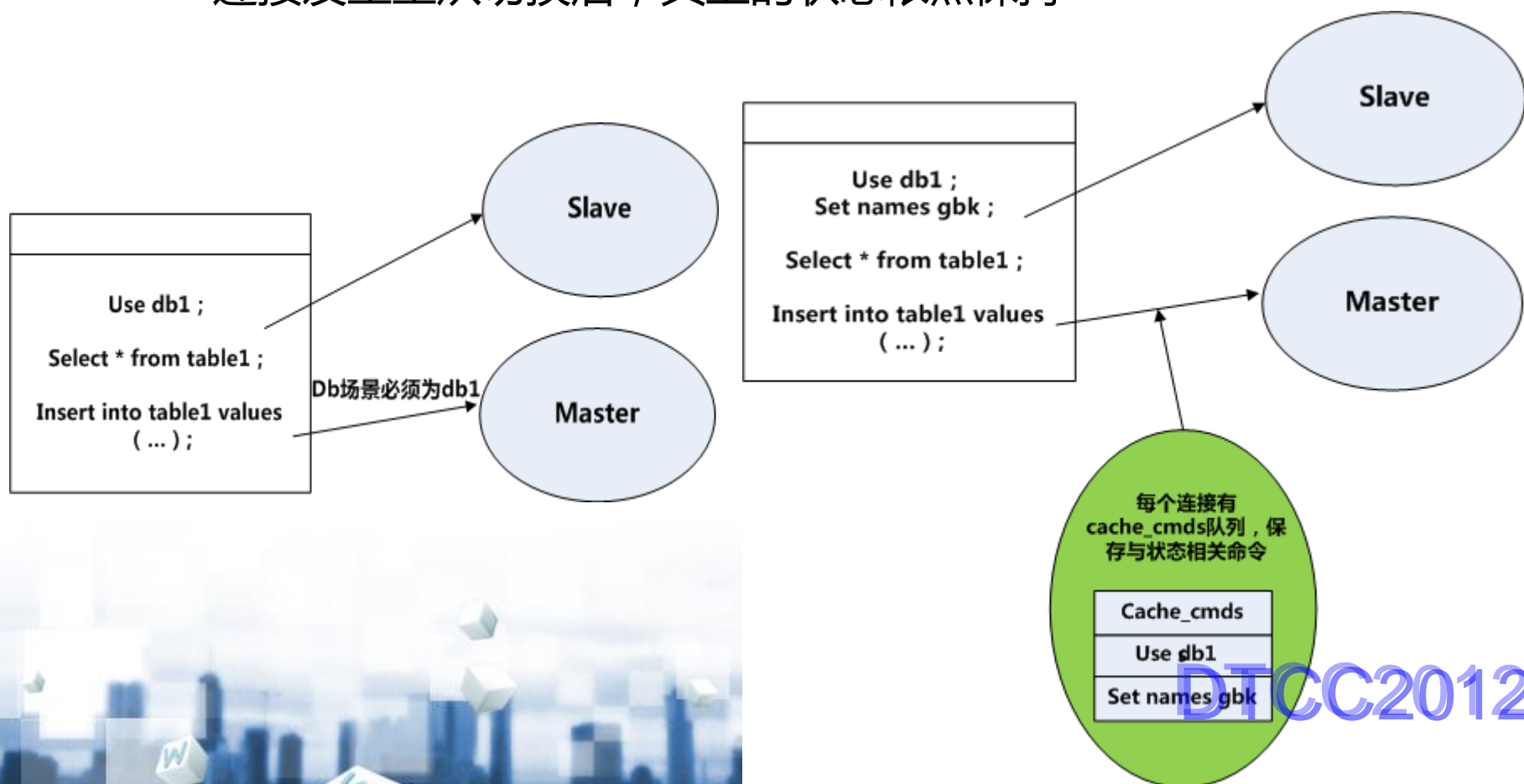
- 会话一致性：一个会话写请求结束后指定时间内的读请求会发往主库

- ✓ 读写分离策略—事务处理
- 支持事务
 - 有关事务的读写分离策略



✓ 读写分离策略—会话状态保持

- 连接发生主从切换后，其上的状态依然保持

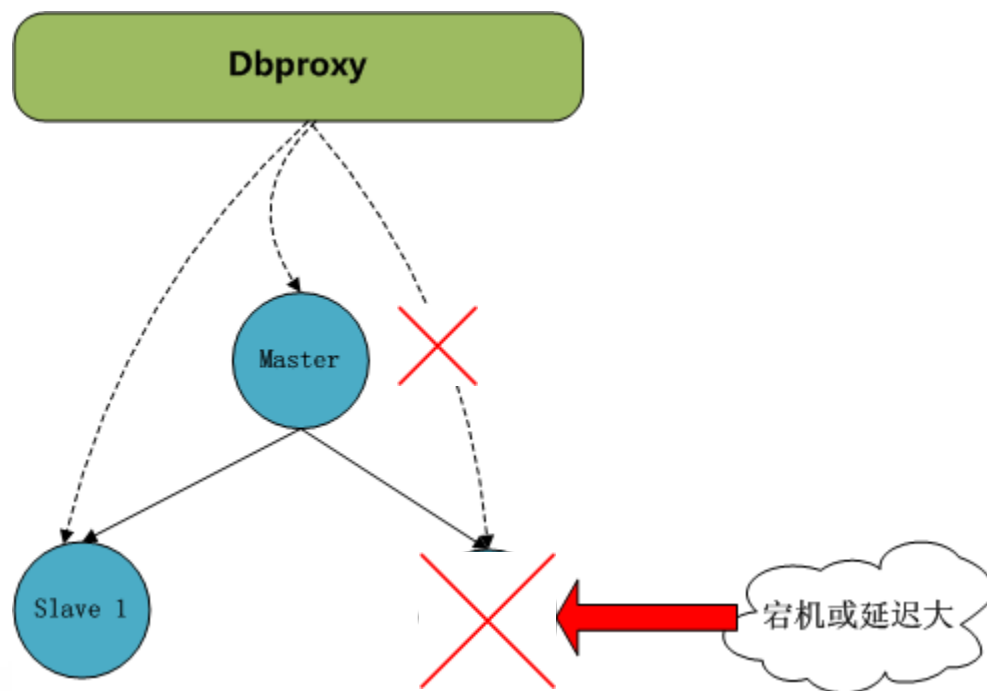


✓ 负载均衡策略

- 基本算法是基于数据库当前连接数的
- 新建连接—选取集群中同类角色中“当前连接数/权重”最小的数据库

✓ 负载均衡策略

- 负载均衡策略考虑MySQL实例的状态

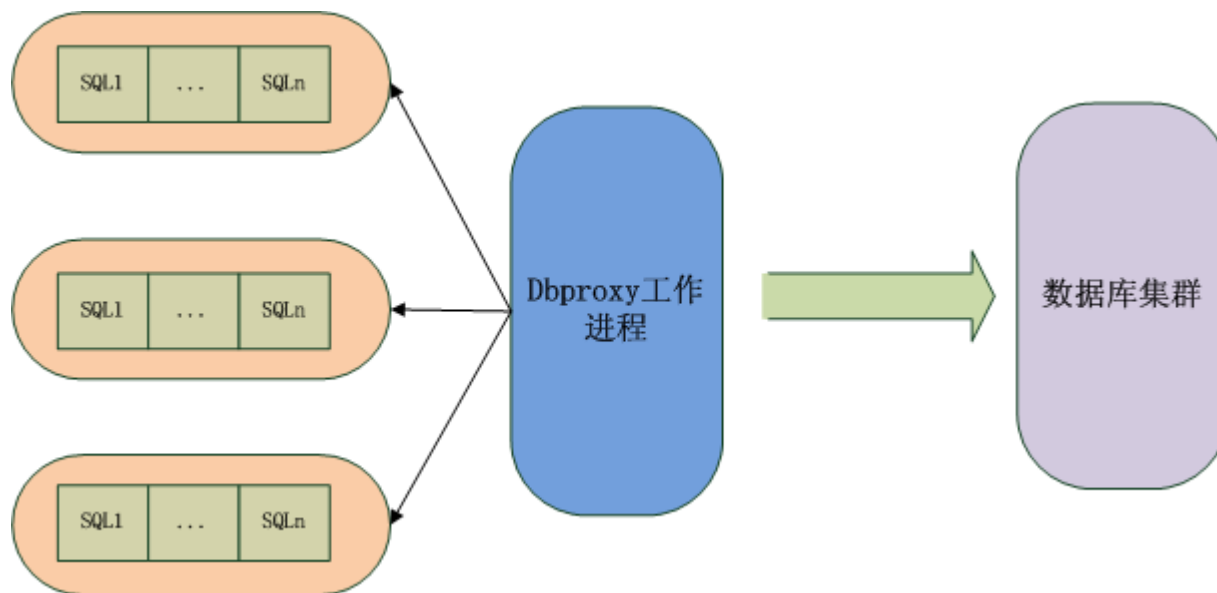


✓ 连接池

- 构建有与MySQL连接池，提升性能
- 连接池有三级hash策略，保证取到的连接可以正确使用
 - read/write
 - 用户权限
 - 连接（ session ）属性，如client_found_rows等（ include/mysql_com.h ）

✓ 流量控制

- 并发数控制
- QPS控制



✓ Dbproxy集群化部署与运维

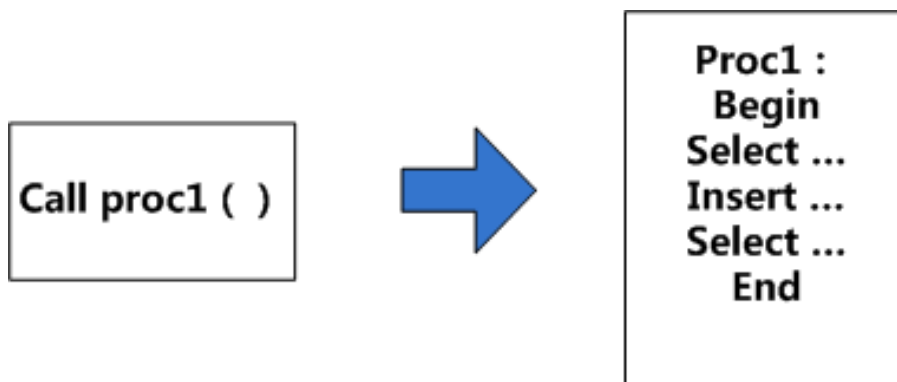
- 配置信息热加载
- 支持连接多个数据库集群
- 自身压力信息准确输出
- 流量控制

✓ Dbproxy对应用程序透明策略

- MySQL ClientServer protocol
- 数据库 (MySQL) 某些特性
- 各种语言的MySQL API , 各种MySQL的连接框架的特殊行为

✓ 存储过程的支持

- 支持 “一条SQL对应多个结果集” 地接收



✓ Dbproxy提高性能策略

- 减少延迟
 - 进程+状态机的模式，无锁，nonblock，网络操作都是就绪后才执行
 - 对于大结果集的接收，采用了非连续接收的模式，减少其他连接的等待
- 提高并发处理能力
 - 支持多个dbproxy工作进程监听一个端口

✓ Dbproxy性能数据

- 单进程QPS可达17000
- 请求经过dbproxy的延迟为百微秒级

dbproxy的延时	数据量	压力值	压力端进程数
149us	100字节	10547次/s	4
126us	800字节	5311次/s	4
205us	800字节	16994次/s	8
257us	2000字节	13730次/s	8
261us	7000字节	3139次/s	2
351us	7000字节	5260次/s	4
388us	20000字节	2000次/s	2
689us	20000字节	2828次/s	4

- 系统资源消耗少
 - 单dbproxy进程每日1200w请求
 - cpu消耗小于1%，内存消耗约200M

MySQL数据库集群面临的问题

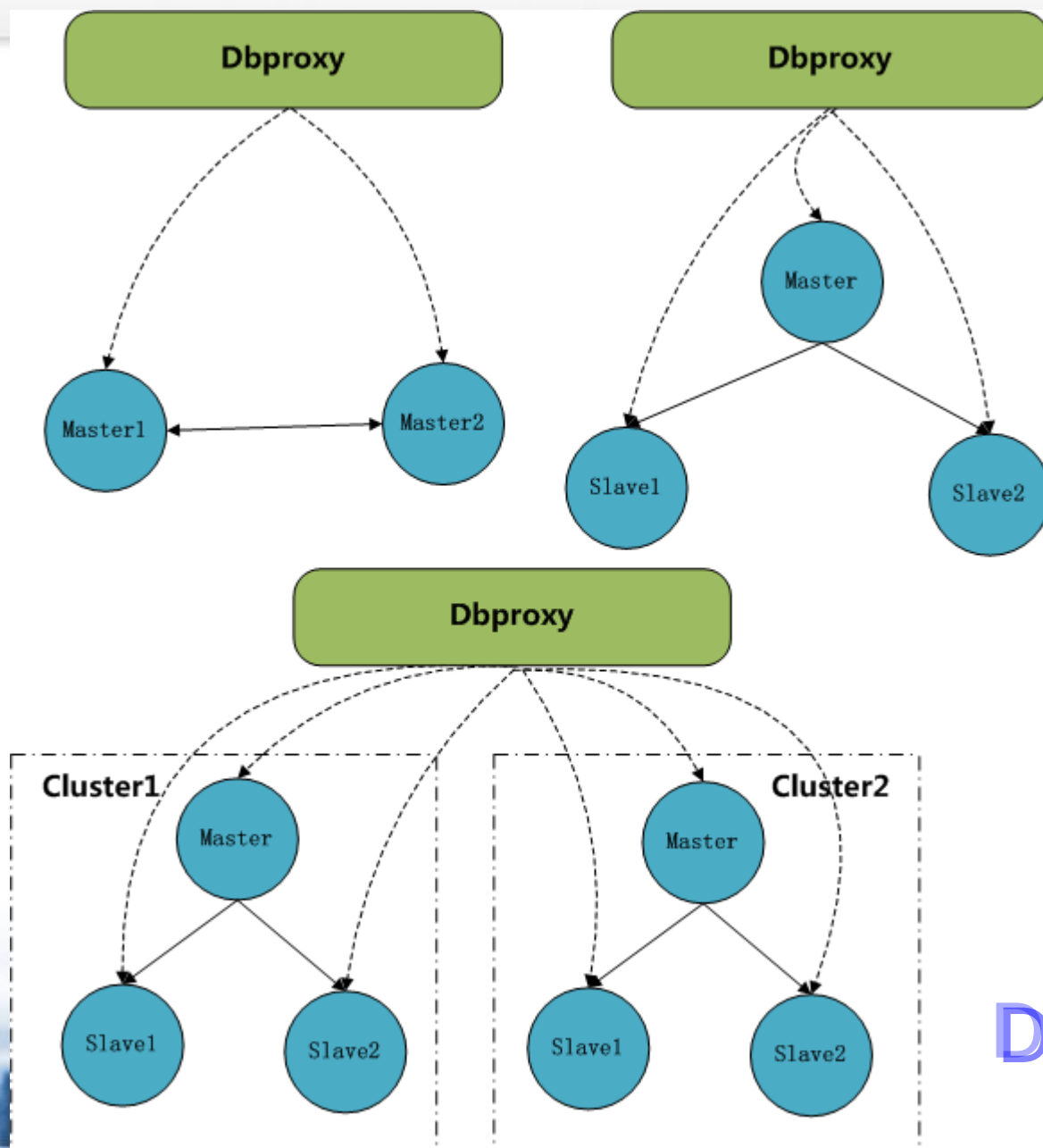
数据中间层整体设计

数据库中间层主要功能模块及策略



数据库中间层应用

数据库中间层发展



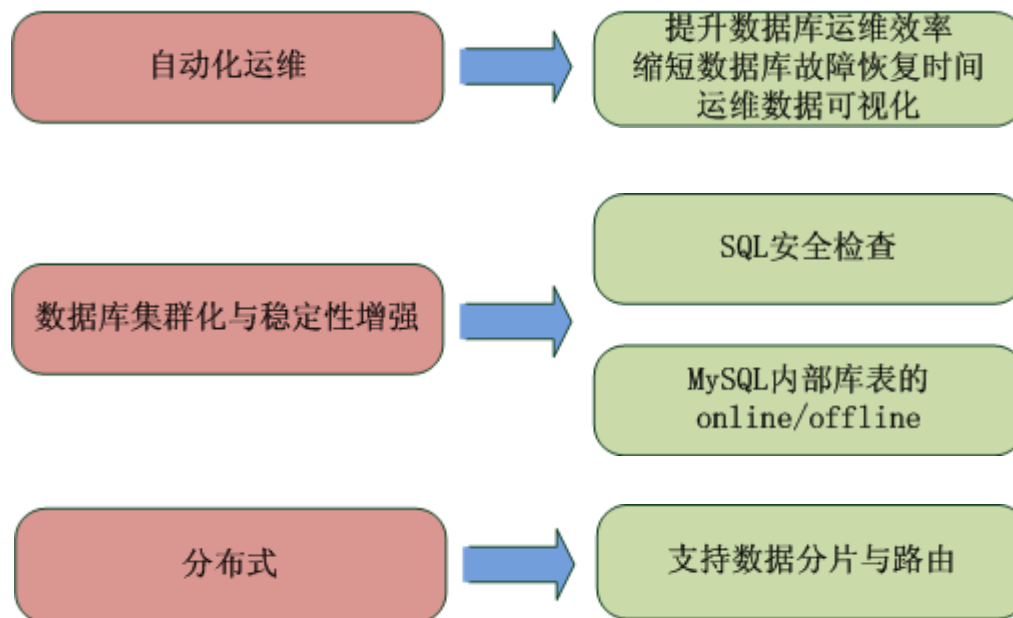
MySQL数据库集群面临的问题

数据中间层整体设计

数据库中间层主要功能模块及策略

数据库中间层应用

数据库中间层发展



谢谢！



DTCC2012