

阿里云Elasticsearch架构设计与运维实践

吴楠 | 阿里云 Elasticsearch 云服务





集群规模

3000+

集群数量

15000+

节点数

5PB+

数据量

性能

环境:

Elasticsearch版本: 5.5.3, 3节点

规格: 2核4G、4核16G、16核64G

磁盘: 1T SSD云盘

压测工具:

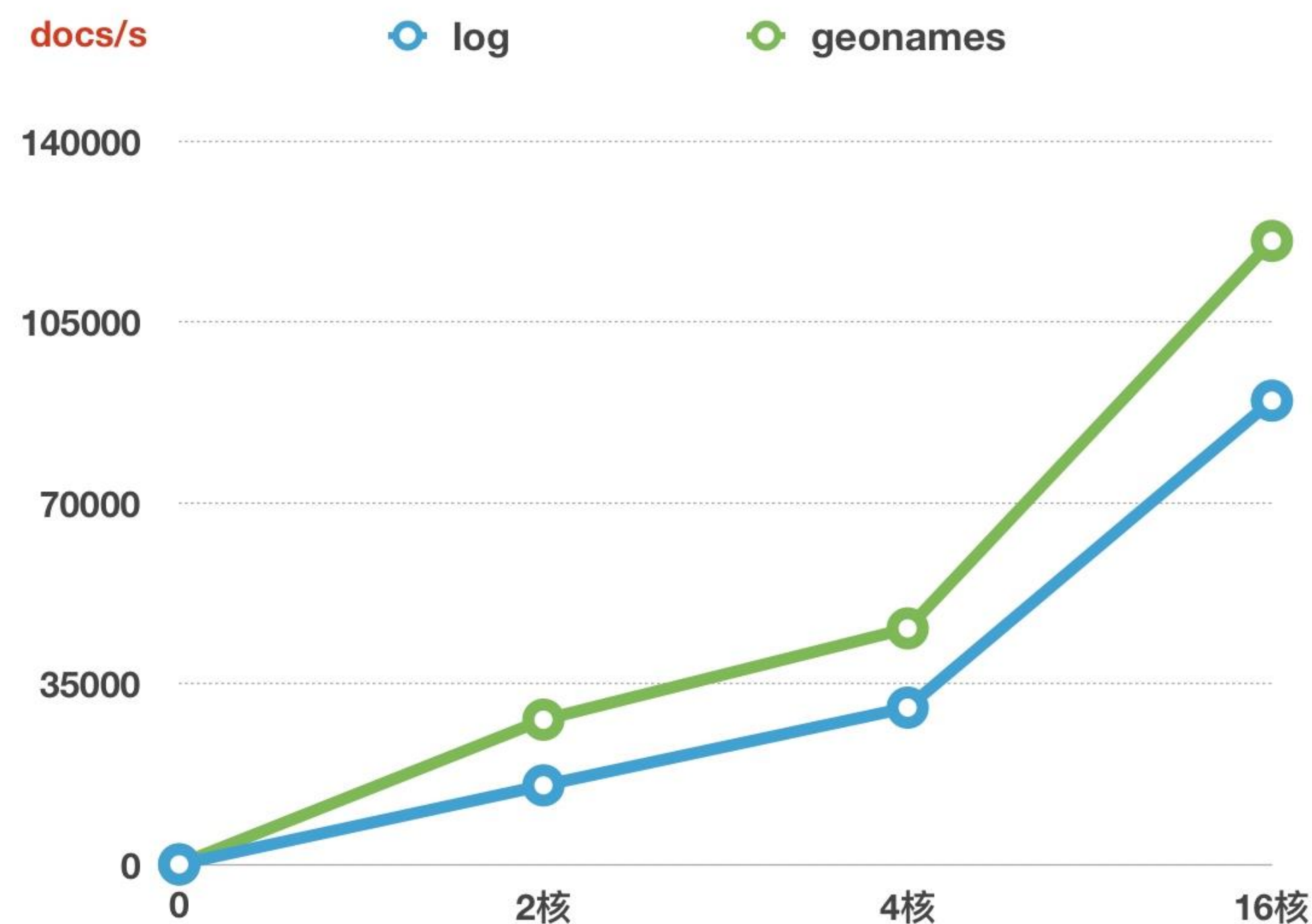
esrally

rest api

压测数据集:

esrally官方数据 geonames 3.3 GB, 单doc 311B

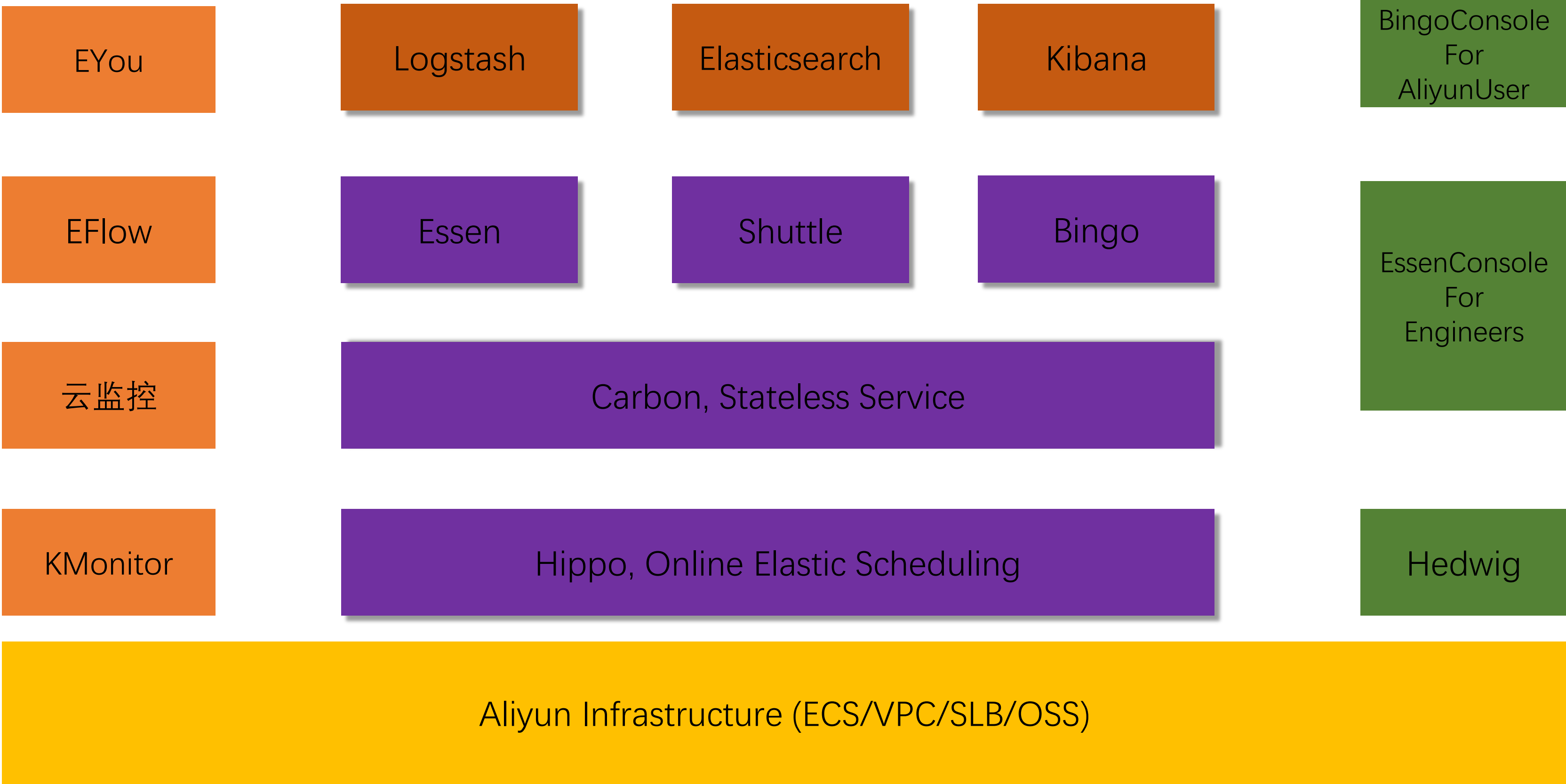
模拟某业务日志数据 80GB, 单doc 1432B



运维Elasticsearch集群曾经的痛

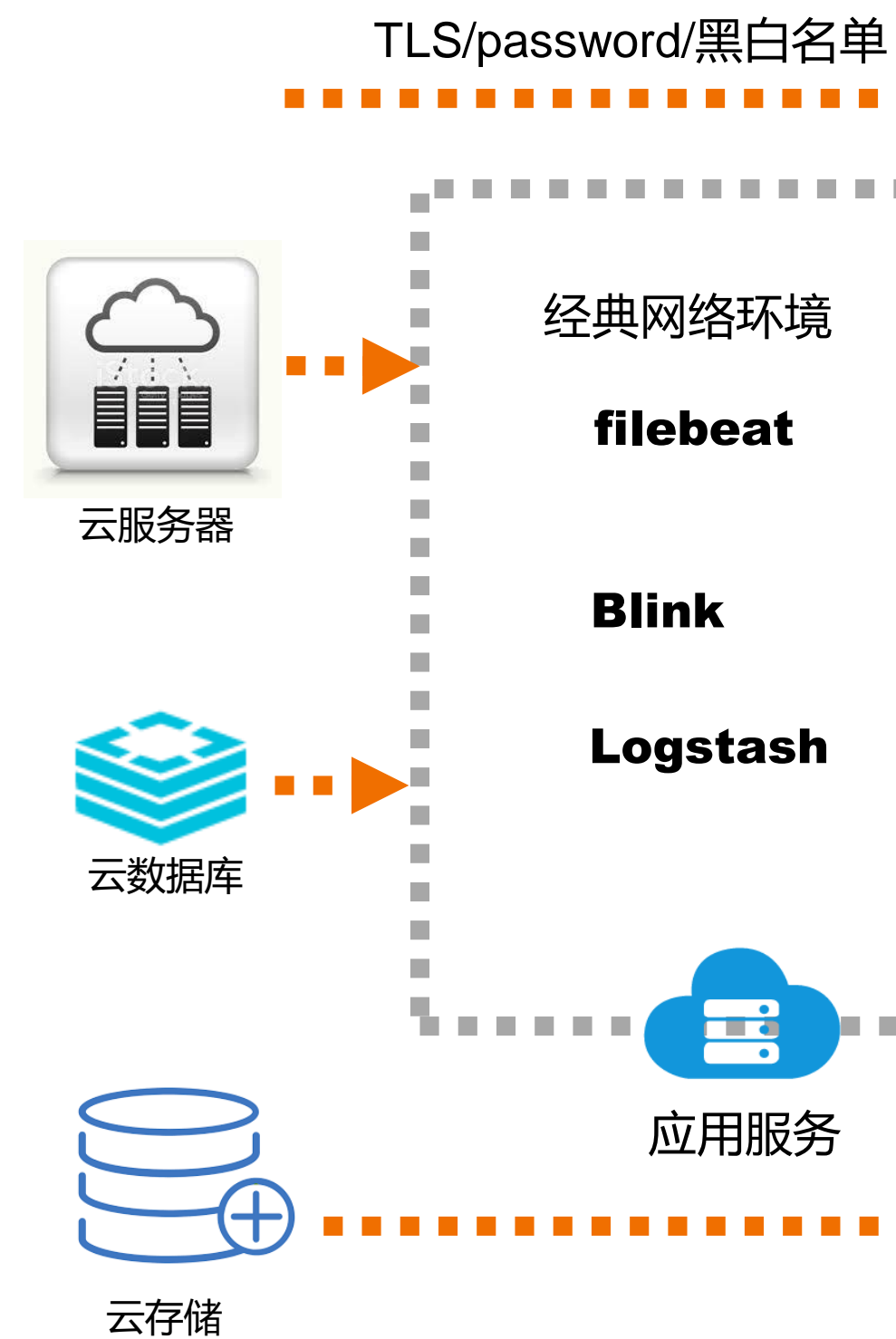
- 1、节点配置差异化的维护（节点角色的差异、磁盘的差异。。。）
- 2、集群升级Rolling耗费大量的时间
- 3、集群的指标的采集（通过脚本上报指标。。。）
- 4、集群的报警（对接公司内部的邮件、短信服务。。。）
- 5、节点进程的守护（supervisor、守护脚本。。。）

软件架构

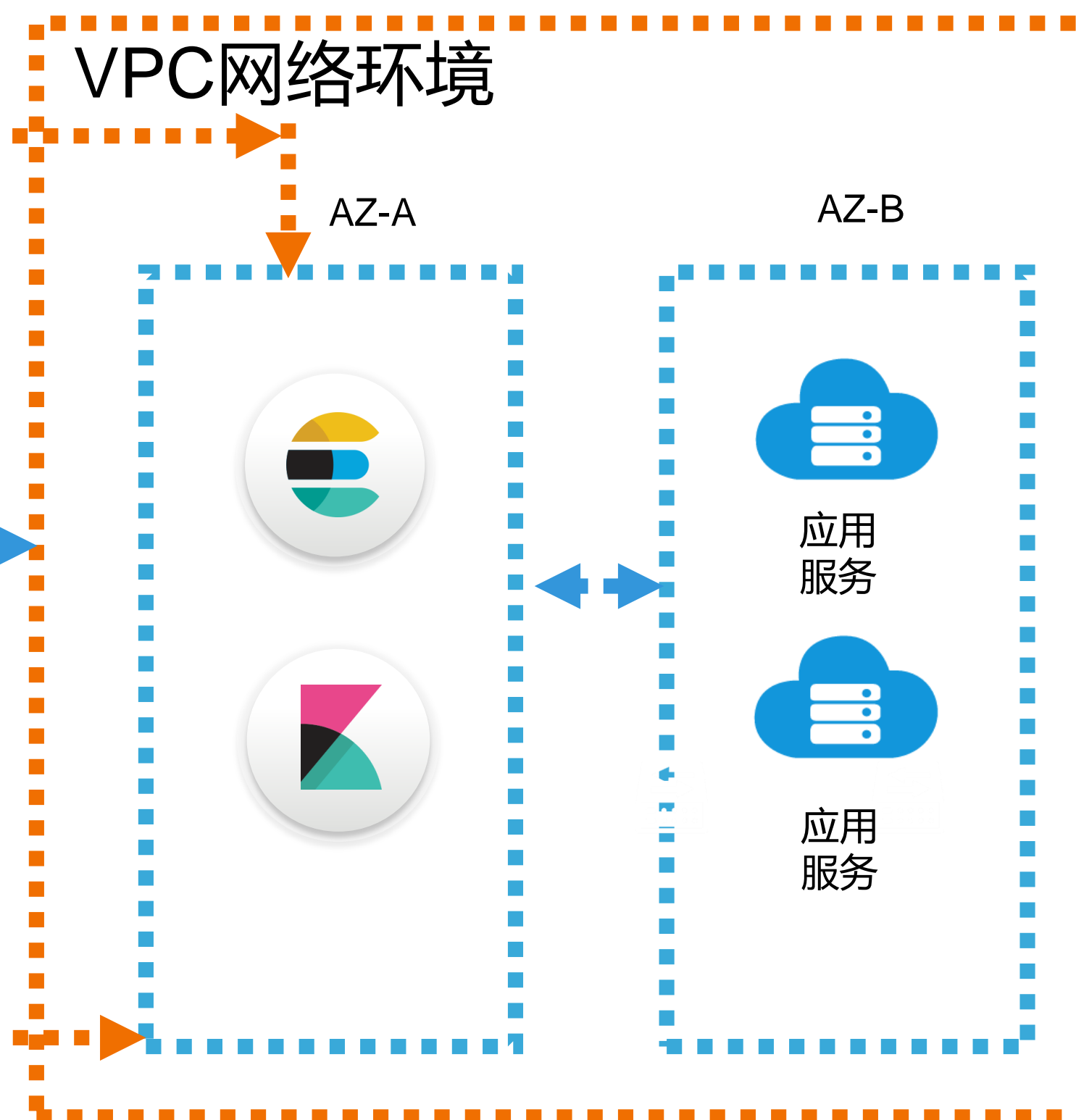


数据安全

公网网络环境



VPC网络环境



管理控制台



实例ID: es-12345678901234567890

Kibana控制台

集群监控

重启实例

刷新



基本信息

ES集群配置

插件配置

集群监控

日志查询

安全配置

数据备份

▼ 智能运维

集群概况

健康诊断

历史报告

基本信息

转包年包月

实例ID: es-12345678901234567890

创建时间: 2019年2月27日 18:15:21

名称: 本地SATA 测试集群-测试 编辑

状态: ● 正常

Elasticsearch 版本: 5.5.3_with_X-Pack

付费类型: 后付费

区域: 华东1

可用区: cn-hangzhou-b

专有网络: vpc-12345678901234567890

VSwitch信息: vsw-12345678901234567890

内网地址: 10.10.10.10

内网端口: 9200

公网地址: 请开启公网访问地址后使用

配置信息

集群升配

数据节点规格: elasticsearch.d1.4xlarge(16核 64GB)

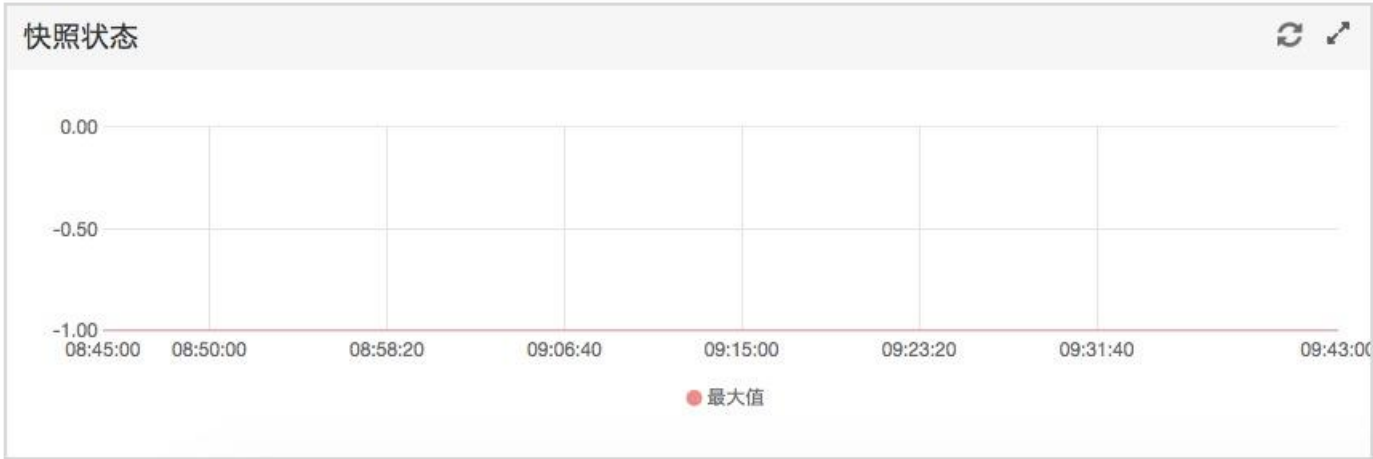
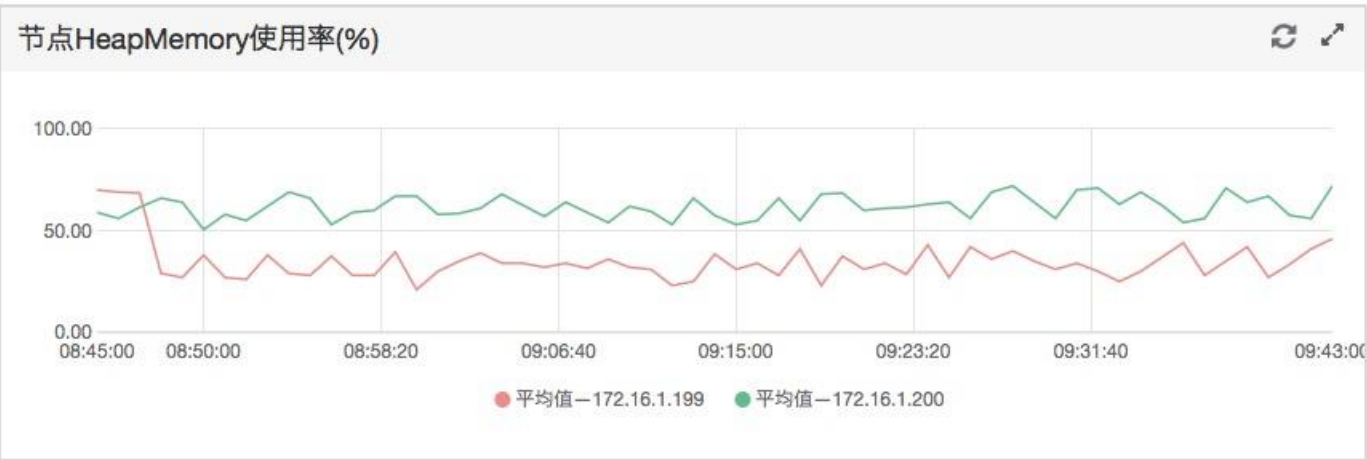
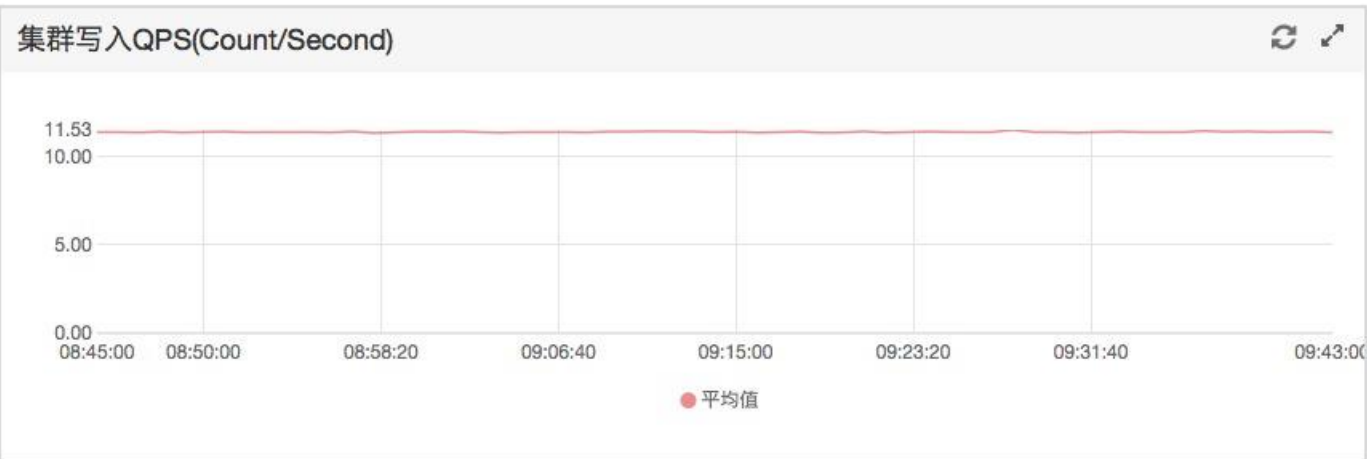
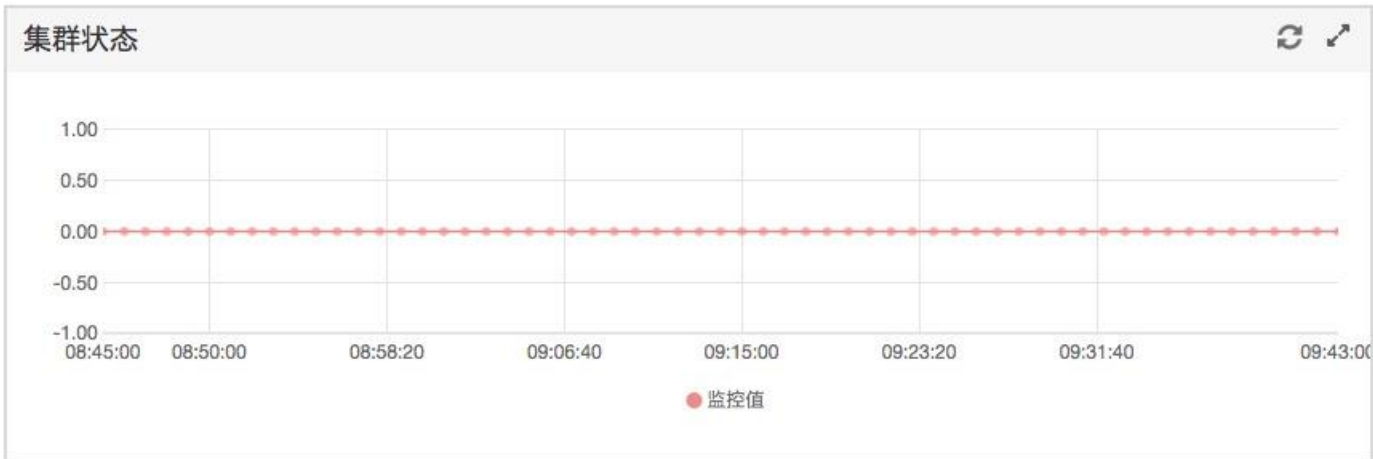
数据节点数量: 2

存储规格: 本地SATA盘

存储容量: 44000 GB

咨询·建议

监控报警



咨询·建议

日志可视化

kibana控制台

集群监控

重启实例

刷新

主日志

searching慢日志

indexing慢日志

GC日志

Search... (e.g. status:200 AND extension:PHP)



2018-05-22 08:00:29



至

2018-05-23 08:00:29



搜索

时间	节点IP	内容
2018-05-23 08:00:05		<div>level : info</div> <div>host : </div> <div>time : 2018-05-23T08:00:05.007Z</div> <div>content : [o.e.c.r.a.AllocationService] [B9Mu1Qd] Cluster health status changed from [YELLOW] to [GREEN] (reason: [shards started [[.monitoring-es-6-2018.05.23][0]] ...]).</div>
2018-05-23 00:30:00		<div>level : info</div> <div>host : </div> <div>time : 2018-05-23T00:30:00.002Z</div> <div>content : [o.e.x.m.a.DeleteExpiredDataAction\$TransportAction] [B9Mu1Qd] Deleting expired data</div>

弹性扩容

- kibana控制台
- 集群监控
- 重启实例
- 刷新

基本信息

- 转包年包月
- 节点扩容

实例ID: elasticsearch-n4-small-1

名称: elasticsearch-n4-small-1 [编辑](#)

规格ID: elasticsearch.n4.small

Elasticsearch 版本: 5.5.3_with_X-Pack

付费类型: 后付费

内网地址: 10.10.10.10

公网地址: [请开启公网访问地址后使用](#)

区域: 华东2

专有网络: vpc-22owb02n7jntj00ckr00a

创建时间: 2018-05-18 23:14:07

Dedicated master: 未开通 [?](#)

规格: CPU: 1核 内存: 2GB 存储: 20GB SSD

节点数: 2

状态: ● 正常

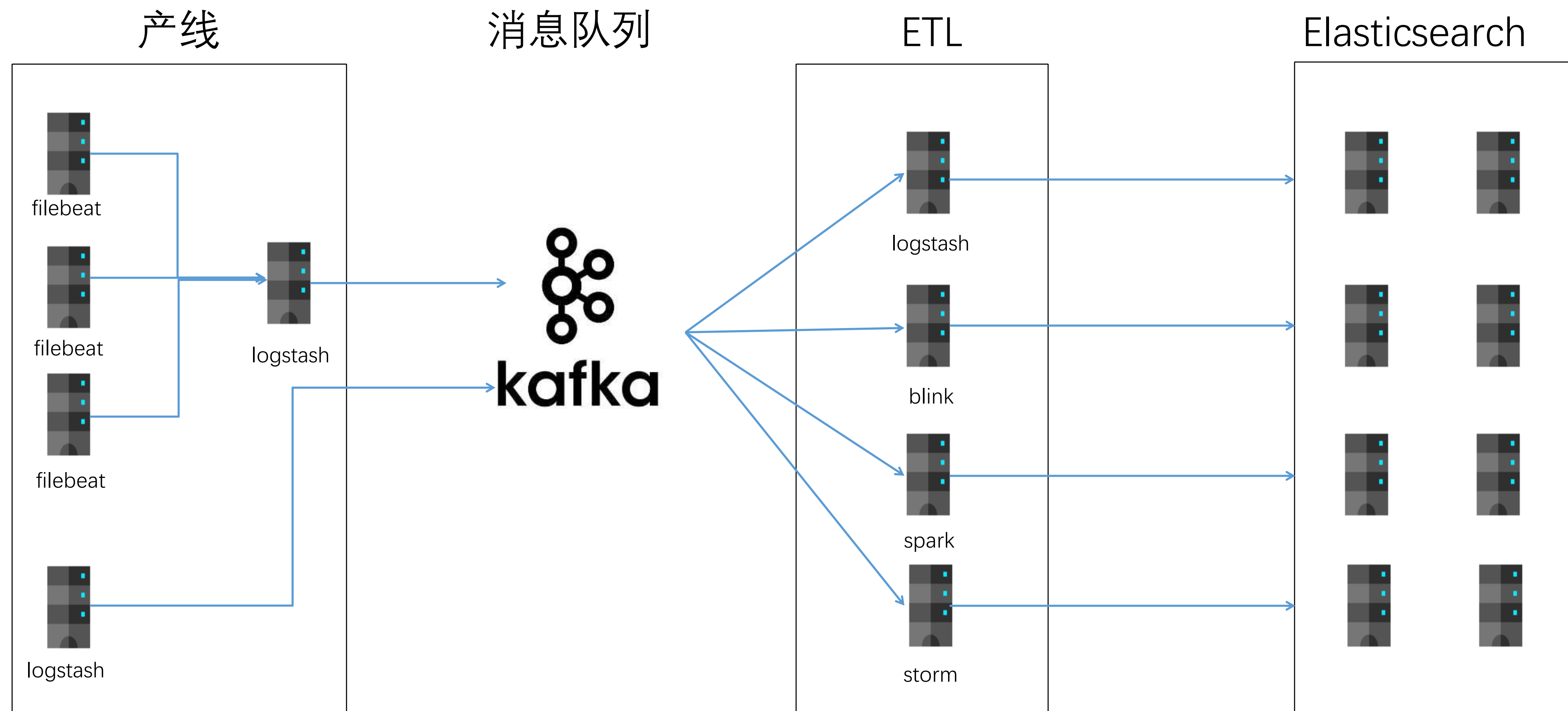
私网端口: 9200

公网端口: 9200

可用区: 华东2

vswitch信息: vsw-22owb02n7jntj00ckr00a





问题现象：

产线上出现服务出现不可用，kafka消费队列延迟，触发延迟报警。

排查问题思路及解决方案：

通过elasticsearch API:GET
/_cat/thread_pool/bulk?v&h=name,host,active,queue,rejected,completed
定位哪个节点比较忙：queue比较大，rejected不断增加。

通过GET /_cat/shards 找到该node上活跃的shard。

通过POST /_cluster/reroute API把shard移到load比较低的node上，缓解该node的压力。

问题现象：

集群触发健康度红色报警，集群为Red。

_cluster/health

```
{  
  cluster_name: "[REDACTED]",  
  status: "red",  
  timed_out: false,  
  number_of_nodes: 4,  
  number_of_data_nodes: 4,  
  active_primary_shards: 187,  
  active_shards: 376,  
  relocating_shards: 0,  
  initializing_shards: 0,  
  unassigned_shards: 2,  
  delayed_unassigned_shards: 0,  
  number_of_pending_tasks: 0,  
  number_of_in_flight_fetch: 0,  
  task_max_waiting_in_queue_millis: 0,  
  active_shards_percent_as_number: 99.47089947089947  
}
```

_cat/allocation

shards	disk.indices	disk.used	disk.avail	disk.total	disk.percent	host	ip	node
86	80.1gb	90.7gb	56.9gb	147.6gb	61			4phke_E
97	63.2gb	73gb	74.5gb	147.6gb	49			azYlNSC
96	93.3gb	103.7gb	43.9gb	147.6gb	70			gRhTBDa
97	85.9gb	96gb	51.5gb	147.6gb	65			uCxVau0
2								UNASSIGNED

_cluster/allocation/explain?pretty

`allocate_explanation: "cannot allocate because allocation is not permitted to any of the nodes that hold an in-sync shard copy",`

“shard copies that can safely be selected as primary, also called the in-sync shard copies ”

_cluster/reroute?retry_failed

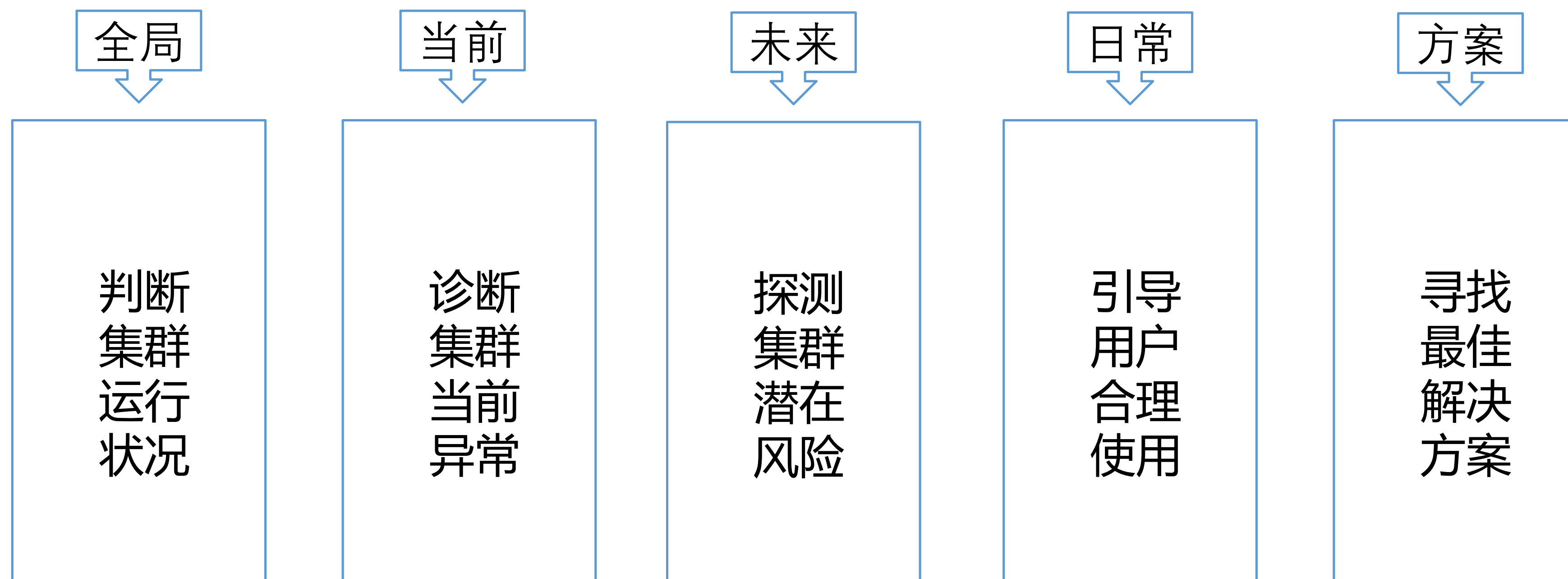
客户的烦恼：

- 1、有一个新的业务如何规划集群大小
- 2、集群咋成Red的了
- 3、集群某个节点负载为什么别其他节点高
- 4、某个节点突然自动重启过
- 5、内存使用很高但CPU使用不高
- 6、节点负载很低但是查询很慢

○ ○ ○ ○ ○ ○ ○ ○

EYou是阿里云Elasticsearch智能诊断系统。

目标：更全方面的了解ES集群健康，寻找更佳的使用方式，更佳智能化的运维集群



诊断项是可以直接反馈ES
集群某一个状态或行为是
否合理的指标。



华东1 (杭州)

es-c

立刻诊断

历史诊断报告

● scheduled__2018-07-16 03:31:13)

● scheduled__2018-07-15 03:31:10)

● scheduled__2018-07-14 03:31:10)

● scheduled__2018-07-13 03:31:10)

● scheduled__2018-07-12T03:30:00 (2018-07-12 03:31:08)

● scheduled__2018-07-11T03:30:00 (2018-07-11 03:31:10)

● scheduled__2018-07-10T03:30:00 (2018-07-10 03:31:06)

● scheduled__2018-07-09T03:30:00 (2018-07-09 03:31:09)

索引名称:

请输入要诊断的索引名, 多个索引用逗号分隔; 默认全部索引

诊断项:

☒ 集群颜色状态诊断

☒ 集群master分配诊断

☒ 集群存储资源诊断

☒ 集群计算资源诊断

☒ 索引shard合理性诊断

索引shard数可能需要调整。按照当前索引大小计算, 给出了参考方案, 但实际操作时且要尽可能匹配节点数。

Action:

参考方案:

2018-07-05 [1GB] [3 -> 1]

2018-07-04 [size < 1GB] [3 -> 1]

2018-07-03 [size < 1GB] [3 -> 1]

3-2018-07-03 [size < 1GB] [3 -> 1]

3-2018-07-02 [size < 1GB] [3 -> 1]

诊断结果

[查看原始文件](#)

集群整体诊断结果

实例ID: es-cv, 诊断时间: 2018-07-03 20:43:10

集群颜色状态诊断

诊断集群基本颜色状态, 检查副本分配情况
颜色不正常的索引会影响数据读写。黄色索引副本丢失, 会影响到数据的可靠性和读写性能; 红色索引会引起数据丢失或者kibana加载异常, 最高优处理

诊断结果及建议:

磁盘空间不足导致集群颜色异常[RED], 丢失索引分片, 如: [2018-07-05, 2018-07-04, 2018-07-03, 2018-07-03, 2018-07-02], .kibana, .security, 2018-07-03, 2018-07-02]。建议扩充磁盘空间至少到 9000GB, 或者参考<集群存储资源诊断>

Action:

GET /_cluster/allocation/explain GET /_cluster/health GET /_cat/indices?v

集群存储资源诊断

诊断集群存储空间是否充足
磁盘使用超过85%的时候将不允许创建新索引, 超过90%就会尝试重新分配分片。空间不足时可能会导致新索引创建不出来, 分片丢失, kibana加载异常, 负载增加等, 高优处理

诊断结果及建议:

磁盘空间不足, 最大使用率为 94.92, 报警次数 WARN[2726次] CRITICAL[1364次]

Action:

建议集群总体磁盘扩容到 9000 GB, 单节点容量[1500 -> 1800 GB], 节点个数[4 -> 5]

为了无法计算的价值 | 阿里云

❌ 集群计算资源诊断

诊断集群节点和规格是否充足
计算资源不足会全方面影响到集群稳定性，读写性能

诊断结果及建议：

计算资源不足。
报警次数 CPU[99次] JVM[0次]
系统资源使用情况：CPU.AVG[12.93] JVM.AVG[57.9] LOAD.AVG[1.8] CPU.MAX[100.0] JVM.MAX[100.0]
LOAD.MAX[16.06]

Action：

建议增加一个数据节点

❌ 集群状态频繁变更诊断

诊断集群状态变更是否合理
短时间频繁变更集群状态会给master节点带来很高的负担，GC频繁，负载突增，甚至阻塞相关索引的读写，影响性能

诊断结果及建议：

集群状态变更频繁，过去24小时内状态发生频繁变更
2018-07-03 19:24 -- 2018-07-03 20:42 连续变更358次
2018-07-02 20:43 -- 2018-07-03 19:21 连续变更1376次
2018-07-03 19:24 -- 2018-07-03 19:25 连续变更30次


Action：

请确认是否有频繁创建，删除，打开，关闭索引，如有请尽量在低峰期操作
请确认是否有频繁增加type，动态增加字段，如有请提前创建完整的mapping，尽量不使用动态映射
请确认是否有频繁修改索引或集群配置，如有请尽量在低峰期操作
请确认是否有集群变更，重启，节点上下线等操作

⚠ 节点负载偏差过大诊断

诊断集群当天节点负载偏差是否过大
节点间负载不一致会使得某个节点成为系统瓶颈，影响集群稳定性

诊断结果及建议：

数据节点 负载相对较高。以下索引可能存在shard不均匀 [[]]



Action：

请试着调整shard数或数据节点数，尽可能保证两者均衡

⚠ 索引segment合理性诊断

诊断索引segment是否合理，是否需要优化
非大量写入情况下，过多的segment会降低查询性能，消耗内存，单条更新下性能极低

诊断结果及建议：

索引segment个数过多，索引列表如下：
 期望segment数45 实际segment数71
 期望segment数50 实际segment数77

Action：

建议可以在负载低峰时执行ES API：`{indexName}/_forcemerge`

⚠ 索引shard合理性诊断





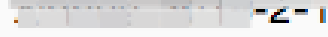

诊断索引的shard数和大小是否合理
shard不合理会极大的影响索引读写性能，meta信息过多会占用较高的系统资源

诊断结果及建议：

索引shard数可能需要调整。按照当前索引大小计算，给出了参考方案，但实际操作时需要考虑后续扩展且要尽可能匹配节点数。

Action：

参考方案：

 [1GB] [10 -> 1]
 [1GB] [10 -> 1]
 [size < 1GB] [10 -> 1]
 3 [size < 1GB] [10 -> 1]
 -2- [size < 1GB] [10 -> 1]
 [size < 1GB] [10 -> 1]

节点shard数过多诊断

诊断集群当前节点shard数是否过多 单节点shard过多会大量消耗系统资源，读写失败，负载增加，索引加载异常等

诊断结果及建议：

部分节点shard个数过多，具体如下：[: 80311]

Action:

请考虑增加集群节点数到 [3 -> 4]，或者提升规格到 [S2C8G -> S4C16G] 或者及时清理关闭无效索引，减少副本个数，减少shard个数

索引recovery过慢诊断

诊断集群当天索引recovery是否过慢

诊断结果及建议：

部分索引recovery过慢，最大耗时[190]min，最大任务数12582个，变更期间尽可能停止写操作，请考虑修改集群配置如Action

Action:

```
PUT _cluster/settings
{
  "transient": {
    "cluster.routing.allocation.node_concurrent_recoveries":4,
    "cluster.routing.allocation.cluster_concurrent_rebalance":2,
    "indices.recovery.max_bytes_per_sec": "200mb"
  }
}
```

休假中

@EYou 北京

EYou 机器人

@北京 (华北2) 诊断报告如下：

最近一次诊断 (2018-07-05 17:13:45) 结果为 -- RED

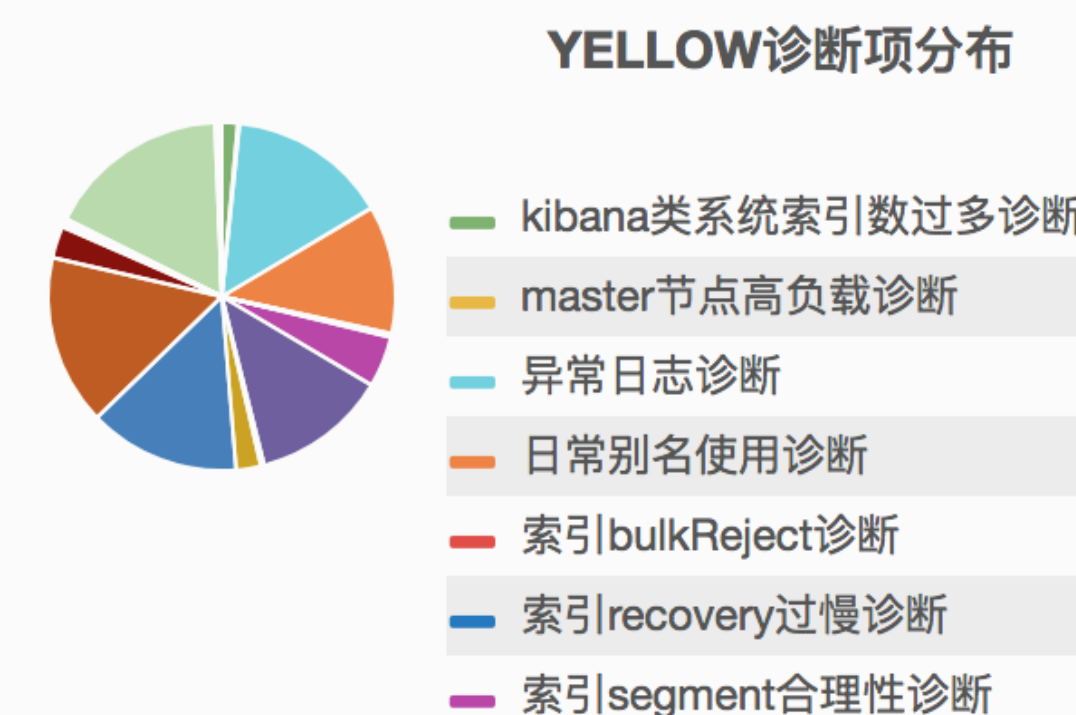
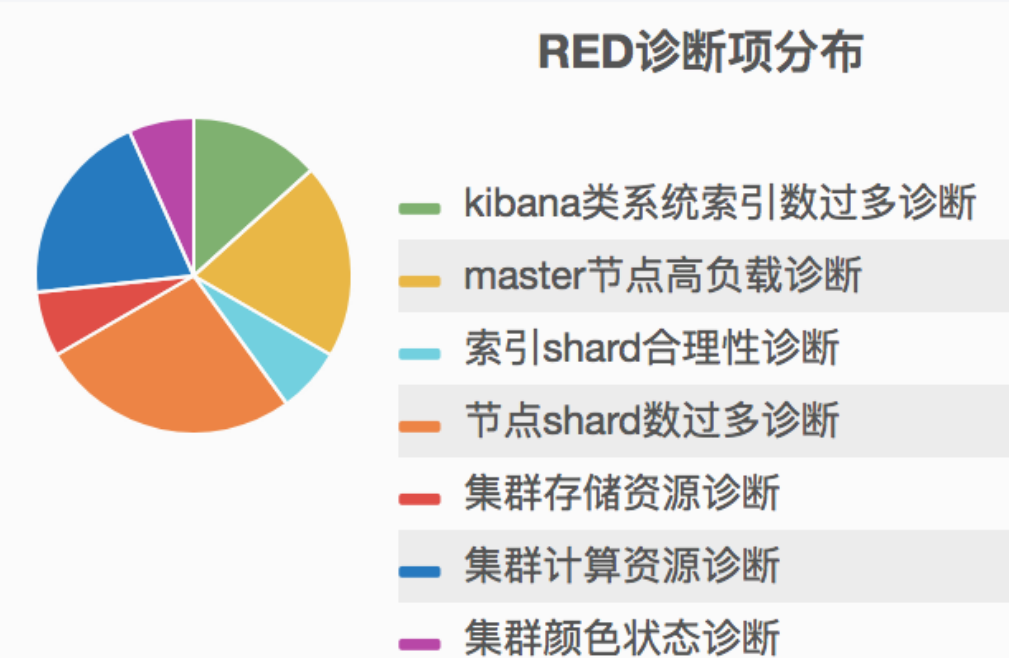
资源详情：

- CPU : 8 核
- MEM : 32 G
- DISK : 1500 G
- NODE : 4 个
- DedicateMaster : 无

结果详情：

1. 集群颜色状态诊断 -- RED

集群颜色异常[RED]，丢失索引分片，如：[.security, 1000000000] 原因可能为：{ "index" : ".security", "shard" : 0, "primary" : false, "current_state" : "unassigned", "unassigned_info" : { "reason" : "NODE_LEFT", "at" : "2018-07-03T11:22:42.267Z", "details" : "node_left[tCE8FR1TBicfzRQZrHjg]", "last_allocation_status" : "no_attempt", "has_attempt" : "no", "allocation_explanation" : "cannot allocate" }



- 1、提前创建索引。
- 2、避免索引稀疏，使用单个type，index中document结构最好保持一致，如果document结构不一致，建议分index。
- 3、批量导入大量数据时可设置refresh_interval=-1，index.number_of_replicas=0，索引完成后再设回来。
- 4、load和io压力不大的情况，用bulk比单条的PUT/DELETE操作索引效率更高。
- 5、调整index buffer(indices.memory.index_buffer_size)。
- 6、不需要score的field，禁用norms；不需要sort、aggregate、脚本的field，禁用doc_value。
- 7、如果heap压力不大，可适当增加node query cache(indices.queries.cache.size)
- 8、定期合并segment。
- 9、增加shard replica 可提高查询并发能力，但要注意node上的shard总量
- 10、使用routing提升某一维度数据的查询速度

节点角色划分:

master node, data node, ingest node, coordinate node, ML node。

heap大小:

不大于32g (26g、28g)

Index管理:

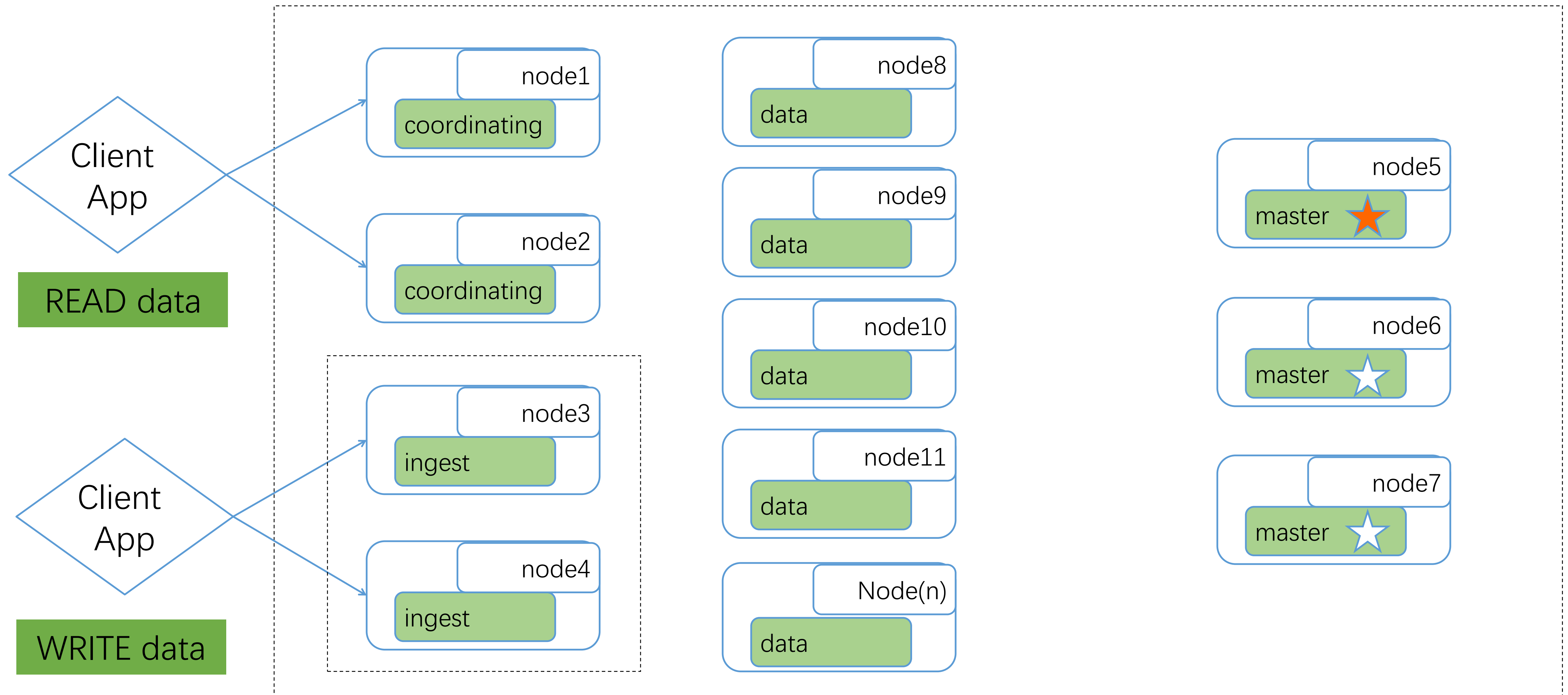
mapping template, alias, create, delete, merge segment, rotation (curator)

Shard、node管理:

每个shard大小不超过40g, 每个node的shard总个数不超过 $\text{memory} \times 20$, 同一个index的shard在同一个节点上尽量少于4个, 每个node节点数据量不要超过5T。

磁盘空间:

合理规划磁盘空间, 一般数据膨胀系数为3.4 (有一个replica)





EFlow:

支持更多的数据源、数据毫秒级实时可见

阿里云Es管控系统:

集群的自动化扩缩容

Eyou:

Machine Learning在EYou上的使用, 更多的诊断项, 更准确的预测, 针对不同场景和行业有相应的模版

开源社区:

积极参与社区活动, 提供更多、更好用的插件

Elasticsearch:

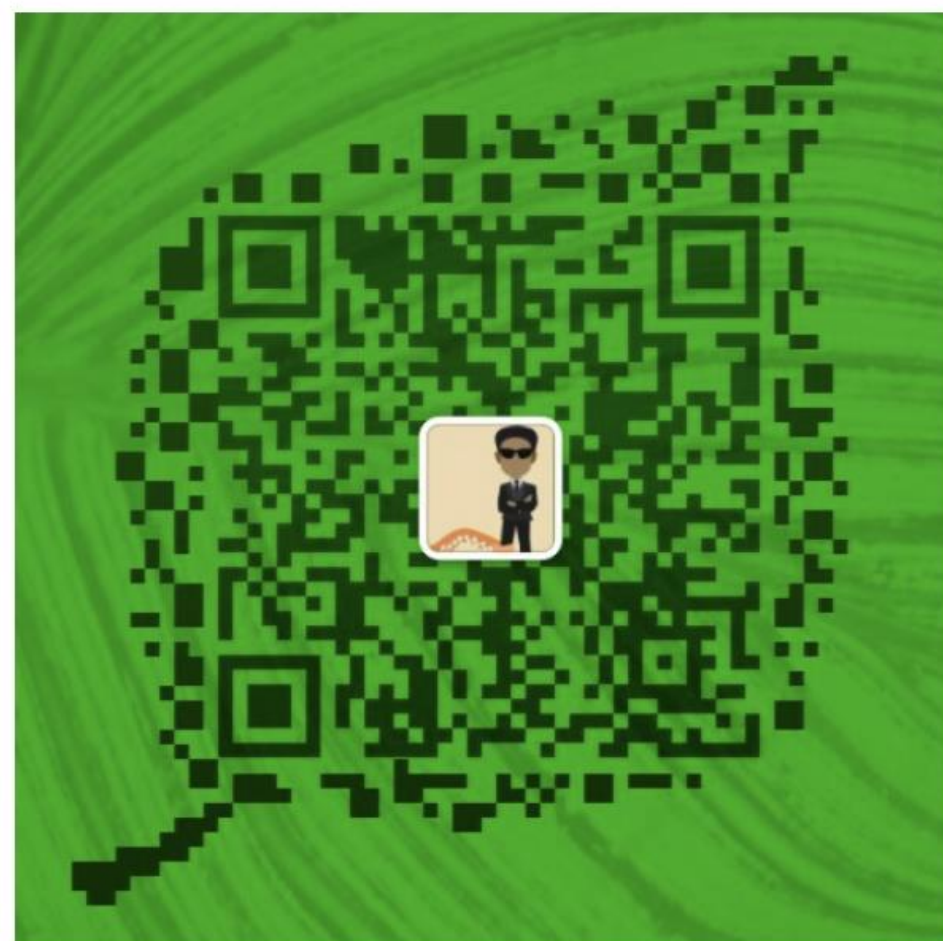
在现有的metric基础上, 丰富更多的metric

解决客户痛点问题，阿里云Elasticsearch一直在路上。。。



楠木可依

中国



扫一扫上面的二维码图案，加我微信

Elasticsearch中文技术...

1253人



扫一扫群二维码，立刻加入该群。



奥运会全球指定云服务商



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