vivo Elasticsearch集群
应用实践

2021年08月21日

刘石林　数据库工程师

elastic

目录
CONTENTS

elastic

当前业务规模介绍

全球服务

俄罗斯

德国

中国

新加坡

印度

elastic

Elasticsearch 业务服务规模

数据规模
- 业务数据约 1 PB+
- 日志数据约 500 TB+

节点规模
- 服务器   700+
- 运行实例   1800+
- 集群数量   600+

数据写入
- 日志实时写入   百万/秒
- 每天存储平均   百亿+
- 稳定性   99.99%
- 单集群写入   30TB/日

elastic

# 调用链应用实践

Elasticsearch 痛点

可用性
高可用
99.99%

写入
高吞吐
百万级/秒

响应
低延迟
1~3秒

elastic

可用性 - 多机房双活

机房 A

web层

存储层

计算层

中间层

汇聚层

采集层

Trace Web

前端请求

ES(span)   ES(metric)

写入

Flink Cluster

读

Kafka

flume

Java-agent

采集指标数据

机房 B

Trace Web

前端请求

ES(span)   ES(metric)

写入

Flink Cluster

读

Kafka

flume

Java-agent

采集指标数据

elastic
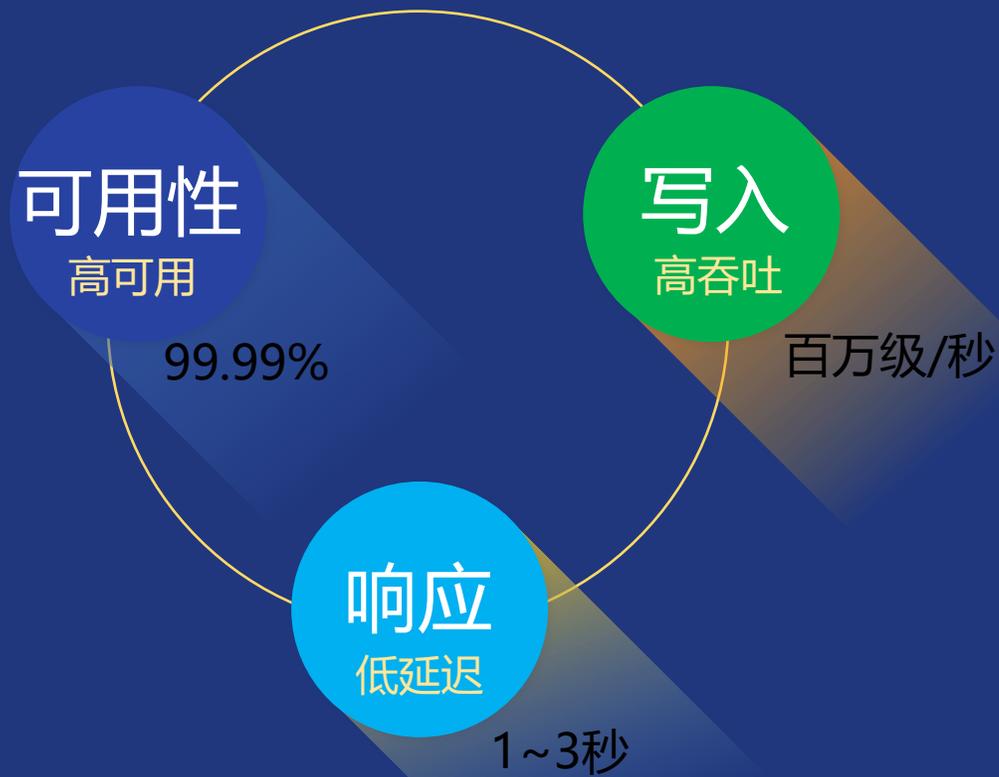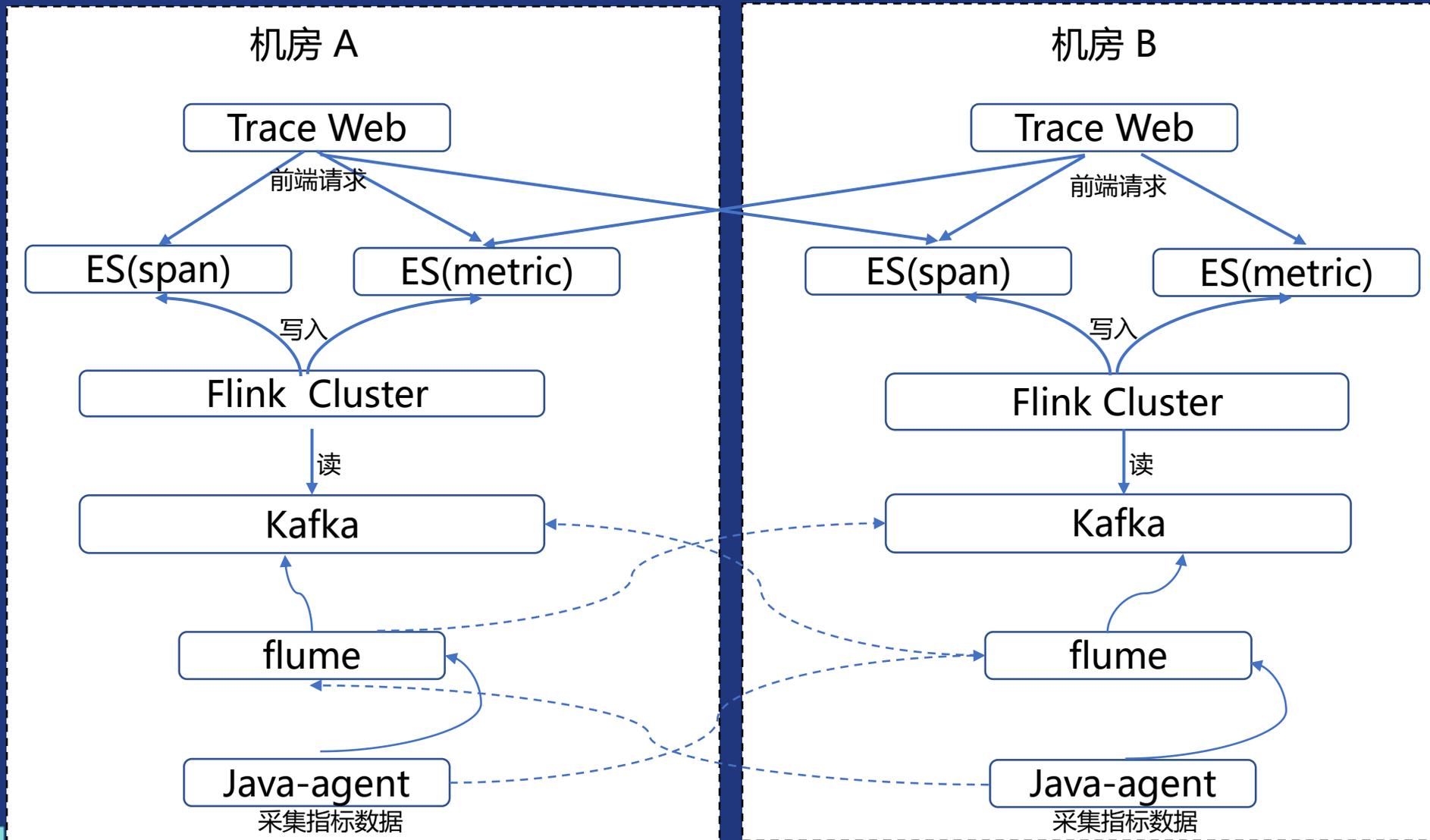
写入 - 索引划分

日期拆分

▫ 减少索引和分片数量

▫ 减少故障影响时间

▫ 方便集群管理

索引拆分

▫ 管理精细化（指标、告警、日志）

▫ 冷热数据分离

▫ 在一定程度上减少磁盘压力

elastic

响应 - 数据采集治理

日志降级	数据降级	频率控制降级
动态日志级别	细粒度采样策略	Kafka发送频率控制
异步日志打印	链路降级	JVM数据采集频率控制
异常日志去重	异常流控	数据聚合频率控制
	全流程span流转监控	云端管控
		配置广播(应用级/IP级)

ES集群配置

- 硬件配置
  - 内存 192GB CPU 32核
  - 磁盘采用nvme ssd 3.7TB 日志类 73TB SATA
  - 安装部署时会预留40%内存作为系统缓存以提升查询效率

- 集群配置
  - 目前现存2个版本：6.3.2 和 7.4.2
  - 部署集群默认最少3个节点，实例 jvm 默认最小4GB
  - 混合部署，基于套餐分配

套餐类型	套餐容量	套餐说明
D	0-400G	介于0G-400G 容量使用,内存4G
C	400-900G	介于400G-900G 容量使用,内存10G
B	900-1536G	介于900G-1.5T 容量使用,内存20G
A	1536G +	大于1.5T 容量使用,内存30G

elastic

ES参数优化调整

- 分片设置
  - 日志和监控指标类索引，默认无副本，业务类默认1副本

- 集群参数设置
  - 业务类默认refresh_interval 1s，部分写入流量大调整为 5s；日志类 30s
  - Mapping字段最大字段 Index. mapping. total_fields. limit 默认 20000
  - 事务日志持久性模式 translog. durability 设置 async
  - 事务日志同步间隔 translog. sync_interval 设置 90s
  - 事务日志内存大小 translog. flush_threshold_size 设置 1000m
  - 推迟分片分配 unassigned.node_left.delayed_timeout 设置 20min

- 插件
  - 默认所有实例安装ik和ik-pinyin

elastic

踩坑之Bug（一）

集群状态为yellow

```
{
  "cluster_name" : "███████████",
  "status" : "yellow",
  "timed_out" : false,
  "number_of_nodes" : 19,
  "number_of_data_nodes" : 19,
  "active_primary_shards" : 60,
  "active_shards" : 110,
  "relocating_shards" : 0,
  "initializing_shards" : 0,
  "unassigned_shards" : 10,
  "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" : 91.66666666666666
}
```

手动多次重试分配副本分片不成功，并且日志提示空指针异常

```
Caused by: org.elasticsearch.index.translog.TranslogException: Failed to write operation [NoOp{seqNo=369645587, primaryTerm=1, reason='null'}]
    at org.elasticsearch.index.translog.Translog.add(Translog.java:506) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.index.engine.InternalEngine.index(InternalEngine.java:820) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.index.shard.IndexShard.index(IndexShard.java:732) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.index.shard.IndexShard.applyIndexOperation(IndexShard.java:701) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.index.shard.IndexShard.applyTranslogOperation(IndexShard.java:1244) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.indices.recovery.RecoveryTarget.indexTranslogOperations(RecoveryTarget.java:400) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.indices.recovery.PeerRecoveryTargetService$TranslogOperationsRequestHandler.messageReceived(PeerRecoveryTargetService.java:454) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.indices.recovery.PeerRecoveryTargetService$TranslogOperationsRequestHandler.messageReceived(PeerRecoveryTargetService.java:445) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.transport.TransportRequestHandler.messageReceived(TransportRequestHandler.java:30) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.xpack.security.transport.SecurityServerTransportInterceptor$ProfileSecuredRequestHandler$1.doRun(SecurityServerTransportInterceptor.java:259) ~[?:?]
    at org.elasticsearch.common.util.concurrent.AbstractRunnable.run(AbstractRunnable.java:37) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.xpack.security.transport.SecurityServerTransportInterceptor$ProfileSecuredRequestHandler.messageReceived(SecurityServerTransportInterceptor.java:317) ~[?:?]
    at org.elasticsearch.transport.RequestHandlerRegistry.processMessageReceived(RequestHandlerRegistry.java:66) ~[elasticsearch-6.3.2.jar:6.3.2]
    at org.elasticsearch.transport.TcpTransport$RequestHandler.doRun(TcpTransport.java:1592) ~[elasticsearch-6.3.2.jar:6.3.2]
    ... 5 more
Caused by: java.lang.NullPointerException
```

踩坑之Bug（二）

修复Case https://github.com/elastic/elasticsearch/pull/43523

修复前 https://github.com/elastic/elasticsearch/blob/v6.3.2/server/src/main/java/org/elasticsearch/index/engine/InternalEngine.java#L820

```
812                                   plan.versionForIndexing, plan.seqNoForIndexing, plan.currentNotFoundOrDeleted);
813                           }
814                       if (index.origin() != Operation.Origin.LOCAL_TRANSLOG_RECOVERY) {
815                           final Translog.Location location;
816                           if (indexResult.getResultType() == Result.Type.SUCCESS) {
817                               location = translog.add(new Translog.Index(index, indexResult));
818                           } else if (indexResult.getSeqNo() != SequenceNumbers.UNASSIGNED_SEQ_NO) {
819                               // if we have document failure, record it as a no-op in the translog with the generated seq_no
820                               location = translog.add(new Translog.NoOp(indexResult.getSeqNo(), index.primaryTerm(), indexResult.getFailure().getMessage()));
821                           } else {
822                               location = null;
823                           }
824                           indexResult.setTranslogLocation(location);
825                       }
826                       if (plan.indexIntoLucene && indexResult.getResultType() == Result.Type.SUCCESS) {
827                           final Translog.Location translogLocation = trackTranslogLocation.get() ? indexResult.getTranslogLocation() : null;
```

修复后 https://github.com/elastic/elasticsearch/blob/v6.4.1/server/src/main/java/org/elasticsearch/index/engine/InternalEngine.java#L858

```
848                   } else {
849                       indexResult = new IndexResult(
850                               plan.versionForIndexing, getPrimaryTerm(), plan.seqNoForIndexing, plan.currentNotFoundOrDeleted);
851                   }
852                   if (index.origin() != Operation.Origin.LOCAL_TRANSLOG_RECOVERY) {
853                       final Translog.Location location;
854                       if (indexResult.getResultType() == Result.Type.SUCCESS) {
855                           location = translog.add(new Translog.Index(index, indexResult));
856                       } else if (indexResult.getSeqNo() != SequenceNumbers.UNASSIGNED_SEQ_NO) {
857                           // if we have document failure, record it as a no-op in the translog with the generated seq_no
858                           location = translog.add(new Translog.NoOp(indexResult.getSeqNo(), index.primaryTerm(), indexResult.getFailure().toString()));
859                       } else {
860                           location = null;
861                       }
862                       indexResult.setTranslogLocation(location);
863                   }
864                   if (plan.indexIntoLucene && indexResult.getResultType() == Result.Type.SUCCESS) {
865                       final Translog.Location translogLocation = trackTranslogLocation.get() ? indexResult.getTranslogLocation() : null;
866                       versionMap.maybePutIndexUnderLock(index.uid().bytes(),
867                           new IndexVersionValue(translogLocation, plan.versionForIndexing, plan.seqNoForIndexing, index.primaryTerm()));
868                   }
869                   if (indexResult.getSeqNo() != SequenceNumbers.UNASSIGNED_SEQ_NO) {
```

踩坑之选举不成功（一）

背景：故障集群版本 7.x，节点数31，存储61TB，业务反馈数据不能写入

org.elasticsearch.cluster.metadata.ProcessClusterEventTimeoutException: failed to process cluster event (put-mapping) within 30s
    at org.elasticsearch.cluster.service.MasterService$Batcher.lambda$onTimeout$0(MasterService.java:134) ~[elasticsearch-7.2.0.jar:7.2.0]
    at java.util.ArrayList.forEach(ArrayList.java:1540) ~[?:?]
    at org.elasticsearch.cluster.service.MasterService$Batcher.lambda$onTimeout$1(MasterService.java:133) ~[elasticsearch-7.2.0.jar:7.2.0]
    at org.elasticsearch.common.util.concurrent.ThreadContext$ContextPreservingRunnable.run(ThreadContext.java:688) ~[elasticsearch-7.2.0.jar:7.2.0]
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1128) ~[?:?]
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628) ~[?:?]
    at java.lang.Thread.run(Thread.java:834) [?:?]

[2021-06-10T00:01:04,942][ERROR][o.e.x.m.c.i.IndexRecoveryCollector] [7.129] collector [index_recovery] timed out when collecting data

[2021-06-10T00:02:57,974][WARN][o.e.c.s.MasterService] [7.129] cluster state update task [shard-started StartedShardEntry{shardId [[elastiflow-2021.06.10][12]], allocationId [KneKNv9ZSsanY21O5y96XQ], primary term [1], message [after peer recovery]][StartedShardEntry{shar
[[elastiflow-2021.06.10][12]], allocationId [KneKNv9ZSsanY21O5y96XQ], primary term [1], message [after peer recovery]}]] took [1.9m] which is above the warn threshold of 30s

集群内主节点和其他节点相互之间会定期进行NodesFaultDetection和MasterFaultDetection，以检测加入集群的节点以及集群的主节点是否正常运行，根据主节点和其他节点日志，两方面的检测都出现了超时

[2021-06-10T00:01:02,174][WARN][o.e.t.TransportService   ] [7.129] Received response for a request that has timed out, sent [11296ms] ago, timed out [1403ms] ago, action [internal:coordination/fault_detection/follower_check], node [{   39.133}{rriM-x8JTImzjX3wqjR9gw}
{g_PUDQd2SY2DWWteDXG-oA}{   39.133}{   39.133:11451}{ml.machine_memory=201651568640, ml.max_open_jobs=20, xpack.installed=true}], id [1765522748]

[2021-06-10T00:01:48,886][WARN][o.e.t.TransportService   ] [.8.137] Received response for a request that has timed out, sent [10203ms] ago, timed out [200ms] ago, action [internal:coordination/fault_detection/leader_check], node [{   7.129}{MBTXGABoQv2Af2CmU5H5Cw
uYmCSd2wL0n18a8OOQ}{   7.129}{   129:11451}{ml.machine_memory=202444304384, ml.max_open_jobs=20, xpack.installed=true}], id [3838237015]

在其他节点中有16个出现了leader_check超时的日志记录，这些节点会选择重新加入这个主节点的集群，但是此时主节点已无法响应

[2021-06-10T00:02:22,250][DEBUG][o.e.a.a.i.m.p.TransportPutMappingAction]   .8.10] no known master node, scheduling a retry
[2021-06-10T00:02:28,228][DEBUG][o.e.a.a.c.h.TransportClusterHealthAction] [8.10] no known master node, scheduling a retry
[2021-06-10T00:02:31,130][WARN][o.e.c.c.ClusterFormationFailureHelper] | .8.10] master not discovered or elected yet, an election requires at least 14 nodes with ids from [...]

因此有超过半数的节点离开了集群，主节点放弃master身份重新开始选举

[2021-06-10T00:02:57,974][INFO][o.e.c.s.ClusterApplierService] [7.129] master node changed {previous [{   7.129}{MBTXGABoQv2Af2CmU5H5Cw}{zHV-uYmCSd2wL0n18a8OOQ}{   7.129}{   129:11451}
{ml.machine_memory=202444304384, xpack.installed=true, ml.max_open_jobs=20}], current []}, term: 53, version: 60834, reason: becoming candidate: joinLeaderInTerm

但是因为ES 7.x版本选举算法有所变化，所有节点都有资格成为主节点且各节点可以多次投票，选出最后当选的主节点。当发起选举的节点收到其他节点的选举请求，是的选举中断，从而导致一直未能完成选举

[2021-06-10T00:02:56,982][INFO][o.e.c.s.MasterService   ] [8.10] elected-as-master ([15] nodes joined)[{   8.10}{DhxVZBAgSmmxHLF-PoBu-g}{RNyMeq0OQDe4X7mDo9VI_g}{   8.10}{   .8.10:11451}
{ml.machine_memory=202444304384, xpack.installed=true, ml.max_open_jobs=20} elect leader, ..., _BECOME_MASTER_TASK_, _FINISH_ELECTION_], term: 525, version: 60835, reason: master node changed {previous [], current [{   .8.10}{DhxVZBAgSmmxHLF-PoBu-g}
{RNyMeq0OQDe4X7mDo9VI_g}{   .8.10}{   .8.10:11451}{ml.machine_memory=202444304384, xpack.installed=true, ml.max_open_jobs=20}]
[2021-06-10T00:02:56,982][WARN][o.e.c.s.MasterService   ] [8.10] failing [elected-as-master ([15] nodes joined)[{   10}{DhxVZBAgSmmxHLF-PoBu-g}{RNyMeq0OQDe4X7mDo9VI_g}{   8.10}{   .8.10:11451}
{ml.machine_memory=202444304384, xpack.installed=true, ml.max_open_jobs=20} elect leader, ..._BECOME_MASTER_TASK_, _FINISH_ELECTION_]]: failed to commit cluster state version [60835]
org.elasticsearch.cluster.coordination.FailedToCommitClusterStateException: node is no longer master for term 525 while handling publication

踩坑之选举不成功（二）

1. 临时解决办法：
    1. 暂时关闭业务写入
    2. 修改问题索引将 number_of_replicas 设置为 0
    3. 重启es节点，待集群恢复后再写入数据
2. 长期解决办法：
    1. 减少master选举节点，根据业务情况拆分master node和data node
    2. 预先创建索引
    3. 使用固定模板或mapping
    4. 关闭动态创建索引

平台化建设实践

Elasticsearch 平台建设能力

Elasticsearch生命周期管理

预算管理	资源部署	安全管控	成本核算

预算管理
- 预算收集
- 预算汇总
- 预算调整
- 预算汇总
- 预算控制

资源部署
- 工单申请
- 部署架构
- 部署规格
- 部署策略

安全管控
- 最小权限
- 权限管控
- 查询审计

成本核算
- 成本核算
- 成本统计
- 成本分摊

集群管理

数据管理	运维管理	安全管理

数据管理
- 数据查询
- 数据变更
- 数据传输
- 数据清理
- 数据导出
- 转换视图

运维管理
- 索引管理
- 模板管理
- 备份恢复
- 集群扩容
- 集群回收
- 集群监控
- 参数调整
- 混合部署
- 别名管理
- 巡检服务
- 性能诊断
- 生命周期
- 快照管理
- 集群搜索
- 组件模板
- 故障分析
- 故障自愈

安全管理
- 账户管理
- 权限管理
- 数据脱敏
- 查询审计

elastic

平台查询

平台索引管理

索引管理　模板管理　别名管理

创建索引　批量清理缓存　　　　　　　　　　　　　　　全部 ⌄　　全部 ⌄　　请输入索引名称　搜索

	索引名称 ⇅	运行状况 ⇅	状态 ⇅	主分片 ⇅	副主分片 ⇅	文档数 ⇅	存储大小 ⇅	操作
☐	mysql_cap_2021_04	green	open	5	1	21200853	13gb	详情　清理缓存
☐	mysql_cap_	green	open	5	1	0	2.5kb	详情　清理缓存
☐	mysql_cap_2021_07	green	open	5	1	24021008	14.7gb	详情　清理缓存
☐	mysql_cap_2020_11	green	open	5	1	14183173	8.7gb	详情　清理缓存
☐	mysql_cap_2020_01	green	open	5	1	8453980	5.2gb	详情　清理缓存
☐	mysql_cap_2020_07	green	open	5	1	11483953	7gb	详情　清理缓存
☐	mysql_cap_2020_12	green	open	5	1	19262275	11.8gb	详情　清理缓存
☐	mysql_cap_2021_01	green	open	5	1	20404499	12.4gb	详情　清理缓存
☐	mysql_cap_2021_02	green	open	5	1	18987463	11.6gb	详情　清理缓存
☐	mysql_cap_2021_08	green	open	5	1	6621139	4.1gb	详情　清理缓存
☐	mysql_cap_2021_03	green	open	5	1	21308150	13.1gb	详情　清理缓存
☐	mysql_cap_2020_08	green	open	5	1	12646047	7.7gb	详情　清理缓存
☐	mysql_cap_2020_03	green	open	5	1	5890659	3.6gb	详情　清理缓存

elastic

平台集群管理

应用详情　　数据查询　　索引管理　　**服务变更**　　账户管理　　服务监控　　实例管理　　集群管理　　数据灾备

扩容　　回收　　配置变更　　JVM调整

* 工单类型　**ES扩容**

* 工单名称　请填写工单名称

* 工单描述　请填写工单描述

* 申请扩容节点数　🔘 端口保持一致　请填写申请扩容节点数

　　　　　　　预算剩余 -7680 G；扩容后剩余 -7680 G

　　　　　　　本次扩容将消耗 0 G

　　　　　　　(预算已超支，将预支 -7680 G)

　　　　　　　扩容的数据存储增量容量＝0 G / 索引副本数

* 执行时间　🕐 立即执行

　紧急扩容　否 🔘 是

* 审批人　可输入搜索，请选择

　　　　　请选择审批人

　审阅人　可输入搜索，请选择

* 执行DBA　可输入搜索，请选择

提 交

elastic

平台集群实例管理

可视化监控

后续改进

故障自愈

节点自动拉起、状态yellow重试等机制

性能分析

系统性能巡检，提供类似Oracle AWR报告

性能优化

针对性能分析结果进行提供建设性建议

日志分析

分析慢查询和gc日志

白屏化

进一步减少DBA黑屏化操作

elastic

扫描关注我

THANK YOU
谢谢观赏

elastic