



# DRDS分库分表—— RDS关系数据库云服务的水平扩容技术

阿里云RDS产品线- 皓庭

[haoting.wq@taobao.com](mailto:haoting.wq@taobao.com)

2014.04.10



# 主要内容

---

总览

如何使用

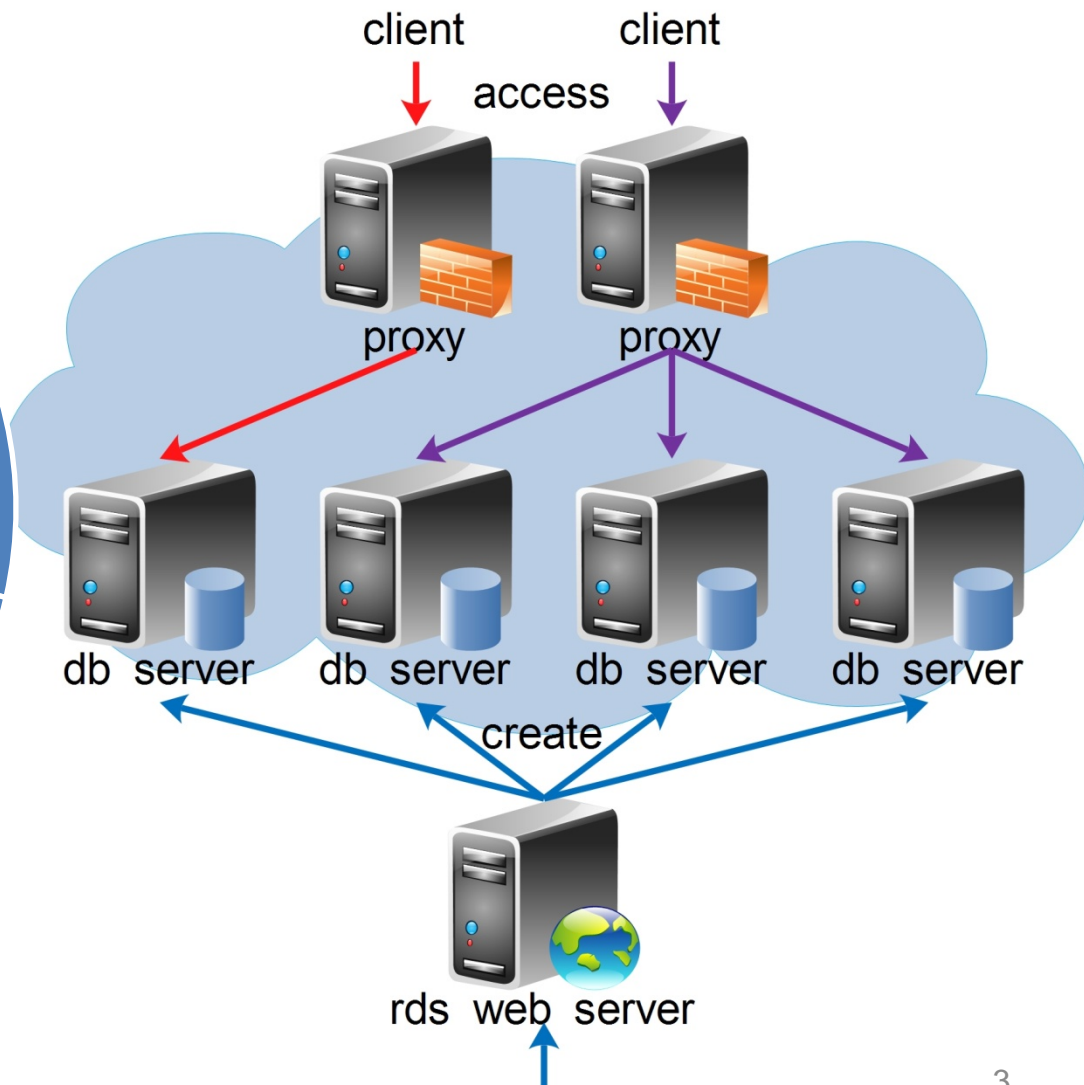
如何工作

目前的限制

将来的工作



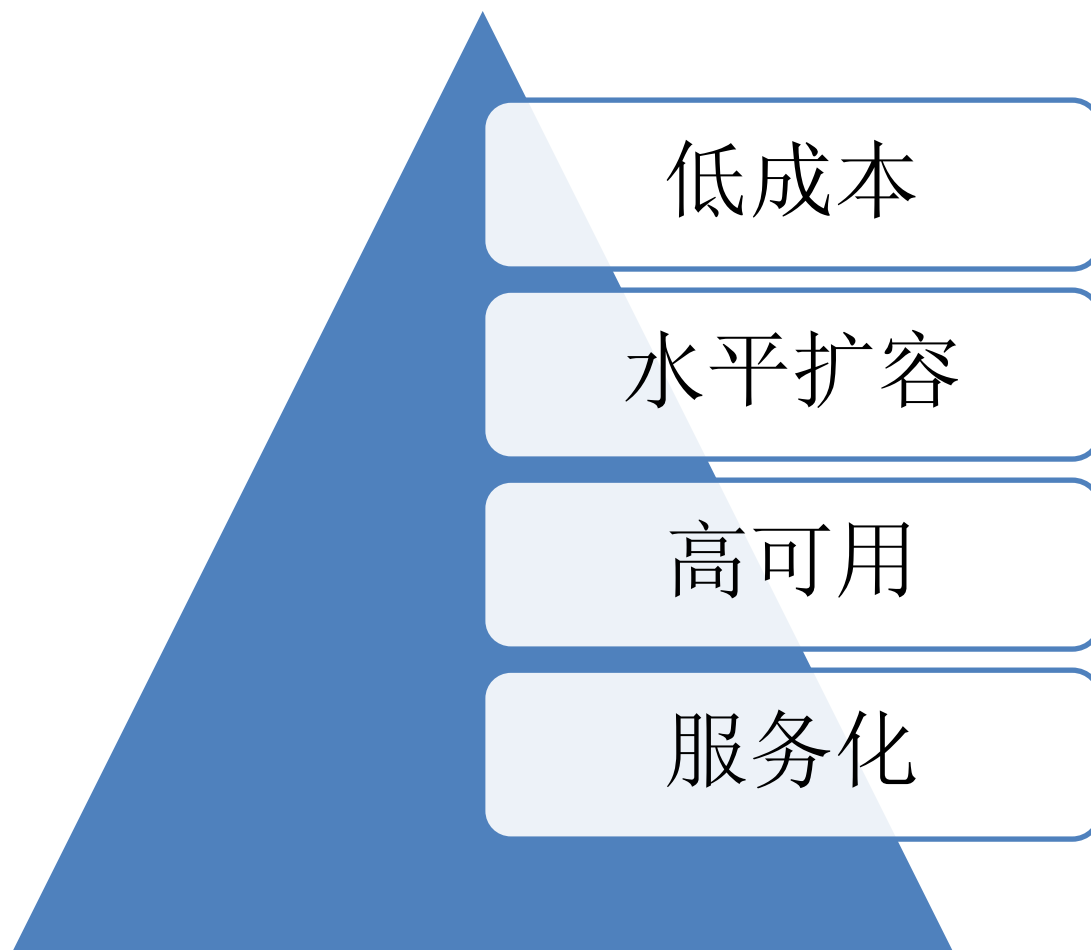
# 总览——RDS关系数据库云服务





# 总览——大数据引入的问题

---





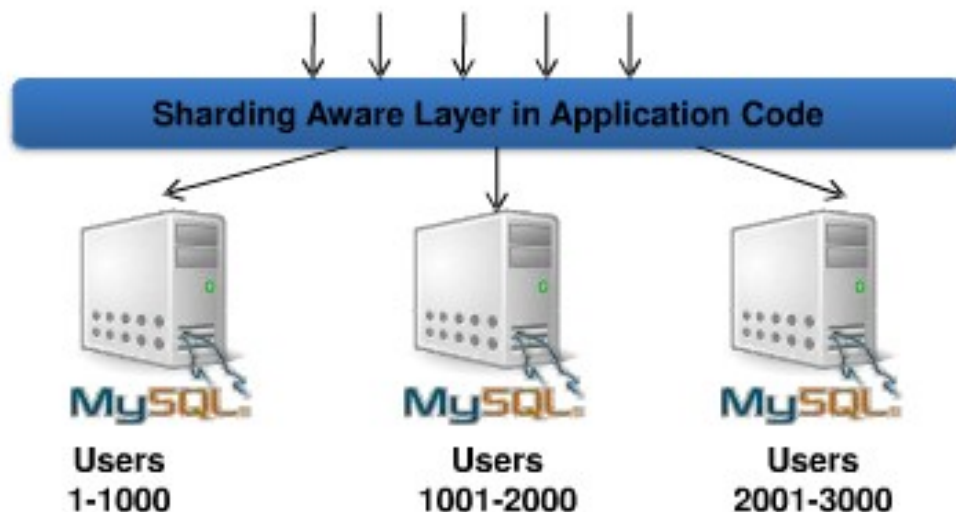
# 总览——DRDS分库分表

RDS为大数据用户提供的分库分表服务

需新建DRDS账户实例，目前仅支持MySQL

用户购买RDS实例作为DRDS实例的分区

满足常用的业务场景，但有一定限制





# 如何使用——必要操作

## 创建实例

- 申请RDS实例作为DRDS的分区
- 创建用户表，指定主键和分区字段
- 数据迁入或直接访问

## 扩容

- 添加新的RDS实例
- 使用扩容服务



# 如何使用——注意事项

创建DRDS实例前，需先申请若干RDS实例作为分区

用户不能通过DDL语句建立表，而需要通过web控制台

创建表需要预留额外的主键和分区键供给DRDS使用

主键来源

```
create table rooler.test_table(  
    primary_key  
        unsigned int64  
        auto_increment not null,  
    partition_key int not null,  
    field_other type,  
    ...)
```

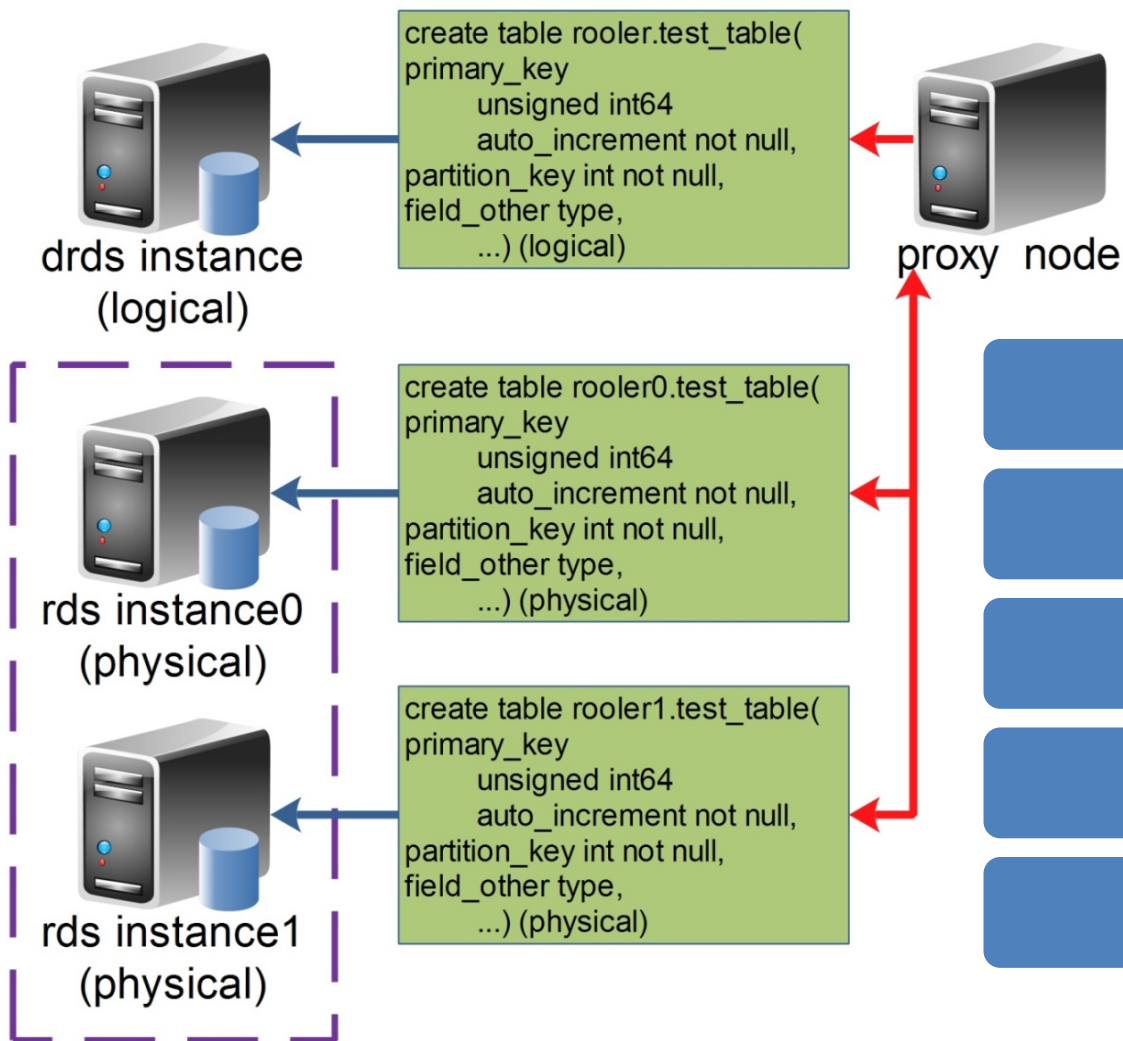


# 如何工作——库表结构

create instance

create table

access table



drds实例vs.rds实例

逻辑库vs.物理库

逻辑表vs.物理表

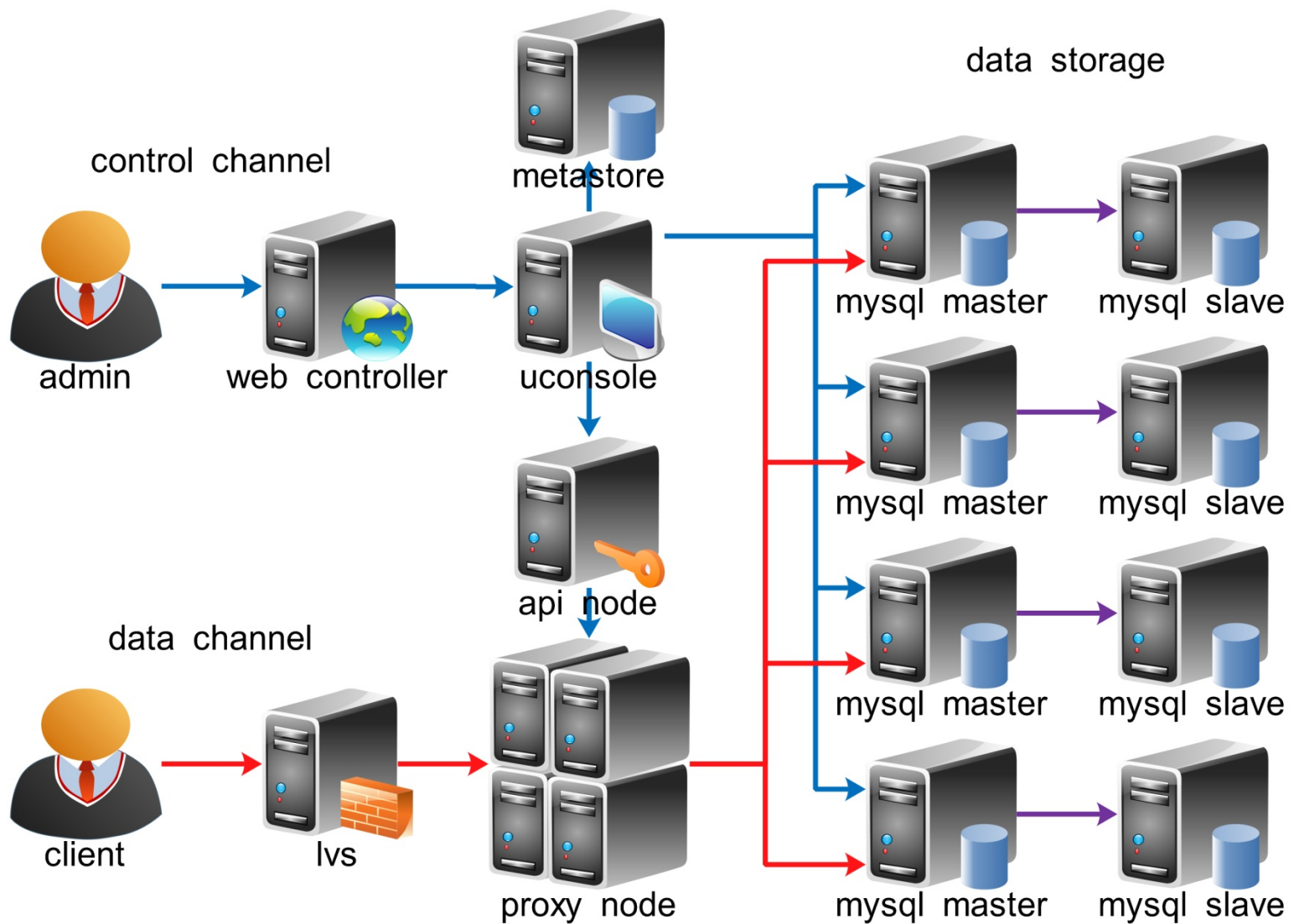
逻辑请求vs.物理请求

逻辑响应vs.物理响应





# 如何工作——架构





# 如何工作——控制通道

## 控制web端:

- 提供DRDS实例创建、销毁、建表等功能

## 实例控制器uconsole:

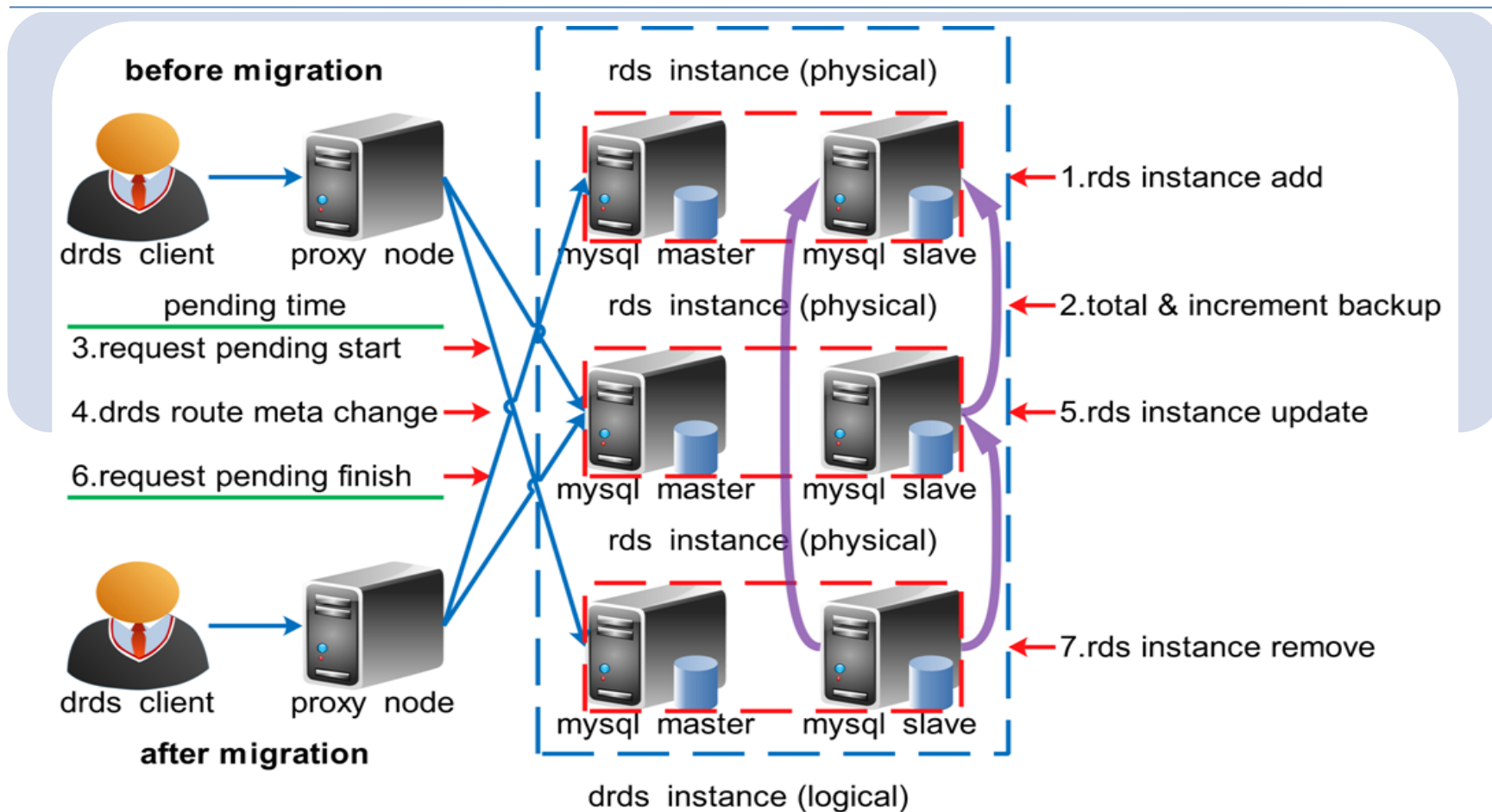
- 批量建立物理库表、账户等
- 数据迁移

## 元数据存储:

- 存储逻辑与物理实例的映射关系等



# 如何工作——迁移过程



物理数据迁入迁出、逻辑与物理映射关系改变的过程，数据扩容等均通过迁移完成



# 如何工作——数据通道

api节点：缓存元数据，控制proxy节点

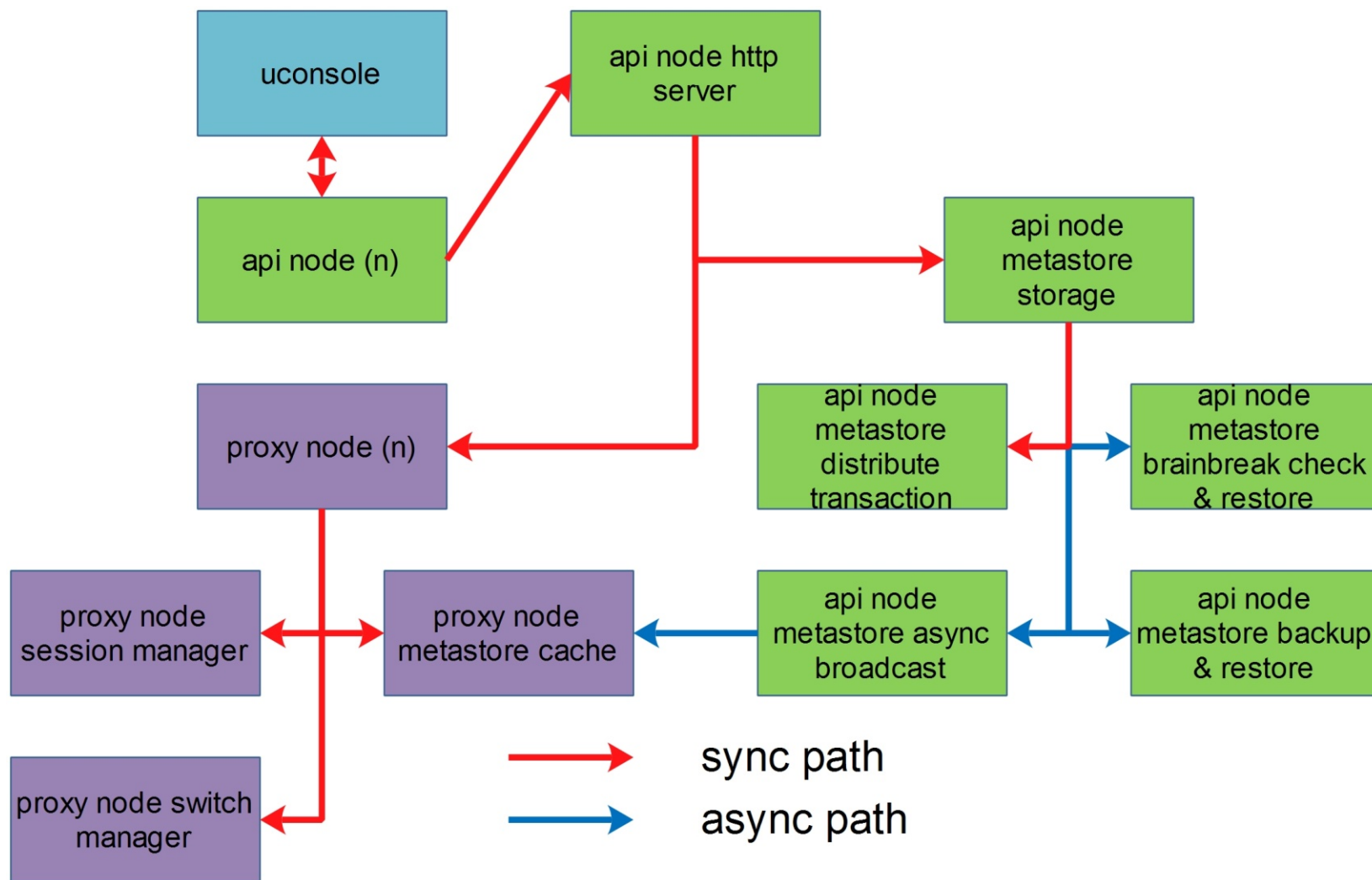
- 高可用的元数据缓存
- 物理实例的切换控制
- 逻辑实例的扩容重映射

proxy节点：与用户、RDS物理分区交互

- 会话管理、切换控制、兼容性
- 语义分析、请求路由、语句改写、执行计划
- 事务、环境变量、请求控制、结果合并
- 热升级

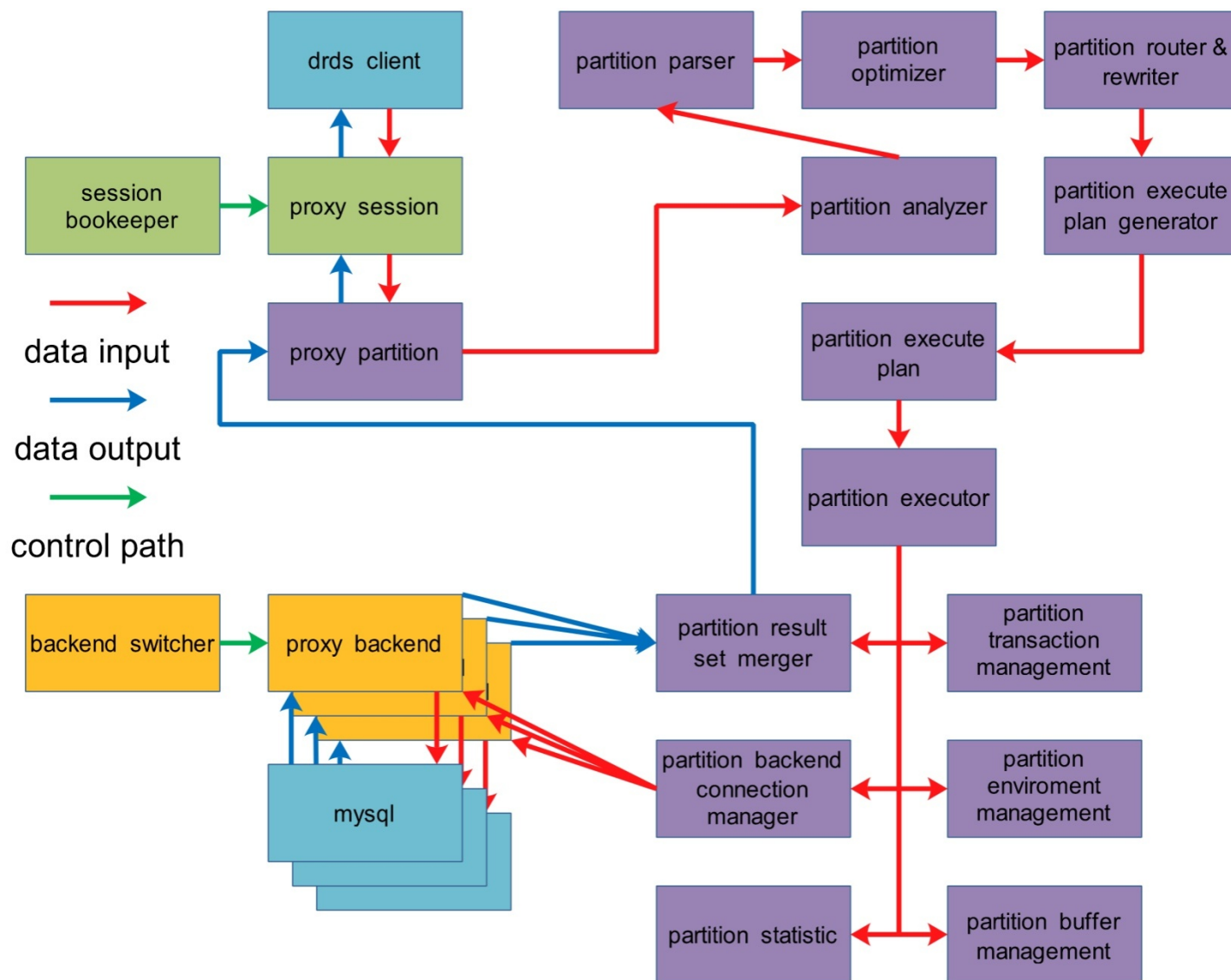


# 如何工作——api节点技术





# 如何工作——proxy节点技术





# 如何工作——数据存储

## 订制的MySQL内核

- 配置参数
- 命令

## 灵活的复制结构

- 主从复制
- 多主复制

## 多样的存储引擎

- innodb
- tokudb

## 强大的运维工具

- transfer
- flashback





# 如何工作——高可用与高可靠

## 无单点

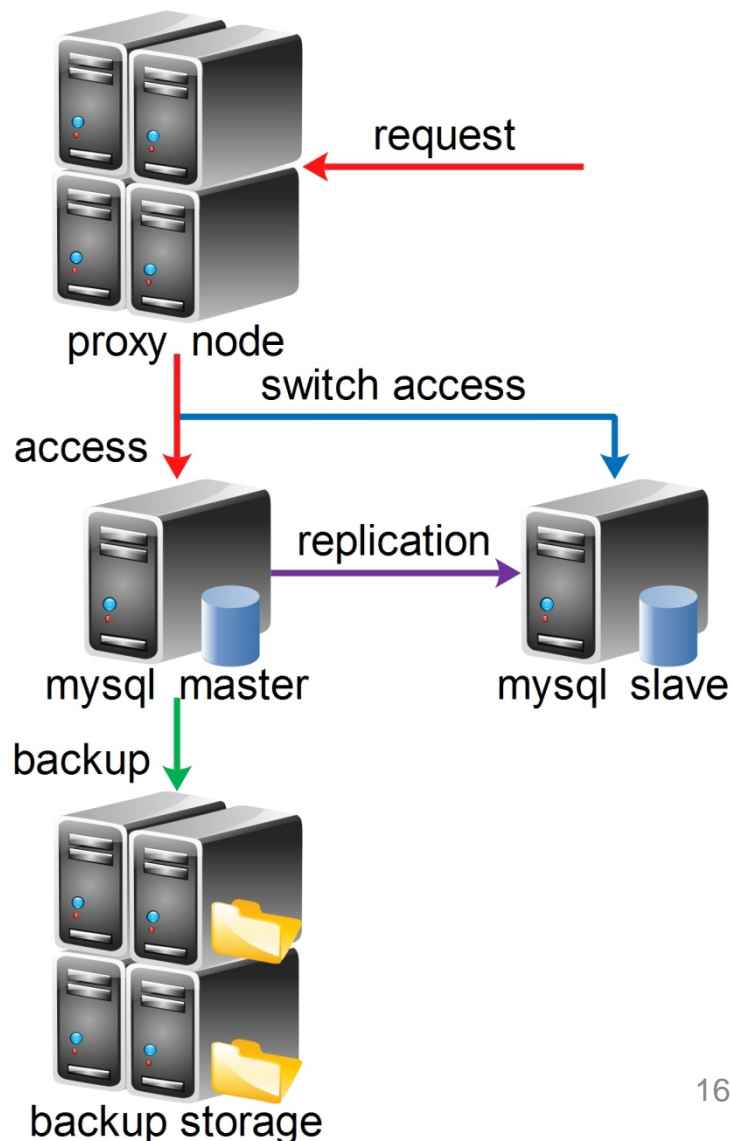
- 复数的元数据存储节点
- 复数的控制节点
- 复数的proxy节点

## 高可用的数据存储

- mysql主从复制结构
- 快速切换

## 数据备份

- 大规模的后备存储







# 目前的限制

## 表

- 字符集、完整性约束、外键限制
- 数据规模有上限

## SQL语句

- 不支持DDL语句
- 支持单表CRUD语句和事务，暂不支持JOIN
- 支持普通语句的数据合并，受限类型的order by、accumulate、group by、distinct、limit
- 有限的跨分区事务支持和环境变量支持



# 将来的工作

---

跨分区两阶段提交事务支持

更多的查询语句支持

更大的数据规模

强化资源监控

# 谢谢

---

[haoting.wq@taobao.com](mailto:haoting.wq@taobao.com)

@皓庭霄度天

