

基于Kibana和ES的苏宁实时日志分析平台

2017-08-14 苏宁云商IT总部 Elasticsearch研究会

苏宁日志平台架构演进



基于Kibana和ES的苏宁实时日志分析平台

苏宁云商IT总部技术总监-彭燕卿
2016.11.21

Elastic{ON} DevChina – Dec 10, 2016

 Elasticsearch研究会

Agenda

- 集群现状
- 日志平台架构演进
- 日常优化总结
- 运维小技巧
- Kibana4二次开发

2

 Elasticsearch研究会 

集群现状

- 126个数据节点，7个cluster，12C/128G/2T SATA、16C/128G/3T SSD、12C/128G/16T
- 接入苏宁近2000个系统的应用日志、web访问、缓存、应用防火墙等日志
- 大促每天新索引25T数据，doc数超过450亿条
- open 1100索引、130T、2500亿数据、20000 shard、7天存储
- 峰值90W/s数据写入
- 平均每个doc 0.6kb

3

Elasticsearch 研究会
elastic

整体架构

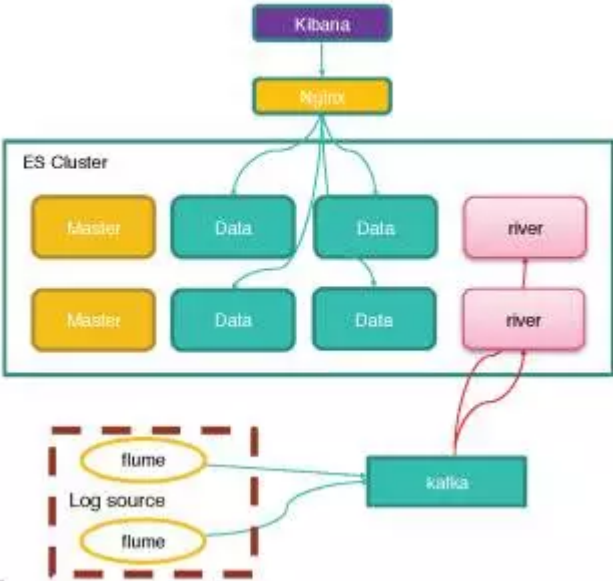
- 实时日志平台采用的是flume+Kafka+Elasticsearch+Kibana的部署架构。
- 与ELK架构有点不同，我司使用flume实时采集日志，Kafka作为数据通道，ES river插件消费Kafka里面的数据，将Kafka中的数据清洗过滤后，index到ES集群中。



4

Elasticsearch 研究会
elastic

日志平台架构演进-①



配置:

- 虚拟机节点
- 按天生成索引

问题:

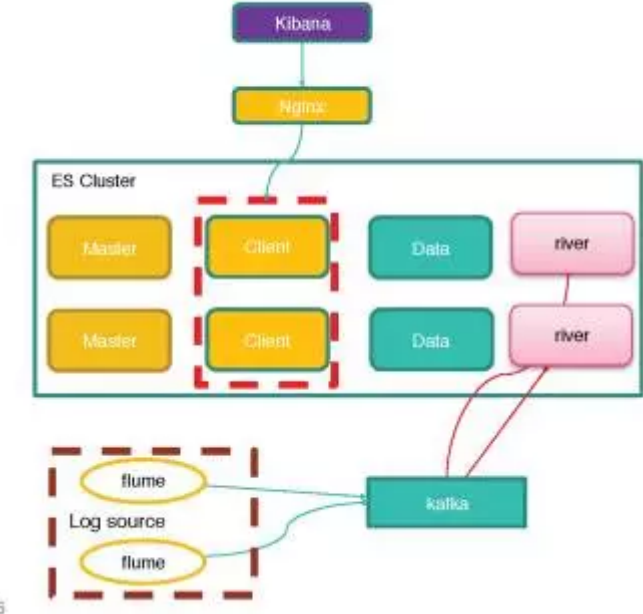
- 部分数据节点负载非常高
- 数据消费延时
- 查询响应慢
- QPS低

原因:

- 同一个主机上存在多个ES 虚拟机Node
- 数据节点同时承担索引和检索, 负荷重
- enabled_all
- 按天索引体量大
- 集群节点少



日志平台架构演进-②



主要优化:

- 增加client节点
- 解决同一物理机上多虚拟机data节点
- 增加部分物理机
- 关闭_all字段,
- 小时生成索引
- 根据日志类型划分不同的索引

运行状况:

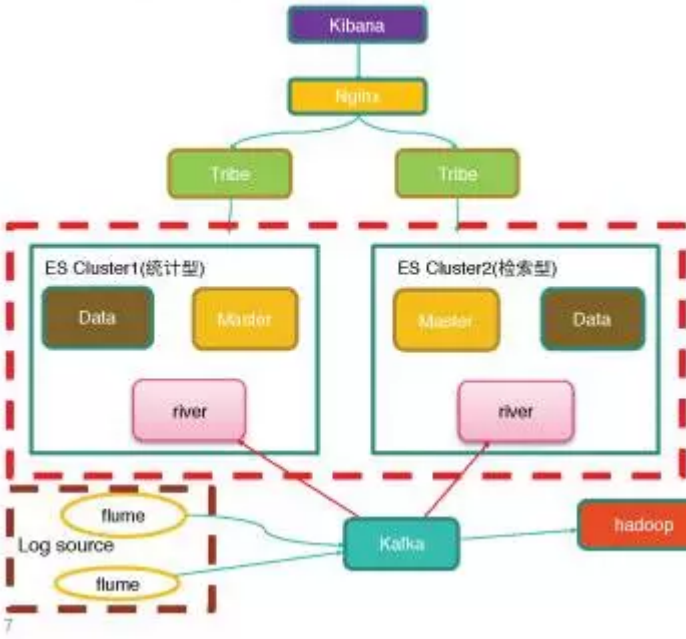
- 非大促期间, 索引和检索速度基本能达到秒级
- 大促期间, 日志量膨胀以及访问人数过多时, 索引和检索速度任然很慢

原因分析:

- 单集群能力受限
- 不同类型的数据混在一个大集群中, 互相影响。
- client 节点性能提升不明显
- 虚拟机和物理机混合, 制约物理机的能力



日志平台架构演进-③



主要优化:

- 使用tribe做多集群路由
- 根据不同的分析类型进行集群拆分
- 将data节点全部替换成物理机
- 提供按照系统、文件路径等应对大促期间的日志洪峰系统进行降级的功能
- 统计型集群使用SSD

运行状况:

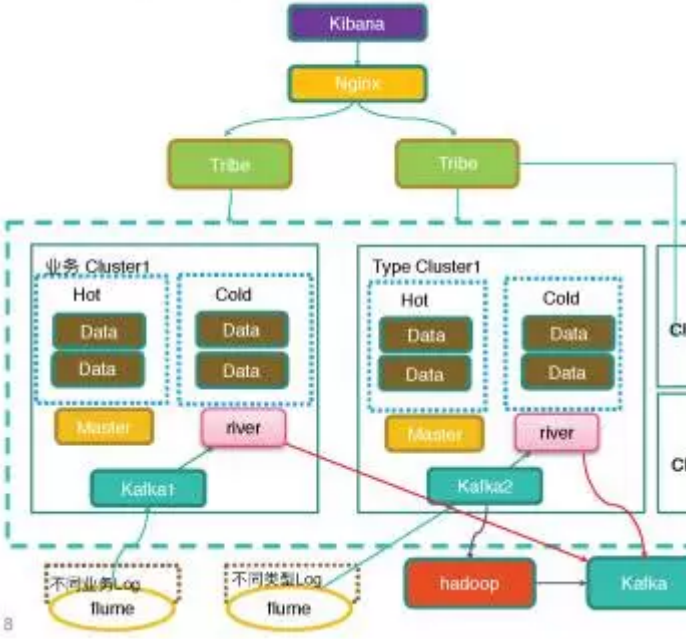
- 流量的剧增通过降级,能够保证核心系统数据的实时性

存在的问题:

- 随着业务量的快速增长,简单的按照分析类型进行集群拆分已不能满足要求,并且增加机器资源容量并不是线性增加。
- 用户希望保留7天以前以及大促期间数据,集群容量以及性能也存在严重问题



日志平台架构演进-现状



主要优化:

- 在③的基础上同时支持按照业务划分以及数据类型划分集群
- 接入流程自动化
- 使用不同的硬件配置划分Hot/Cold节点,业务低峰时将历史数据迁移到Cold节点
- 针对Exception等重点关注且长期保留分析的数据单独存储。
- 多kafka

问题:

- 跨多索引进行统计分析性能差以及对集群压力任然很大



日常优化总结-硬件

- 优先独立物理机
- 对于实时性要求非常高的需求, 优先SSD
- 适当调整OS的max_file_descriptors,解决Too many open files 异常
- 单服务器运行多个node时,调整max user processes, 否则容易native thread OOM.
- 关闭swap交换或锁内存 `ulimit -l unlimited/bootstrap.mlockall: true`

9



日常优化总结-ES

- 根据数据量合理的规划索引pattern和shard数
- disabled _all 节省存储空间、提升索引速度
- 不需要分词的字段设成 not_analyzed
- 对于不要求100%高可用的内部系统, 可不设置副本, 提升index速度和减少存储

10



日常优化总结-ES

- 设置合理的refresh时间

`index.refresh_interval: 300S`

- 设置合理的flush间隔

`index.translog.flush_threshold_size: 4g`

`index.translog.flush_threshold_ops: 50000`

- 合理配置throttling

`indices.store.throttle.max_bytes_per_sec: 200mb`

- 适当调整bulk队列

11 `threadpool.bulk.queue_size: 1000`



日常优化总结-ES

- 有时可能因为gc时间过长，导致该数据节点被主节点踢出集群的情况，导致集群出现不健康的状态。为了解决这样的问题，我们适当的调整ping参数。(master)

`discovery.zen.fd.ping_timeout: 40s`

`discovery.zen.fd.ping_interval: 5s`

`discovery.zen.fd.ping_retries: 5`

- 调整数据节点的JVM新生代大小

数据节点young gc频繁,适当调转新生代大小 (-Xmn3g)，降低young gc的频率。

- 在进行检索和聚合操作时，ES会读取反向索引，并进行反向解析，然后进行排序，将结果保存在内存中。这个处理会消耗很多Heap，有必要进行限制，不然会很容易出现OOM。

Disabled analyzed field fielddata

限制Field Data的Heap Size的使用

`indices.fielddata.cache.size: 40%`

12 `indices.breaker.fielddata.limit: 50%`



ES运维小技巧

增加节点

- 调整shard数

```
index.routing.allocation.total_shards_per_node: 2
```

index在每个node的shard数据

(如果后期需要移除节点,保证每个node有可分配的shard)

移除节点

- 移除node前可以先exclude要移除的node

```
cluster: cluster.routing.allocation.exclude._name: node1
```

```
index: index.routing.allocation.exclude._name: node1
```

```
index.routing.allocation.require.node_type: hot
```

以上参数可根据_ip、_host等来进行配置

以上参数可实现hot-cold cold数据的自动迁移

13

Elasticsearch 研究会
elastic

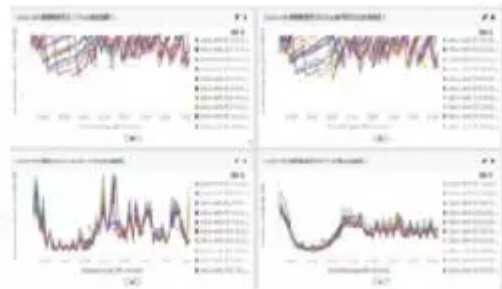
ES运维小技巧-工具

- 监控插件满天飞,各有千秋

Head、Kopf、bigdesk、elasticsearch-sql(NLPchina)

- 定时关闭和删除index: curator

- 基于python脚本实现采集ES集群指标数据,并使用kibana展示



- 日志平台自监控ES日志,重点对以下关键词进行监控告警

- OutOfMemoryError
- removed AND cluster.service node 退出cluster
- unable to create new native thread
- master_left master 退出cluster

- Slow log监控,重量级query可能把整个集群拖慢,对slow log重点监控分析(待实现)

14

Elasticsearch 研究会
elastic

Kibana二次开发-从汉化开始



15

Kibana二次开发-权限

- nodejs实现cas单点登录
- 授权数据权限
- 不同业务可查询的数据范围
- 禁用通过kibana访问_plugin、_shutdown等
- 用户关联 仪表盘、检索、统计分析

```
app.use('/', rat.hausman, rtaa.cyclelist, rtaa.cyclelist);
```

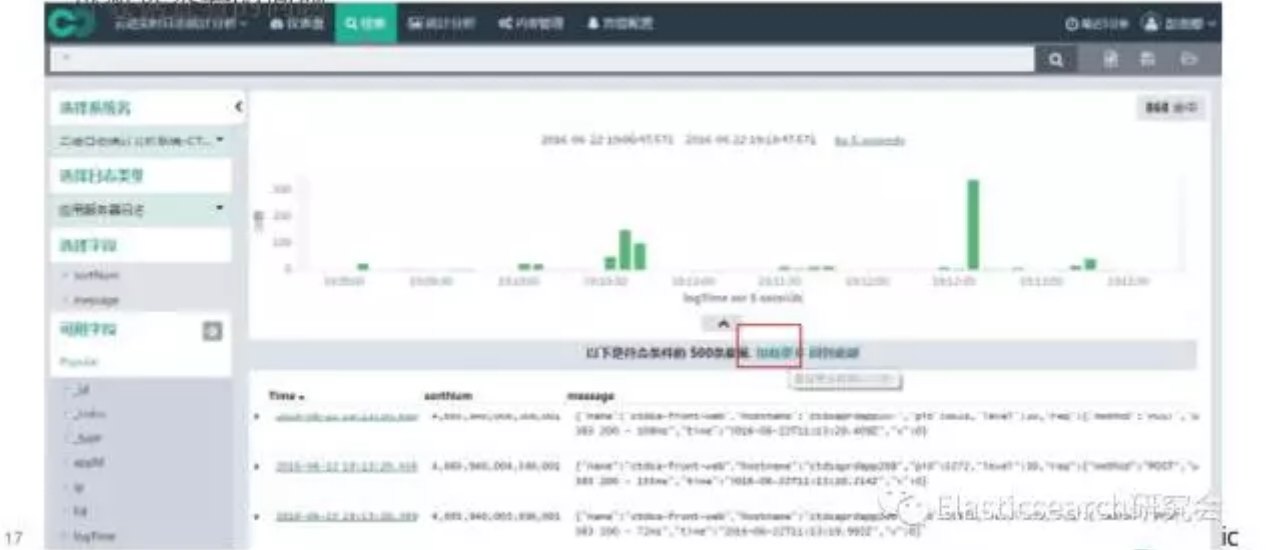


16

Kibana二次开发-discover可查询更多

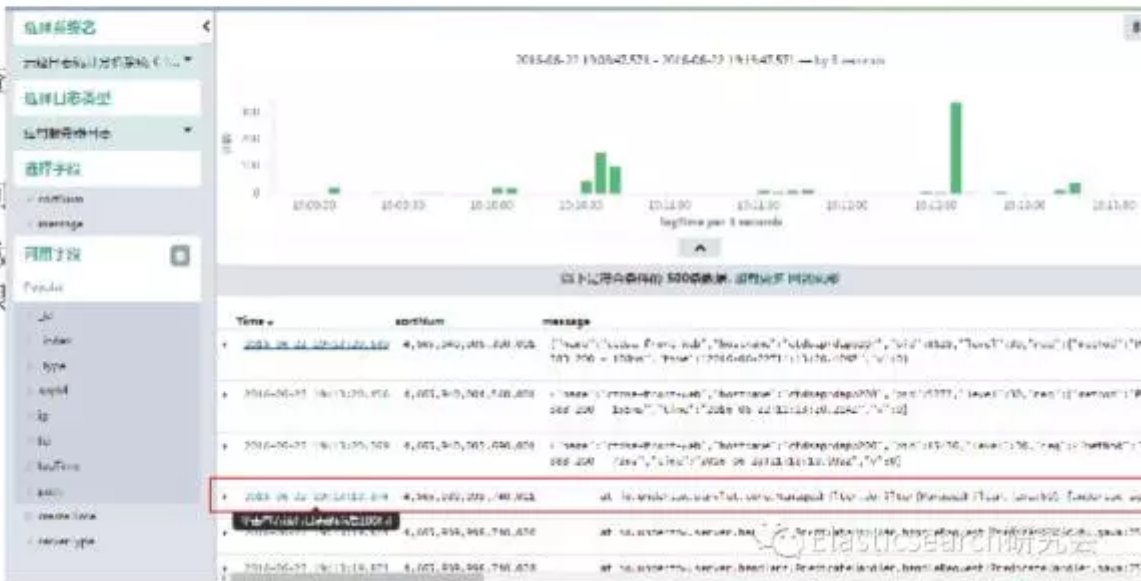
- Discover可以前后查询更多数据，解决查询discover:sampleSize行数限制

外数据本看的问题



Kibana二次开发-数据上下文查询

- 可查
- 时间
- 选取
- 序会很



Kibana二次开发-禁用check es version

- Kibana

- 当使用
得不到

- 可去掉

```

7  var esClient = new ElasticClient({
8    location: 'http://elasticsearch:9200'
9  });
10
11  return esClient.transport().request('GET', '/_cluster/health?pretty');
12
13  var esClient = new ElasticClient({
14    location: 'http://elasticsearch:9200'
15  });
16
17  // remove client nodes (logstash)
18  // remove client nodes (logstash)
19  // remove client nodes (logstash)
20  // remove client nodes (logstash)
21  // remove client nodes (logstash)
22  // remove client nodes (logstash)
23  // remove client nodes (logstash)
24  // remove client nodes (logstash)
25  // remove client nodes (logstash)
26  // remove client nodes (logstash)
27  // remove client nodes (logstash)
28  // remove client nodes (logstash)
29  // remove client nodes (logstash)
30  // remove client nodes (logstash)
31  // remove client nodes (logstash)
32  // remove client nodes (logstash)
33  // remove client nodes (logstash)
34  // remove client nodes (logstash)
35  // remove client nodes (logstash)
36  // remove client nodes (logstash)
37  // remove client nodes (logstash)
38  // remove client nodes (logstash)
39  // remove client nodes (logstash)
40  // remove client nodes (logstash)
41  // remove client nodes (logstash)
42  // remove client nodes (logstash)
43  // remove client nodes (logstash)
44  // remove client nodes (logstash)
45  // remove client nodes (logstash)
46  // remove client nodes (logstash)
47  // remove client nodes (logstash)
48  // remove client nodes (logstash)
49  // remove client nodes (logstash)
50  // remove client nodes (logstash)
51  // remove client nodes (logstash)
52  // remove client nodes (logstash)
53  // remove client nodes (logstash)
54  // remove client nodes (logstash)
55  // remove client nodes (logstash)
56  // remove client nodes (logstash)
57  // remove client nodes (logstash)
58  // remove client nodes (logstash)
59  // remove client nodes (logstash)
60  // remove client nodes (logstash)
61  // remove client nodes (logstash)
62  // remove client nodes (logstash)
63  // remove client nodes (logstash)
64  // remove client nodes (logstash)
65  // remove client nodes (logstash)
66  // remove client nodes (logstash)
67  // remove client nodes (logstash)
68  // remove client nodes (logstash)
69  // remove client nodes (logstash)
70  // remove client nodes (logstash)
71  // remove client nodes (logstash)
72  // remove client nodes (logstash)
73  // remove client nodes (logstash)
74  // remove client nodes (logstash)
75  // remove client nodes (logstash)
76  // remove client nodes (logstash)
77  // remove client nodes (logstash)
78  // remove client nodes (logstash)
79  // remove client nodes (logstash)
80  // remove client nodes (logstash)
81  // remove client nodes (logstash)
82  // remove client nodes (logstash)
83  // remove client nodes (logstash)
84  // remove client nodes (logstash)
85  // remove client nodes (logstash)
86  // remove client nodes (logstash)
87  // remove client nodes (logstash)
88  // remove client nodes (logstash)
89  // remove client nodes (logstash)
90  // remove client nodes (logstash)
91  // remove client nodes (logstash)
92  // remove client nodes (logstash)
93  // remove client nodes (logstash)
94  // remove client nodes (logstash)
95  // remove client nodes (logstash)
96  // remove client nodes (logstash)
97  // remove client nodes (logstash)
98  // remove client nodes (logstash)
99  // remove client nodes (logstash)
100 // remove client nodes (logstash)

```

Elasticsearch研究会

19

Kibana二次开发-dash-board

- 管理员可分享dashboard模板给普通用户使用

- dashboard可选择系统查看数据



Elasticsearch研究会

20