滴滴Elasticsearch Query DSL分析系统

基础平台部-葛晨鑫

滴滴
滴滴一下 美好出行

话题讨论

你们在使用ES过程中是否遇到过这样的问题?

- 查询耗时突增，ES节点上查看SlowLog?
- 如何拦截有危害的查询语句?
- 如何优化日志数据存储成本?
- 如何做到数据共享且每个应用方查询不相互影响?
- 如何应对用户查询量突增，导致ES CPU USE很高?

怎么解决这些问题呢

大纲

目录

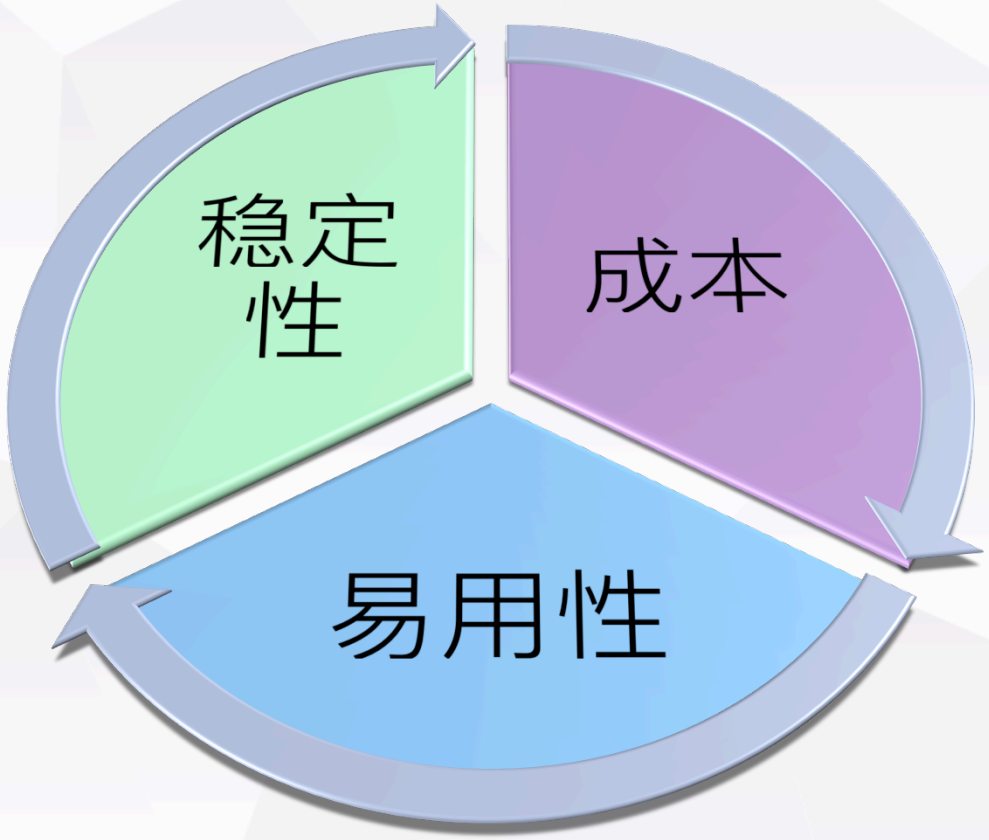
01

第一章

问题与挑战

问题与挑战

- Elasticsearch查询自我保护能力不足
- Elasticsearch用户多，公共索引查询语句不受控
- Elasticsearch默认对所有字段建索引
- Elasticsearch用户视角查询信息不完善



问题分解

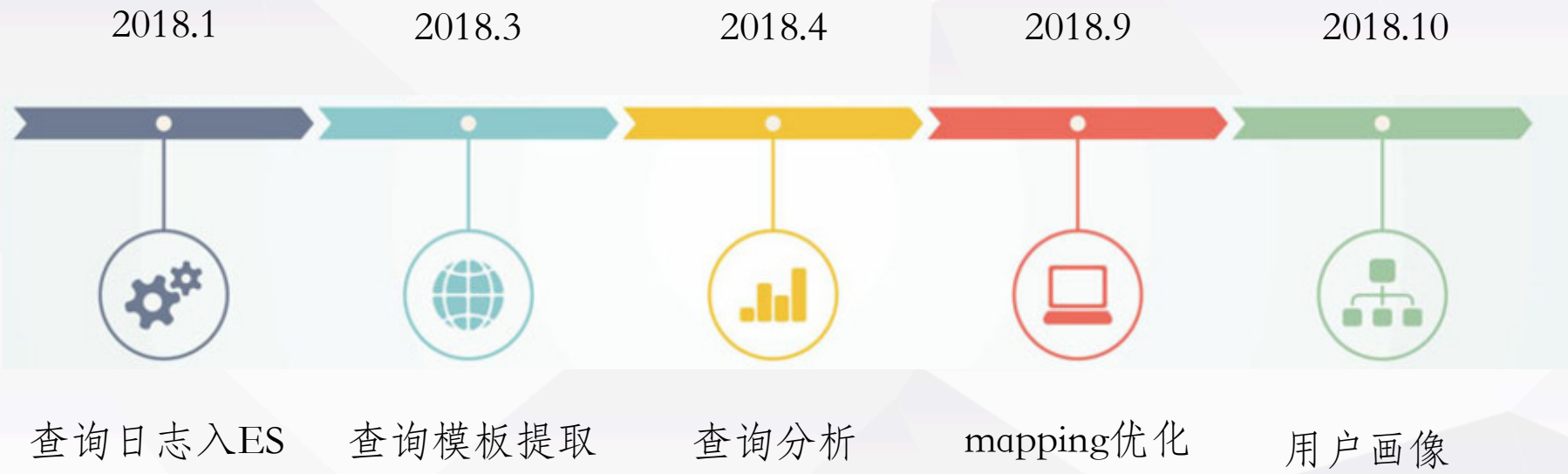

重要索引查询语句审核
合理限流
查询故障快速定位
查询语句抽象化
精细化索引mapping
分析用户查询特点
查询维度数据聚合统计

02

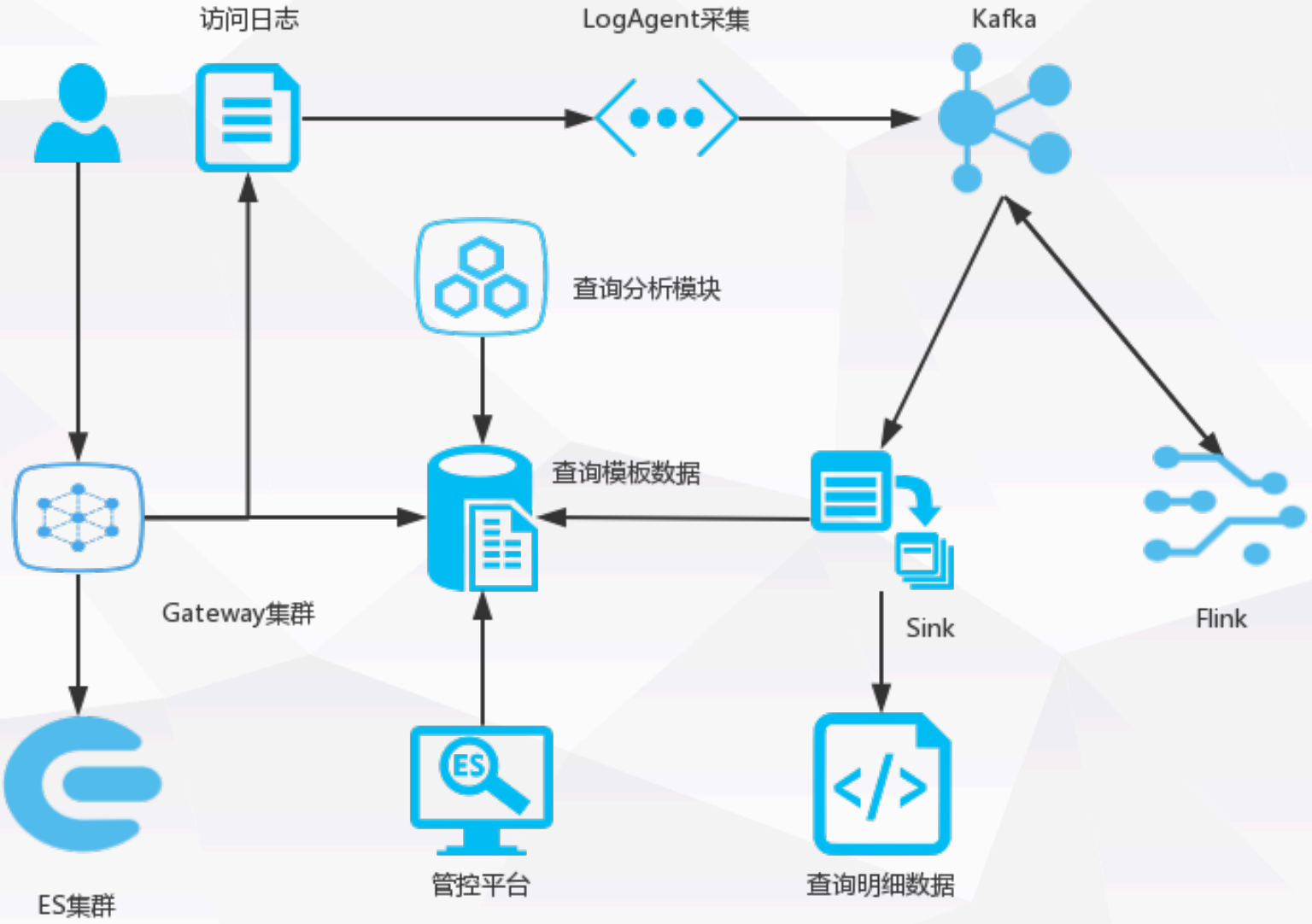
第二章

架构与收益

架构实现：发展历程

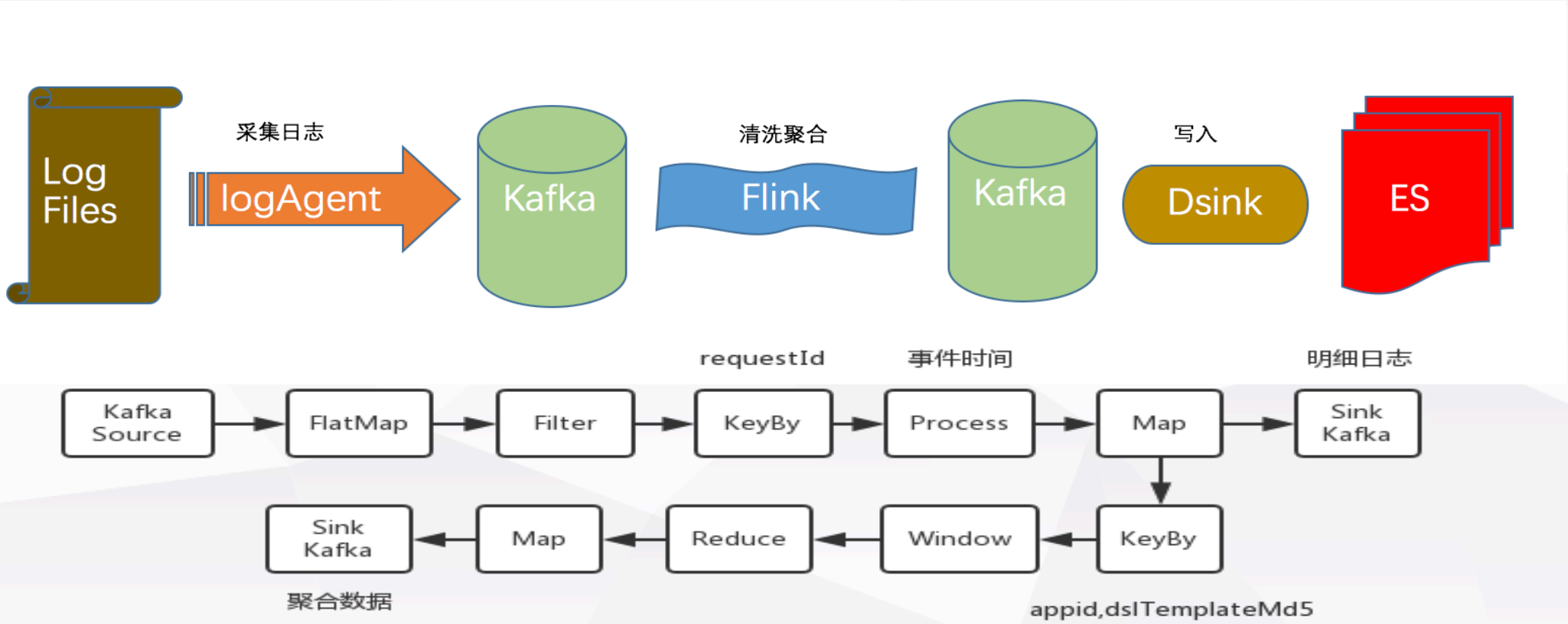

2018.1
2018.3
2018.4
2018.9
2018.10
查询日志入ES
查询模板提取
查询分析
mapping优化
用户画像

架构实现：架构图


访问日志
LogAgent采集
Kafka
查询分析模块
查询模板数据
Gateway集群
Sink
Flink
ES集群
管控平台
查询明细数据

架构实现：查询日志入ES

用户访问gateway进行查询，gateway会记录用户一次查询行为并写入本地日志文件中，需要将用户一次查询产生的多条日志，关联聚合成一条访问记录写入ES。

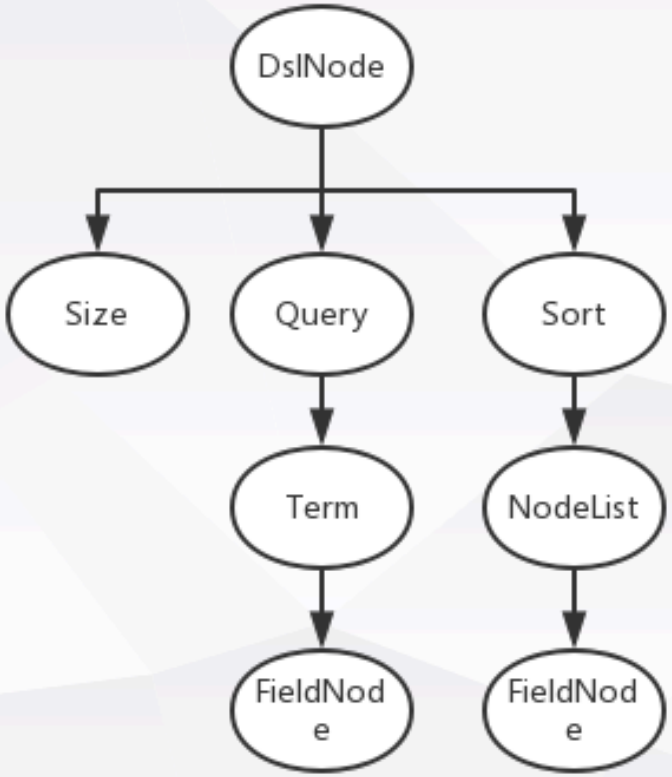


发展历程：DSL查询模板提取

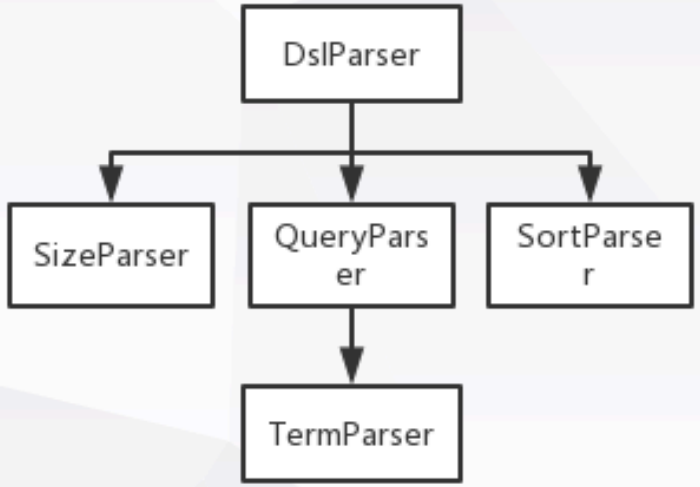
查询语句

```
{
  "size": 10,
  "query": {
    "term": {
      "traceId": {
        "value":
"123"
      }
    }
  },
  "sort": [
    {
      "logTime": {
        "order":
"desc"
      }
    }
  ]
}
```

DSL语法树



DSL解析器



查询模板

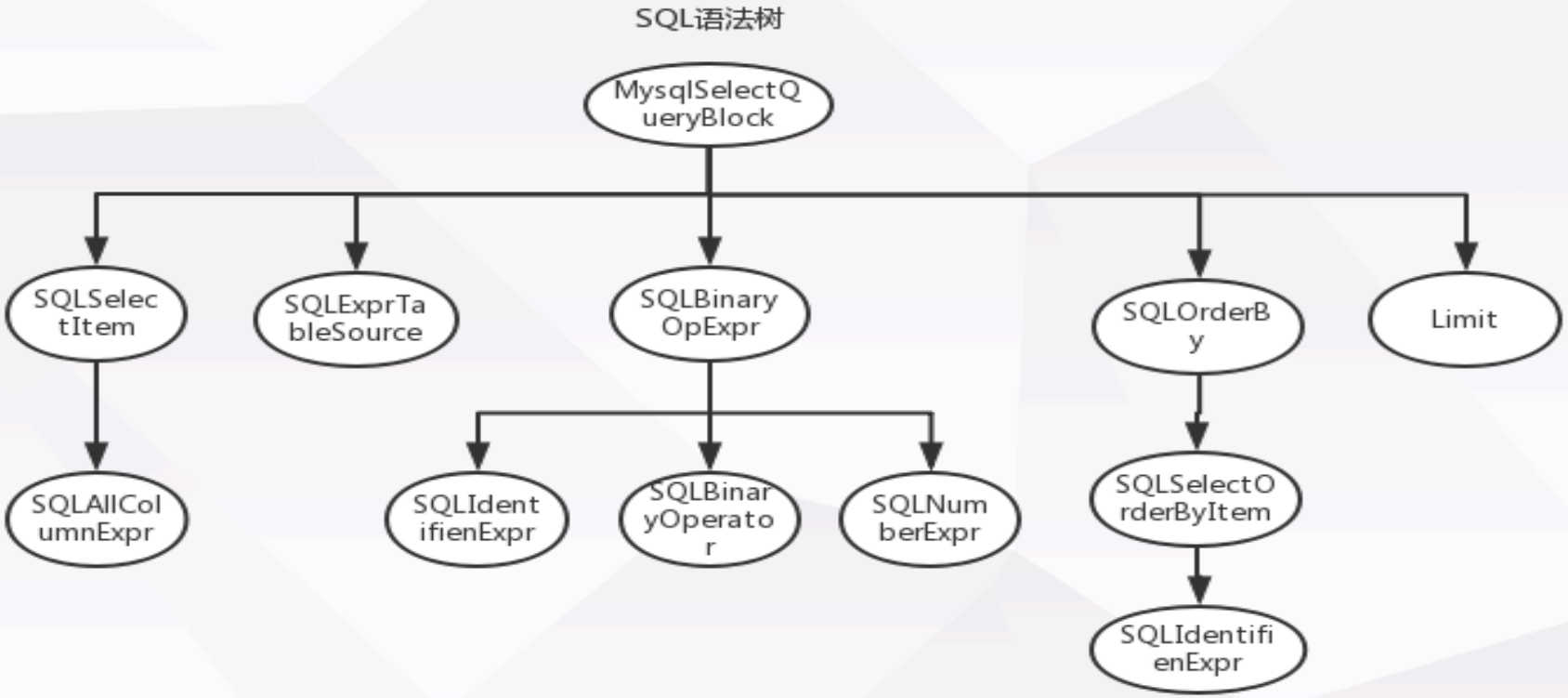
```
{
  "size": ?,
  "query": {
    "term": {
      "traceId": {
        "value": ?
      }
    }
  },
  "sort": [
    {
      "logTime": {
        "order": ?
      }
    }
  ]
}
```

字段提取

过滤字段:
traceId

排序字段:
logTime

发展历程：SQL查询模板提取

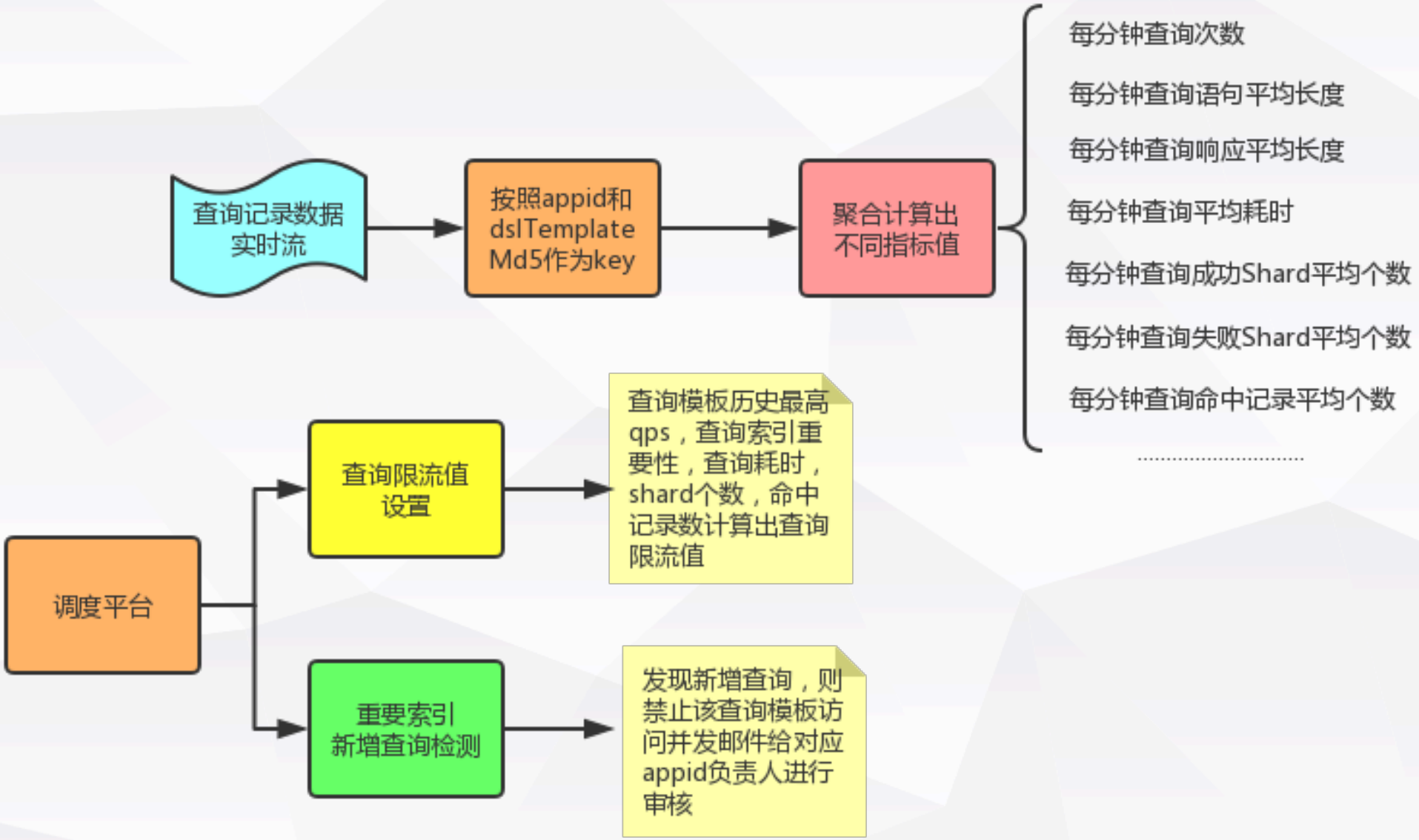



```
select * from table where traceId=123 order by logTime limit 10
```

⇩

```
SELECT ? FROM ? WHERE traceId=? ORDER BY logTime LIMIT ?
```

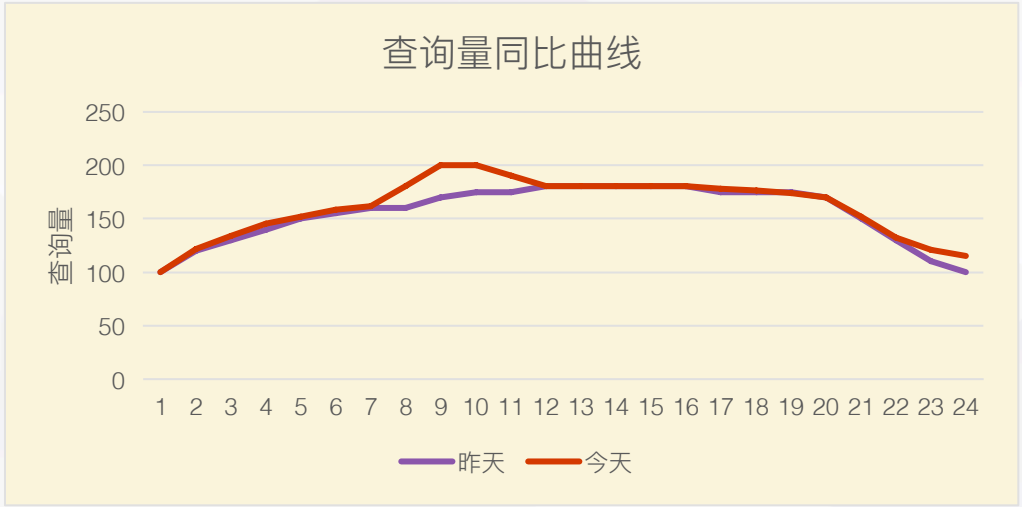
架构实现：查询分析模块

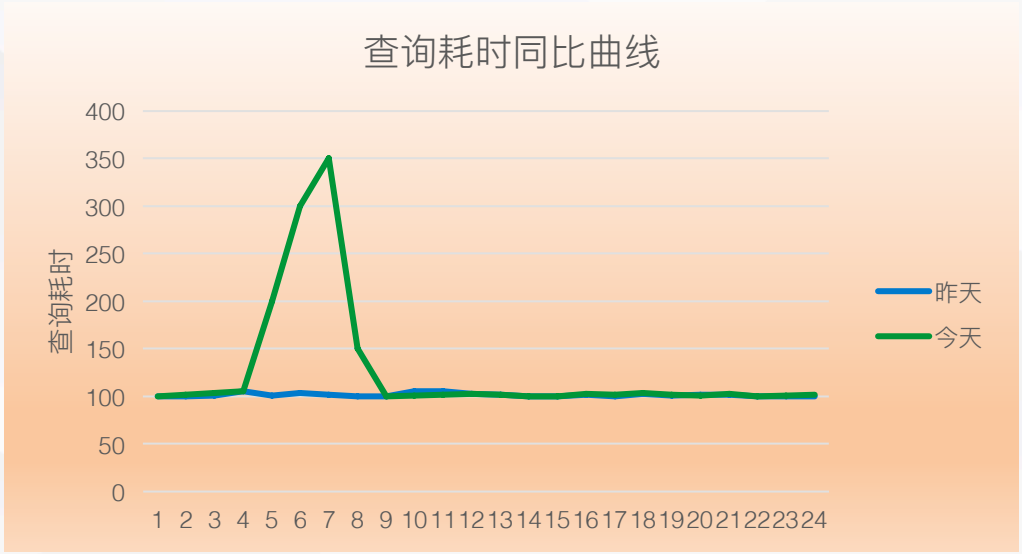

查询记录数据
实时流
按照appid和
dslTemplate
Md5作为key
聚合计算出
不同指标值
每分钟查询次数
每分钟查询语句平均长度
每分钟查询响应平均长度
每分钟查询平均耗时
每分钟查询成功Shard平均个数
每分钟查询失败Shard平均个数
每分钟查询命中记录平均个数
.....
调度平台
查询限流值
设置
查询模板历史最高qps，查询索引重要性，查询耗时，shard个数，命中记录数计算出查询限流值
重要索引
新增查询检测
发现新增查询，则禁止该查询模板访问并发邮件给对应appid负责人进行审核

架构实现：查询问题定位工具

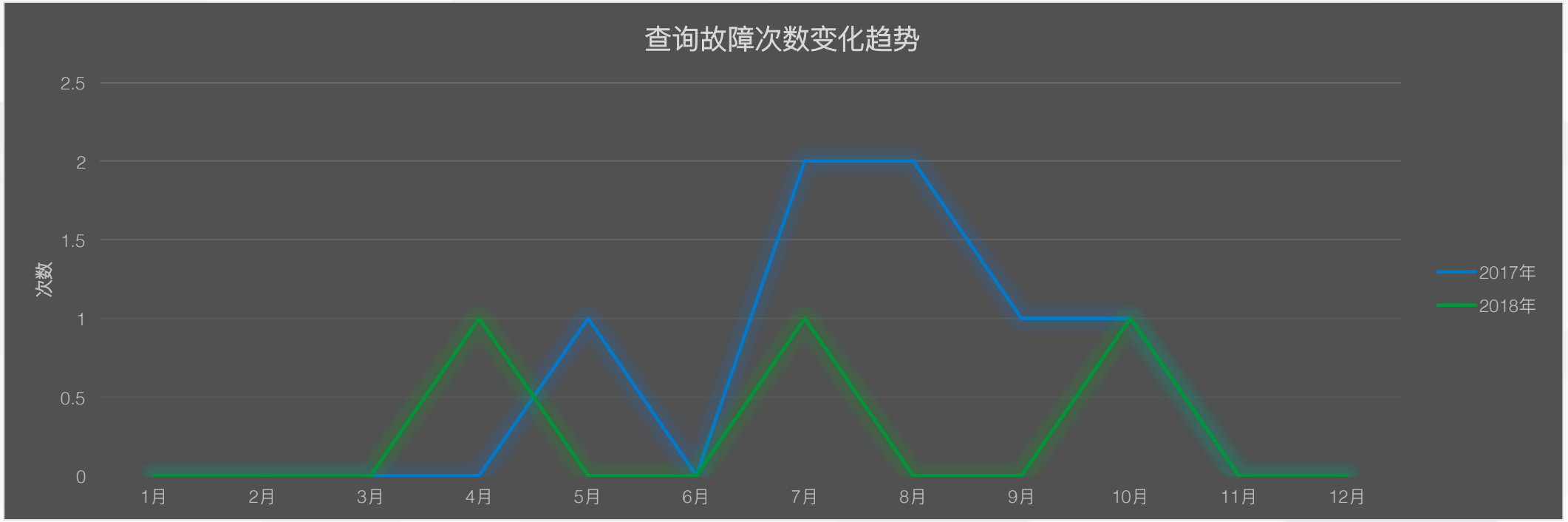
故障发生时刻前一小时：

1. 和昨天同一时间段对比，新增的DSL
2. 和昨天同一时间段对比，查询量有明显增加
3. 和昨天同一时间段对比，查询耗时有明显增加



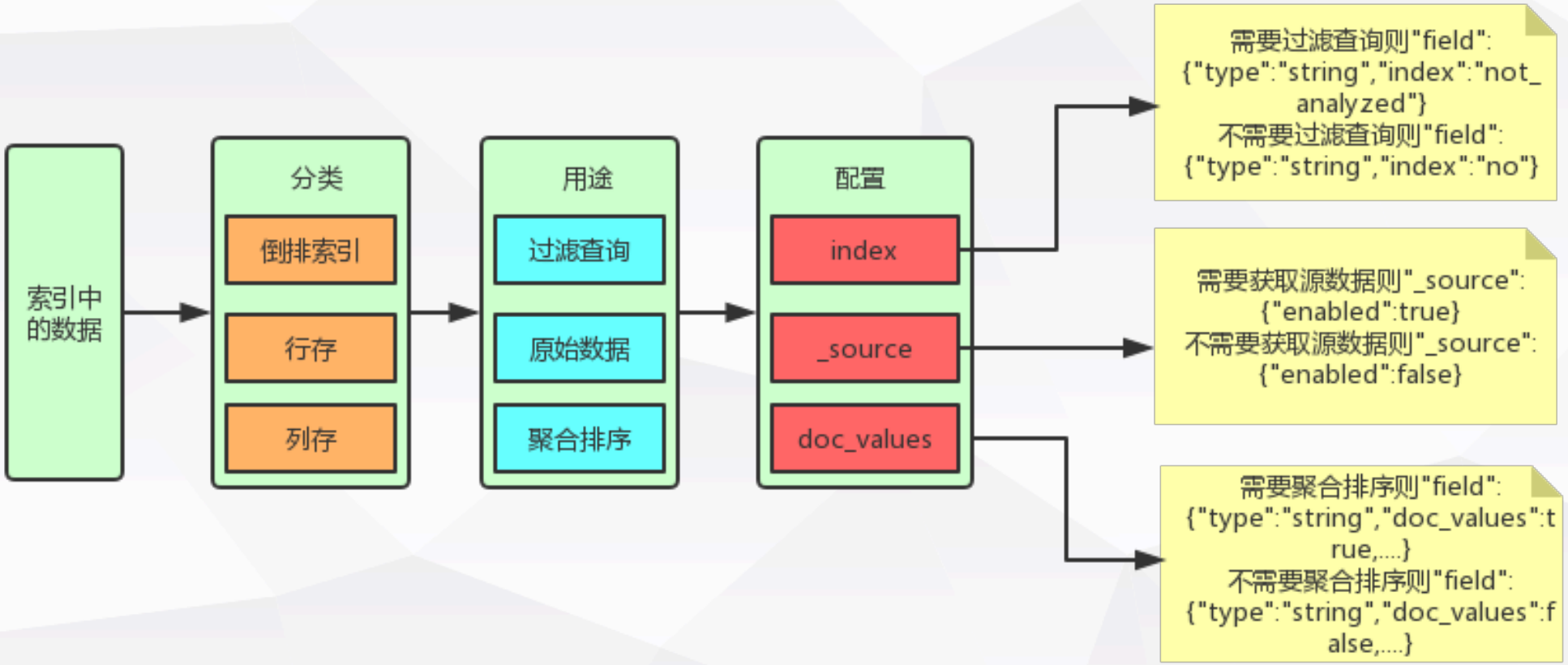


收益：稳定性

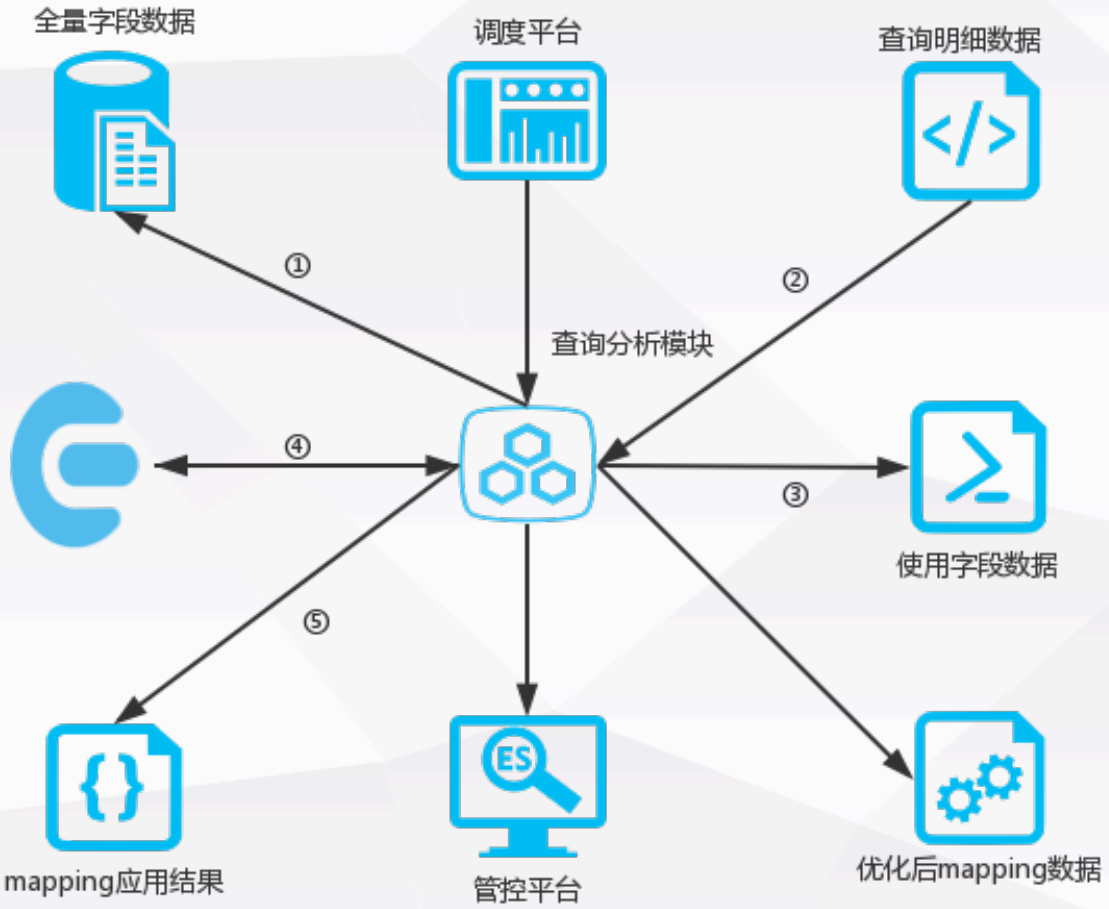


年份	问题定位手段	问题定位耗时
2017	以人工定位为主	半小时以上
2018	人工定位及查询问题定位工具辅助	10分钟左右

架构实现：mapping优化



架构实现：mapping优化



① 从es集群索引mapping中获取全量字段并回写es索引A

② 从查询明细数据中分析使用字段数据并回写es索引B

③ 从索引模板全量字段和索引模板使用字段分析出不使用字段，优化索引mapping并回写es索引C

④ 根据索引模板mapping优化启用标识，使用新老mapping双写部分数据回放查询语句进行验证，通过则将新mapping同步更新到es集群

⑤ 将优化使用的mapping及日期并回写es索引D

架构实现：mapping优化

优化前mapping

```
{
"clientHost": {
"ignore_above": 2048,
"index": "not_analyzed",
"type": "string
},
"role": {
"ignore_above": 2048,
"index": "not_analyzed",
"type": "string"
},
"traceid": {
"index": "not_analyzed",
"type": "string"
},
.....
}
```

优化后mapping

```
{
"clientHost": {
"ignore_above": 2048,
"index": "no",
"type": "string",
"doc_values": false
},
"role": {
"ignore_above": 2048,
"index": "no",
"type": "string",
"doc_values": false
},
"traceid": {
"index": "not_analyzed",
"type": "string"
},
.....
}
```

优化前Lucene文件(993M)

行存:
736M _21.fdt
350K _21.fdx
1.3K _21.fnm

倒排索引:
73M _21_Lucene50_0.doc
90M _21_Lucene50_0.tim
2.4M _21_Lucene50_0.tip

列存:
92M _21_Lucene54_0.dvd
407 _21_Lucene54_0.dvm

优化后Lucene文件(787M)

行存:
736M _s.fdt
350K _s.fdx
1.1K _s.fnm

倒排索引:
6.4M _s_Lucene50_0.doc
22M _s_Lucene50_0.tim
945K _s_Lucene50_0.tip

列存:
23M _s_Lucene54_0.dvd
400 _s_Lucene54_0.dvm

收益：成本优化

日志集群索引mapping优化，目前节省磁盘空间440TB。

架构实现：用户画像功能

分析用户正常查询(查询总次数、查询qps、查询耗时分位图)情况，慢查情况，异常查询情况。

一、正常查询

本周查询总数:**738553169** (上周723208992) ;同比增长2.12%

本周最大qps:**7369** (上周4491) ;同比增长64.08%

指标	2018-10-29	2018-10-30	2018-10-31	2018-11-01	2018-11-02	2018-11-03	2018-11-04
查询总次数	109676755	109534615	110070664	110391599	113273073	76647462	108959001
查询qps	最大:3909; 平均:1269;	最大:4105; 平均:1267;	最大:4418; 平均:1273;	最大:4755; 平均:1277;	最大:7369; 平均:1313;	最大:2419; 平均:887;	最大:3999; 平均:1261;
查询耗时分位	50分位:6.00ms; 75分 位:13.00ms; 95分 位:42.27ms; 99分 位:135.88ms	50分位:6.00ms; 75分 位:10.00ms; 95分 位:34.00ms; 99分 位:107.47ms	50分位:6.00ms; 75分位:9.44ms; 95分 位:33.00ms; 99分位:94.92ms	50分位:5.00ms; 75分位:9.00ms; 95分 位:30.00ms; 99分位:63.66ms	50分位:5.00ms; 75分位:7.00ms; 95分 位:25.00ms; 99分位:69.23ms	50分位:5.00ms; 75分位:7.00ms; 95分 位:20.00ms; 99分位:49.38ms	50分位:6.00ms; 75分位:9.00ms; 95分 位:28.00ms; 99分位:63.38ms

二、慢查

本周慢查总数:**28608** (上周17855) ;同比增长60.22%

指标	2018-10-29	2018-10-30	2018-10-31	2018-11-01	2018-11-02	2018-11-03	2018-11-04
慢查总次数	5617	5271	4914	2387	8349	799	1271

三、异常查询

本周异常查询总数:**4346** (上周98) ;同比增长4334.69% 异常处理手册

指标	2018-10-29	2018-10-30	2018-10-31	2018-11-01	2018-11-02	2018-11-03	2018-11-04
异常总次数	4329	4	1	5	5	1	1

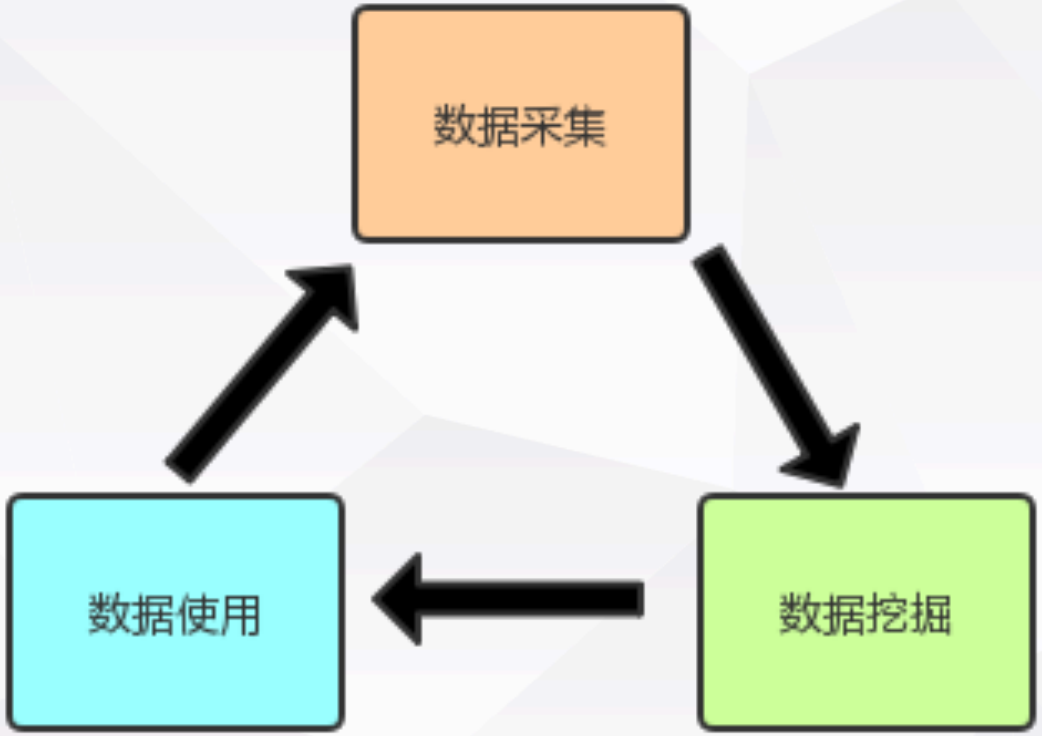
收益：易用性


Discover
Visualize
Dashboard
Settings
Last 7 days
appid画像
appid:2
appid...
Dashboards:
用户查询画像
慢查异常查询详情
appid每天查询总次数
查询总次数
查询总次数
6000000
4000000
2000000
0
2018-10-11
2018-10-15
2018-10-19
2018-10-23
2018-10-27
date per day
appid每天查询分位图
查询50分位(ms)
查询75分位(ms)
查询95分位(ms)
查询99分位(ms)
600
400
200
0
appid查询qps
最大qps
最小qps
平均qps
1500
1000
500
0
appid每天慢查总次数
慢查总次数
40000
30000
20000
10000
0
appid每天异常总次数
异常总次数
80000
60000
40000
20000
0

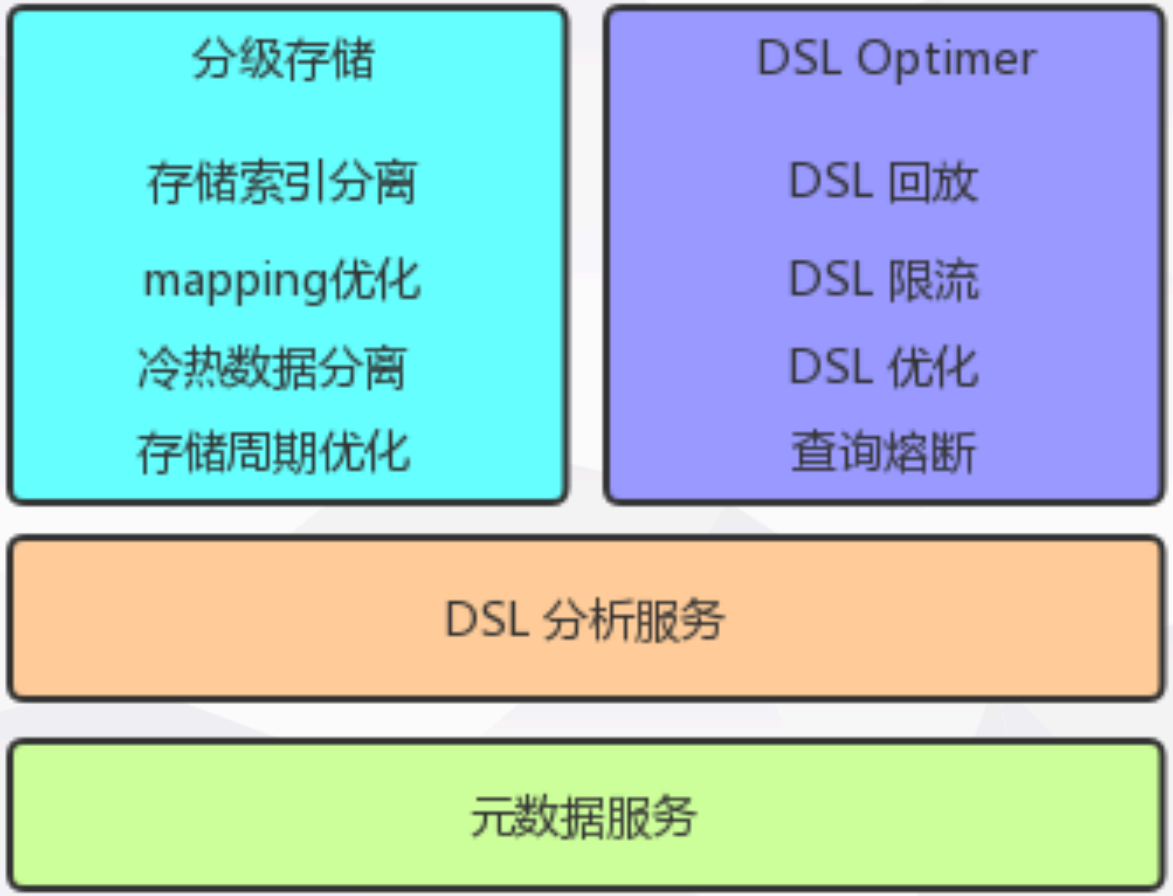
03 第三章

总结与规划

总结



规划


分级存储
存储索引分离
mapping优化
冷热数据分离
存储周期优化
DSL Optimer
DSL 回放
DSL 限流
DSL 优化
查询熔断
DSL 分析服务
元数据服务

问答

Thank you !

elastic
中文社区

专业、垂直、纯粹的 Elastic 开源技术交流社区
https://elasticsearch.cn/