

苏宁的数据库自动化运维实践

籍天扬

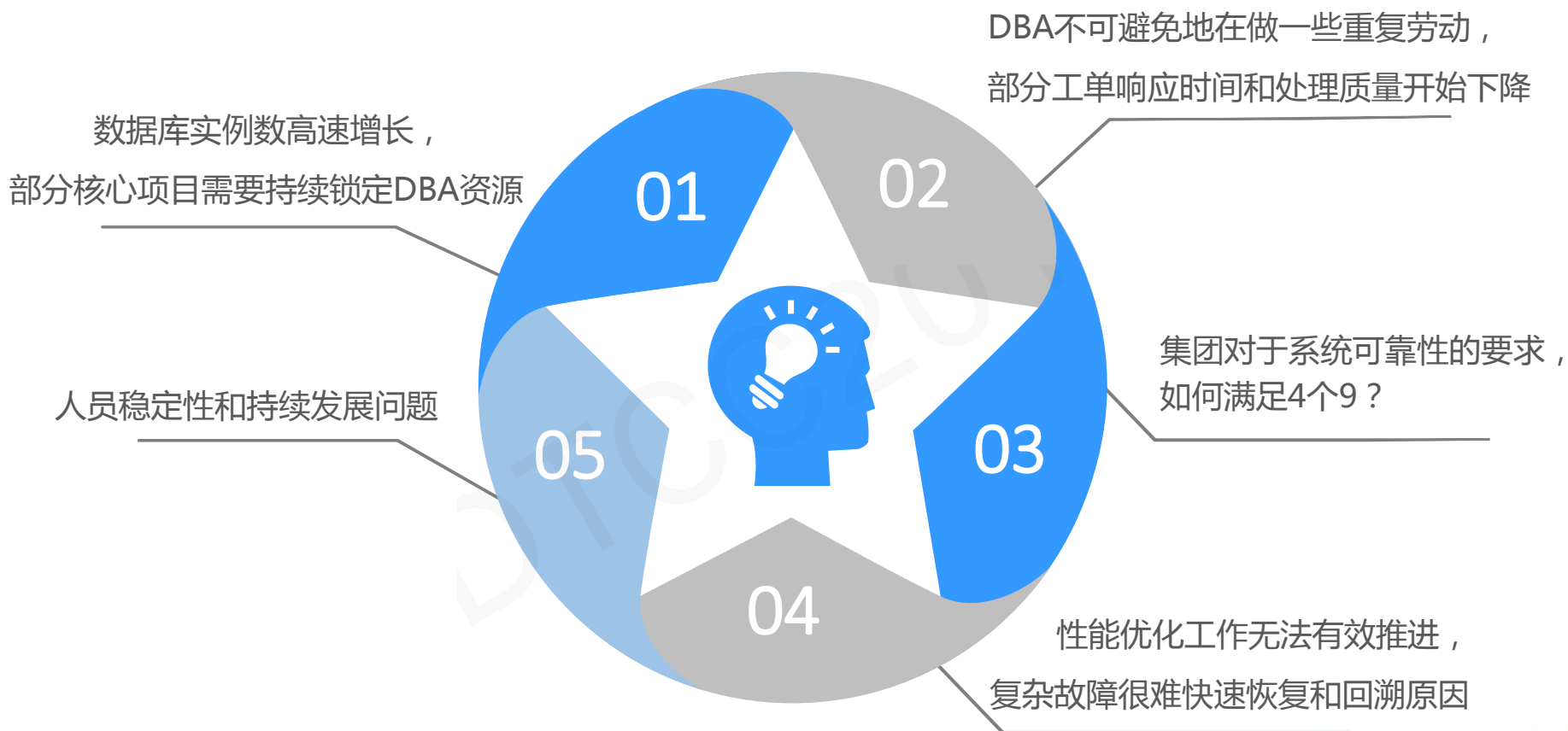
1

传统数据库运维在苏宁面临的问题

2

DBA团队和研发中心的痛点

传统数据库运维在苏宁面临的问题



痛

完全与生产环境隔离，不了解数据库健康状况，对数据库的优化无从下手

痛

传统监控手段只能看到告警信息，无法获得有效的指标预警和专业建议 ...

痛

各种诸如数据导出、SQL发布、元数据/配置查询等数据库操作需求需要DBA人肉配合，沟通和等待成本高

痛

运维事件对研发中心不透明，如备份时段变化、参数历史版本、迁移记录、拓扑变化、近期错误信息等等 ...

痛

测试环境运行良好的SQL在生产环境中可能会出现性能问题



前期设计思路



解决了哪些问题



智能化探索

苏宁 DBMS

AIOps



离线计算

实时计算

事件预警

智能决策



事件感知

基于多场景分析，智能聚合报警及趋势相关信息，实时预警事件。



大数据决策

实时事件分析辅以历史事件分析，即时定位问题，预测未来事件，生成止损方案。



自动化执行

不断完善的事件处理引擎，提供全场景变更和故障的自动化执行能力。



离线计算 — 历史

- 历史数据分析，多维度数据挖掘
- 故障分析，生成时间轴，知识库生成
- 报表展示，问题回溯，热点分析
- 数据库健康风险模型训练
- 机器学习，规则自我优化



实时计算 — 当前

- 秒级监控，告警聚合，事件生成
- 异常判断，邻近比对，决策发起



事件预警 — 未来

- 容量预测
- 故障预测
- 趋势预测

1

平台定位

2

开发资源

3

基础环境标准化

4

平台架构



开发资源

为什么要DBA自研发？



- 只有DBA清楚自己的**痛点**，也只有DBA清楚研发中心的痛点
- DBA知道双方未来需要什么，更需要在数据库自治趋势下思考转型

! 重要的是能绝对把控需求实现的细节和开发进度，沟通成本低、资源利用率高，并且前期快速实现必然带来后期重构的问题，对于这一点单个团队内部看法容易统一。

存在的问题

架构不一，新老环境并存
OS环境不统一
数据库规模庞大，且仍在快速增长



自身对于环境的适应

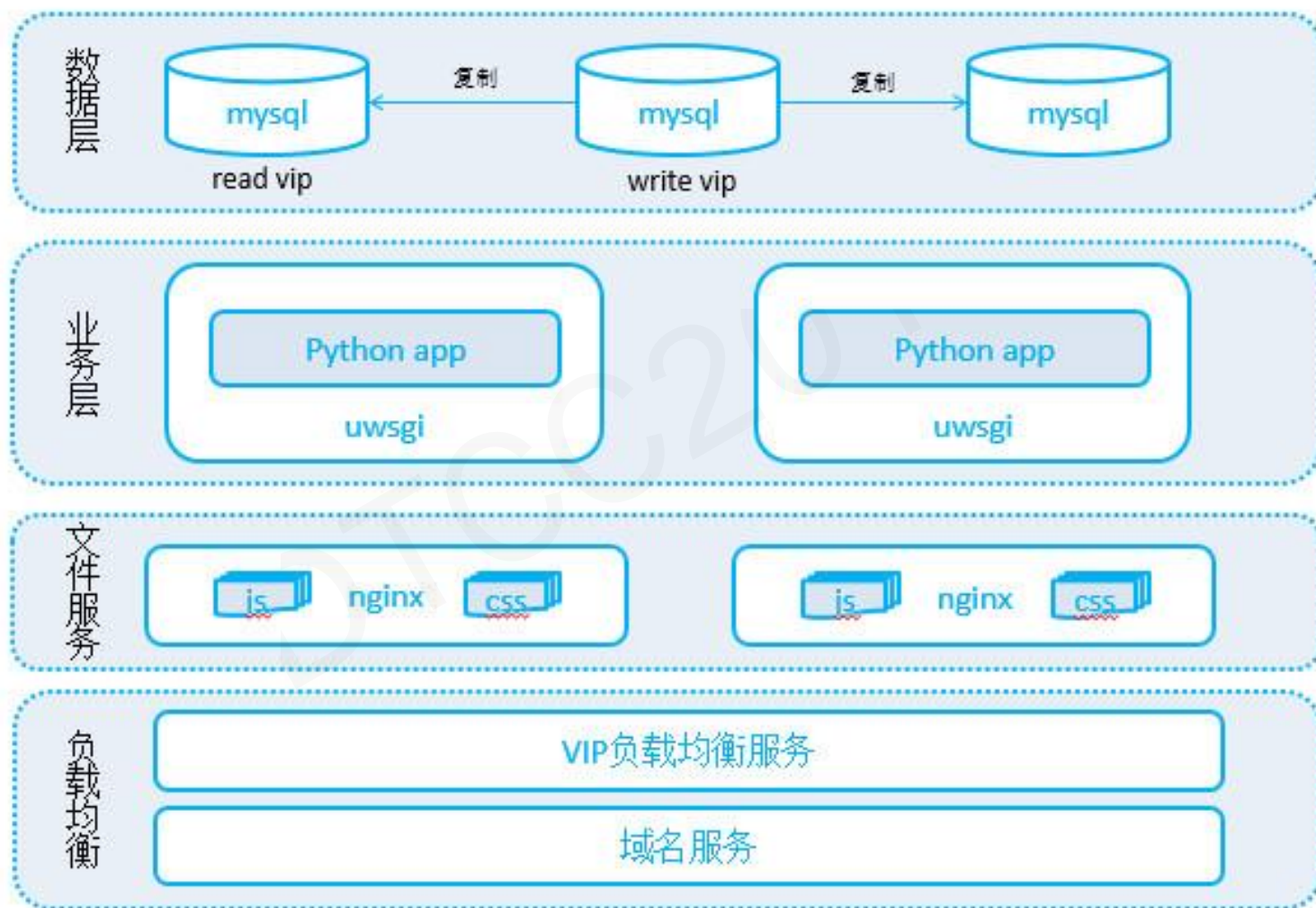
agent要求自动部署，自动升级，脚本实现自动分发

兼容现网各类数据库的主流版本，OS环境自动修复

监控数据要求分级分频、高准确性，对系统无侵入

要能和苏宁现有平台联动，且提供标准化接口服务

平台架构





前期设计思路



解决了哪些问题



智能化探索

苏宁 DBMS

1

面向DBA

2

面向研发中心

3

对大促的支持

支持多数据库类型



MySQL 5.5
MySQL 5.6
MySQL 5.7



DB2 v9.7
DB2 v10.5



PostgreSQL 9.6
PostgreSQL 10.2

DTCC
2018

数领先机 智赢未来 (9)

IT168.com

ChinaUnix

ITPUB





管理工具



综合大盘



精准监控



智能追踪



管理工具

参数管理

智能调度&限速/断点续传

备份管理

自动切换

一键迁移

一键扩缩容

自动告警处理

数据库性能指标秒级监控



综合大盘

趋势分析

TopSQL分析



精准监控

大数据辅助运维决策

变量监控

故障预测

自动化资源调整

自研解析工具SSE
解决PT的效率问题

SQL语句全量追踪

执行效率变化捕获

自动优化建议呈现



智能追踪



元数据



性能诊断



智能解析



SQL发布



元数据

实时数据字典

系统级元数据API



性能诊断

全面的数据库性能分析报告

智能优化建议



智能解析

SQL执行计划实时解析

异常SQL报表

索引和规则优化建议



SQL发布

预发布SQL审计

行级数据备份

自动发布/回滚

对大促的支持



秒级监控

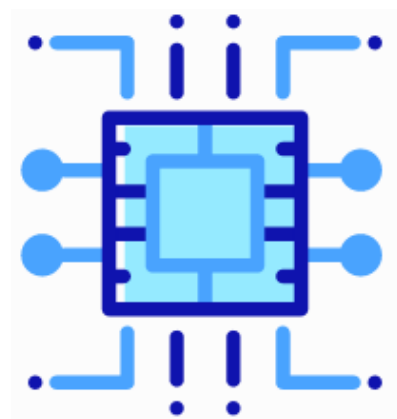
海量数据分布存储

使用最新大数据分布式技术，自开发数据转发层服务，解决数据采集秒级频度和海量存储的问题

实时智能分析

性能问题第一时间精准定位

依托storm实时流式计算，实时性能问题告警；辅以离线性能分析，资源预警定时送达



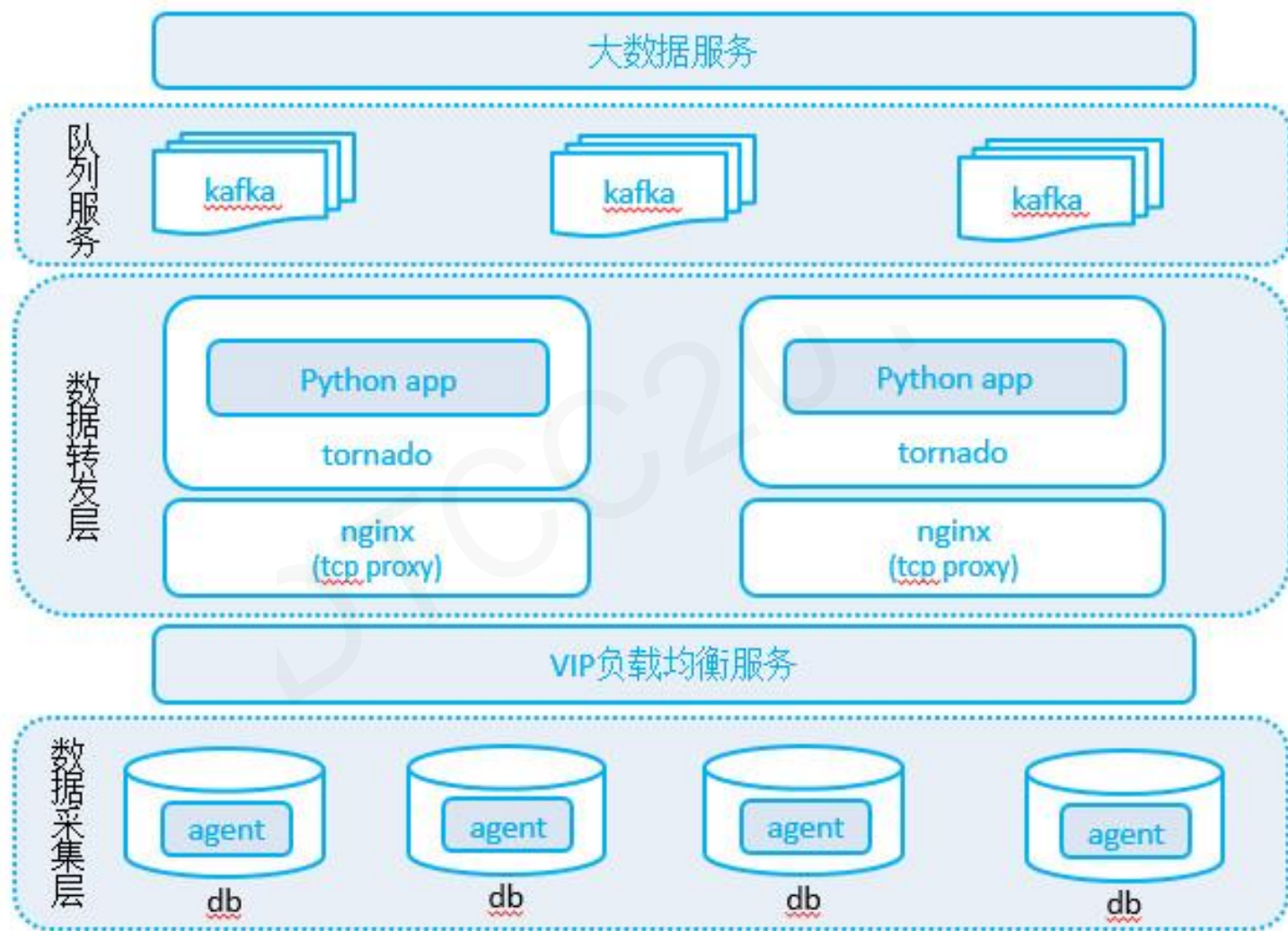
DTCC
2018

数领先机 智赢未来 (9)

IT168.com

ChinaUnix

ITPUB





DBMS平台的设计思路



DBMS解决了什么问题



智能化探索

苏宁 DBMS

- 1 智能大盘/事件感知
- 2 SQL全量追踪
- 3 自动优化建议
- 4 故障预测/告警自动处理
- 5 AIOps决策探索

DBRTMP 突变捕获



DBRTMP-Analyser

所有系统(ALL)

NET-RECV:37077756

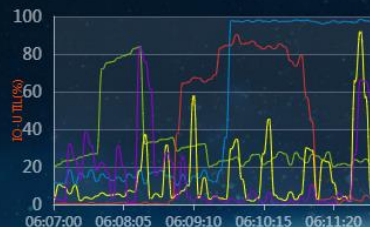
QPS:14765083

NET-SEND:115760263

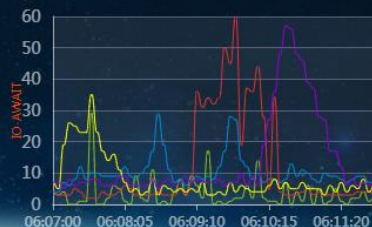
CPU突变Top5统计



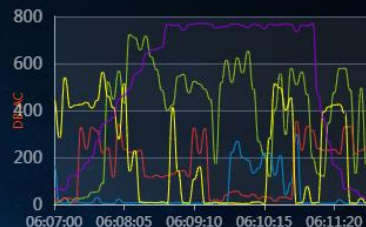
IO-UTIL突变Top5统计



IO-AWAIT突变Top5统计



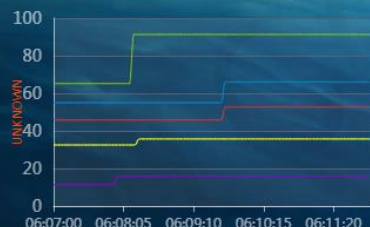
DB-AC突变Top5统计



QPS突变Top5统计



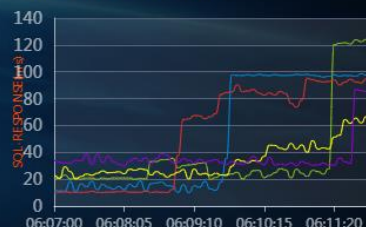
MEM-USAGE突变Top5统计



LOG-INCREASE突变Top5统计



SQL-RESPONSE突变Top5统计



QPS突变Top5统计

QPS>100000(1)



CPU-USER Top5统计

CPU>90(2)



IO-UTIL Top5统计

IO>80(3)



MEM-USAGE Top5统计

MEM>90(5)

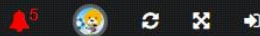


DBRTMP 自动告警处理平台



DBRTMP-Alarm

所有系统(ALL)



实时告警统计 告警系统(3)/所有系统(925) | 未处理告警(5)/未处理事件(5)

刷新频率: 60s

告警分布



订单查询系统(O MSQ)

未处理事件	2
警告 1	
严重 1	
灾难 0	

业务明细账系统(BUDS)

未处理事件	2
警告 1	
严重 1	
灾难 0	

远程服务调用框架(RSF)

未处理事件	1
警告 1	
严重 0	
灾难 0	

会员账户管理平台(CAMP)

未处理事件	0
警告 0	
严重 0	
灾难 0	

云台魔方系统(IDAS)

未处理事件	0
警告 0	
严重 0	
灾难 0	



近1小时系统告警Top5统计



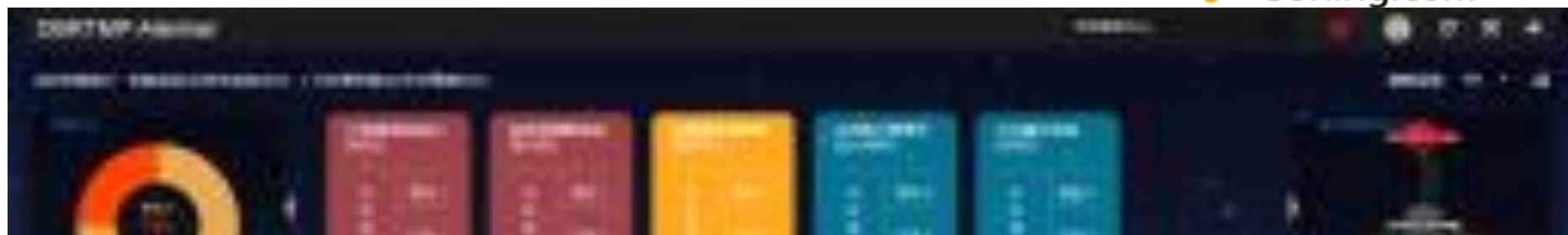
近1小时告警类型Top5统计



实时告警处理

今日告警总量(62) | 自动处理(45)/手动处理(17) | 已恢复(56)/未恢复(6)

序号	告警内容	处理结果
1	主机10.100.111.115/mysql/bak分区剩余空间为9 %小于阈值10%	自动处理: 判断失败
2	主机10.102.26.3 Mysql日志产生ERROR信息	未自动处理
3	主机10.102.146.100/mysql/data分区剩余空间为4.29 %小于阈值10%	自动处理: 已自动请求扩容
4	主机10.100.111.114 SLAVE延迟1S, 端口3306	警告
5	主机10.102.26.3 SLAVE延迟20S, 端口3306	警告
6	主机10.94.22.3Mysql大于1小时的session数为1	警告
7	主机10.103.59.196空闲CPU为1.68 %持续15分钟小于10%	自动处理: SQL慢
8	主机10.94.19.135/mysql/binlog分区剩余空间为9.99 %小于阈值10%	自动处理: 日志清理成功
9	主机10.102.158.229/mysql/data分区剩余空间为0 %小于阈值10%	自动处理: 临时表导致
10	主机10.94.22.3Mysql大于2小时的session数为1	自动处理: 业务长查询, 已发邮件提醒
11	主机10.102.146.132/mysql/binlog分区剩余空间为9.99 %小于阈值10%	自动处理: 日志清理成功



实时告警统计 告警系统(3)/所有系统(925) | 未处理告警(5)/未处理事件(5)

告警分布



订单查询系统(O MSQ)

未处理事件	警告 1
	严重 1
	灾难 0
2	

业务明细账系统(BUDS)

未处理事件	警告 1
	严重 1
	灾难 0
2	

远程服务调用框架(RSF)

未处理事件	警告 1
	严重 0
	灾难 0
1	

会员账户管理平台(CAMP)

未处理事件	警告 0
	严重 0
	灾难 0
0	

云台魔方系统(IDAS)

未处理事件	警告 0
	严重 0
	灾难 0
0	



监视器 连接数 查询 复制 BINLOG INNODB引擎 INNODB-MUTEX 关注状态

指标	描述	当前值	变化	检查时间	
Innodb_buffer_pool_pages_data		747808	+82	12:01:00	数据图
Innodb_buffer_pool_pages_dirty		904	+156	12:01:00	数据图
Innodb_buffer_pool_pages_flushed		1525536334	-12532	12:01:00	数据图
Innodb_buffer_pool_pages_free		0	+3	12:01:00	数据图
Innodb_buffer_pool_pages_misc		38616	-85	12:01:00	数据图
Innodb_buffer_pool_pages_total		786424	-	12:01:00	数据图
Innodb_buffer_pool_reads		53805274	-5838	12:01:00	数据图
Innodb_buffer_pool_read_ahead		0	-	12:01:00	数据图
Innodb_buffer_pool_read_requests		5404842685220	-9914212	12:01:00	数据图
Innodb_buffer_pool_wait_free		39	-	12:01:00	数据图
Innodb_buffer_pool_write_requests		22117805957	-77274	12:01:00	数据图
Innodb_log_waits		0	-	12:01:00	数据图
Innodb_os_log_written		2521782957056	-11692032	12:01:00	数据图
Innodb_pages_created		27936458	-131	12:01:00	数据图
Innodb_pages_read		55684506	-5843	12:01:00	数据图
Innodb_pages_written		762768439	-6266	12:01:00	数据图
Innodb_page_size		16384	-	12:01:00	数据图
Innodb_rows_deleted		471667592	-1119	12:01:00	数据图
Innodb_rows_inserted		1329650448	-5528	12:01:00	数据图
Innodb_rows_read		2842034861512	-22172974	12:01:00	数据图
Innodb_rows_updated		1574319731	-1742	12:01:00	数据图

SQL全量追踪



格式化代码 解析 关注 屏蔽

-- SQL_ID:1533241727961052212

```

SELECT a.db_id,
       UNIX_TIMESTAMP(ifnull(c.sync_timestamp, date_add(now(), INTERVAL -5 MINUTE))),
       UNIX_TIMESTAMP(b.update_time),

```

校验码	首次截获		最新截获		执行次数		查询总耗时(s)		查询最短耗时(s)		查询最长耗时(s)		最大查询行数		扩展		明细	
1533241727961052212	18-05-04 13:21:00		18-05-11 11:13:31				40.601		0.027		0.077		6687					
开始时间	结束时间	用户	执行次数	查询总耗时(s)	查询最短耗时(s)	查询最长耗时(s)	锁总耗时(s)	锁最短耗时(s)	锁最长耗时(s)	传输总行数	最少查询行数	最多查询行数	检查总行数	最少				
18-05-04 13:21:00	18-05-04 13:21:00	fabu	null	0.028	0.028	0.028	0	0	0	6658	6658	6658	20342					
18-05-04 13:46:00	18-05-04 13:46:00	fabu	null	0.041	0.041	0.041	0	0	0	6658	6658	6658	20342					
18-05-04 14:21:00	18-05-04 14:21:00	fabu	null	0.046	0.046	0.046	0	0	0	6658	6658	6658	20342					
18-05-04 15:41:02	18-05-04 15:41:02	fabu	null	0.034	0.034	0.034	0	0	0	6658	6658	6658	20342					
5119468243444413991	18-05-04 13:21:00		18-05-11 11:13:31				46.3		0.028		0.088		6687					
7658952322571191363	18-05-04 14:06:00		18-05-11 11:13:31				36.551		0.172		0.33		571					
14758480409770795287	18-05-04 14:06:00		18-05-11 11:13:31				2869.64		0.17		0.409		667					
4081024690202860173	18-05-04 14:06:00		18-05-11 11:08:31				3.346		0.006		0.021		0					
4174642471673772653	18-05-04 14:21:00		18-05-11 11:08:31				15.363		0.011		0.074		0					
8716941608199230049	18-05-07 09:27:00		18-05-11 11:08:31				21.729		0.242		0.355		10081					
8775975579636802510	18-05-07 09:27:00		18-05-11 11:08:31				24.655		0.237		0.793		10081					
10126360866398807494	18-05-05 16:11:23		18-05-11 11:08:31				144.39		0.25		1.43		10081					

1天 1周 半月 1月 执行耗时 搜索: 校验值. sql

首页 上一页 1 2 3 4 5 6 下一页 尾页

概览@110 连接管理 语句优化器@110

系统名、IP、数据库名

探查器
innotop 模拟器
性能报告
SQL分析
全量SQL性能分析诊断
诊断分析
问题语句
TOP查询
慢查询
语句优化器
优化建议、执行计划

执行 停止 解析 优化 格式化代码 导出

编辑器

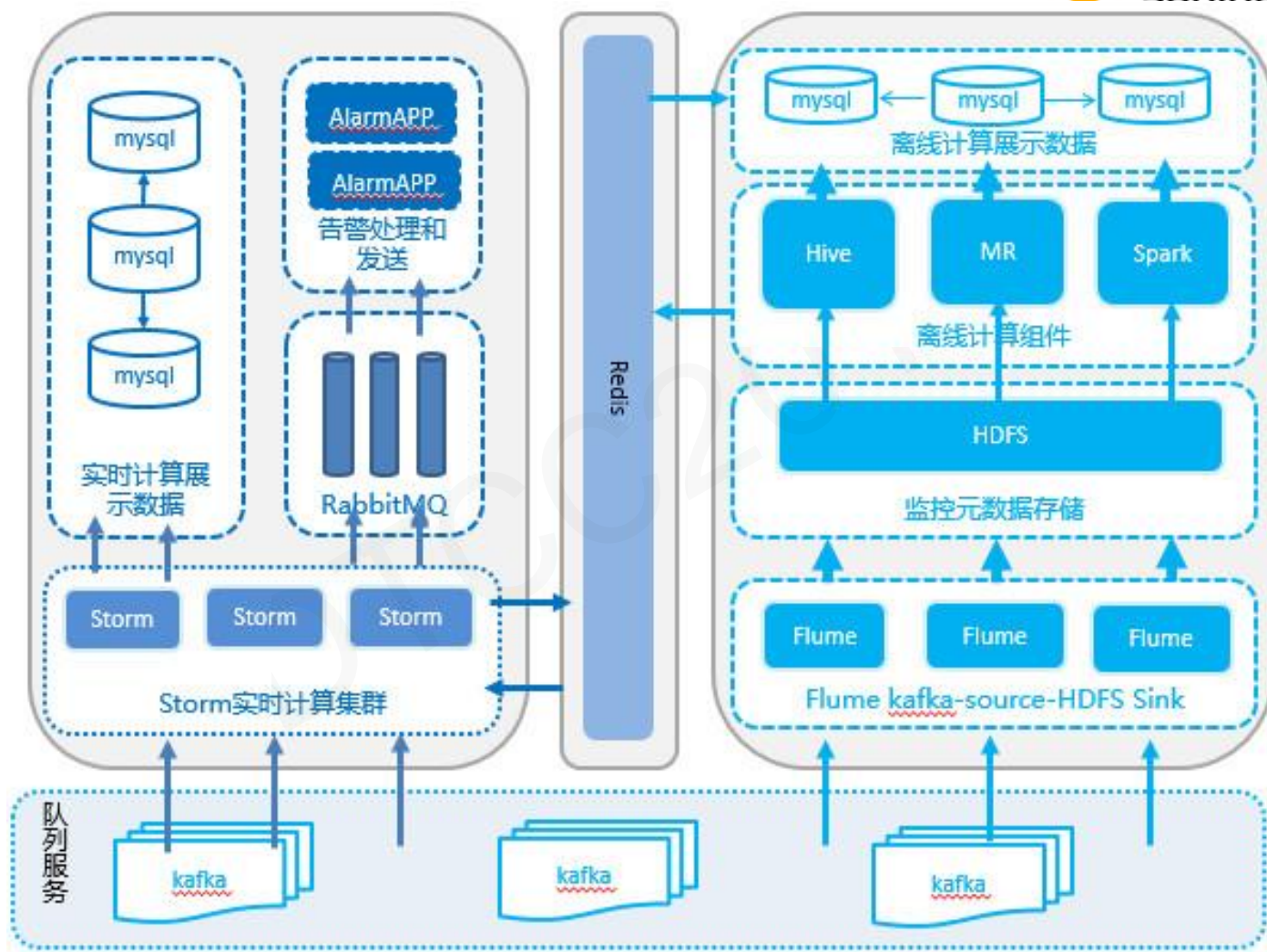
```

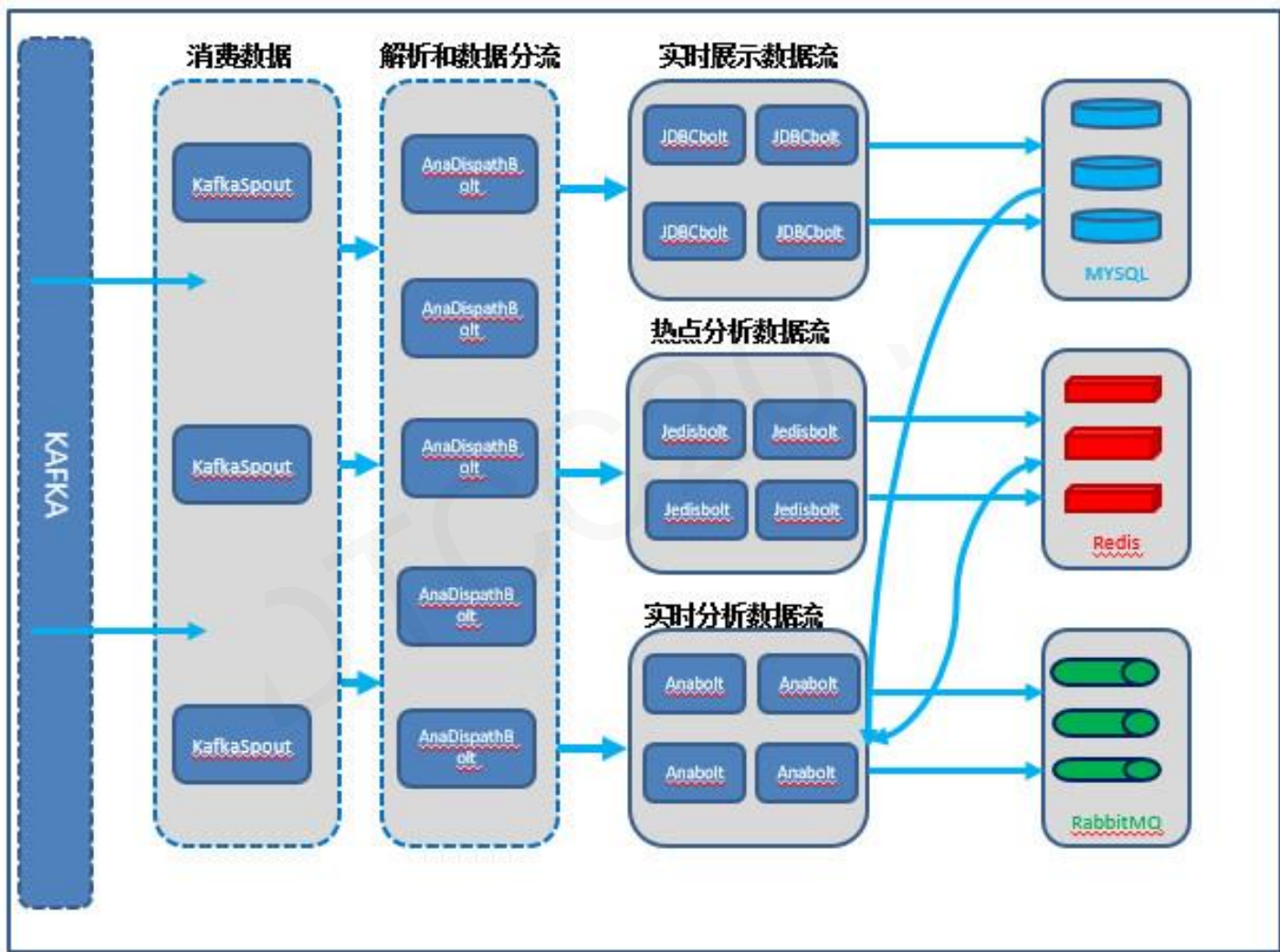
1 select JobID,a.Status,a.Type,State,Policy,Schedule,ipaddr,
2 Started,Ended,Elapsed,speed_KB,Kilobytes,Policy_Type,Retention,
3 Schedule_Type,a.Owner,b.sys_name, group_concat(b.db_name), date_format(update_timestamp, '%Y-%m-%d %T')
4 from nbu_monitorwork a
5 left join (select sys_code, sys_name, 1 as ser_type, ser_host, db_code, db_name from t_info_database union all
6 select sys_code, sys_name, 2 as ser_type, ser_host, db_code, db_name from t_info_database_oracle union all
7 ) b on a.ipaddr = b.ser_host where Policy = 'day' and Kilobytes > '10000000'
8 group by JobID,a.Status,a.Type,State,Policy,Schedule,ipaddr,
9 Started,Ended,Elapsed,speed_KB,Kilobytes,Policy_Type,Retention,
10 Schedule_Type,a.Owner,update_timestamp,b.sys_name
  
```

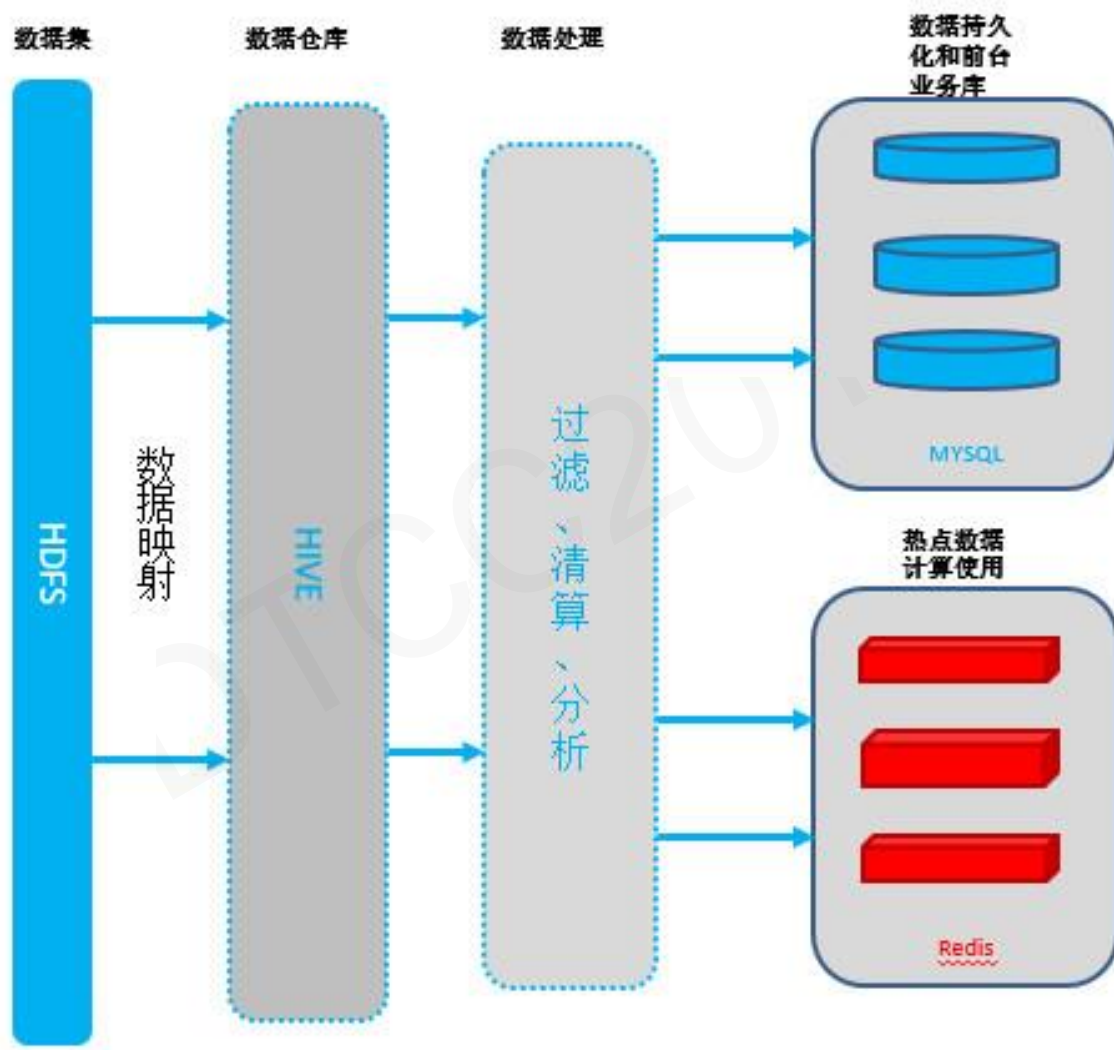
信息 规则优化 索引优化

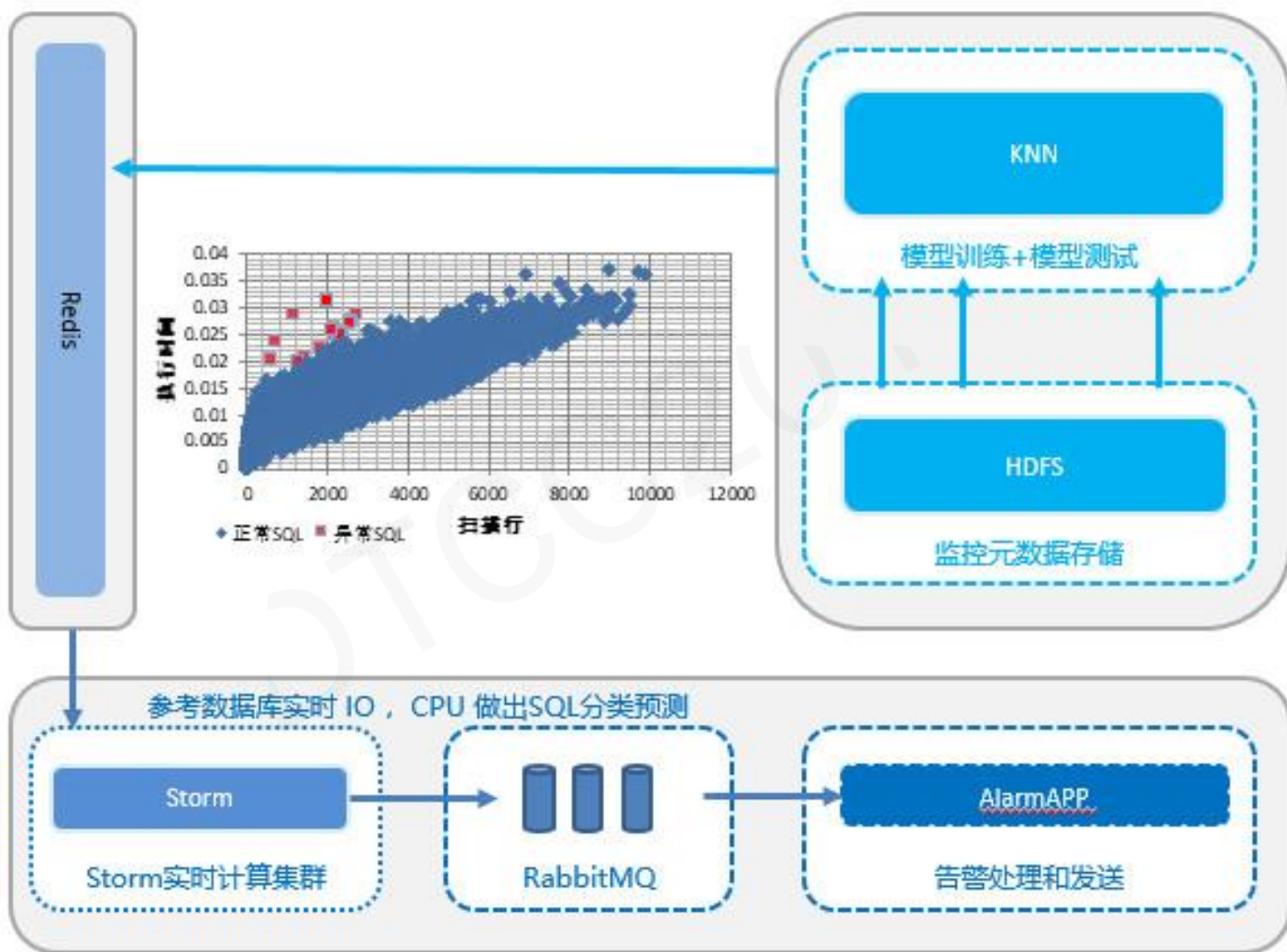
IDX	TABLE	COLUMN	RULE	VIOLATIONS	AFFECT	ADVICE
1	t_info_database		union all	发现该规则	性能损耗	优化业务逻辑, 规避'union all'
2	nbu_monitorwork		left join	发现该规则	性能损耗	优化业务逻辑, 规避'left join'
3	nbu_monitorwork	Kilobytes	隐式转换	发现该规则	索引失效	避免使用隐式转换
4	nbu_monitorwork		index number	大于5	数据插入效率降低	删除多余无效索引
5	b		subquery	发现该规则	效率和内存的损耗	优化业务逻辑, 规避使用'子查询'

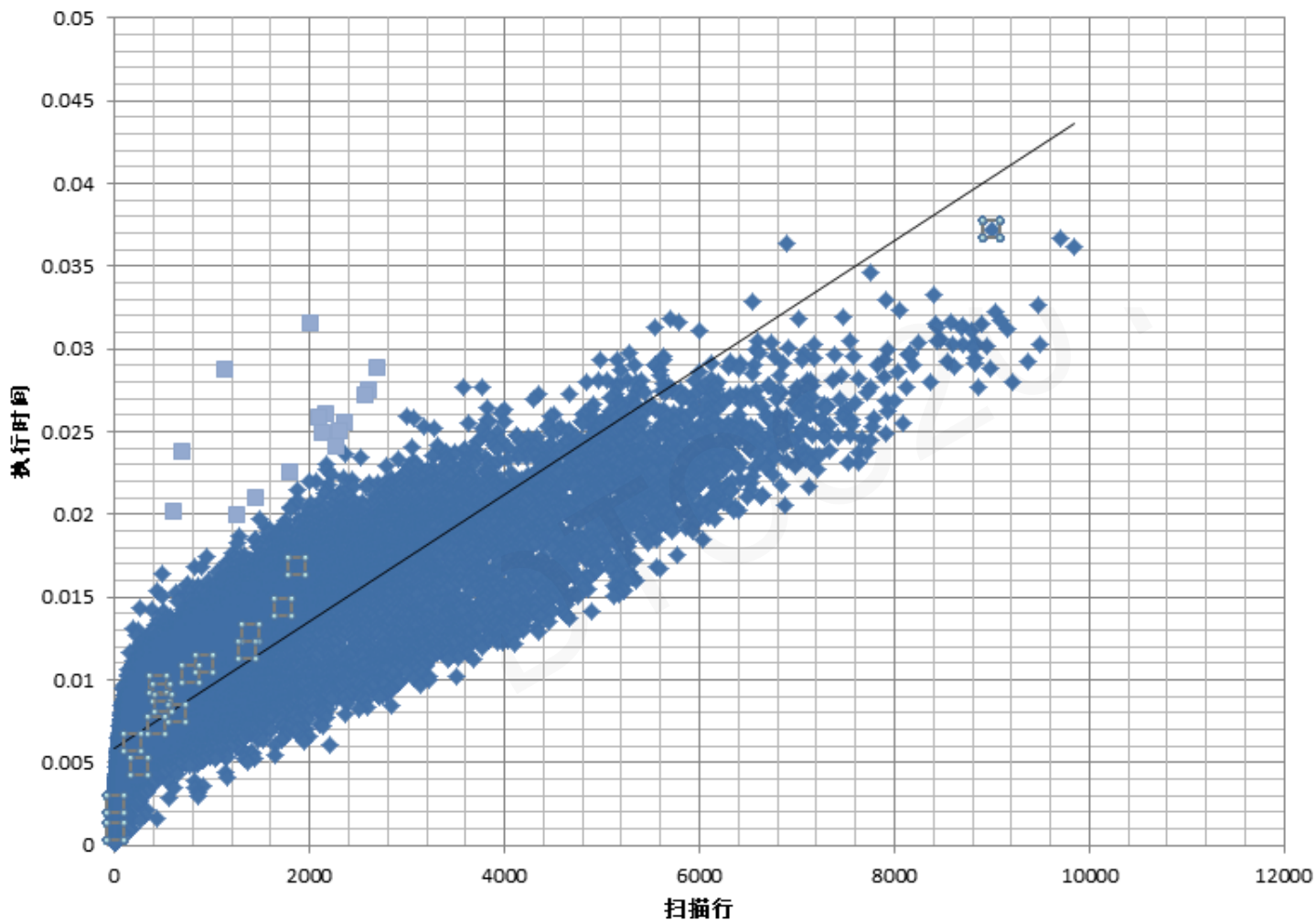
首页 上一页 1 下一页 尾页











离线计算中，利用海量SQL扫描行数和执行时间的数据不断进行训练，得到单个SQL ID的历史效率模型；

捕获到**新**SQL时，实时计算提供出最近的效率模型和历史模型进行邻近对比，同时参考同期IO与CPU消耗突变，精准预判异常SQL，告别传统慢SQL定义。



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