

异构系统链路追踪

——滴滴trace实践



促进软件开发领域知识与创新的传播



关注InfoQ官方信息
及时获取QCon软件开发者
大会演讲视频信息



扫码，获取限时优惠



全球架构师峰会 2017 [深圳站]

2017年7月7-8日 深圳·华侨城洲际酒店

咨询热线: 010-89880682



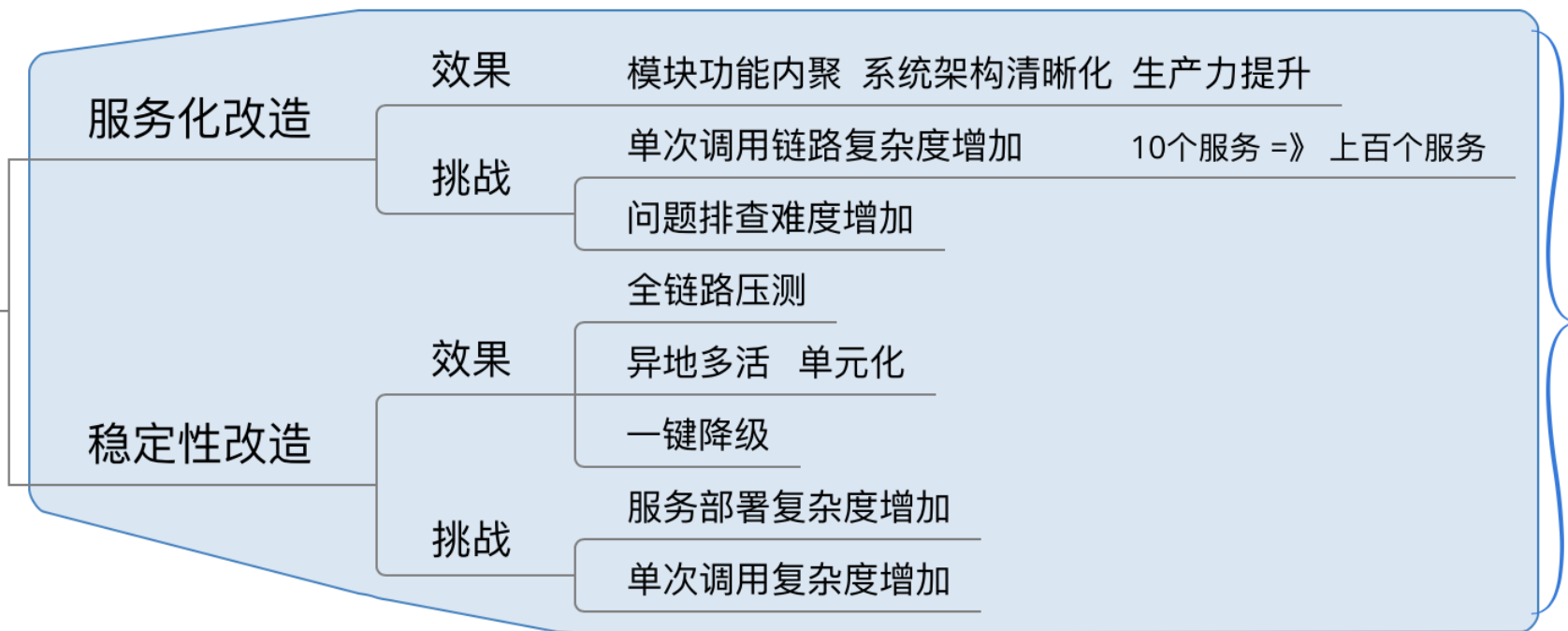
全球软件开发大会 [上海站]

2017年10月19-21日

咨询热线: 010-64738142

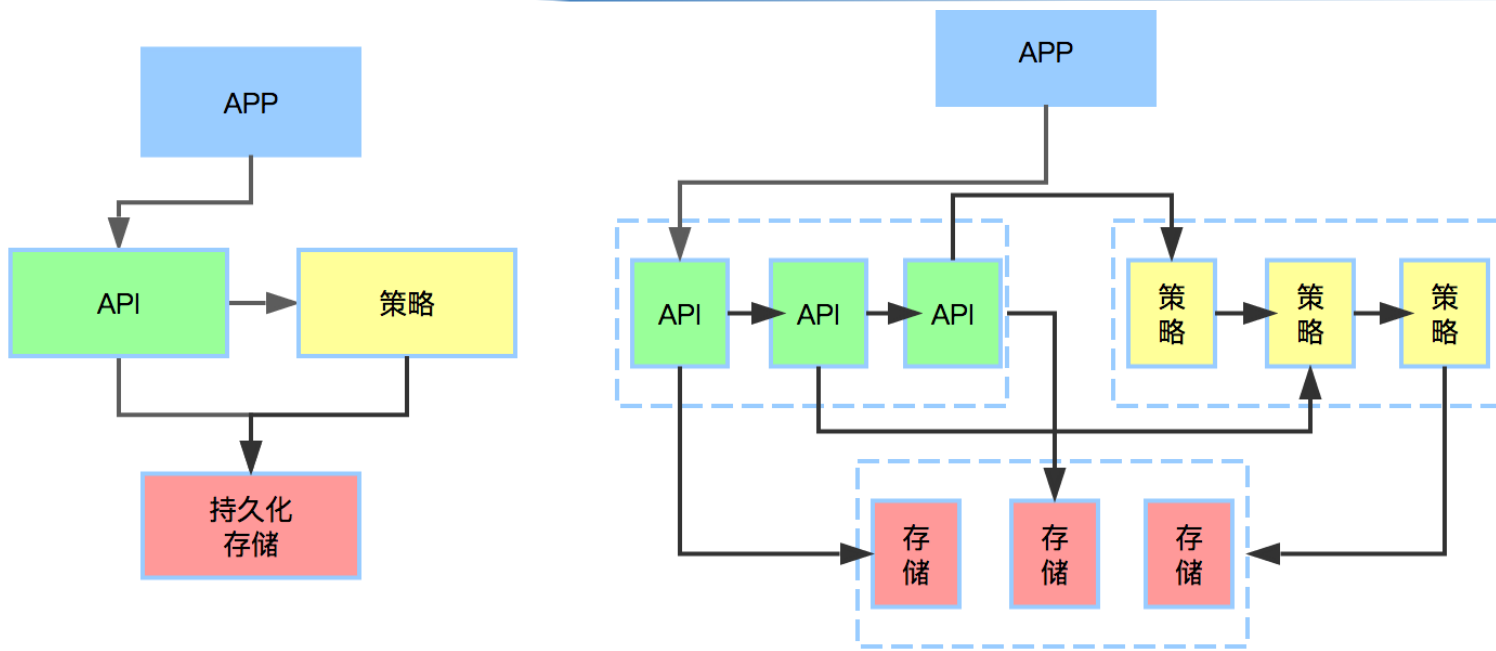
背景

背景

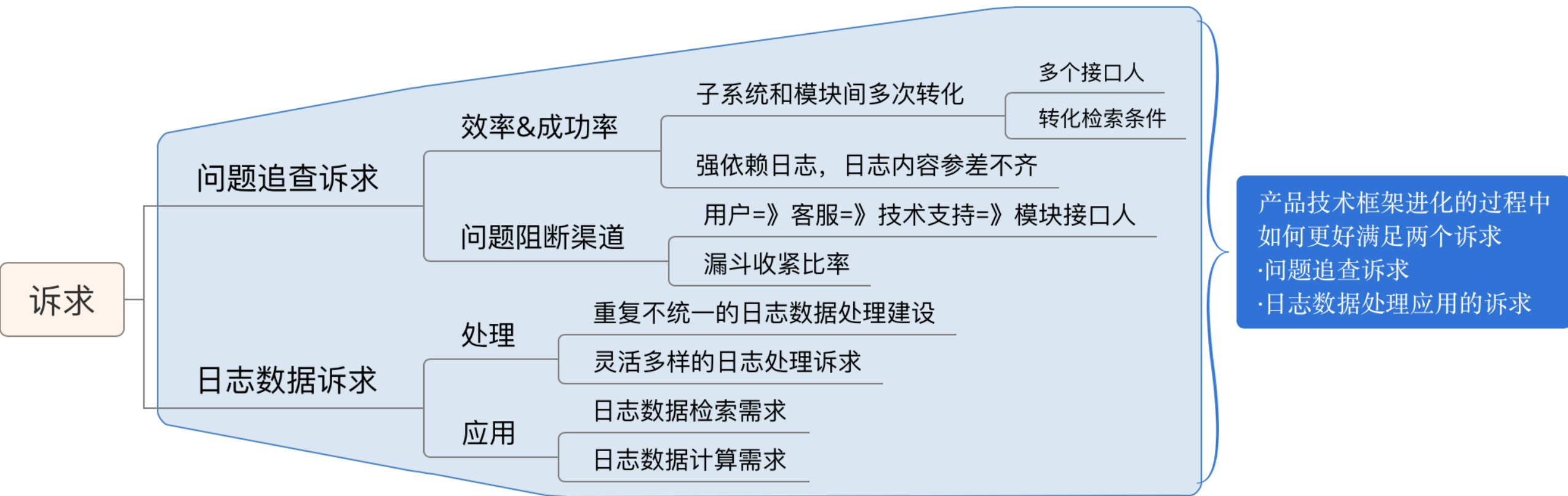


服务化改造和稳定性改造带来

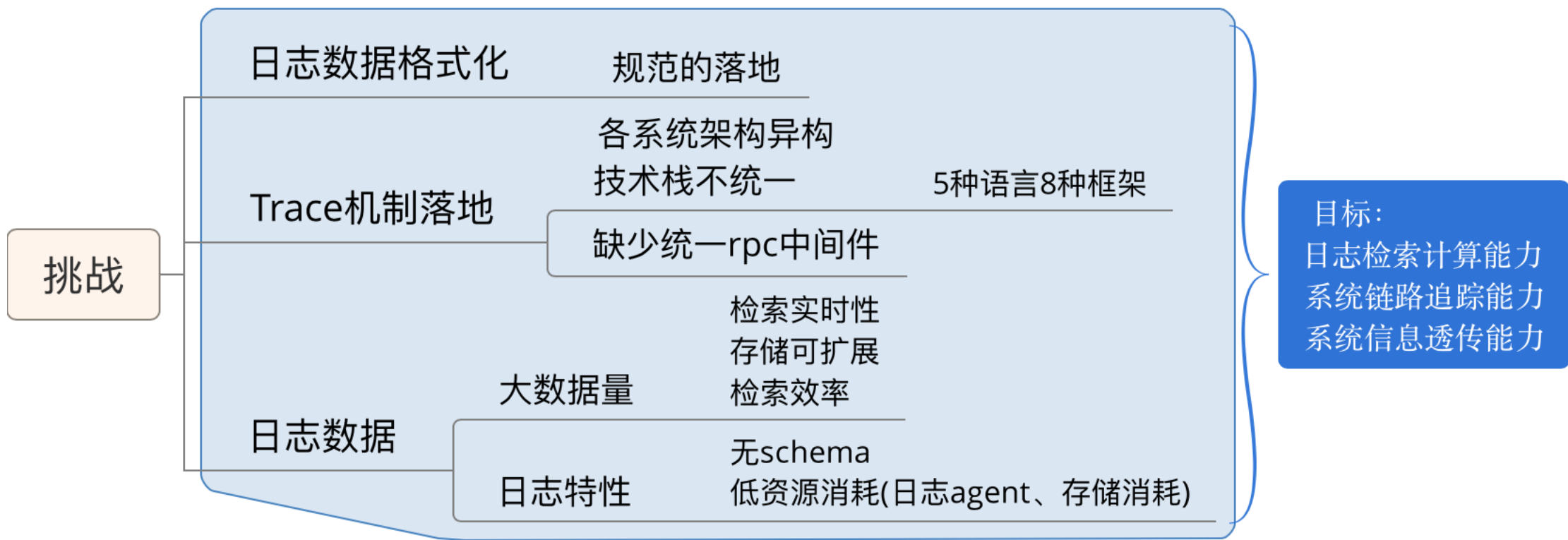
- 调用链路复杂度增加
- 问题排查难度增加



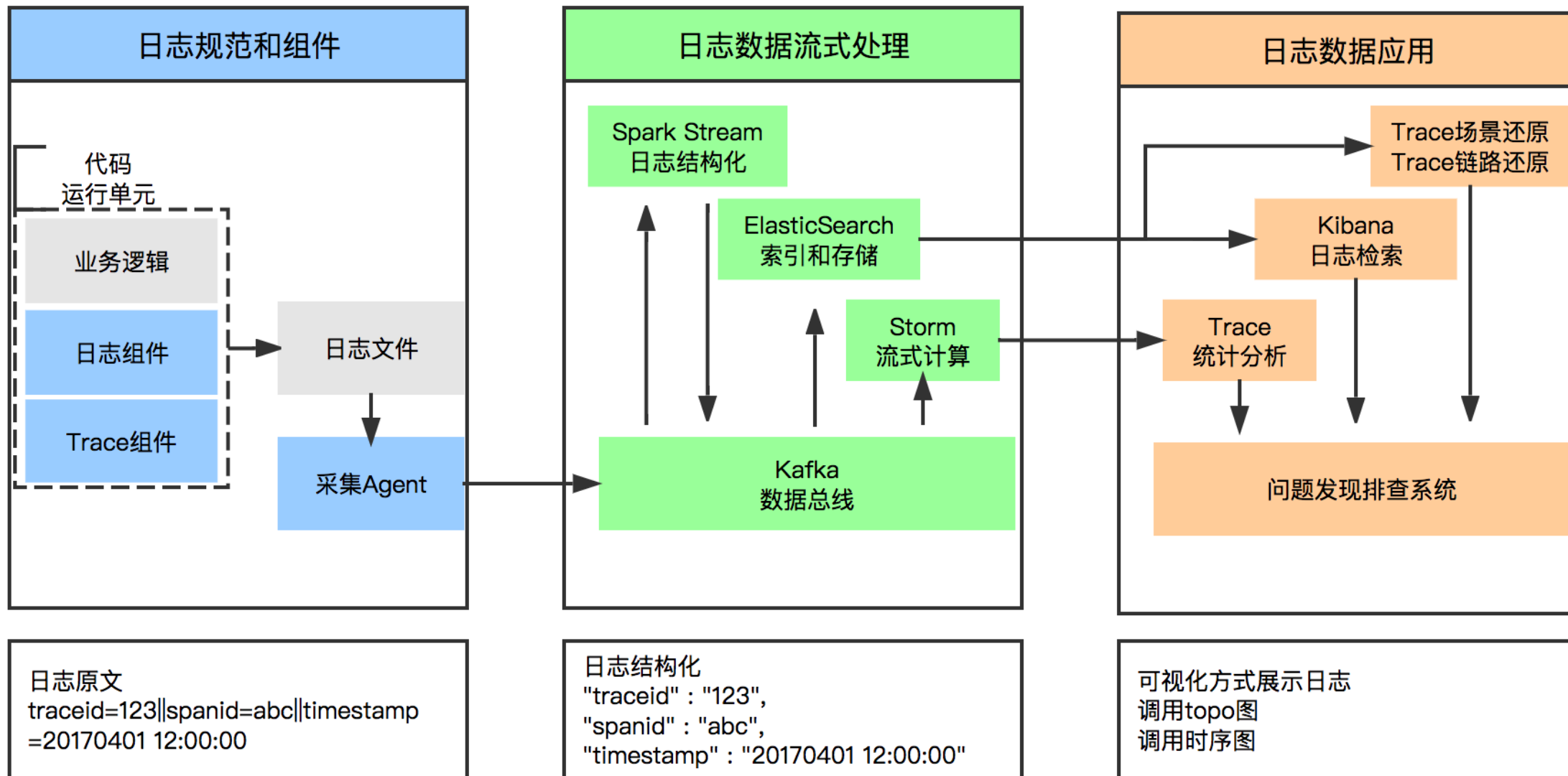
诉求



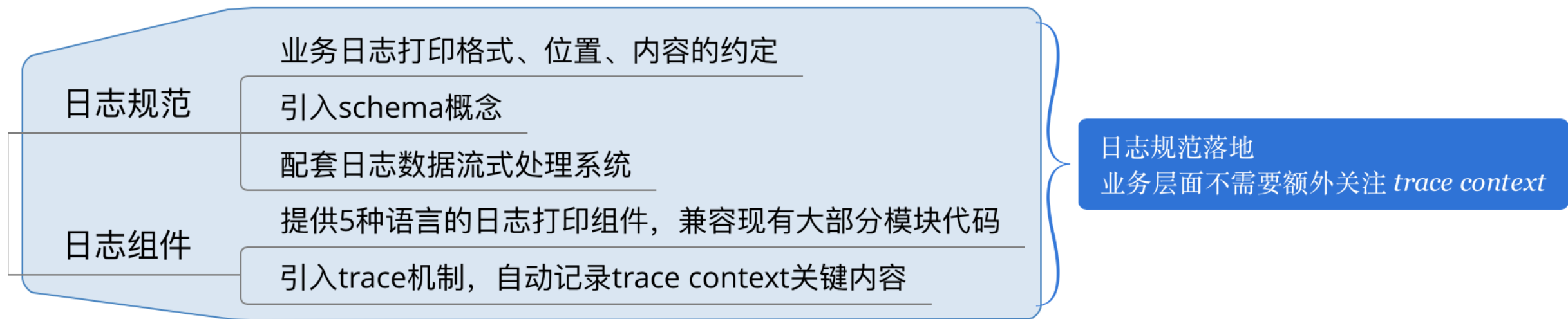
挑战—在滴滴落地面临的挑战



方案



日志规范和组件



[NOTICE][2017-03-16T21:18:41.473+0800][line=]

固定日志头：等级、时间戳、代码行

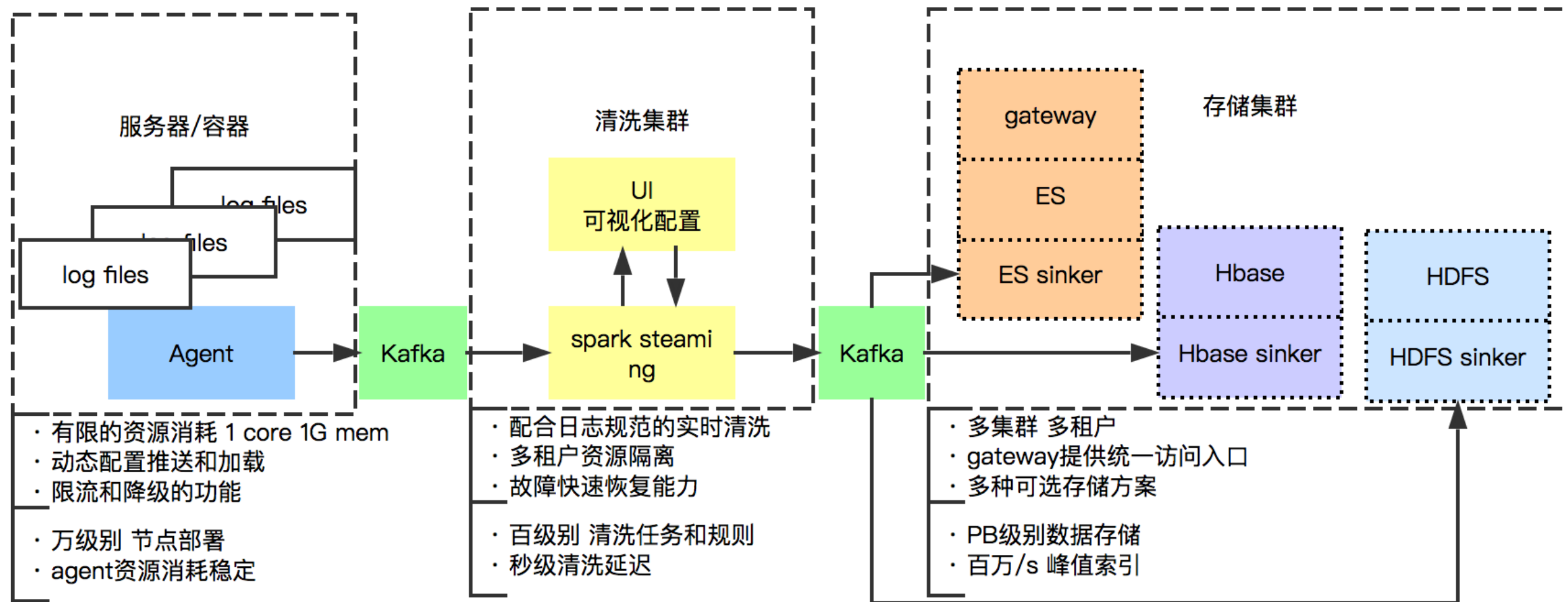
_com_request_in

tag：描述日志打印时机和内容

||traceid=0a61355fe111f8bfed17419410310303||sampling=0
||spanid=248b0b7136fe0b77||order_id=1001

内容：k=v的格式 ||分隔

日志数据处理



日志应用

格式化日志

日志流式处理

错误日志监控

上线观察、主动报警

波动曲线可以下钻到具体trace链条

接口性能监控

分位接口性能曲线和主动报警

波动曲线可以下钻到具体trace链条

日志批处理

MapReduce

业务日志保存入HDFS

ETL 用于离线反作弊

日志实时检索

通过ES进行日志检索

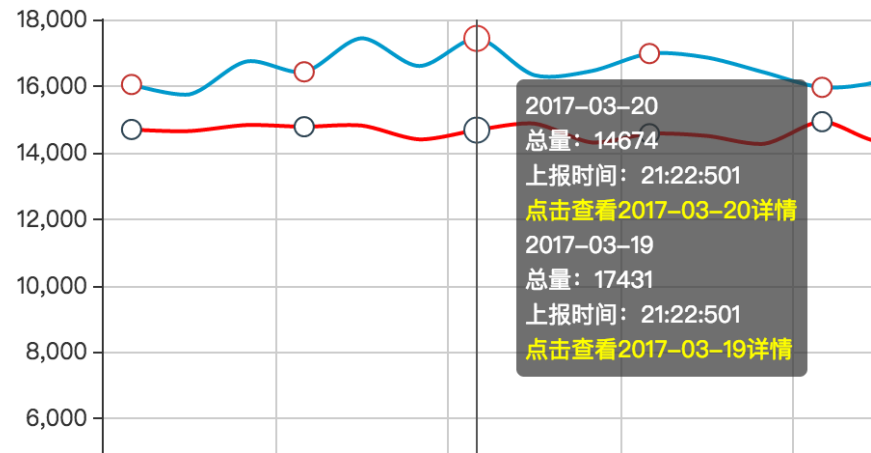
接入trace进行可视化展示

多维度检索和过滤

进行简单聚合和统计

规范化日志的应用
流式处理
批处理
实时检索

错误上报趋势图

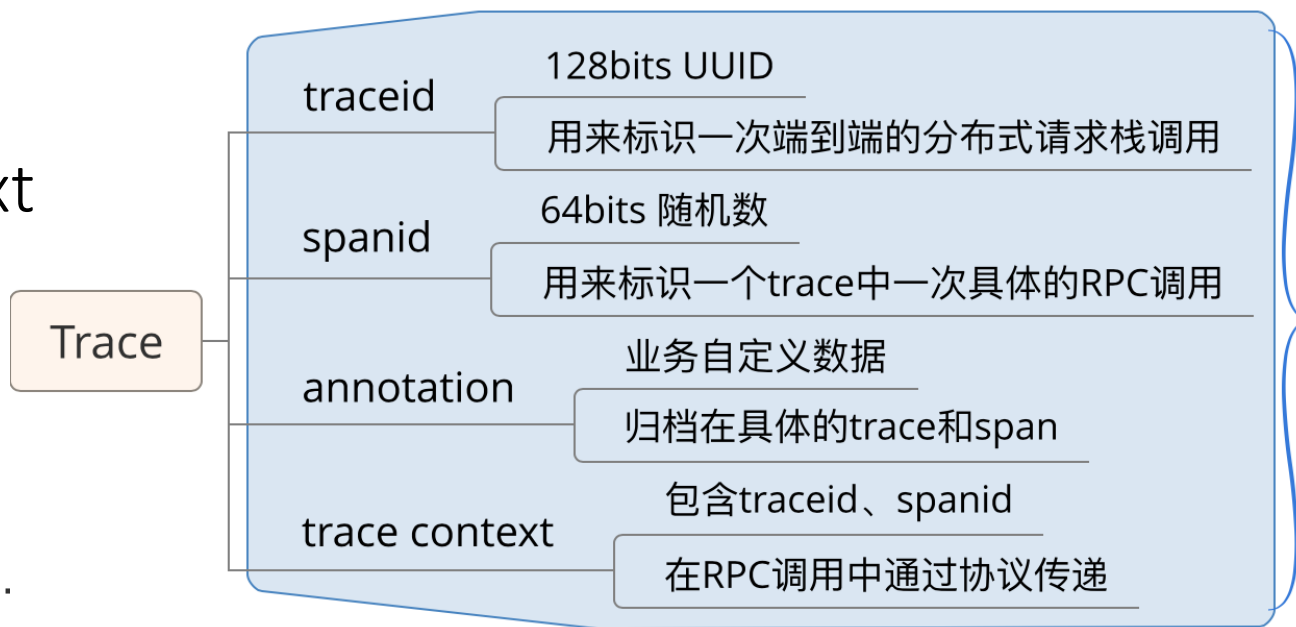


响应时间趋势图



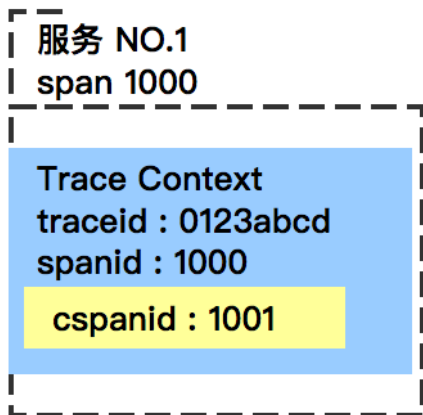
Trace机制

—传递trace context

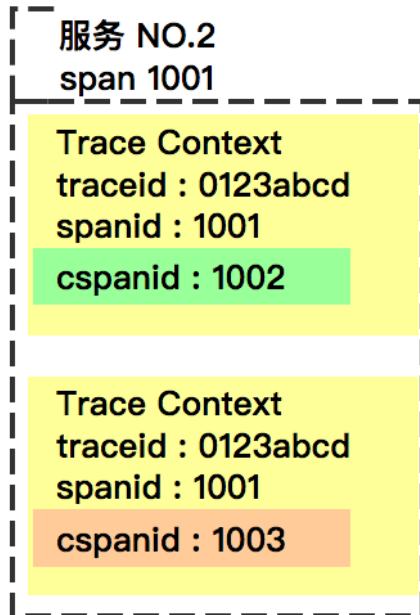


google dapper论文的实践
在分布式调用中引入trace context
并在每次RPC请求中传递和转化

trace 01234abc



①



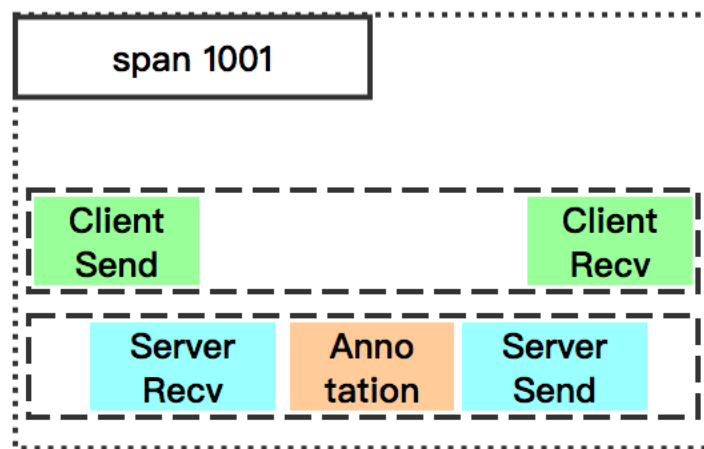
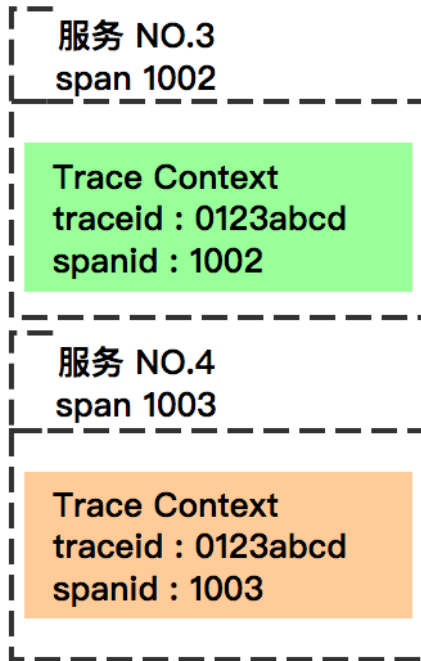
②

③

④

⑤

⑥



Trace落地—结合滴滴情况落地

因地制宜

trace信息 记录位置

★ 在RPC框架统一记录

★ 在AOP、通信LIBs记录

spanid编码

★ 有编码 点分方法记录spanid: 0.1.1.2

★ 随机数 spanid: 52286bce599eff8

trace信息 记录方式

★ 单独记录RPC请求信息

★ 业务日志和RPC通信日志都附带trace信息

采样逻辑

★ 采样记录调用链条, 采样采集调用数据

★ 全量记录关键调用, 采样记录细节调用, 选择性采集

业务id和traceid关联

★ 无关联关系

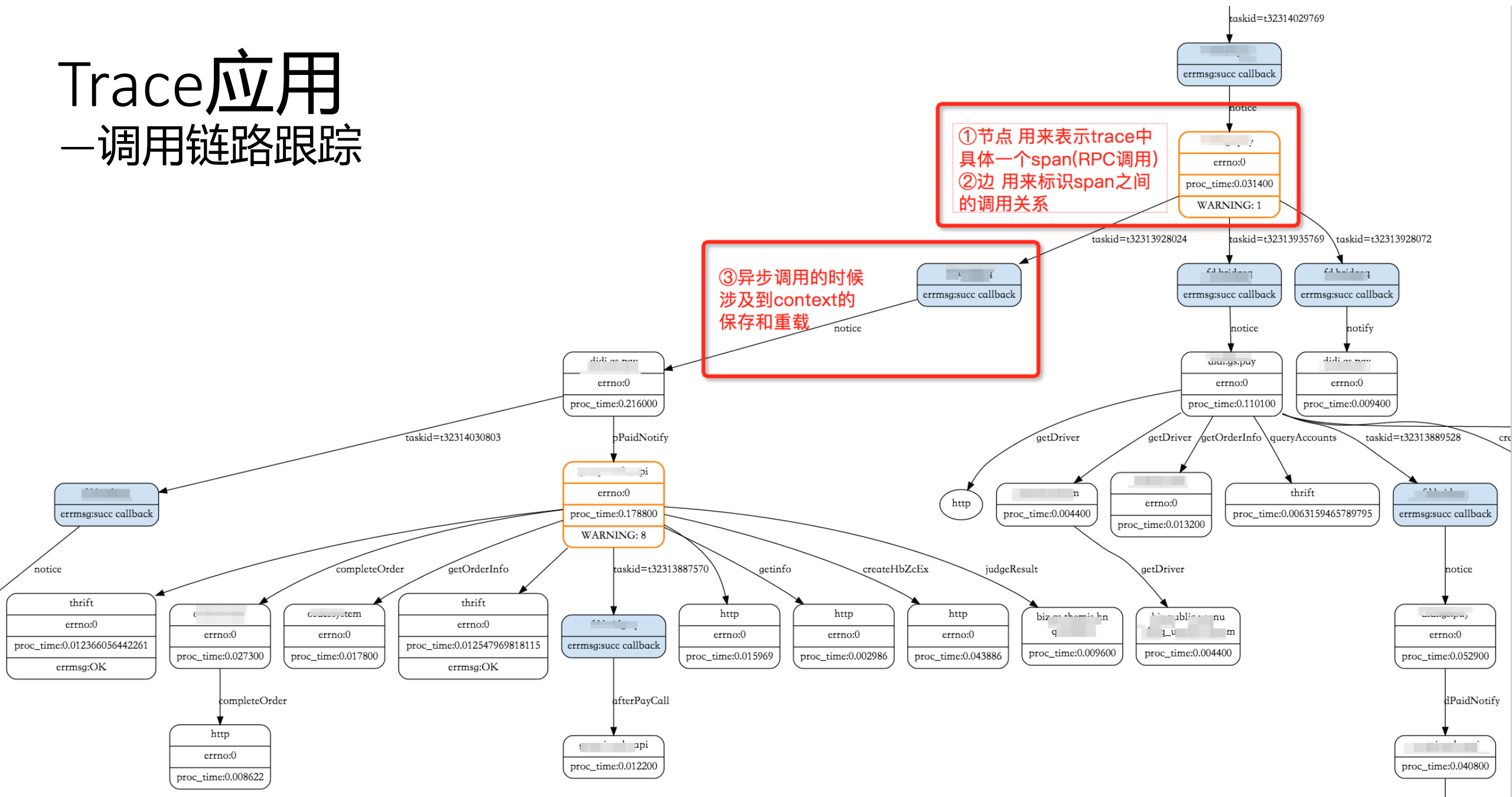
★ 建立业务id和traceid的关联关系

缺少统一RPC中间件
系统架构异构, 技术栈不统一

问题定位诉求
可以追踪单次调用的业务异常

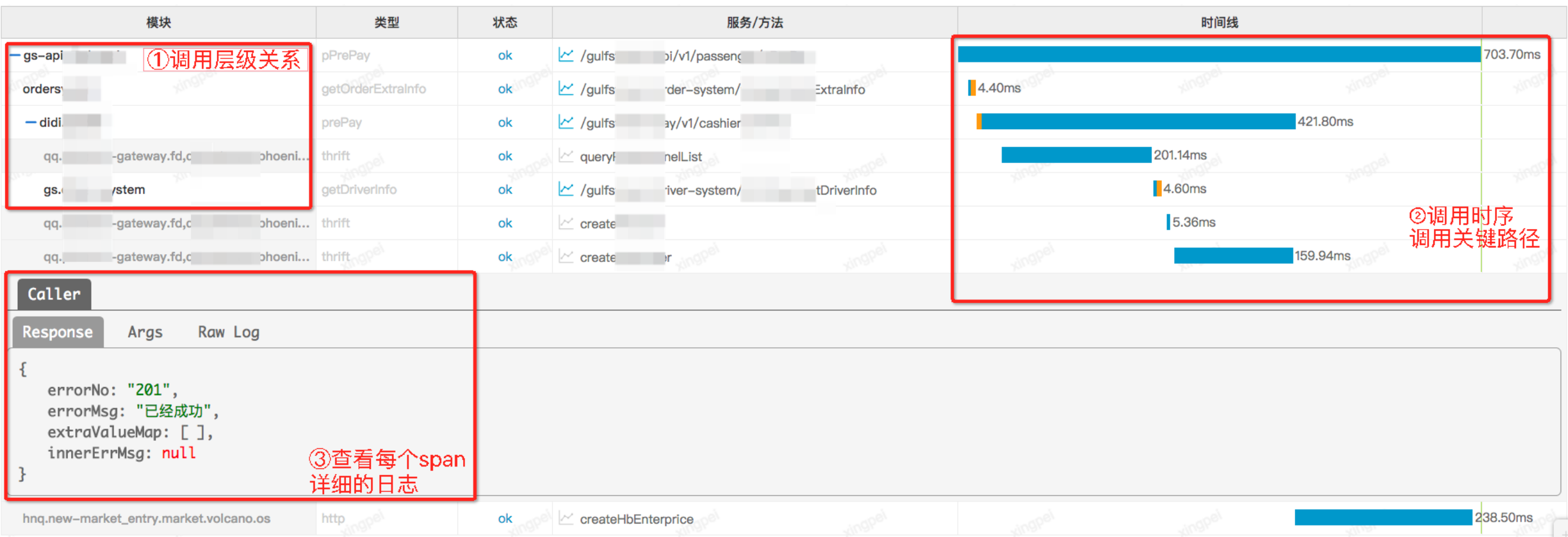
Trace应用

—调用链路跟踪



Trace应用

—调用链路跟踪



Trace应用

—场景还原

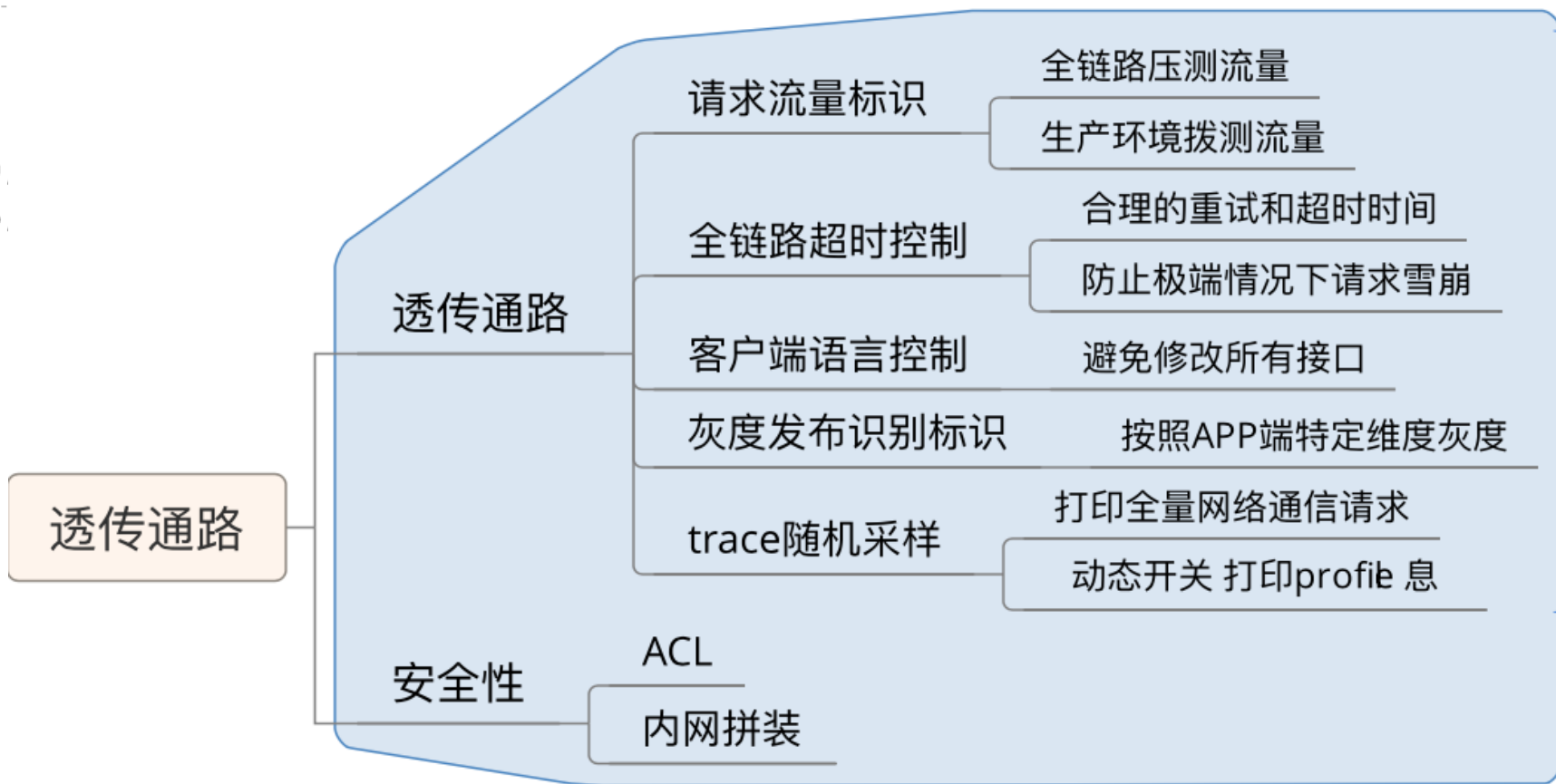
- ① 业务id和traceid关联
- ② 根span具有业务代表性
- ③ 时间顺序还原操作序列
- ④ 可下钻到具体trace

时间	Url	操作	订单号	详情
2017-03-15 22:14:44.943100	/gulfstream/passenger/v2/core/...wOr...	乘客创建订单	3061283	查看详情
2017-03-15 22:15:04.545100	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:15:06.890400	/gulfstream/hermesapi/v1/driver...ueryS...reStatus	查询司机抢单状态	3061283	查看详情
2017-03-15 22:15:07.412100	/gulfstream/api/v1/driver/dOrder...ail	司机端获取订单详情	3061283	查看详情
2017-03-15 22:15:08.569400	/gulfstream/driver/v2/driver/dO...Ca...	司机展示订单卡片	3061283	查看详情
2017-03-15 22:16:33.799800	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:16:37.867200	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:16:51.705400	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:17:29.694000	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:18:54.481500	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:20:59.453400	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:21:03.303900	/gulfstream/driver/v2/driver/dAr...d	司机达到上车点	3061283	查看详情
2017-03-15 22:21:42.208900	/gulfstream/driver/v2/driver/dBe...Char...	开始计费	3061283	查看详情
2017-03-15 22:22:20.510500	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:24:30.034400	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:32:21.019500	/gulfstream/driver/v2/driver/dEr...harge...	司机确认账单	3061283	查看详情
2017-03-15 22:32:23.428900	/gulfstream/driver/v2/driver/dFi...Orde...	结束计费	3061283	查看详情
2017-03-15 22:32:24.435000	/gulfstream/api/v1/driver/dOrder...ail	司机端获取订单详情	3061283	查看详情
2017-03-15 22:32:38.844100	/gulfstream/api/v1/passenger/p...Order...ail	乘客订单详情	3061283	查看详情
2017-03-15 22:32:45.197900	/gulfstream/api/v1/passenger...Colle...n	日志收集	3061283	查看详情
2017-03-15 22:32:45.470300	/gulfstream/api/v1/passenger...ay	乘客发起支付	3061283	查看详情
2017-03-15 22:32:46.141100	/gulfstream/pay/v1/phoenixin.../con...nNotify	支付通知	3061283	查看详情

透传通路

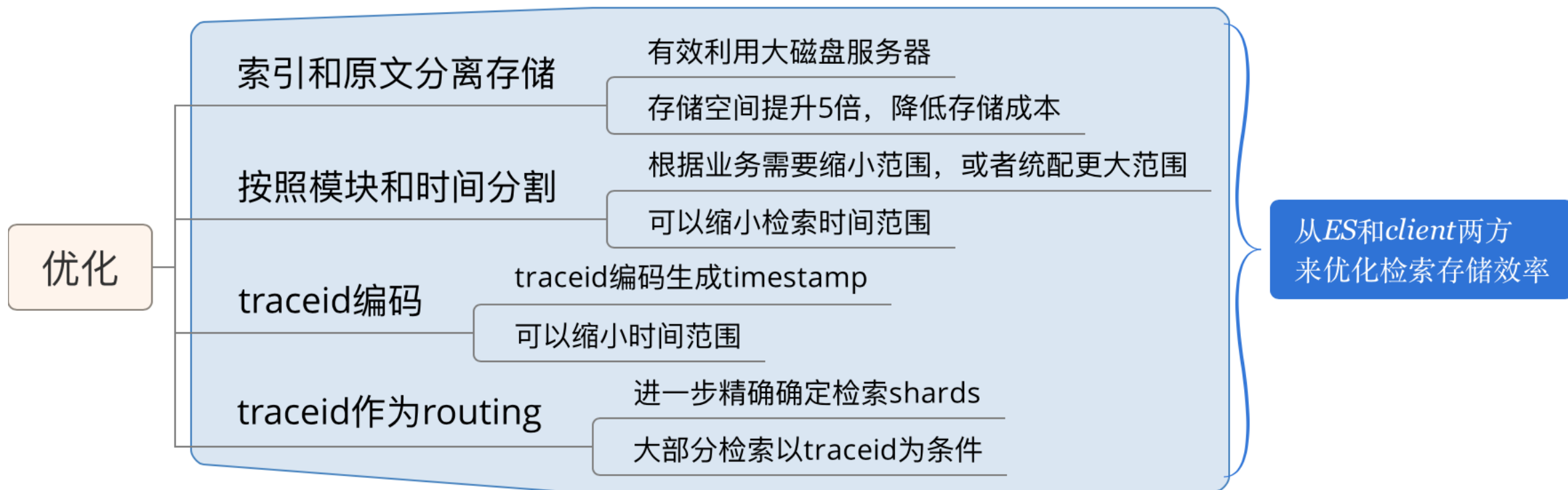
```
struct Trace {  
    1: required string traceid;    // 透传traceid  
    2: required string spanid;    // 透传spanid  
    3: optional i64 TCode;        // 分bit复用 用来标识不同的应用需求  
    4: optional string TContent;  // 和TCode配合使用。  
    // TContent结构体 用来保存需要透传的更多内容。  
}
```

```
TCode: 1  
TContent:{  
    "d_timeout":  
    {  
        "app_timeout_ms":20000,  
        "timestamp":1490022725  
    }  
    "d_lang":"zh-CN",  
    "d_gate_launch":  
    {  
        "cityid":1  
    }  
    "d_sampling":  
    {  
        "code":1  
    }  
}
```

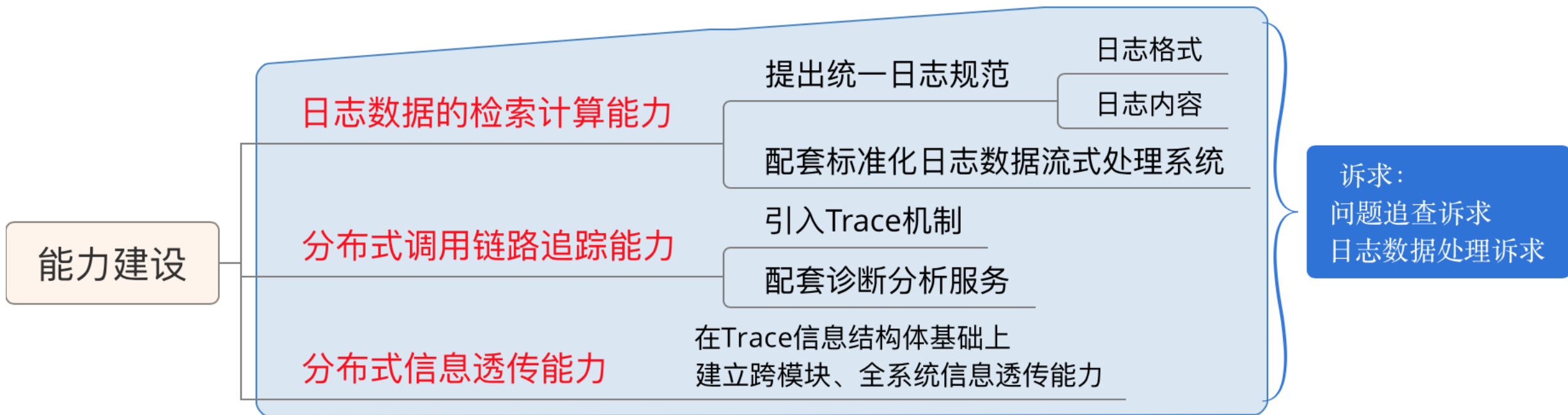


存储检索优化

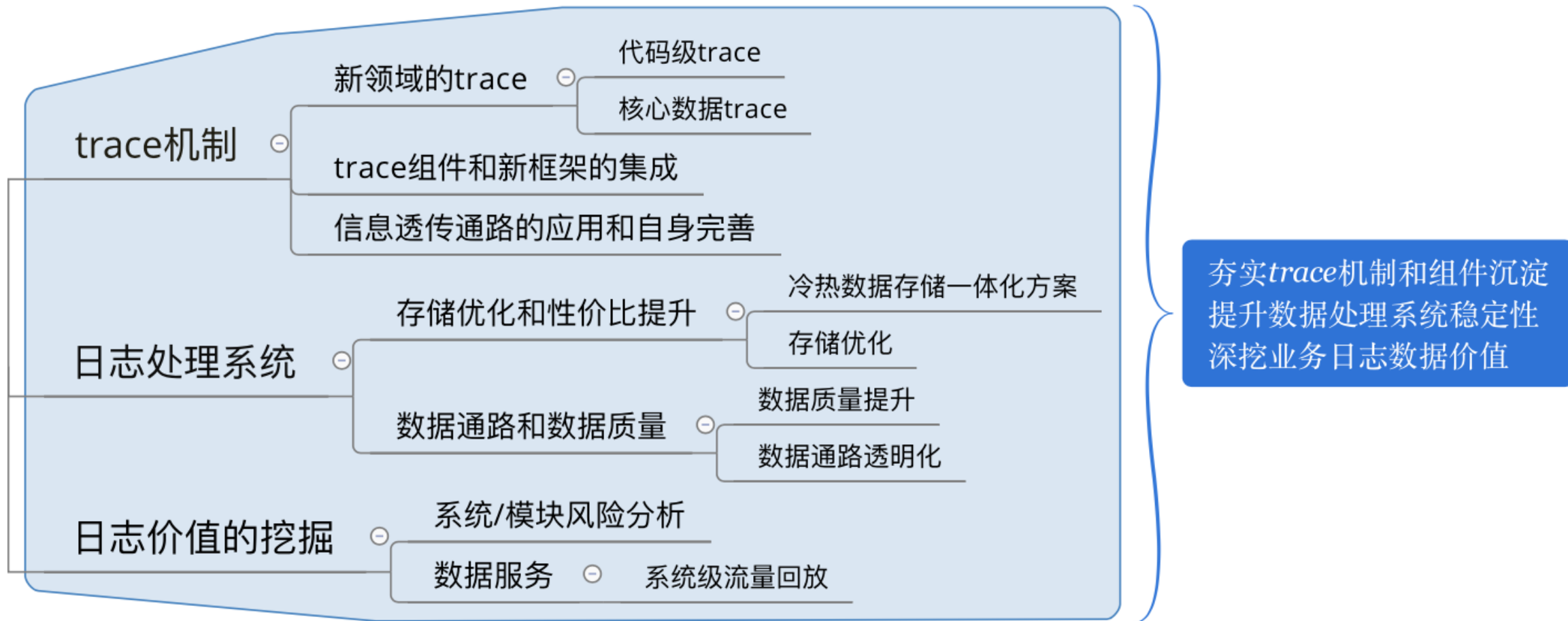
	存储容量	写入	读取
ElasticSearch	3.5TB	5万qps	mget优势
Hbase	5.8TB	6万qps	get优势
基于rocksdb自研存储	6.5TB		



能力建设



规划



Thanks