



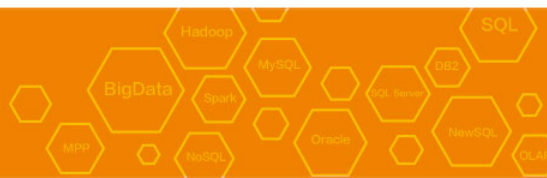
第九届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2018

美团点评数据库智能运维探索与实践

赵应钢

DTCC
2018

2018.05.10 – 12 北京国际会议中心



IT168.com

ChinaUnix

ITPUB

大纲

1

数据库平台的演变

2

现状和面临的挑战

3

从自动化到智能化

大纲

1

数据库平台的演变

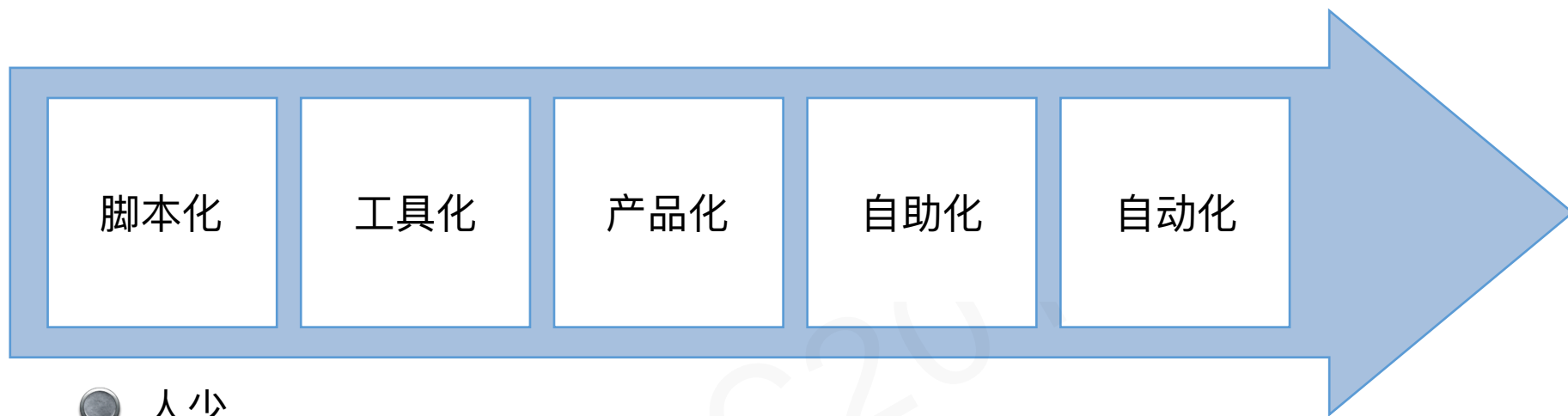
2

现状和面临的挑战

3

从自动化到智能化

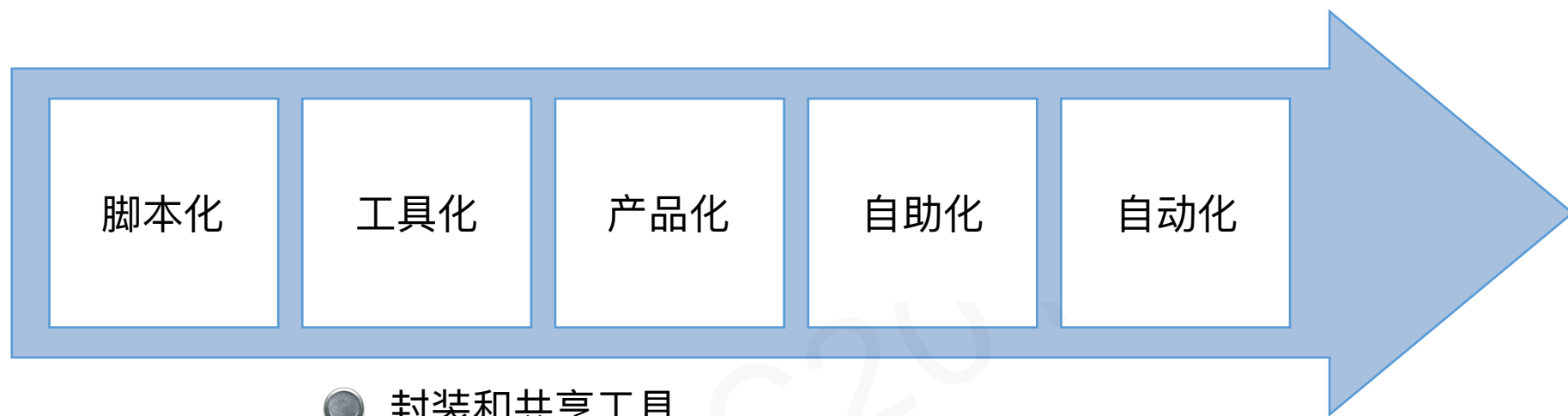
演变历程-纯手动模式



- 人少
- 集群数少
- 流量小



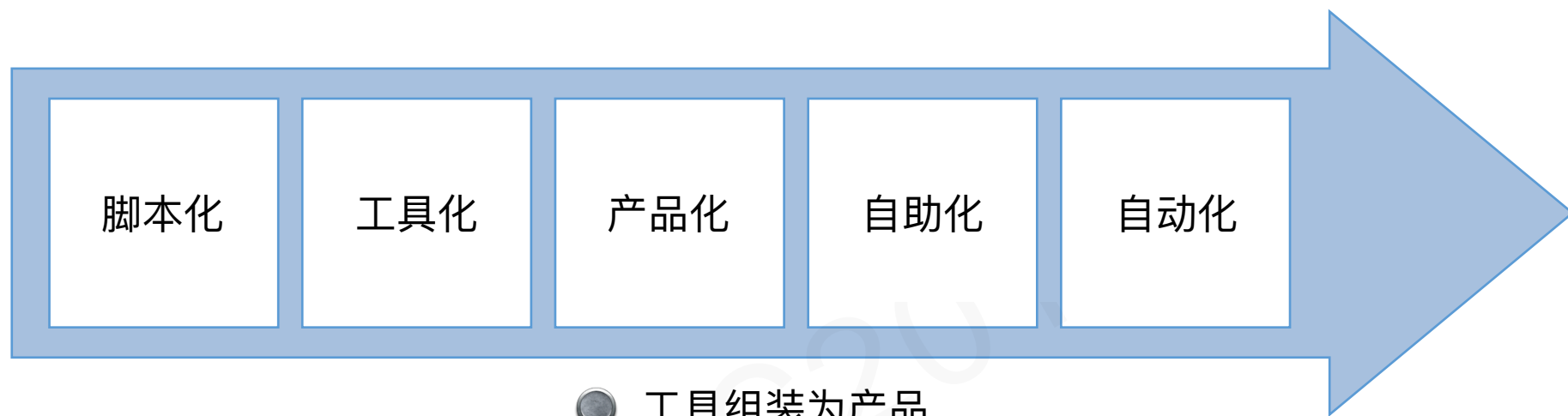
演变历程-利用工具



- 封装和共享工具
- 围绕CMDB管理资产与服务
- 完善监控
- 工具箱逐渐丰富：
 - DDL变更工具
 - SQL Review工具
 - 慢查询采集分析工具
 - 备份闪回工具

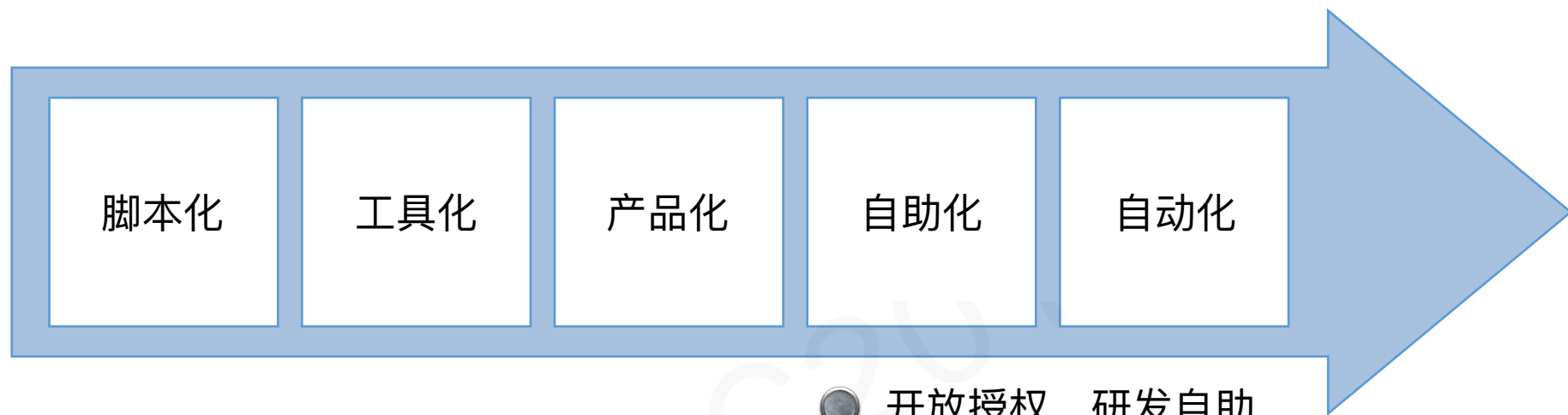


演变历程-工具产品化



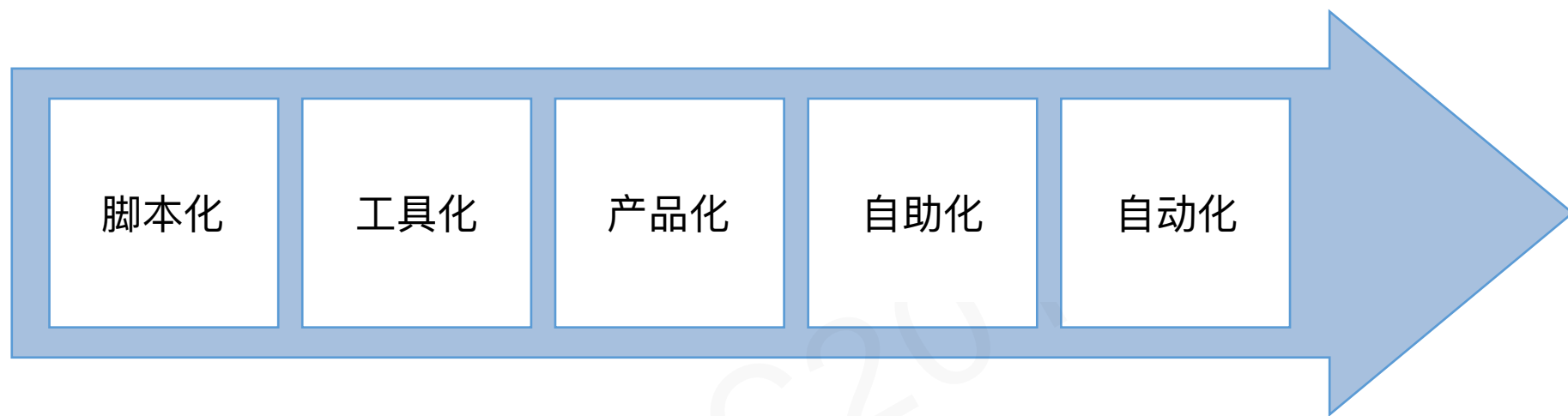
- 工具组装为产品
- 易用性、安全性上不断打磨
- 产品用户为DBA
- 定位为提升运维效率

演变历程-私有云平台



- 开放授权，研发自助
- DBA从操作中解放：
 - 研发每天自助改表300+次
 - 自助查询10000+次
 - 自助申请账号、授权、调整监控
 - 自助定义业务高低峰等
 - 自助下载、查询日志

演变历程-半自动到全自动



- 主从自动切换
- 过载自我保护
- 容量自动评估
- 集群自动扩缩容



大纲

1

数据库平台的演变

2

现状和面临的挑战

3

从自动化到智能化

平台现状

RDS平台

主从HA	CMDB	初始化	MGW管理	DNS管理
扩容	升级	备份	拆分	表变更
DB创建	账号管理	任务管理	数据归档	机器管理

日志服务

慢日志

BINLOG

Processlist

审计日志

监控服务

链路监控

Falcon

深度监控

实时监控

优化诊断

SQL评审

SqlAdvisor

容量管理

报警自动处理

运营服务

DOM

平台质量

服务化治理

日报巡检

按用户区分

- RDS自助平台
- DBA管理平台
- 测试环境管理平台

按功能聚合

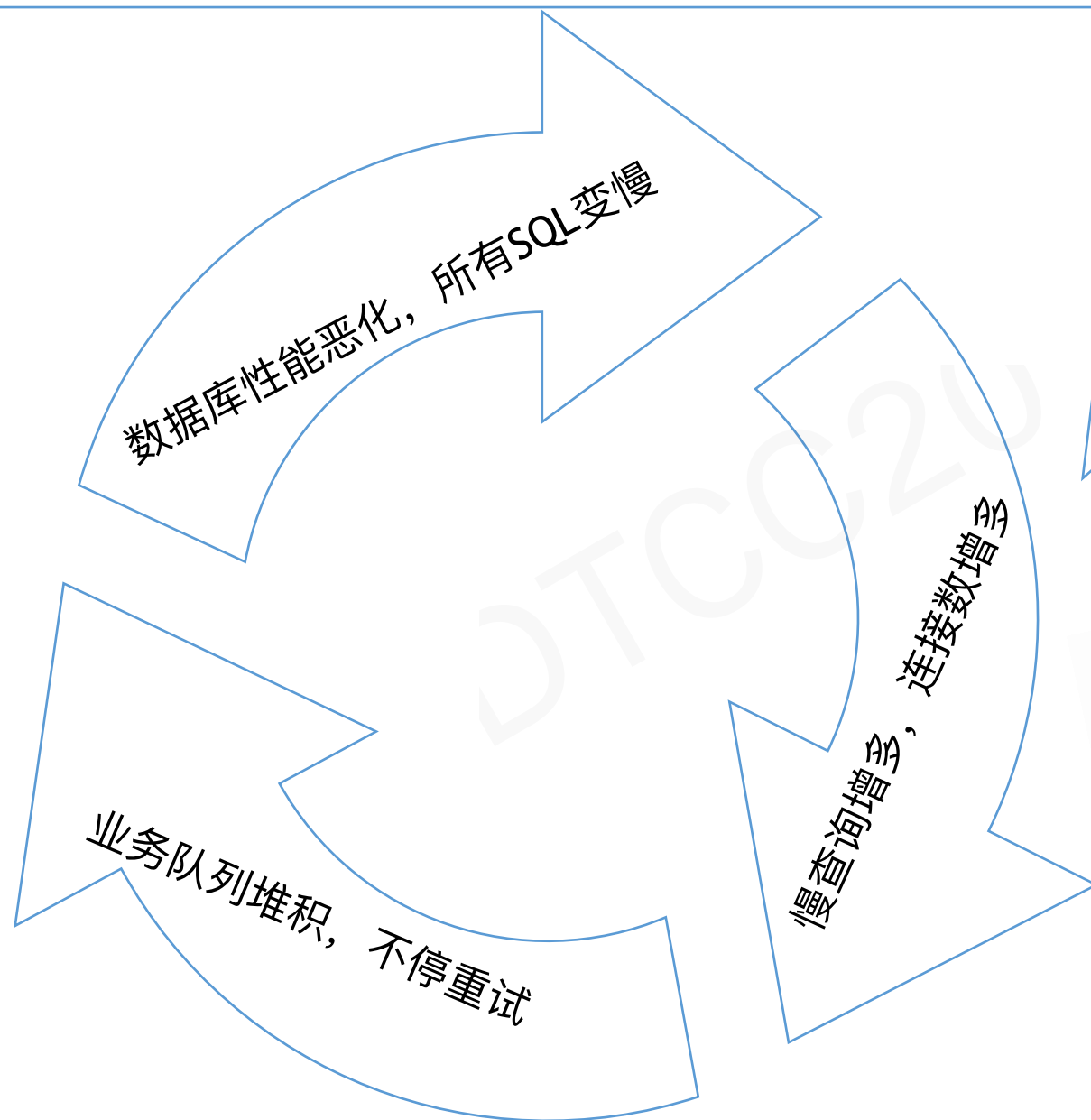
- 监控
- 运维
- 运营

按服务拆分

- MySQL关系型存储
- 分布式KV缓存
- 分布式KV存储
- 分布式关系型数据库

打造MySQL+NoSQL+NewSQL，存储+缓存的一站式服务平台

挑战1：RootCause定位难



1. 数据库是结果还是原因?
2. 如何杜绝再次发生?



挑战2：人力和发展困境

成长的烦恼

- 流量成倍增长
- 集群不断拆分
- 开发诉求多样化
- 实时性要求不断提高

人力的困境

- 时间被碎片化
- 成长遇到瓶颈
- 业务咨询暴增
- 专业DBA匮乏

突破困境，转型智能化

传统运维

故障触发

被动接受

DBA发起和解决

人肉救火

DBA亲临事故现场

智能运维

隐患驱动

主动出击

系统发起，RD自助

智能决策执行

DBA隐身幕后

大纲

1

数据库平台的演变

2

现状和面临的挑战

3

从自动化到智能化

背景目标

高效

稳定

快速

从重复、
繁琐的工作中解
放，转型AI

更稳定、
高可用，专注业
务开发

提升技术
保障的效率，降
低成本

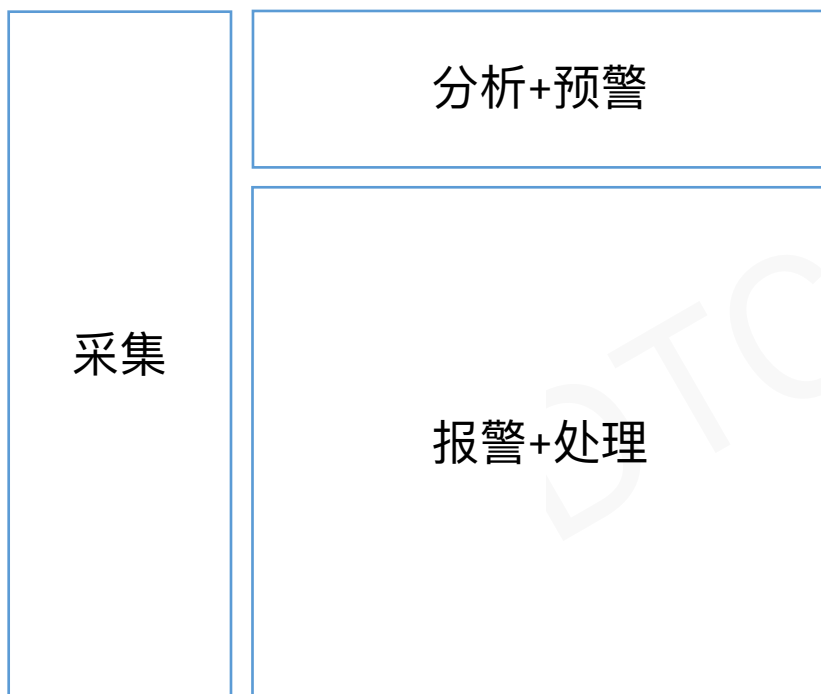
DBA

业务

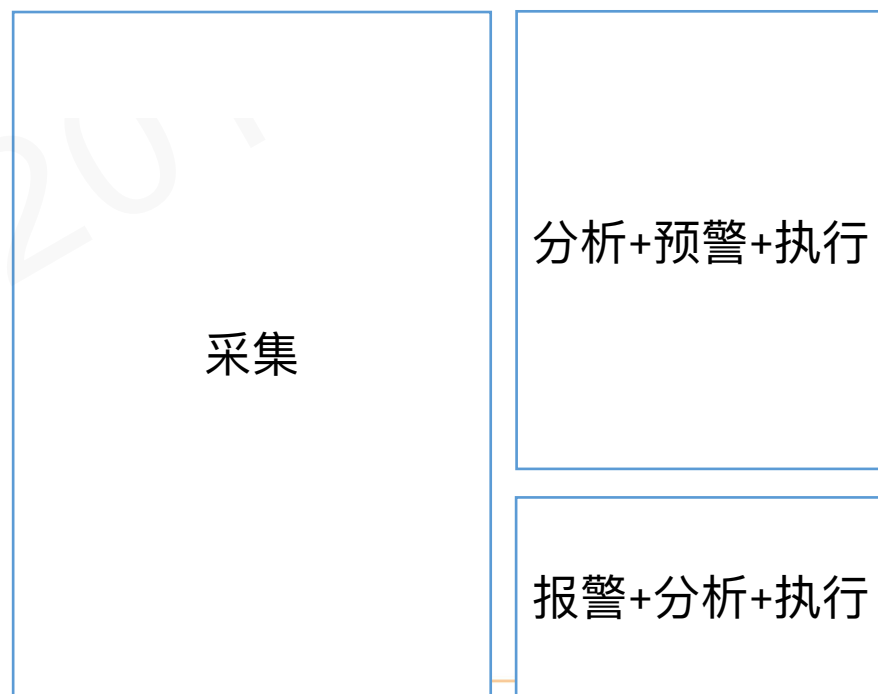
公司

特点分析

传统运维

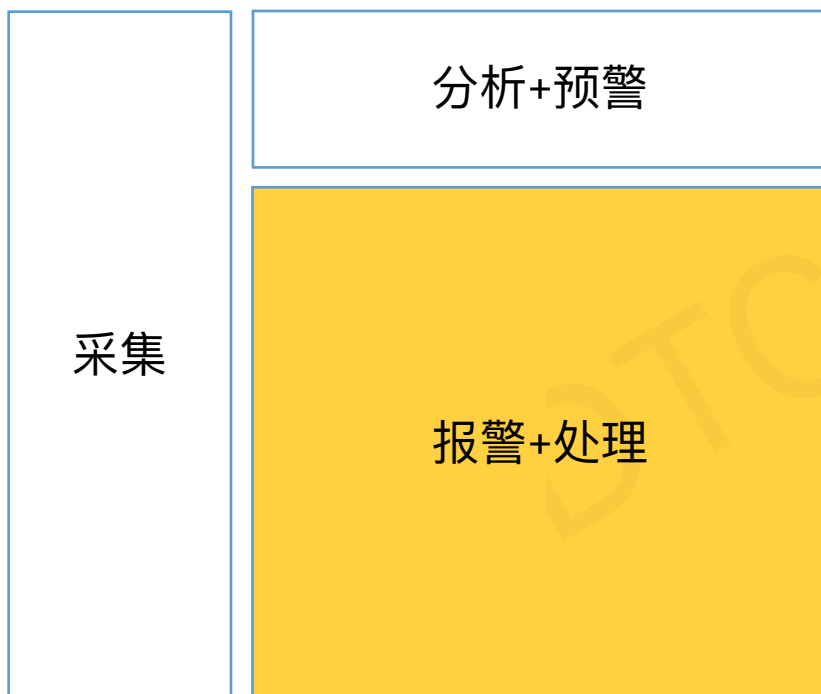


智能运维

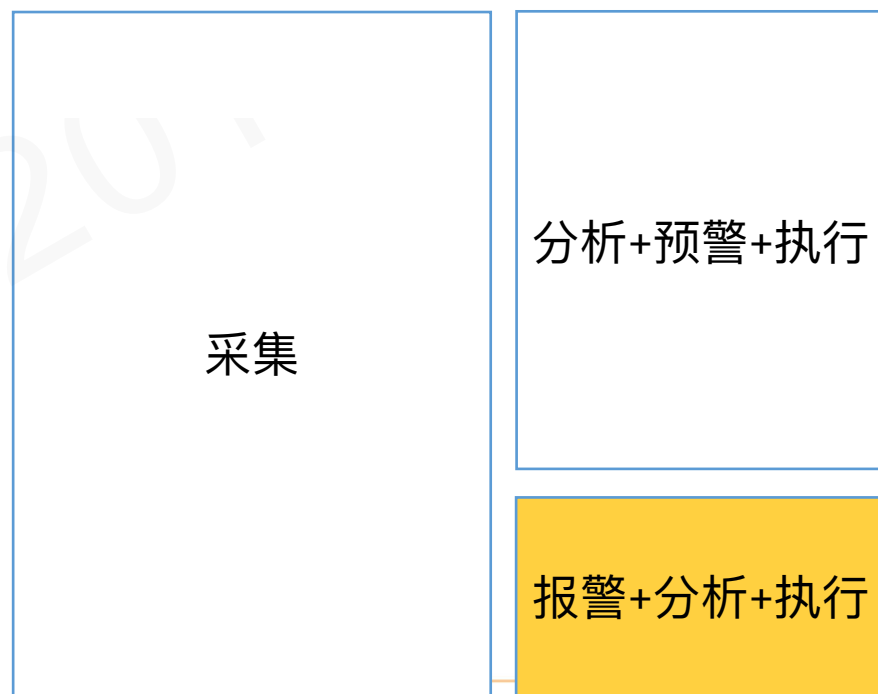


特点分析

传统运维



智能运维



决策执行

预警重要不紧急；报警紧急且重要

预警问题域明确，决策执行思路清晰

报警依赖事故现场链路分析，变量多，决策难，风险大



全面采集；增多预警，减少报警

先实现预警部分智能化；再实现报警部分智能化

数据采集

- 端到端成功率
- 响应时间95线、99线
- 错误日志
- 吞吐量

应用

系统

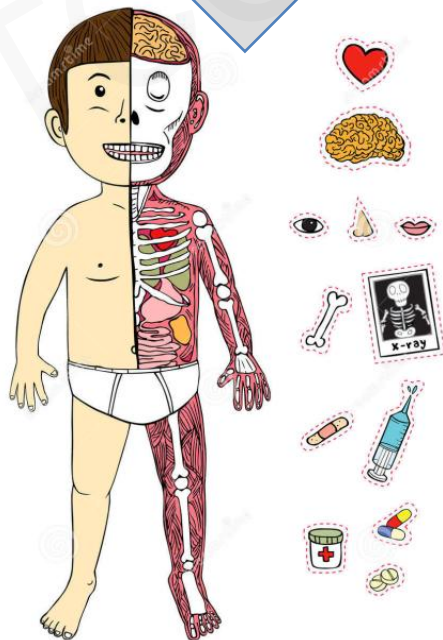
- 支持秒级采样
- 操作系统各项指标

- global status、variables
- Processlist、innodb status
- slow、error、general log
- binlog

数据库

变更

- 集群拓扑调整
- 在线DDL、DML变更
- DB平台操作日志
- 应用端发布记录



数据分析

集群

- 每年新增多少个集群?
- 哪些集群写入峰值大?
- 哪些集群?
 - 内部版本不一致
 - binlog格式不一致
 - 主从复制规则不一致
 - 不支持GTID

实例

- 5.5 5.6 5.7占比如何?
- 同一集群QPS是否倾斜?
- 网卡使用率是否有倾斜?
- 读写比如何?
- 容量是否足够?
- 平均响应时间多少?

库

- 哪些库表已经被废弃?
- 哪些表被最频繁访问?
- 哪些DB被跨业务线访问?

表

- 多少个冷、热、大的表?
- 多少个无主键表?
- 多少个myisam表?
- 哪些废弃表、索引?
- 哪些自增ID即将溢出?

数据库画像-整体-资源规划治理

1主20+从的集群增多

服务器需要scale up

读增加迅猛，读写比变大

考虑存储KV化

利用率和分布情况

影响到采购和预算制定

哪几类报警最多

专项治理，各个击破

数据库画像-部分-集群健康体检

支付核心集群，不能降级

历史上最大写入QPS 2000，从库开启并行复制，无延迟风险，无需拆分

已开启GTID，支持自动切换

10ms慢查询占比小于万分之三



所有表都有主键；全部为innodb引擎

磁盘IO稳定，数据全部在内存中

topN的查询都基于主键，性能良好

最近一个月新增3类查询语句，有一个是慢查询

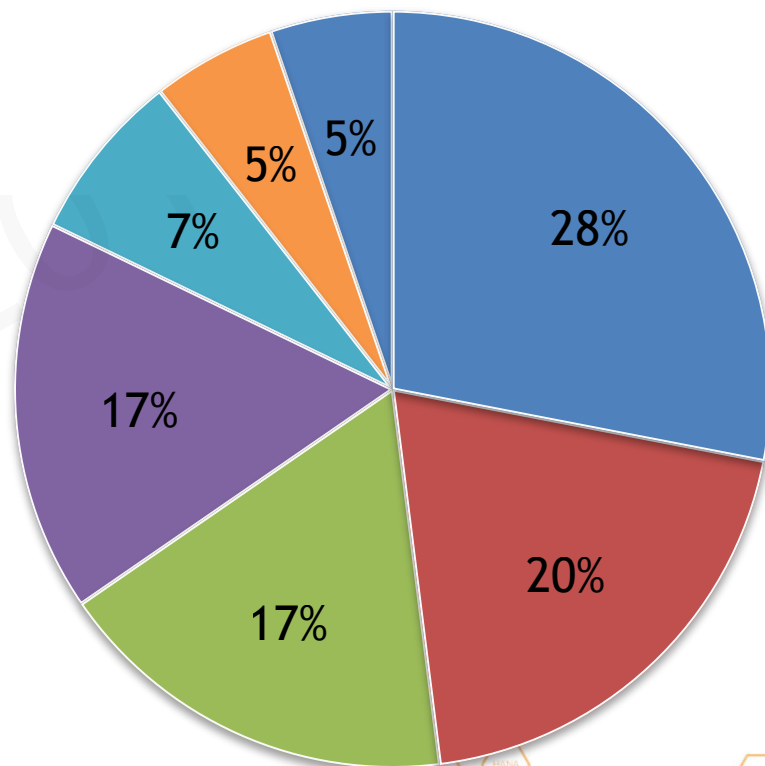
有大事务，可能导致并发问题

数据库预警

- 通过分析发现隐患
- 把报警转化为预警

数据库报警TopN

- Seconds_Behind_Master
- load.1minPerCPU
- df.bytes.free
- Slave_IO_Running
- Slave_SQL_Running
- net.if.out
- Threads_running



容量预警

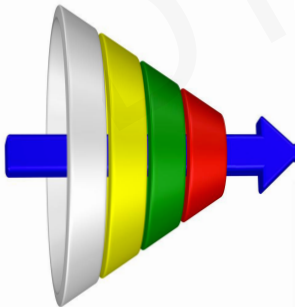
容量状态 状态判断的优先级：有“人工设置”，以“人工设置”为判断标准；如果没有，则以“falcon历史最大值”&“系统评估”两者取大者作为判断标准；否则以“默认配置”为准。

业务组【 】						
服务组	主库信息	指标项	当前值	阈值	类型	状态
		diskIops	1601.3	16017.0	falcon获取的历史最大值	满足
	主库类型 : 物理机	load	0.2	21.89	falcon获取的历史最大值	满足
	主库磁盘类型 : SSD	network	5099.24	8000.0	默认配置	不满足
	容量倍数 : 2	diskFree	<div></div>	10000	falcon获取的历史最大值	满足
	查看详情	sqlSelectQps	15812.6	20000.0	默认配置	不满足
		sqlWriteQps	3273.4	9107.47	falcon获取的历史最大值	满足

MT-Falcon

CAT

Others



- 平均响应时间与：
 - QPS的关系
 - 慢查询比例的关系
 - 高峰、低峰时间段关系
 - 客户端可用性关系

空间预警

机器名	业务组	服务组	磁盘大小	磁盘空间 剩余比例	预测剩余 日期	binlog 文件个数	binlog 文件大小
[REDACTED]	商家销售	[REDACTED]	493G	25.61	4	99	98G
[REDACTED]	商家销售	[REDACTED]	493G	26.54	4	100	99G
[REDACTED]	数据应用	[REDACTED]	296G	13.51	5	41	40G
[REDACTED]	云计算	[REDACTED]	296G	13.37	5	52	51G
[REDACTED]	外卖研发	[REDACTED]	296G	10.44	6	21	20G
[REDACTED]	外卖研发	[REDACTED]	493G	9.43	6	43	42G
[REDACTED]	外卖研发	[REDACTED]	296G	11.73	7	22	20G
[REDACTED]	外卖研发	[REDACTED]	296G	11.34	7	21	20G
[REDACTED]	外卖研发	[REDACTED]	493G	9.11	7	36	35G
[REDACTED]	酒店后台研发	[REDACTED]	5.9T	25.56	7	301	303G



慢查询预警

红 榜

出轨率

BG	出轨应用数	出轨应用占比
[REDACTED]	7	0.000%
[REDACTED]	1	0.000%
[REDACTED]	0	0.000%

3S慢查询

BG	慢查询数量
[REDACTED]	118
[REDACTED]	56
[REDACTED]	1854

低利用率机器

BG	低利用率机器数
[REDACTED]	0
[REDACTED]	1
[REDACTED]	2

黑 榜

出轨率

BG	出轨应用数	出轨应用占比
[REDACTED]	13	7.927%
[REDACTED]	7	3.955%
[REDACTED]	37	2.735%

3S慢查询

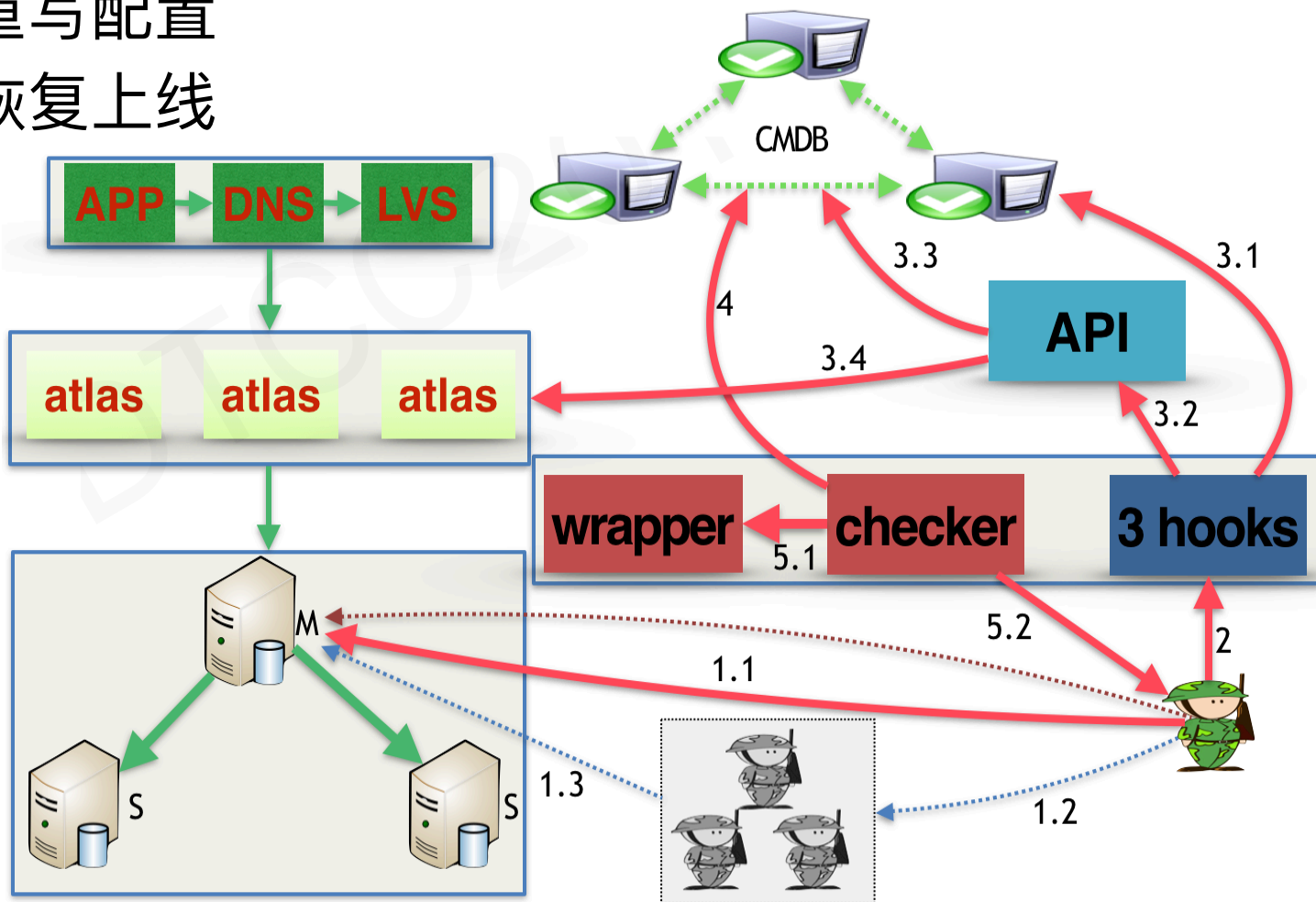
BG	慢查询数量	
[REDACTED]	31663	↑
[REDACTED]	3176	↑
[REDACTED]	20277	↑

低利用率机器

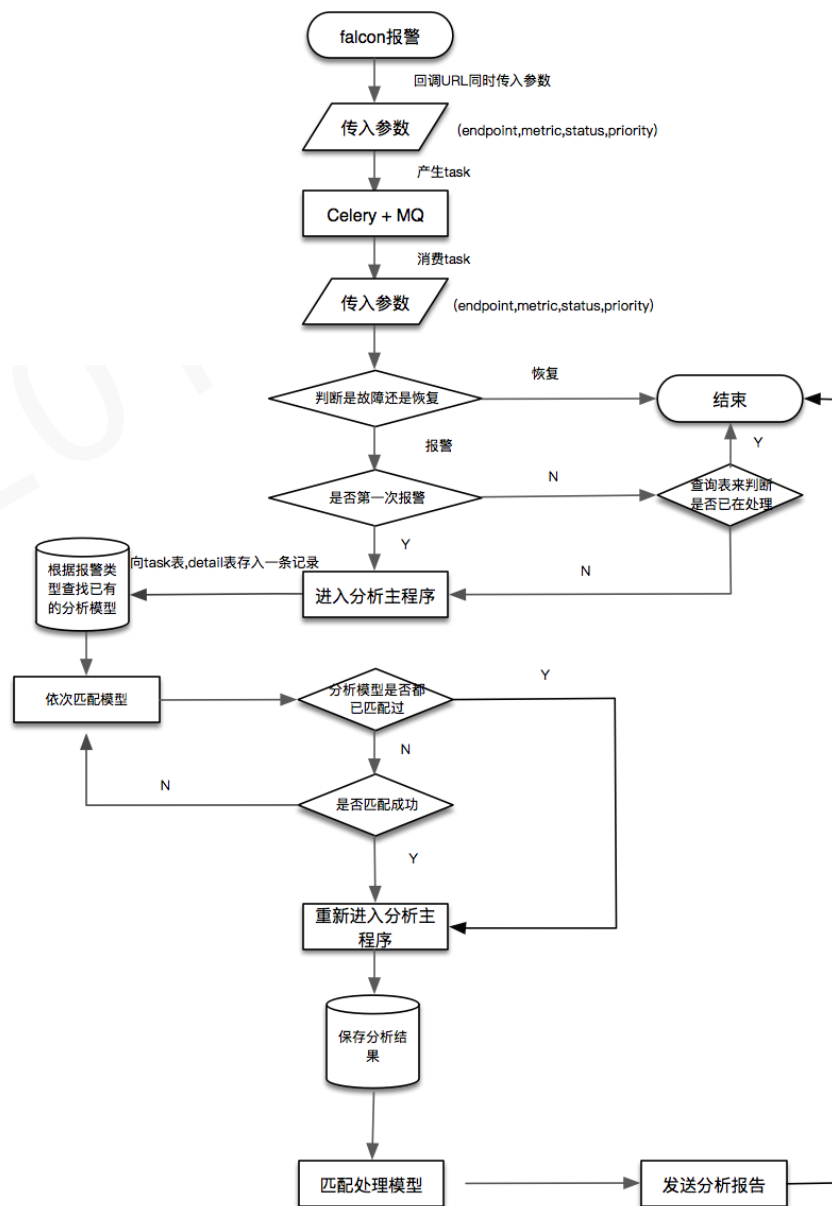
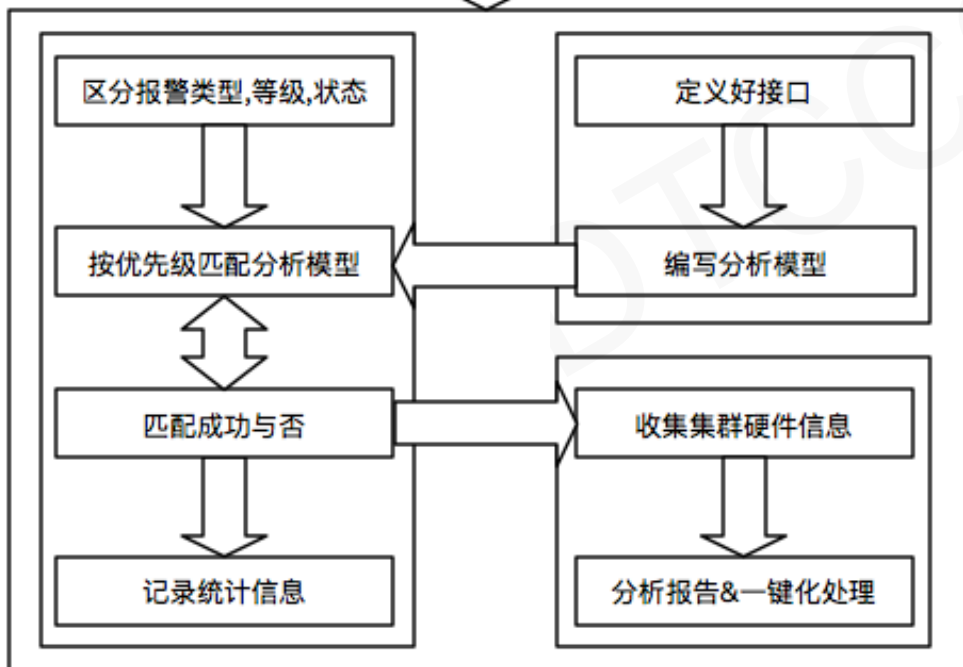
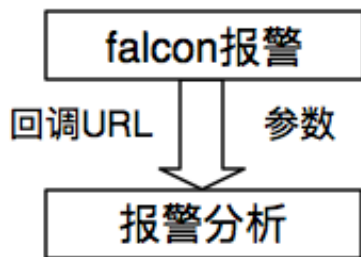
BG	低利用率机器数
[REDACTED]	48
[REDACTED]	20
[REDACTED]	13

故障快速转移

- 自动生成配置文件
- 自动判断是否启用监控
- 切换后自动重写配置
- 从库可自动恢复上线



报警自动处理



报警自动处理

[报警分析结果报告][p2][redacted][Slow_queries][2017-03-17 15:18:30]

[报警处理方案]

- [1: kill file sort的连接(5s,1min)]
- [2: kill tmp table的连接(5s,1min)]
- [3: kill正在运行的select连接(5s,1min)]

[报警分析结果报告][p0][redacted][Threads_running][2017-03-12 07:35:00]

[报警处理方案]

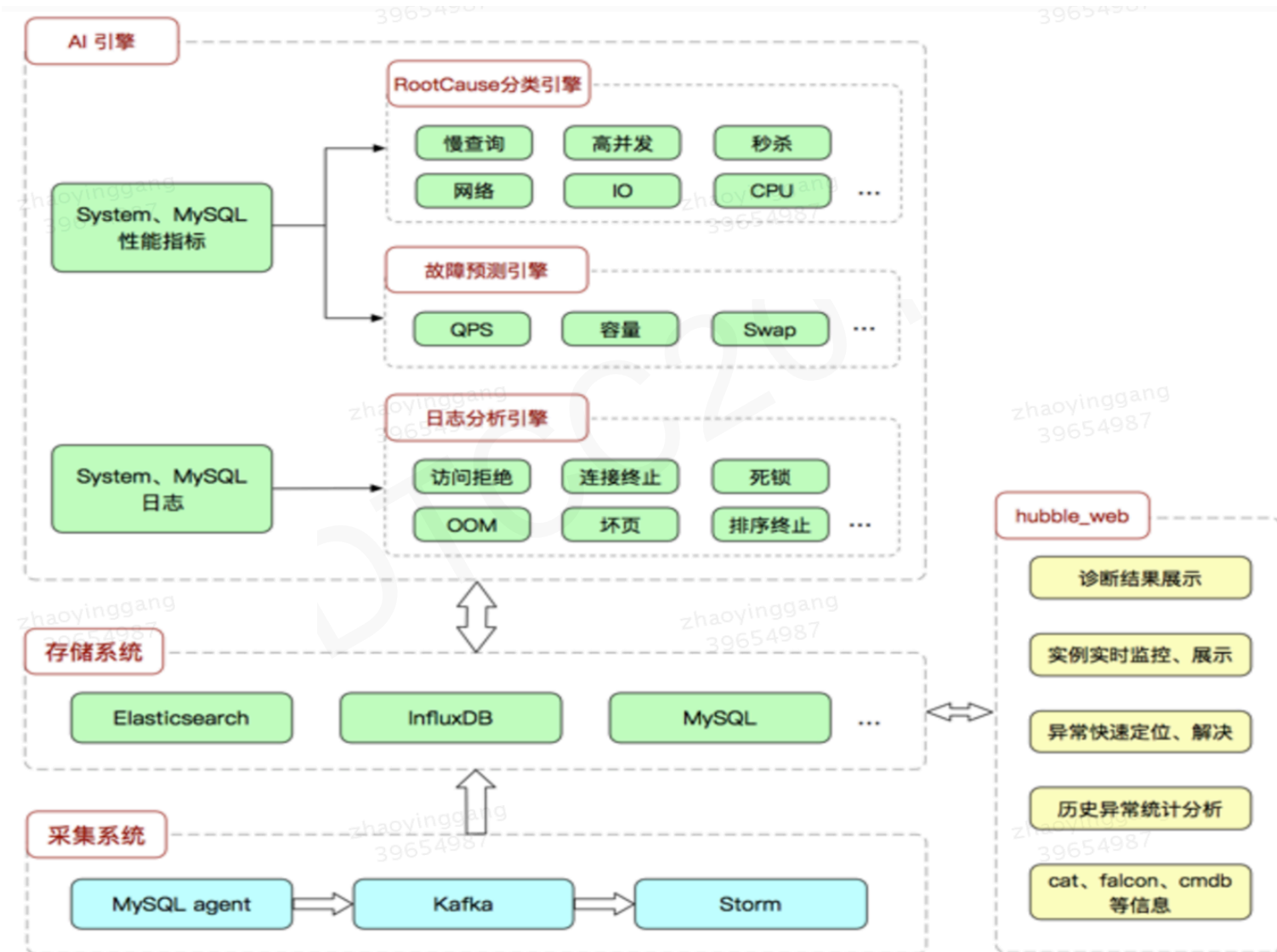
- [1: kill正在运行的select连接(5s,1min)]
- [2: kill file sort的连接(5s,1min)]
- [3: kill sleep的连接30s,1min]
- [4: kill tmp table的连接(5s,1min)]
- [5: kill Creating sort的连接(5s,1min)]
- [6: kill Copying to group table的连接(5s,1min)]
- [7: kill converting HEAP to MyISAM 的连接(5s,1min)]

[报警分析结果报告][p0][redacted][load.1minPerCPU][2017-03-14 16:03:00]

[报警处理方案]

- [1: kill file sort的连接(5s,1min)]
- [2: kill tmp table的连接(5s,1min)]
- [3: kill正在运行的select连接(5s,1min)]
- [4: kill Creating sort的连接(5s,1min)]
- [5: kill Copying to group table的连接(5s,1min)]
- [6: kill converting HEAP to MyISAM 的连接(5s,1min)]
- [7: kill 处于Waiting for 的连接(5s,1min)]
- [8: kill sleep的连接(30s,1min)]

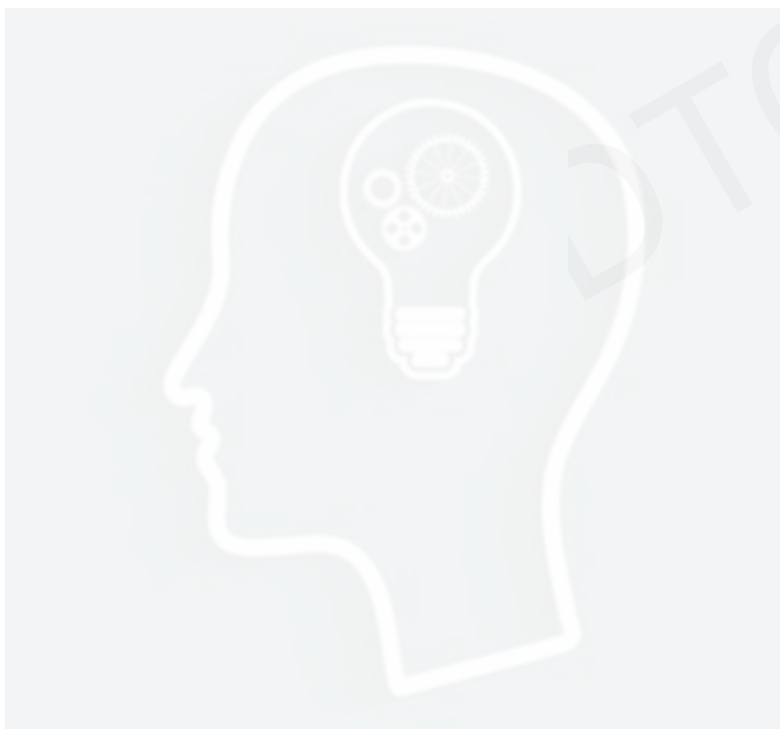
故障诊断平台



数据库智能运维展望

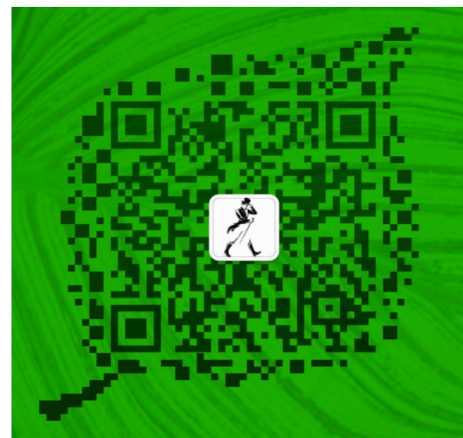
●三个方向交叠演进

●ABC深入融合



招聘多名DBA专家、运维开发专家

- 为众多开源项目贡献patch: **gh-ost**、**MHA**
- 开源**DBproxy**、**SQLAdvisor**、**Zebra**、**MyFlash**等数据库周边产品
- 自研分布式**KV**存储系统**Cellar**, 已日趋成熟
- 自研分布式关系型数据库: **Blade**, 正招贤纳才
- 负责外卖、配送、金融、酒旅、猫眼、打车等业务线
- 邮箱: **zhaoyinggang@meituan.com**



THANKS





讲师申请

联系电话（微信号）：18612470168

关注“ITPUB”更多
技术干货等你来拿~

与百度外卖、京东、魅族等先后合作系列分享活动



让学习更简单

微学堂是以ChinaUnix、ITPUB所组建的微信群为载体，定期邀请嘉宾对热点话题、技术难题、新产品发布等进行移动端的在线直播活动。

截至目前，累计举办活动期数60+，参与人次40000+。

ITPUB学院

ITPUB学院是盛拓传媒IT168企业事业部（ITPUB）旗下
企业级在线学习咨询平台
历经18年技术社区平台发展
汇聚5000万技术用户
紧随企业一线IT技术需求
打造全方式技术培训与技术咨询服务
提供包括企业应用方案培训咨询（包括企业内训）
个人实战技能培训（包括认证培训）
在内的全方位IT技术培训咨询服务

ITPUB学院讲师均来自于企业
一些工程师、架构师、技术经理和CTO
大会演讲专家1800+
社区版主和博客专家500+

培训特色

无限次免费播放
随时随地在线观看
碎片化时间集中学习
聚焦知识点详细解读
讲师在线答疑
强大的技术人脉圈

八大课程体系

基础架构设计与建设
大数据平台
应用架构设计与开发
系统运维与数据库
传统企业数字化转型
人工智能
区块链
移动开发与SEO



联系我们

联系人：黄老师
电话：010-59127187
邮箱：edu@itpub.net
网址：edu.itpub.net
培训微信号：18500940168