

2013

# 阿里技术沙龙

享技术·聚朋友

D2



## DBFree-阿里数据库自动化运维平台

aDev

iDevOps

陈旭 @ Alibaba  
2013-12

# 个人简介

- 2009年加入阿里
- 负责阿里广告、CRM、商品、社区等业务线
- 经历阿里去IOE，推动广告、商品等业务从O到M的转变
- orzdba/orztop工具的开发维护者
- 微博：@Orz\_dba
- 目前关注数据库自动化运维

# 目录

- 背景
- 平台功能
  - 资源管理分配
  - 在线扩容
  - 其他
- 实现自动化过程
- 平台架构
- 平台使用情况

# 背景

## • Oracle -> MySQL运维方式的转变

	Oracle	MySQL
服务器	小型机+存储	PC Server
软硬件成本	高	低
管理方式	集中	分散
实例数量	少	多
集群运维复杂度	低	高

- 业务飞速发展
- 每年双十一、双十二活动都要兴奋一下

## • MySQL服务器集群规模化

## • 常见的日常实例运维

- 业务新项目资源分配和环境初始化



- 主机硬件问题、过保下线、机房迁移

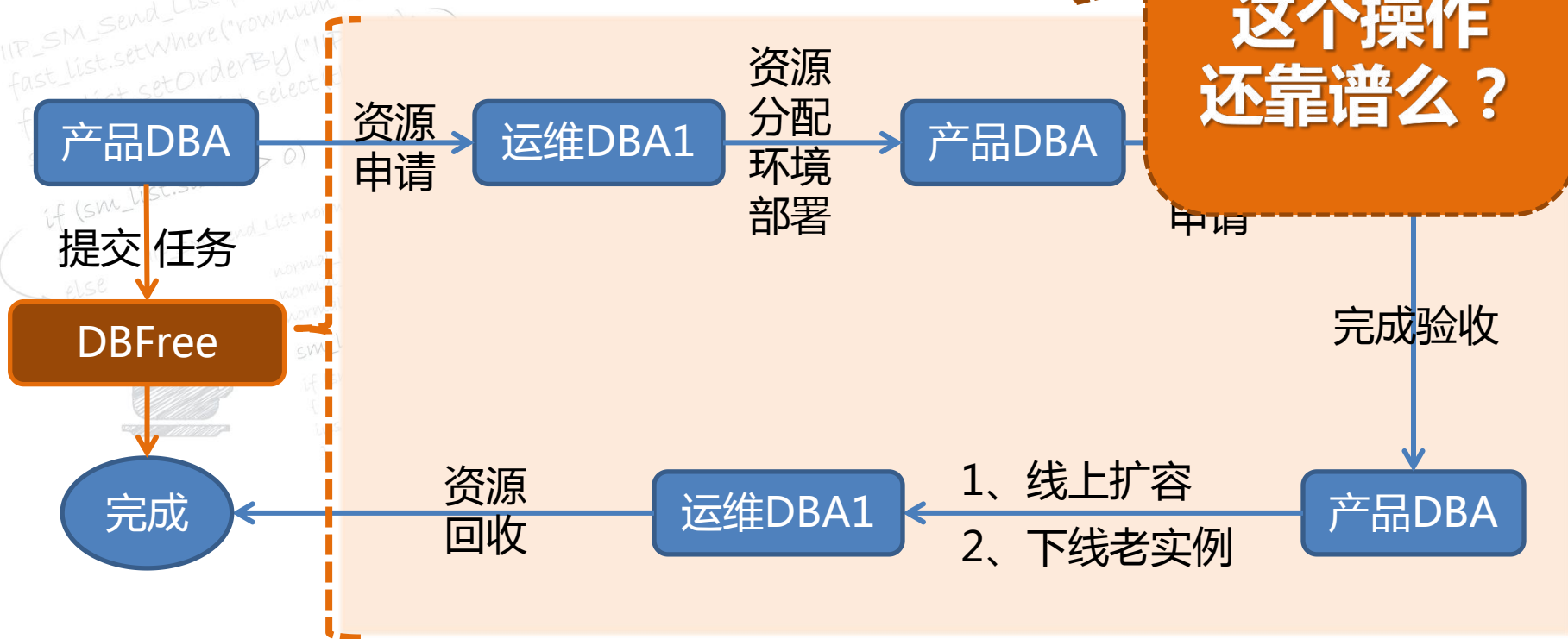
- 业务大促（1111、1212）扩容支持

- 活动后，低水位机器收缩

- 合并，多实例共用，实例迁移

- 重启，减少分配的实例资源

## 扩容流程



### 问题：

- 一次扩容需要至少3个DBA进行至少5个环节的沟通，低效。
- 扩容变更繁琐，人工操作易出错，耗时耗力。



# 平台功能

- 资源管理

- 部署、监控、分配、回收下线

- 在线扩容

- 替换主/备库、添加备库、主备库实例迁移
- 下线、拆分
- 数据迁移

- 备份恢复管理

- 异常容灾

- ADHA

- 在线分析诊断

- 异常性能诊断
- 水位分析监控

# 资源管理分配

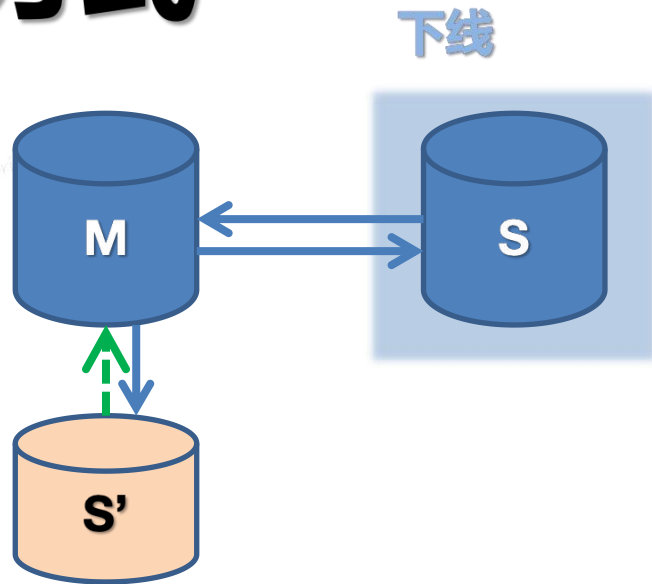
- 资源池分组
- 资源初始化部署 ( puppet )
  - 环境初始化、部署crontab定时调度
  - 模块脚本同步推送
  - MySQL、ADHA、dbagent、stargate等基础服务软件安装升级
- 资源监控
  - 本地dbagent定时采集软硬件信息将元数据中心汇报
  - 监控中心分析资源水位，异常告警，进行相应诊断任务处理
- 资源自动分配
  - 如何才能合理重复利用资源？
- 资源回收下线



# 实例基本操作方式

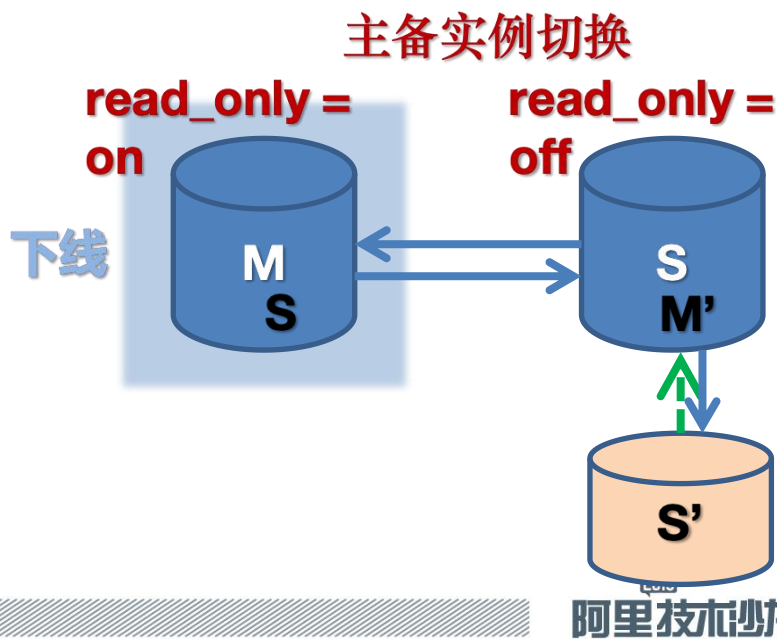
## 场景一（替换备库实例）

- ① 主机新实例资源分配
- ② 备份恢复新备库，无延时
- ③ 替换备库TDDL
- ④ 下线老备库实例



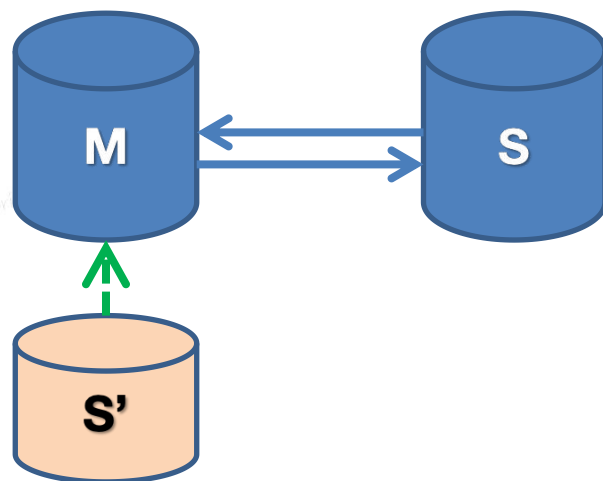
## 场景二（替换主库实例）

- ① 主机新实例资源分配
- ② 备份恢复新备库，无延时
- ③ 老的主备实例切换
- ④ 替换TDDL
- ⑤ 下线老主库实例



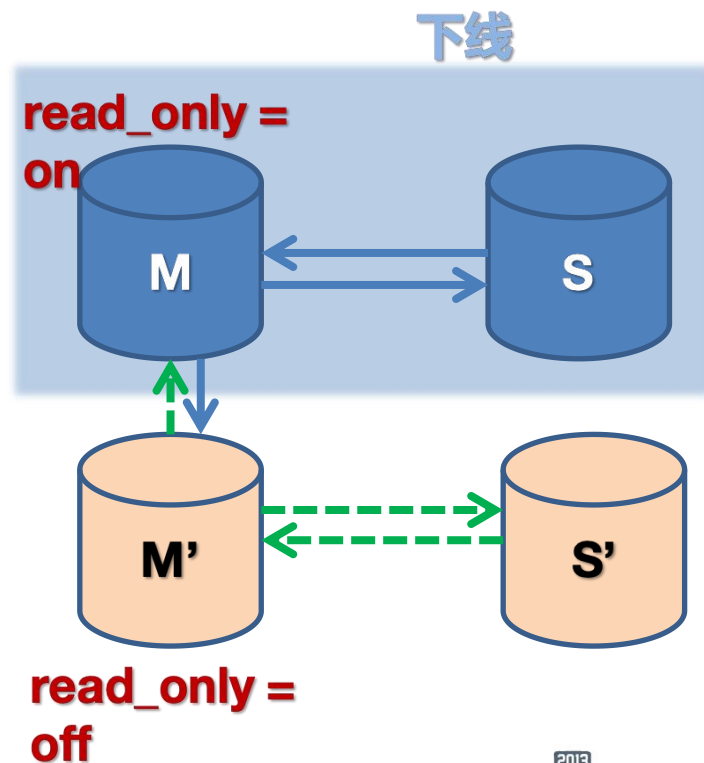
### 场景三（添加备库）

- ① 主机实例资源分配
- ② 备份恢复备库
- ③ TDDL配置添加新备库  
TDDL均衡读流量



### 场景四（主备实例迁移）

- ① 主机实例资源分配
- ② 备份恢复新主备库
- ③ DB+TDDL切换到新主备库
- ④ 老主备库实例下线

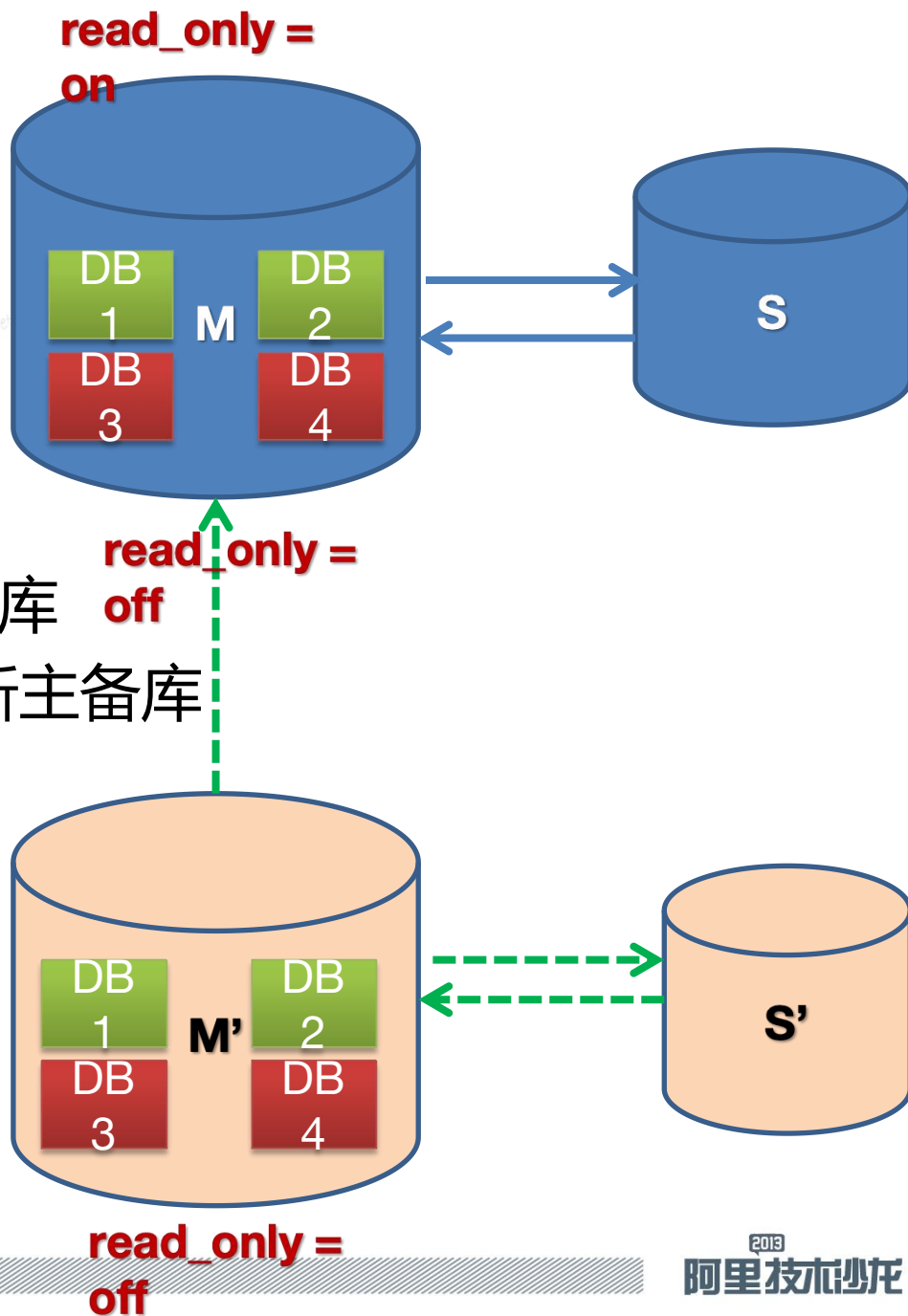


## 场景五（实例下线）

## 场景六（实例分库拆分迁移）

- ① 主机实例资源申请
- ② 备份恢复新拆分DB的主备库
- ③ 新拆分DB+TDDL切换到新主备库
- ④ 老拆分DB下线

## 场景七（愚公数据迁移）



# 实现自动化过程

- 运维变更抽象模块化

- 资源上下线

- 备份恢复

- 数据迁移

- HA

- 外部系统接口

- 环境标准统一

- 监控、基础运维脚本等

- 流程标准化、模块组合自动化

# 平台架构

- WEB Server

- 前台：自动化任务流程创建等  
在线分析诊断和服务器性能指标水位展现等

- 后台：提供API服务

- 后台调度运行系统

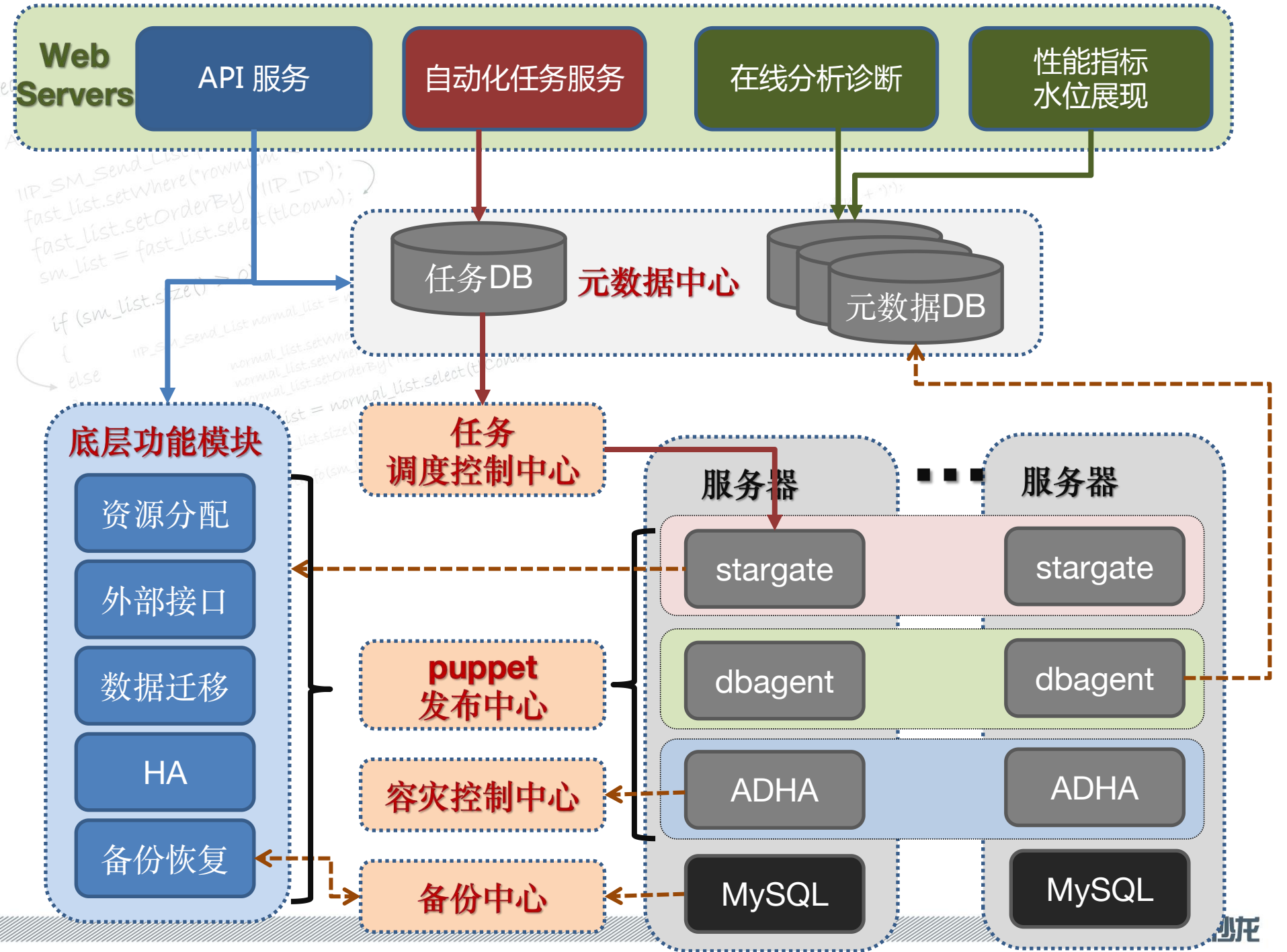
- 任务控制中心调度agent

- 底层功能模块

- 负责原子性功能：资源分配、备份恢复、数据迁移、HA、外部接口

- 基础服务组件

- puppet发布中心、备份中心、监控元数据中心、容灾控制中心
- 任务运行agent：stargate
- 性能采集agent：dbagent
- 异常容灾切换：ADHA





# 平台使用情况

## • 下半年数据

类型	能力
资源实例分配	5000 次以上
实例操作（扩容、拆分、下线等）	5000 次以上
ADHA机房异常容灾演练（ <b>5分钟</b> ）	1000+ 机器
puppet <b>30分钟</b> 资源部署初始化	500+ 机器
双11 <b>一周</b> 扩容完成	1000+ 机器
双12 活动结束 <b>一天</b> 下线	400+ 机器

```

can_is_empty = false;
ArrayList<IP_SM_Send_List> sm_list;
IP_SM_Send_List fast_list = new IP_SM_Send_List();
fast_list.setWhere("rownum <= " + row_count);
fast_list.setOrderBy("IP_ID");
sm_list = fast_list.select(tlConn);
if (sm_list.size() > 0)
{
    IP_SM_Send_List normal_list = new IP_SM_Send_List();
    normal_list.setWhere("IP_ServiceID in (" + m_service_handler.getServiceList(m_int_normal_priority) + ")");
    normal_list.setWhere("rownum <= " + row_count);
    normal_list.setOrderBy("IP_ID");
    sm_list = normal_list.select(tlConn);
    if (sm_list.size() > 0)
    {
        insertInfo(sm_list, m_service_handler);
    }
}
else
{
}

```



# 谢谢