

美团点评数据库运维自动化实践与发展

美团点评高级DBA

翁宁龙

技术架构未来



新美大技术团队

自我介绍

➤ 我是谁?

- 翁宁龙
- 2012年毕业参加工作成为一名DBA
- 2015年加入美团，负责美团数据库运维自动化系统、数据库生态的建设工作

➤ 联系方式

- 邮箱: 446978574@qq.com
- 微信:



目 录

➤ 美团点评数据库生态系统

- 为什么要打造数据库生态系统?
- 美团点评数据库生态是什么样的?

➤ 案例分享

- 集群初始化
- 在线表变更
- 高可用方案

➤ 极简运维

- 极简运维理念
- 场景化、定制化、智能化

➤ Q & A

目 录

➤ 美团点评数据库生态系统

- 为什么要打造数据库生态系统?
- 美团点评数据库生态是什么样的?

➤ 案例分享

- 集群初始化
- 在线表变更
- 高可用方案

➤ 极简运维

- 极简运维理念
- 场景化、定制化、智能化

➤ Q & A

美团点评数据库生态系统 - 为什么要打造数据库生态系统?

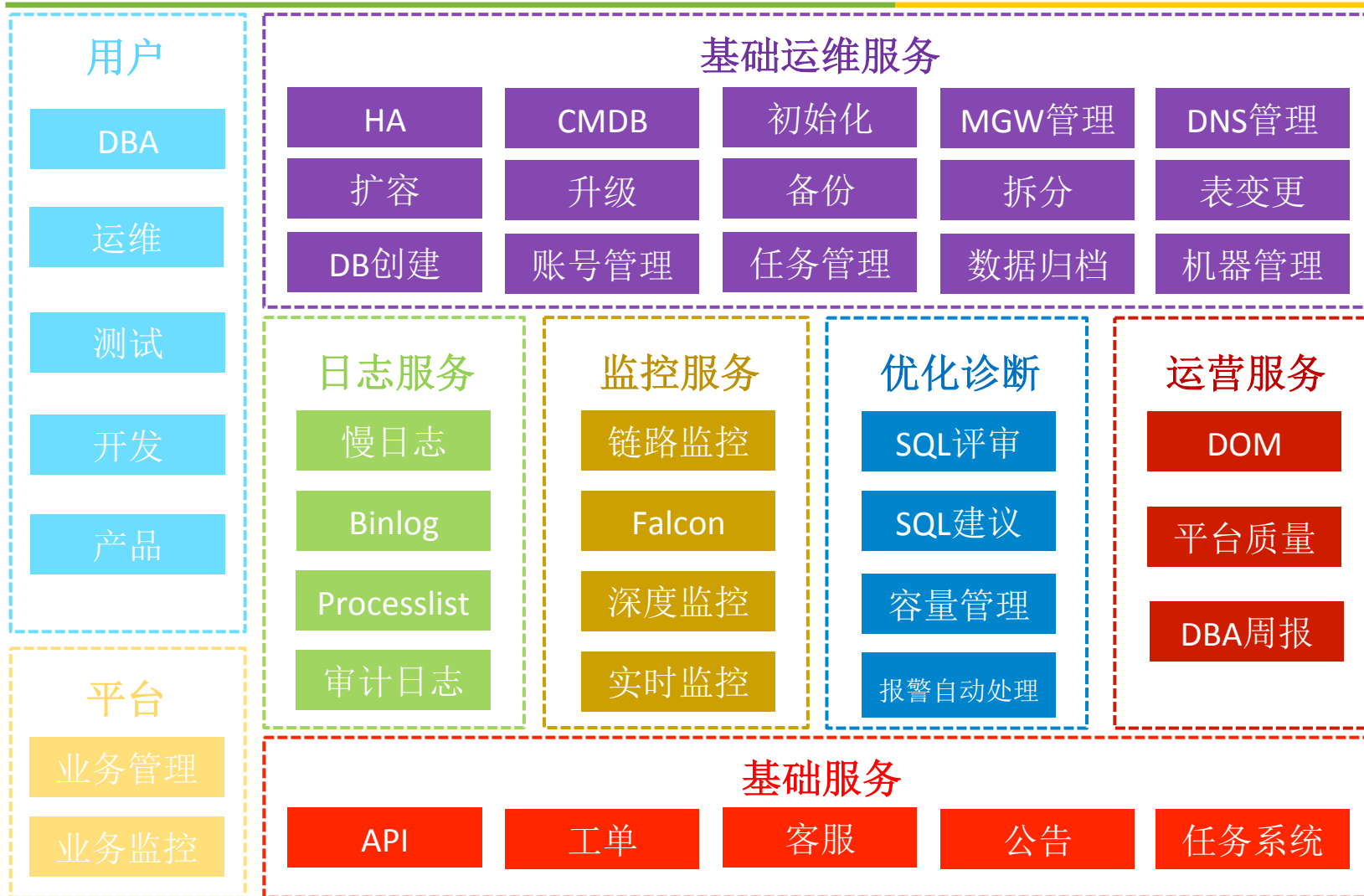


.....

数据库产品&服务



美团点评数据库生态系统 - 美团点评数据库生态是什么样的?



标准化

- 接口标准
- 数据标准
- 流程标准

自助化

- 操作自助
- 问题定位自助

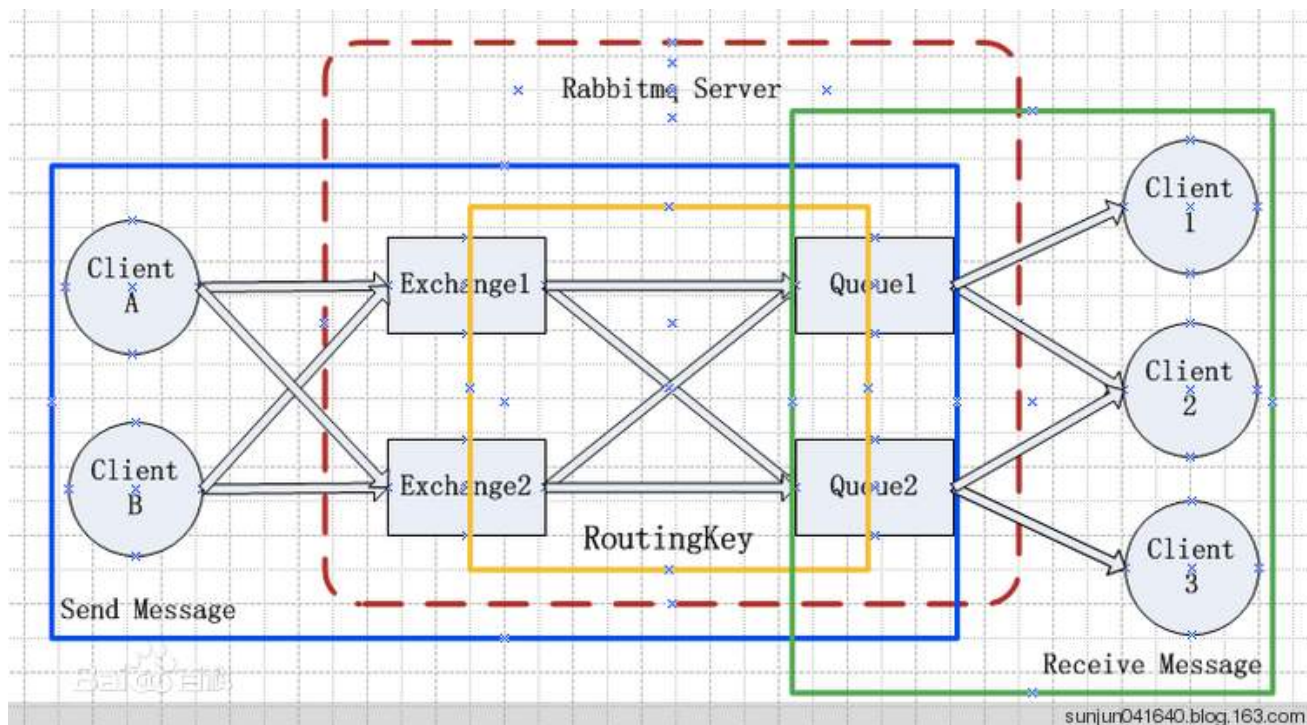
自动化

- 高可用
- 报警处理

美团点评数据库生态系统 - 任务管理系统的构建

✓ RabbitMQ

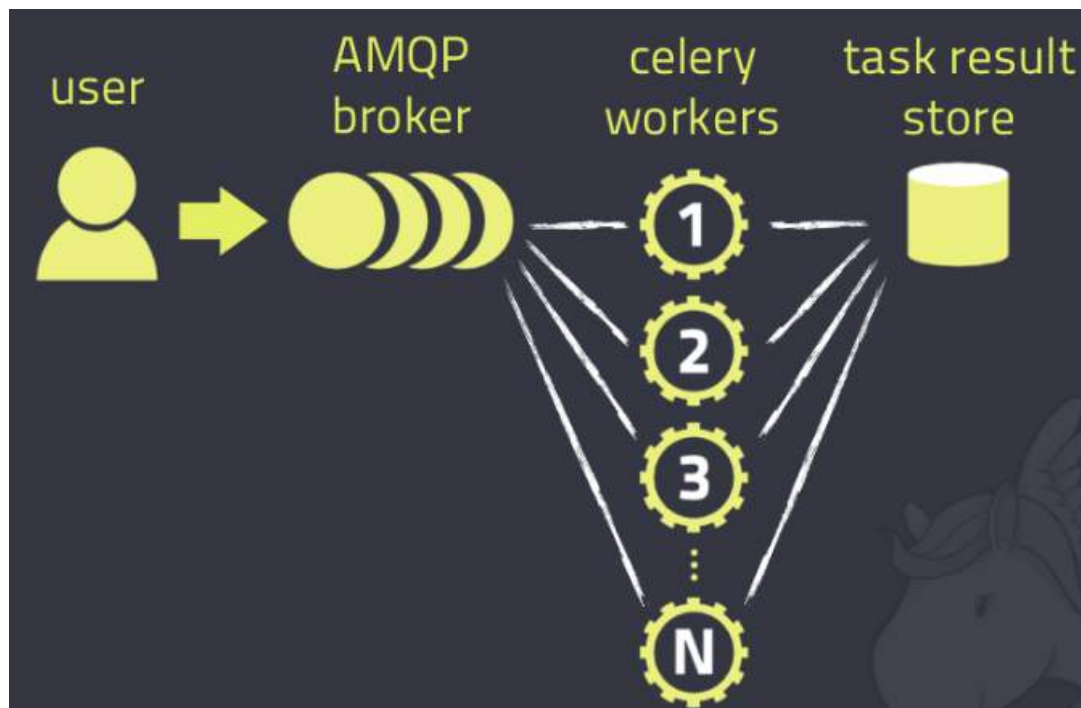
消息队列 (MQ) 是一种应用程序对应用程序的通信方法。应用程序通过读写出入队列的消息 (针对应用程序的数据) 来通信, 而无需专用连接来链接它们。



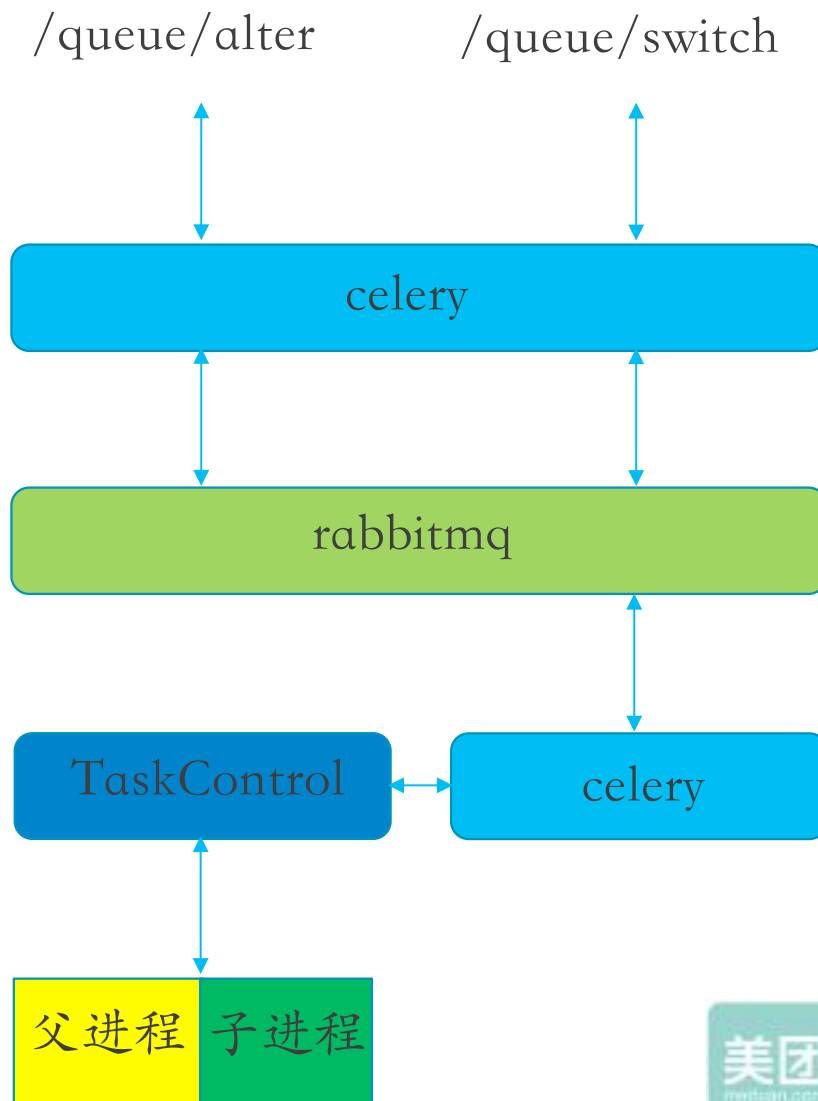
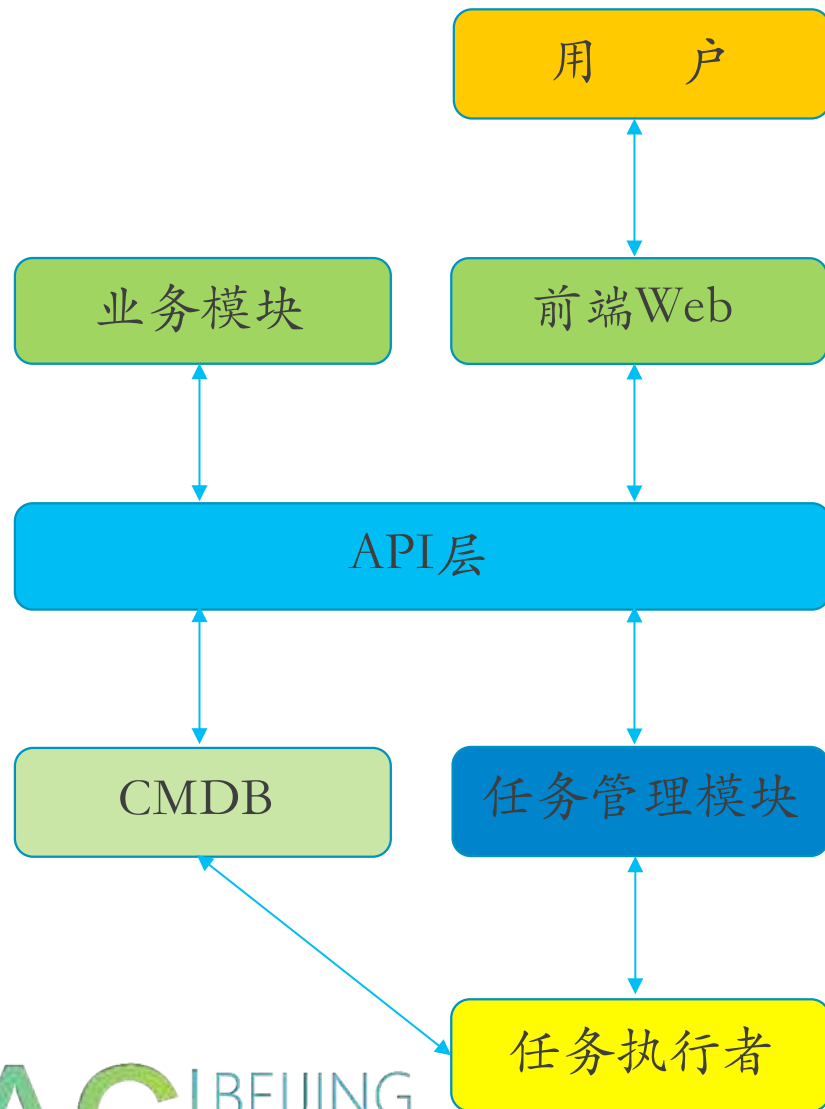
美团点评数据库生态系统 - 任务管理系统的构建

✓ Celery

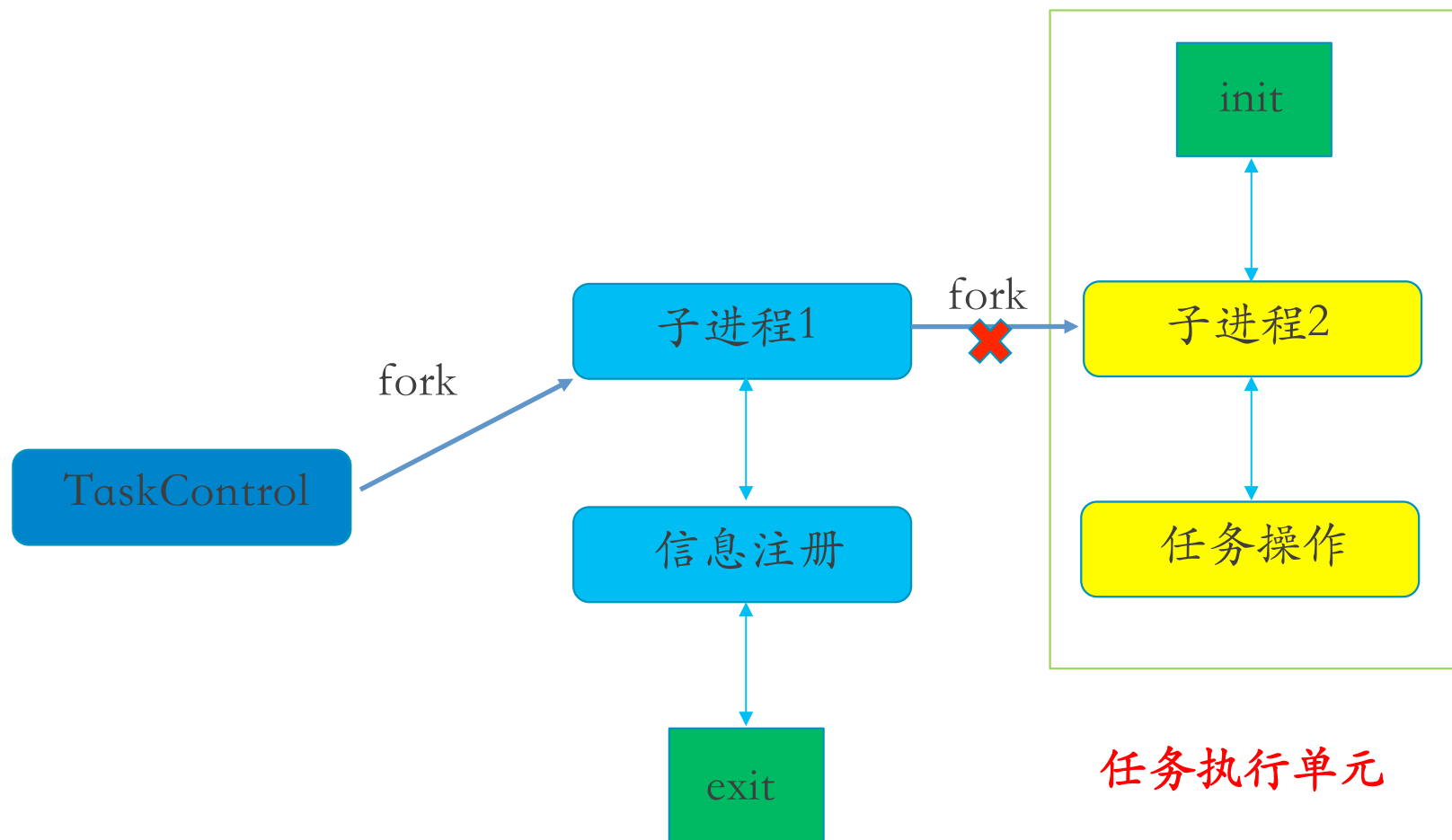
Celery（芹菜）是基于Python开发的分布式任务队列。它支持使用任务队列的方式在分布的机器/进程/线程上执行任务调度。Celery的架构由三部分组成，消息中间件（message broker），任务执行单元（worker）和任务执行结果存储（task result store）组成。



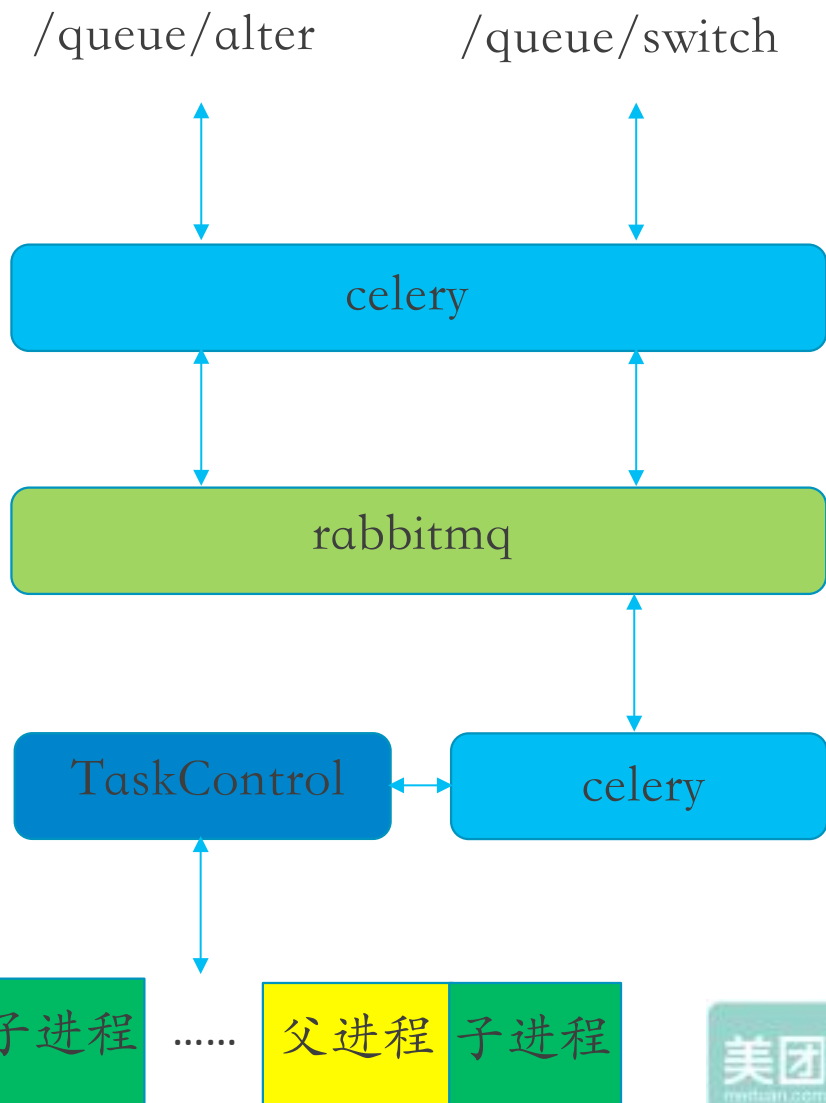
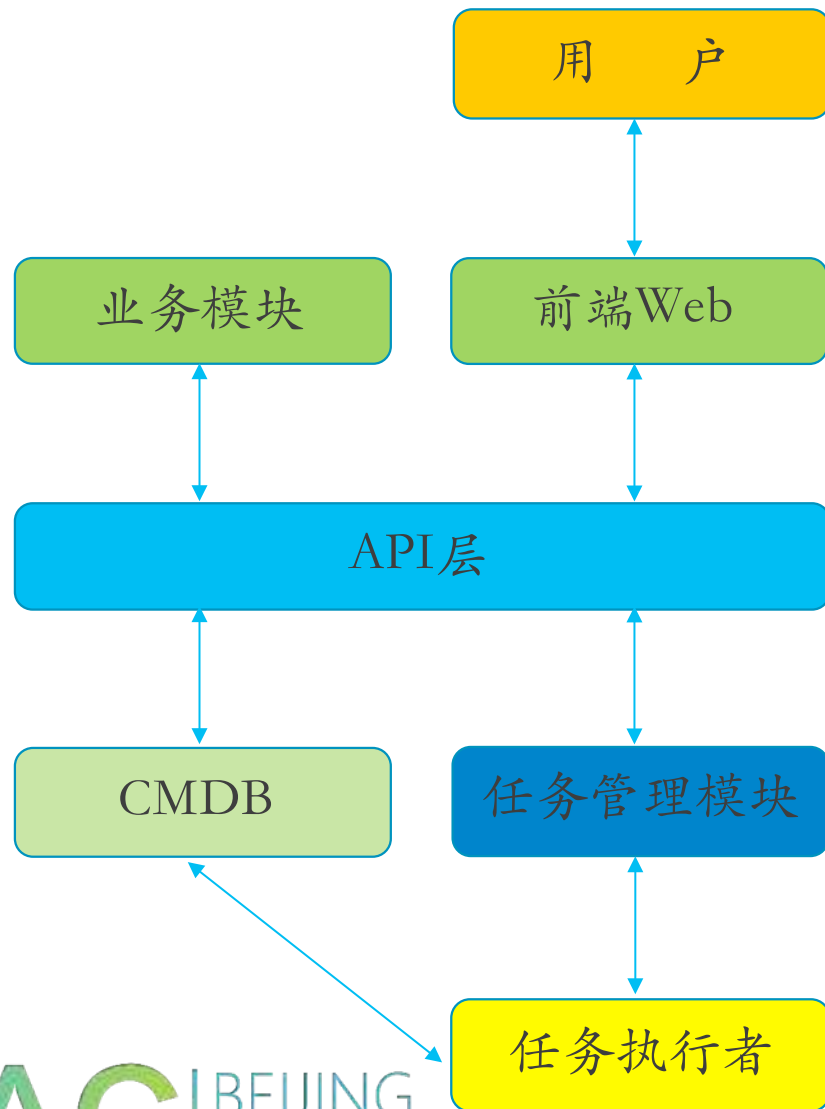
美团点评数据库生态系统 - 任务管理系统的构建



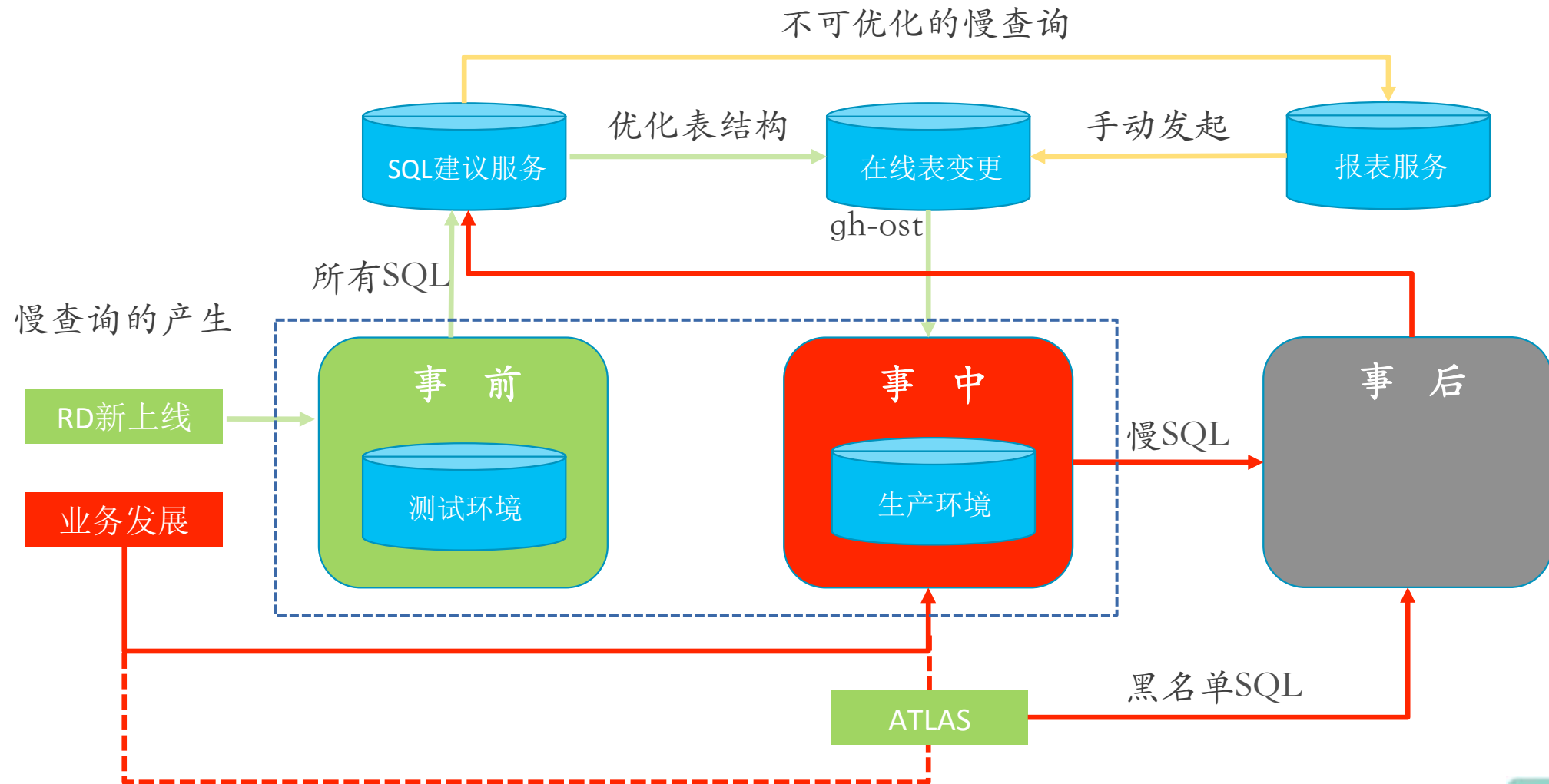
美团点评数据库生态系统 - 任务管理系统的构建



美团点评数据库生态系统 - 任务管理系统的构建



美团点评数据库生态系统 - 美团点评数据库生态是什么样的? - 慢查询



美团点评数据库生态系统 - 美团点评数据库生态是什么样的? - 量化运营

慢查询

业务组: [redacted] ; 服务组: -全部- ; 开始时间: 2016-12-01 ; 结束时间: 2016-12-15 ; 查询

SQL: [redacted]

SQL建议: 无SQL优化建议

索引建议: Create_Index_SQL: alter table [redacted] add index [redacted]

分析日志: [打开详情](#)

[redacted]	select [redacted]	1124	2.60	优化建议	慢查询系统链接
[redacted]	[redacted]	890	2.61	优化建议	慢查询系统链接
[redacted]	[redacted]	16	0.16	优化建议	慢查询系统链接

目 录

➤ 美团点评数据库生态系统

- 为什么要打造数据库生态系统?
- 美团点评数据库生态是什么样的?

➤ 案例分享

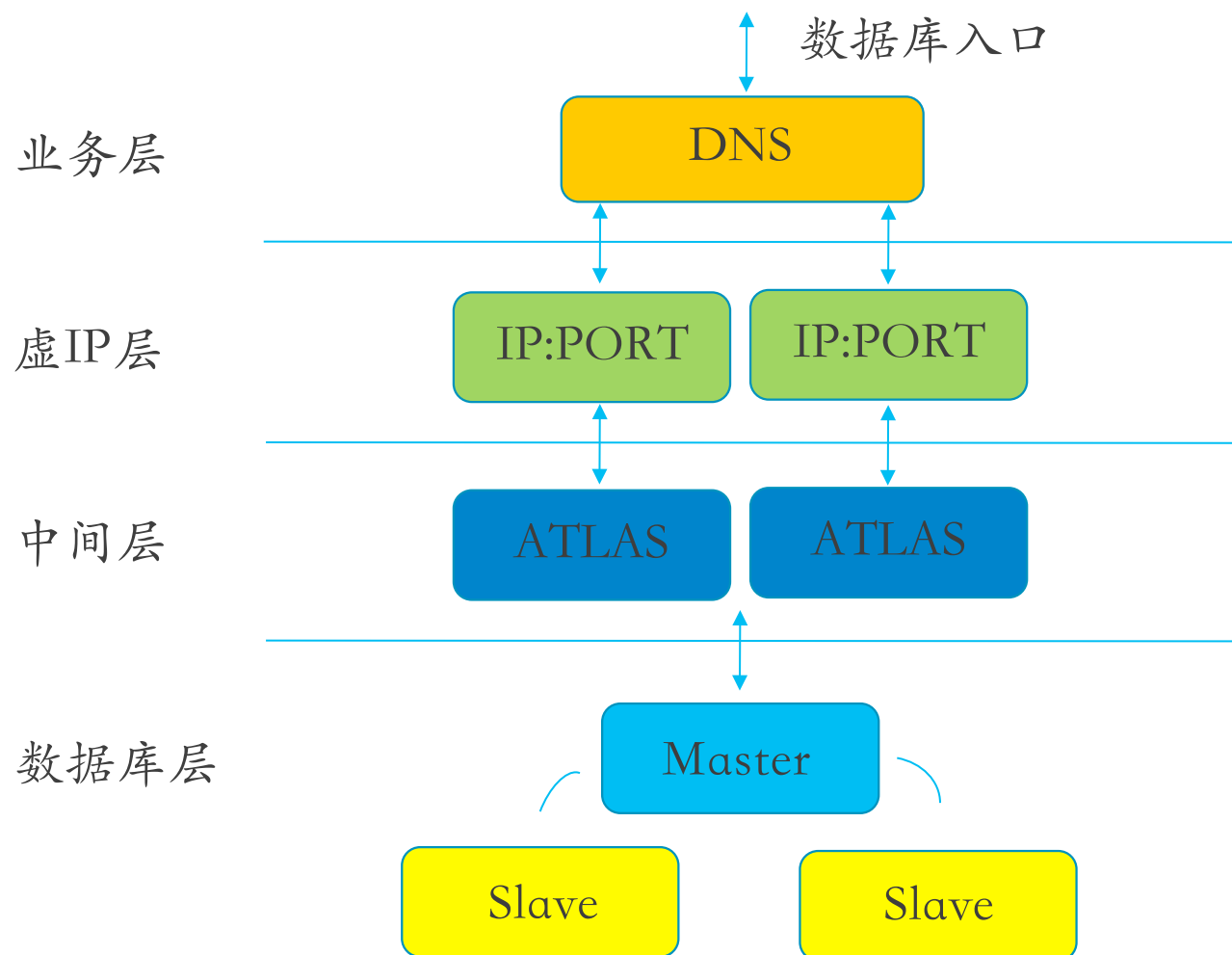
- 集群初始化
- 在线表变更
- 高可用方案

➤ 极简运维

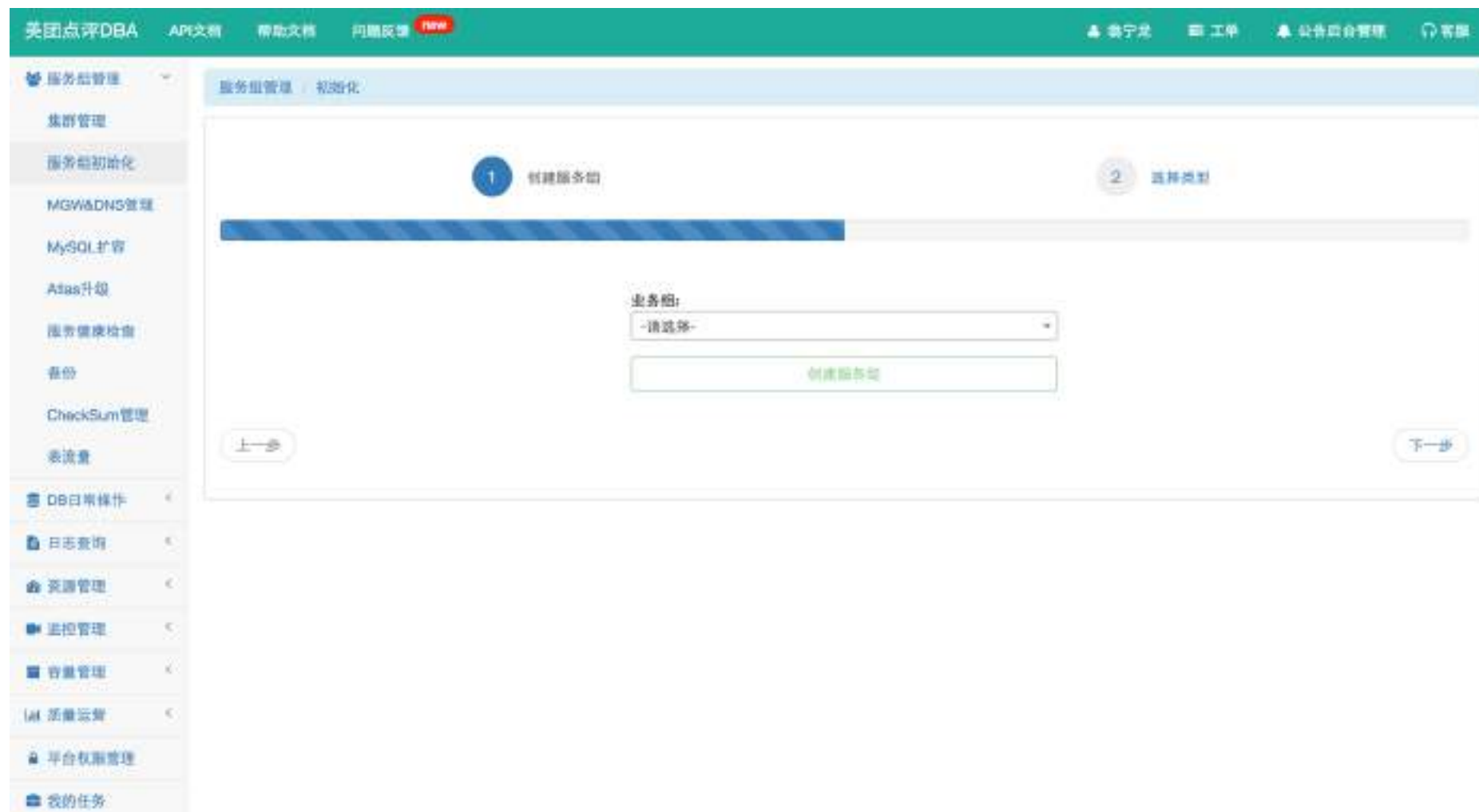
- 极简运维理念
- 场景化、定制化、智能化

➤ Q & A

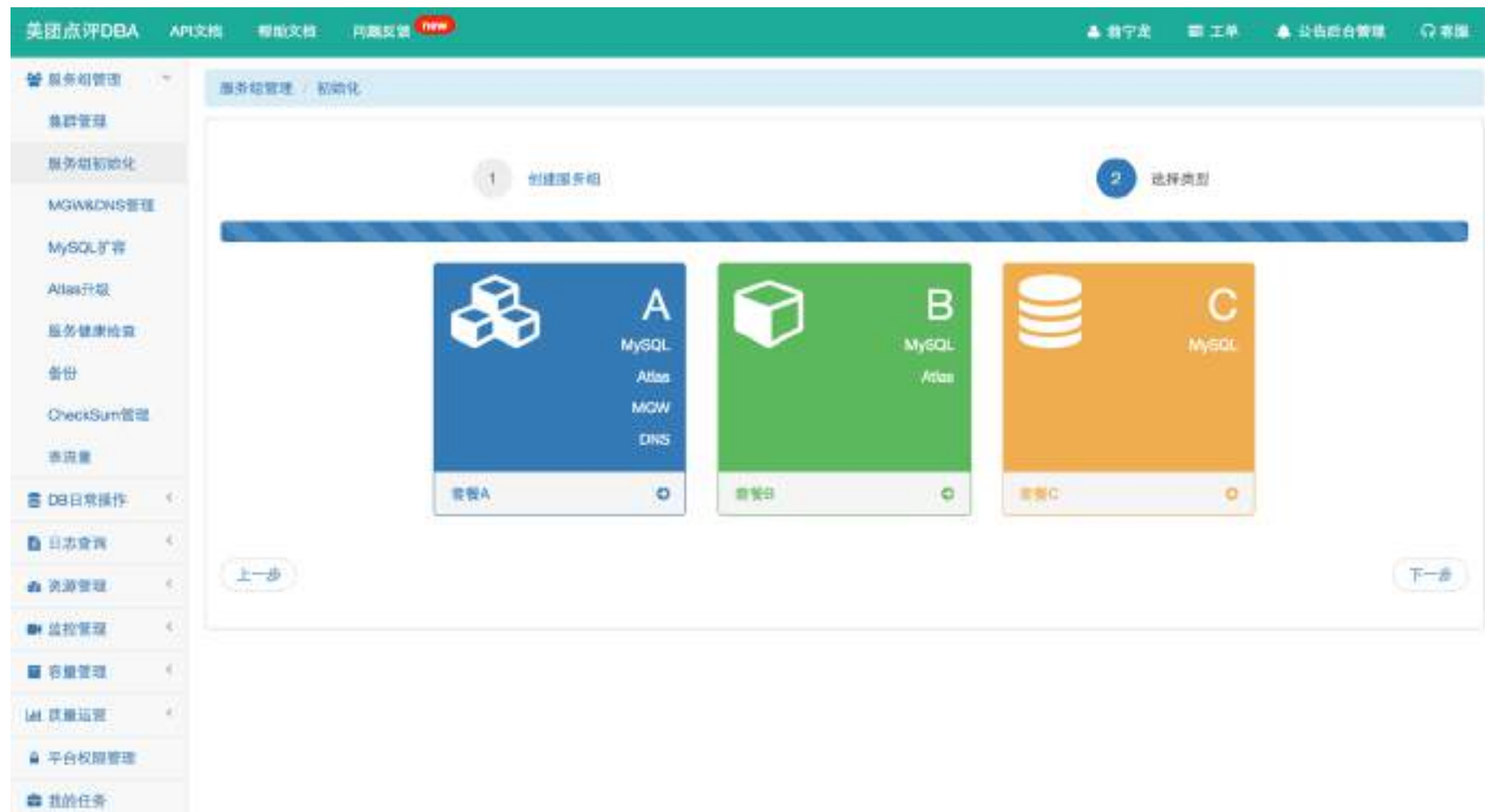
案例分享 - 集群搭建



案例分享 - 集群搭建



案例分享 - 集群搭建



案例分享 - 集群搭建

美团点评DBA API文档 帮助文档 问题反馈 new 翁宁龙 工单 公告后台管理 客服

服务组管理 / 初始化

1 创建服务组 2 选择类型 3 MySQL搭建 4 Atlas搭建 5 MGW&DNS 6 创建DB 7 确认

机器来源

我有机器 新申请机器(暂不支持)

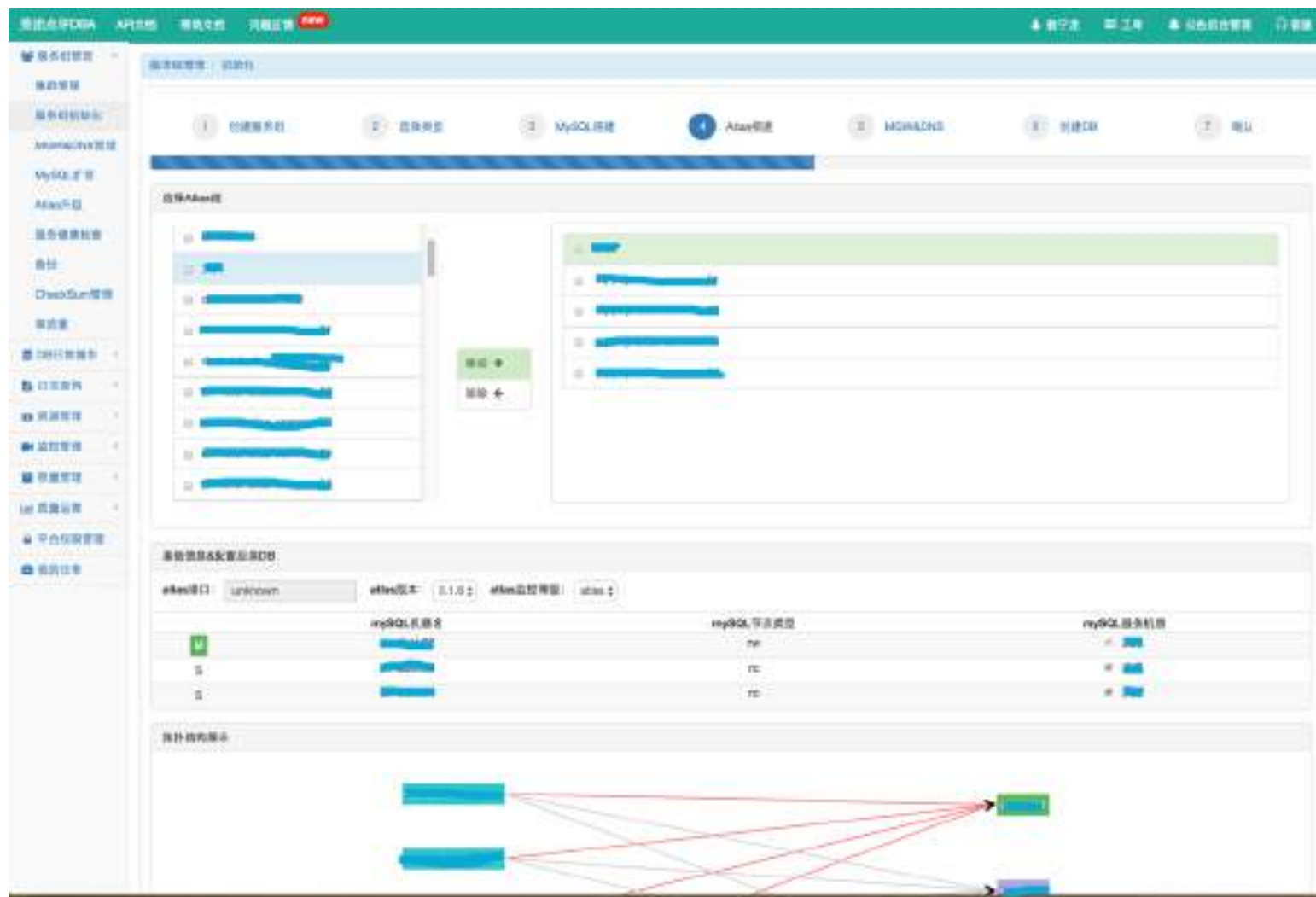
MySQL机器配置

#	机器	机房	BufferPool	单位	版本	监控等级	
master	[redacted]	[redacted]	80	GB	5.6	mysql_b	
slave	[redacted]	[redacted]	80	GB	5.6	mysql_b	
slave	[redacted]	[redacted]	80	GB	5.6	mysql_b	

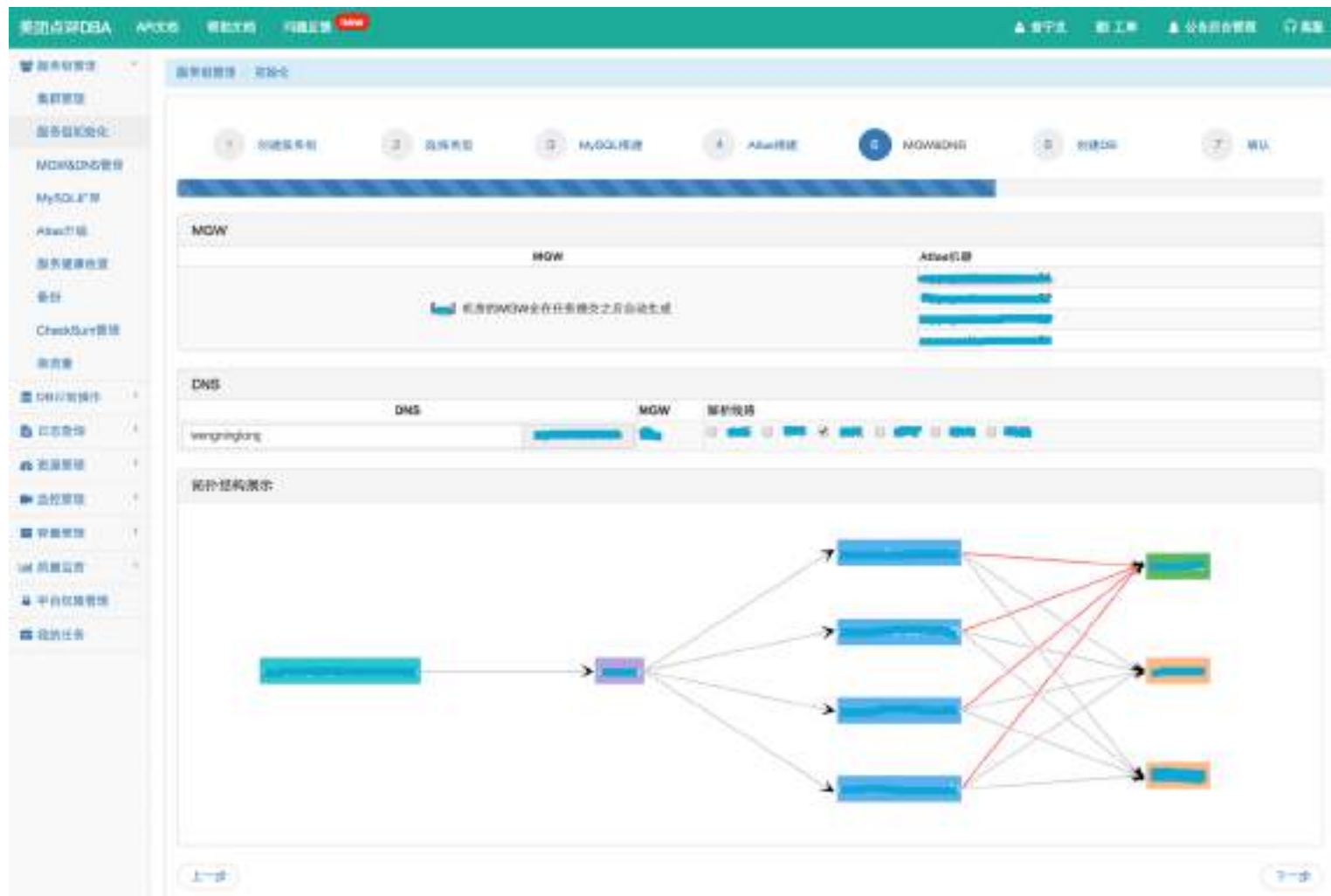
拓扑结构展示

上一步 下一步

案例分享 - 集群搭建



案例分享 - 集群搭建



案例分享 - 集群搭建

美团点评DBA API文档 帮助文档 问题反馈 new

翁宁龙 工单 公告后台管理 客服

服务组管理 / 初始化

1 创建服务组 2 选择类型 3 MySQL搭建 4 Atlas搭建 5 MGW&DNS 6 创建DB 7 确认

数据库名: 填写前, 请仔细核对数据库名

添加账号

账户名	密码	权限	
<input type="text"/>	<input type="text"/>	select,insert,update,delete	+
			✖

上一步 下一步

案例分享 - 集群搭建

服务部署管理 / 集群配置

1 创建数据库 2 选择类别 3 MySQL集群 4 AIO集群 5 MGW/DNS 6 创建IDC 7 确认

部署名称: [redacted]
主类名称: [redacted]
子类名称: [redacted]
实例: [redacted]
业务高峰期: 10:45-12:45;17:15-19:15
部署人: [redacted]
DNS: [redacted]
备注: [redacted]

MySQL:

#	名称	所在机房	DataPath	版本	监控等级	服务机房
mysql	[redacted]	[redacted]	80 GB	5.6	mysql_3	[redacted]
mysql	[redacted]	[redacted]	80 GB	5.6	mysql_3	[redacted]
mysql	[redacted]	[redacted]	80 GB	5.6	mysql_3	[redacted]

AIO:

角色	名称	端口	版本	监控等级
[redacted]	[redacted]	artoo/raft	D.1.8	artoo
[redacted]	[redacted]	artoo/raft	D.1.8	artoo
[redacted]	[redacted]	artoo/raft	D.1.8	artoo
[redacted]	[redacted]	artoo/raft	D.1.8	artoo

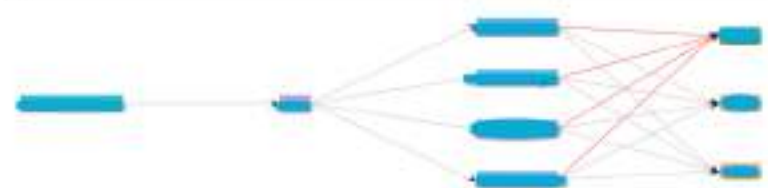
MGW & DNS

DNS	MGW	监控等级
[redacted]	配置为MGW服务之非活动节点	[redacted]

创建IDC:

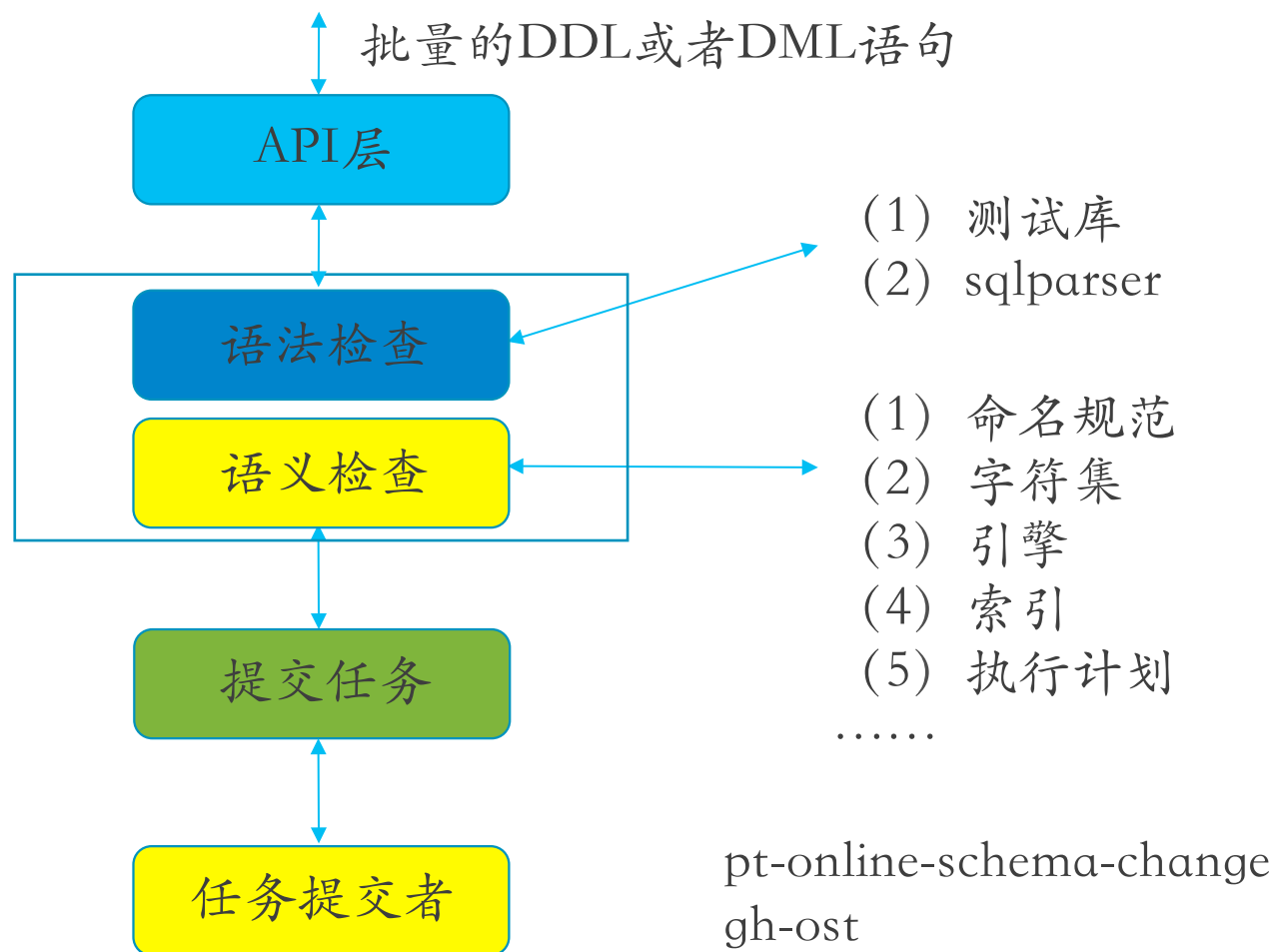
账号	密码	权限
[redacted]	[redacted]	select,insert,update,delete

设计模式展示

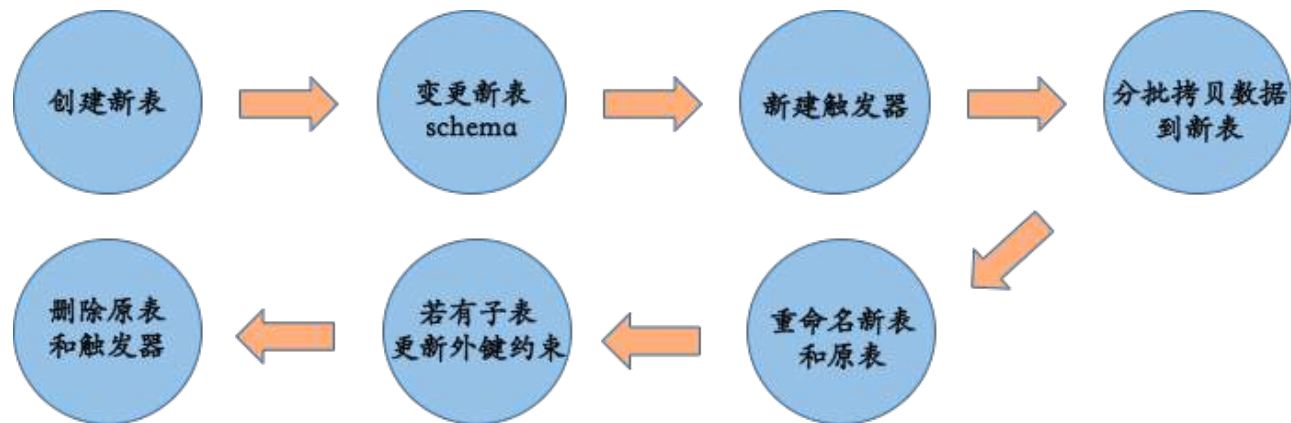


The diagram illustrates a network topology. On the left, a single blue node is connected to a central blue node. This central node is then connected to a group of four blue nodes. These four nodes are further connected to a final group of four blue nodes on the right, representing a distributed or multi-tier architecture.

案例分享 - 在线表变更



案例分享 - 在线表变更之pt-online-schema-change工具介绍



如上图所示，pt-osc的执行步骤：

步骤 1：创建一个新表，跟需要更改的表的结构保持一致

步骤 2：将alter语句作用于新建的表，那么新建表的结构跟最终需要的表结构保持一致

步骤 3：创建触发器来捕捉原始表的改表（insert -> replace; update -> replace; delete -> delete）；保证原表的更新能够作用到新表

步骤 4：分chunk复制数据，将原来中的数据分成若干个chunk并复制到新表

步骤 5：复制完成之后，将新表rename成原始表名；原始表rename成其他名称

步骤 6：更新外键，如果是子表 # 步骤 7：删除原始表和触发器

案例分享 - 在线表变更之gh-ost工具介绍



gh-ost migration:

- creates *ghost* table on migrated server
- alters *ghost* table
- hooks up as a MySQL replica, streams binary log events
- interchangeably:
 - applies events on *ghost* table
 - copies rows from original table onto *ghost* table
- cut-over

Preferred setup:

- connects to replica
- inspects table structure, table dimensions on replica
- hooks as replica onto replica
- apply all changes on master
- writes internal & heartbeat events onto master, expects them on replica

案例分享 - 在线表变更

The screenshot shows the 'Meituan DBA' online table change interface. The top navigation bar includes '美团点评DBA', 'API文档', '帮助文档', '问题反馈', and user information. The left sidebar lists various database management functions. The main content area is titled 'DB日常操作 / 自助表变更' and contains several sections: '基本信息' with dropdowns for '业务组' and '服务组'; '表变更操作' with dropdowns for '数据库' and '操作类型'; '可执行时间' with '业务高峰' and '当前任务执行时间区间要求' fields and associated notes; '批量改表' with radio buttons for '生成规则' and a '修改表SQL' text area; and a '数据库所包含的表' list on the right. A '生成批量修改SQL' button is located at the bottom right of the main content area.

案例分享 - 在线表变更

执行结果

第【1】条SQL存在问题: 【MySQL error】CREATE TABLE `mts_...` (`id` int(11) NOT NULL AUTO_INCREMENT COMMENT='主键id', `trace_id` char(32) COMMENT='产品中心-调用记录号' NOT NULL DEFAULT '' COMMENT='产品中心-调用记录号' ENGINE=InnoDB DEFAULT CHARSET=utf8 COMMENT='产品中心-调用记录号') ENGINE=InnoDB DEFAULT CHARSET=utf8 COMMENT='产品中心-调用记录号' 产品中心-调用记录号

错误信息



SQL长度超限

SQL长度限制不能超过1M;该限制适用于:alter | create | delete | update | insert | select



SQL语法错误

SQL存在语法错误,请确保提交的SQL语句是在线下测试环境测试通过的



表名和字段名规范错误

DBA要求表名需要符合规范:表名只能由字母、数字、下划线组成,且新建表的表名包含的字母只能为小写字母;字段名只能含有字母、数字和下划线,并不能以数字和下划线开头;该限制适用于:alter | create



表名长度超限

为了兼容后台相关的工具,要求表名的长度不能超过50个字符;该限制适用于:alter | create



不能修改表字段名称

案例分享 - 在线表变更

任务ID: 3308

任务操作者: [REDACTED]

任务观察者: [REDACTED]

任务所属业务组: [REDACTED]

任务所属服务组: [REDACTED]

任务类型: autosqlAlter

任务状态: succeeded

任务开始执行时间: 2016-06-16 17:07:05

任务执行结束时间: 2016-06-16 17:09:16

任务计划执行时间: 2016-06-16 17:07:05

任务提交时间: 2016-06-16 17:07:05

任务相关数据:

数据库名字: [REDACTED]

执行时间: 00:00-23:59

备注: [REDACTED]

创建的SQL:

表名	创建SQL
[REDACTED]	1 [REDACTED]

任务详情:

[17:07:05][1/4]任务开始执行,相关操作信息获取

[17:07:05][1/4]获取相关操作信息成功

[17:07:05][2/4]开始将改表任务加入执行队列

[17:07:05][2/4]改表任务加入执行队列成功[总共:1条执行语句]

[17:07:05][3/4]开始获取执行队列中语句的执行状态

[17:09:06][3/4]第[1]条sql[REDACTED]执行成功

目 录

➤ 美团点评数据库生态系统

- 为什么要打造数据库生态系统?
- 美团点评数据库生态是什么样的?

➤ 案例分享

- 集群初始化
- 在线表变更
- 高可用方案

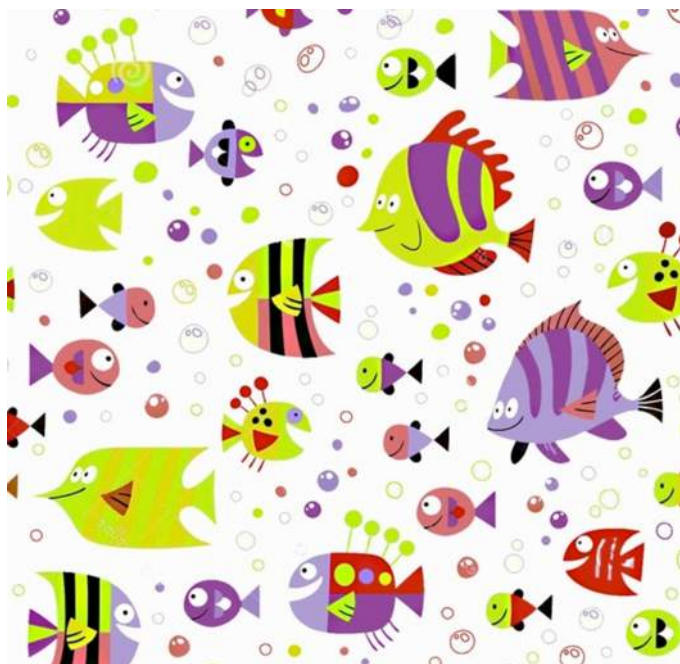
➤ 极简运维

- 极简运维理念
- 场景化、定制化、智能化

➤ Q & A

极简运维 - 极简运维理念

- 以最简单的方式达到目标，满足业务需求
- 是一个由简到繁，再由繁到简的过程



极简运维 - 场景化、定制化、智能化





GIAC | BEIJING
Dec.12.16-17

技术架构未来

架构
ARCHNOTES
设计 应用 实践



新美大技术团队