



ES服务平台化在小米的 应用与实践

郑钧元

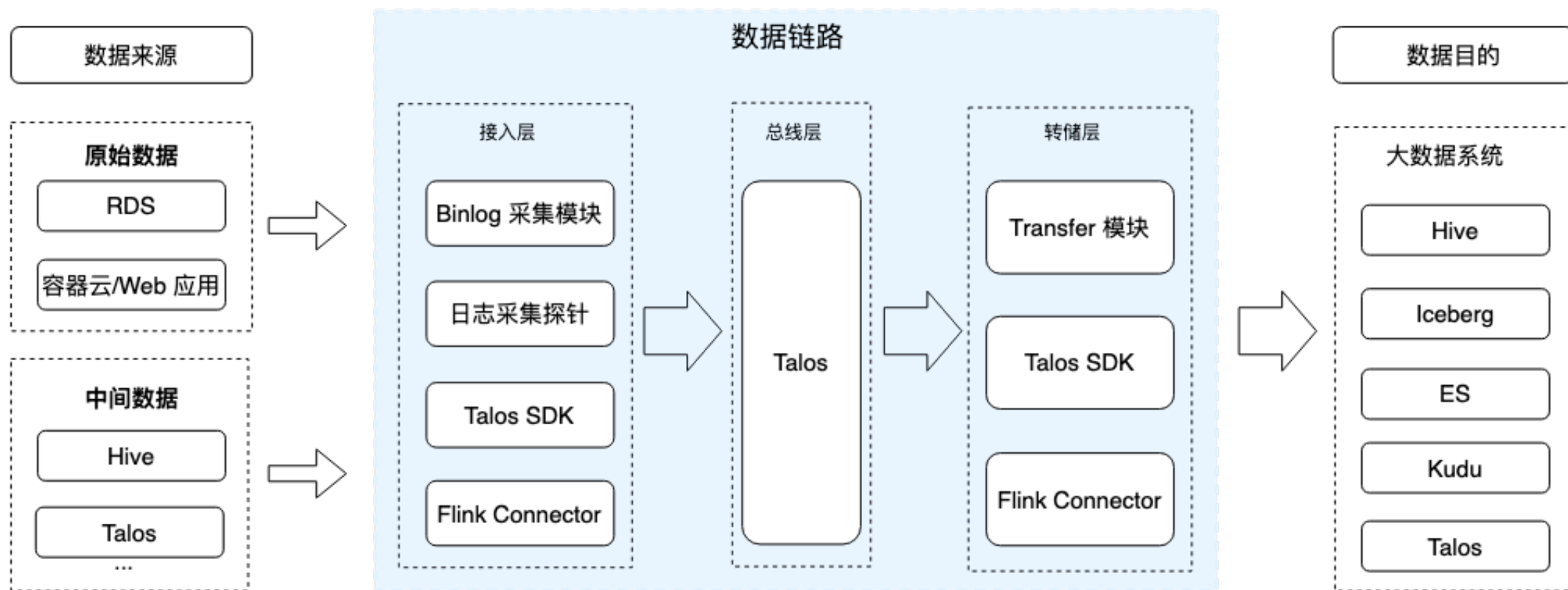
小米高级软件开发工程师



Contents

- 系统介绍
- 场景应用
- 运营升级
- 未来可期

系统介绍 - 计算平台

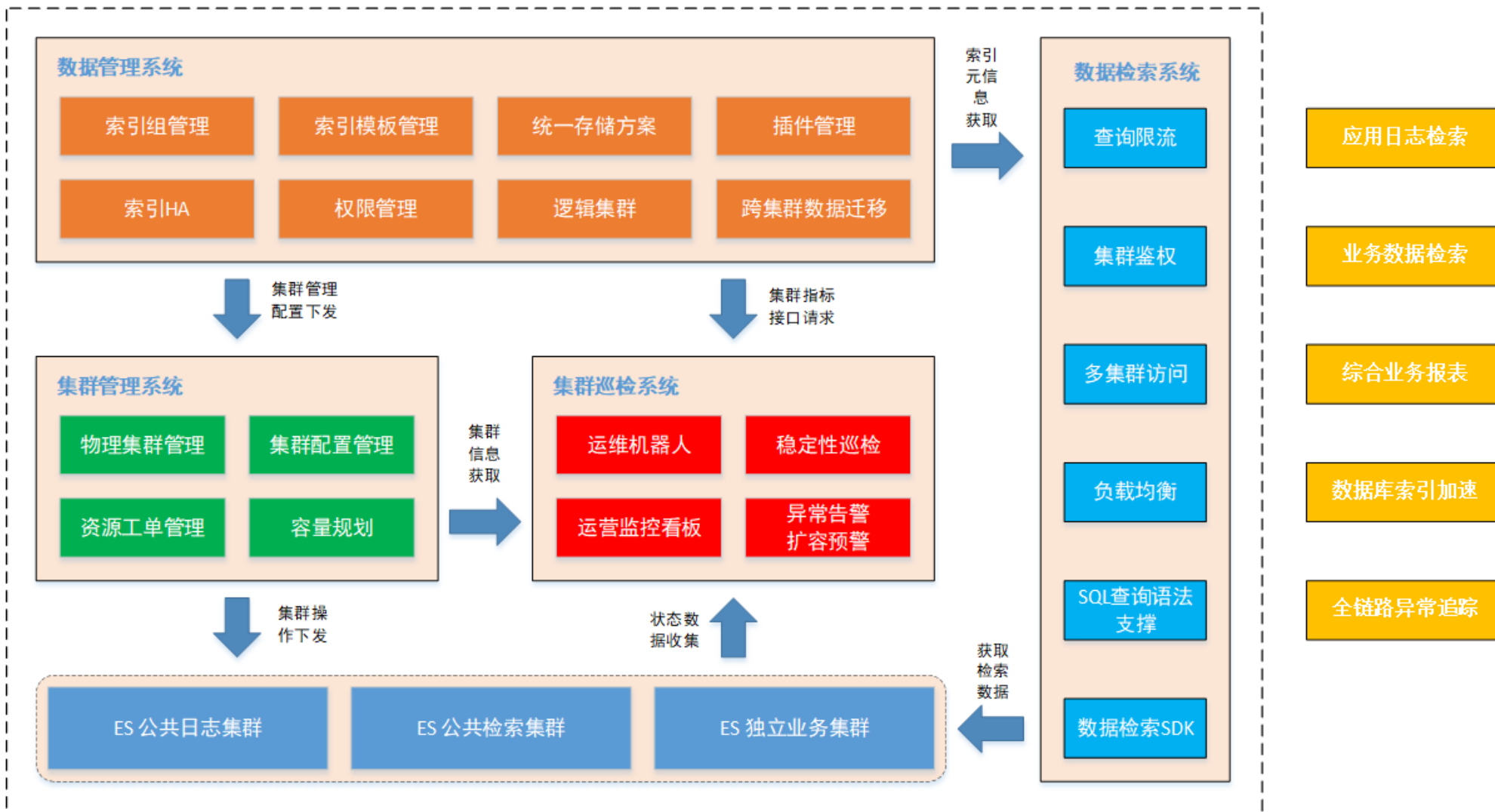


系统介绍 – ES的问题与挑战

- 业务日益增长的检索需求与稳定性低的 ES 系统之间的矛盾
- 逐步扩大的集群规模与人工运维效率低下之间的矛盾
- 日志实时写入量级大，ES 集群大面积节点429 write_reject写入请求拒绝
- 元数据压力大，凌晨大量auto createindex及 put_mapping请求导致master节点pending_task过多OOM
- 集群CPU负载居高不下，心跳同步请求超时，多个数据节点被同时踢出集群，分片恢复慢
- 每日上百个ES集群告警，超过30% P1以上级别告警
- 电商实时检索场景，可用性无法保证99.99%
- ...



系统介绍 – 架构升级



系统介绍 – 服务现状



50+

集群数

1000万+

每秒写入文档数

9000 亿

单日文档增量

750+

节点数

10万亿

总文档数

7+ PB

总存储量



Contents

- 系统介绍
- 场景应用
- 运营升级
- 未来可期

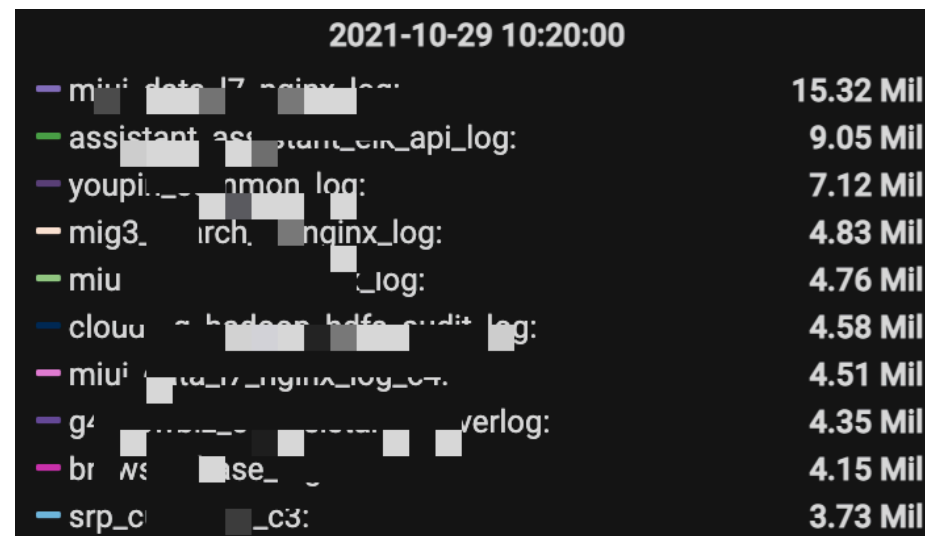
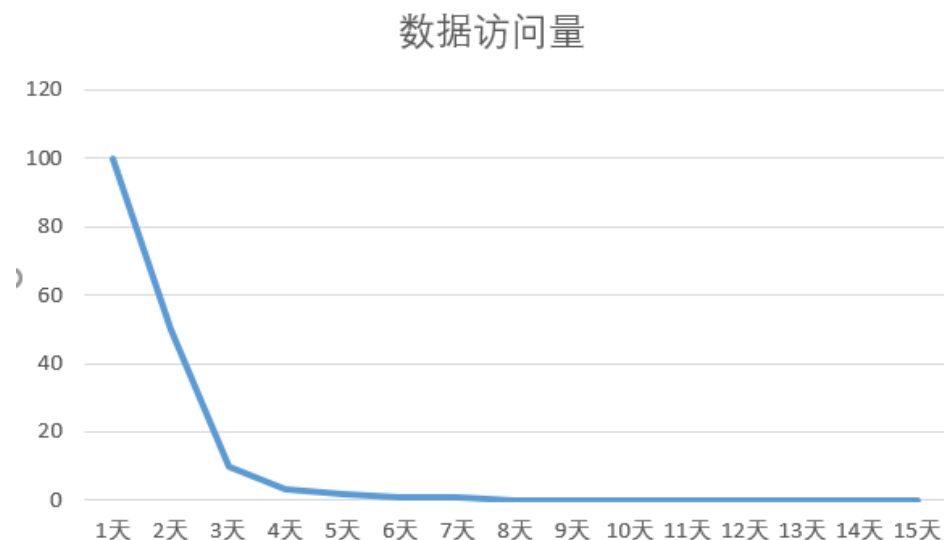


场景应用 – 日志场景

场景一：全链路业务系统日志接入，支撑业务系统问题定位排查

场景特点：数据流量大、实时性要求高、日志数据存在时序性特征、具备冷热数据特征

问题一：实际场景中，部分业务数据加副本单日数据增量20TB+，单日索引存在超过680个分片且分片全部处于可读可写状态。在大量接入，同时有检索服务时，CPU Load高，内存压力大，如何解决这些问题？持续扩容集群？





场景应用 – 日志场景

冷热分离

- 集群多机型支持
- 集群多角色节点部署隔离
- 冷热数据存储隔离
- 冷热数据索引配置

依据数据特点
降低存储成本

统一存储

- 索引RollOver滚动策略
- 索引Mapping/settings预计算
- 索引生命周期管理
- total_shards_per_node
- Force merge 线程数控制
- Frozen / Delete

控制资源并发
降低集群负载

事务日志

- 异步translog同步机制
- 优化translog同步时间间隔及缓存大小
- 写入性能提升10%

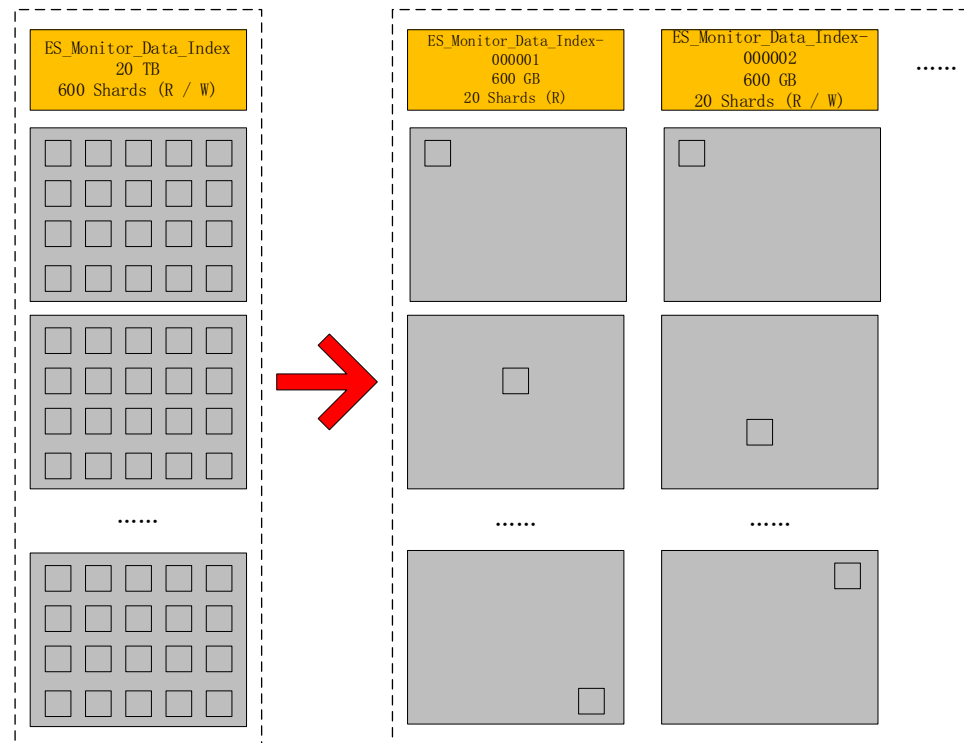
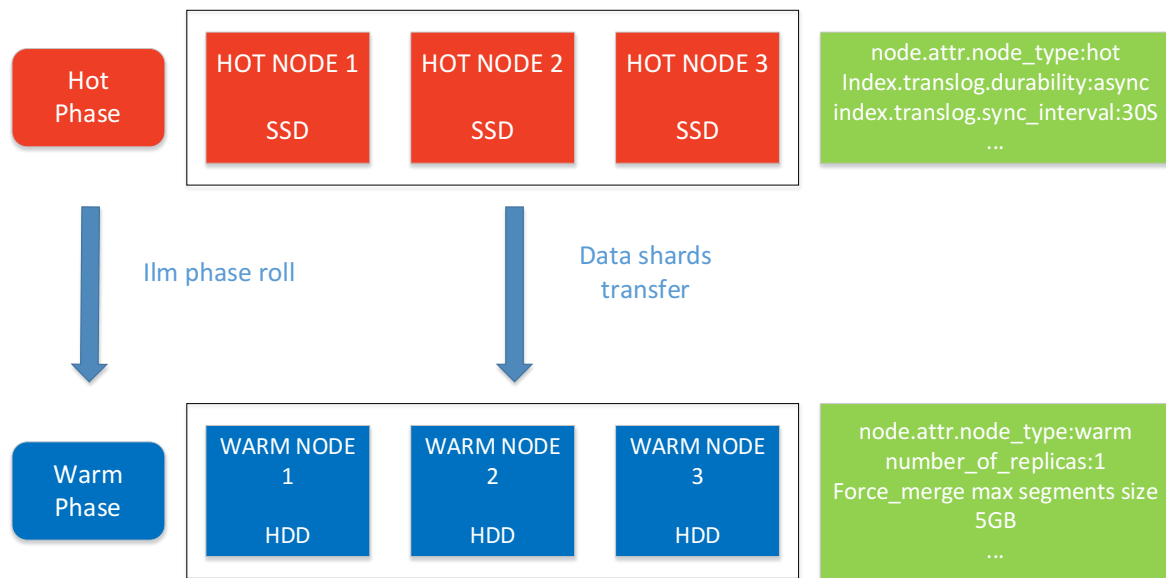
异步写入事务日志
提升写入性能

其他优化

- JDK 13 之前版本, CMS有着比G1更优的稳定性
- 解决Log4j ConsoleAppender引发的死锁问题
- 解决7.6.2版本在并发较大时出现Too Many ScrollContext问题代码Bug

代码优化
解决稳定性异常

场景应用 – 日志场景



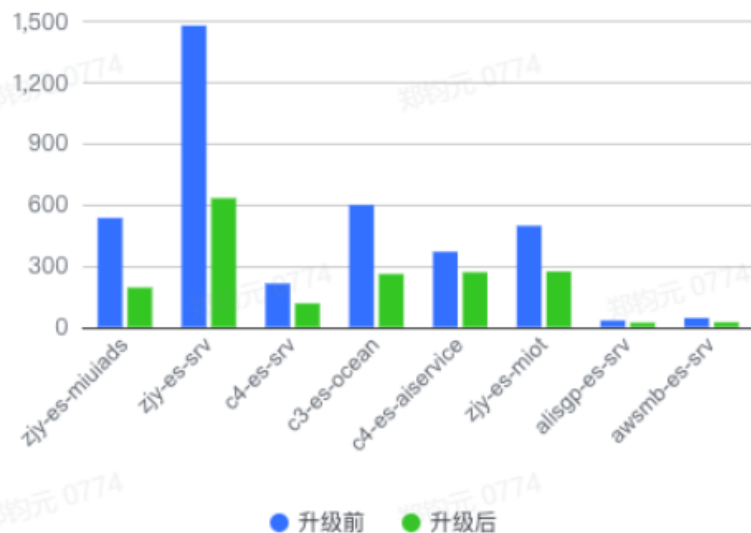
冷热数据架构，充分结合数据特征节约成本

统一存储方案，采用索引滚动与精细化切片策略。充分利用技术手段保证性能

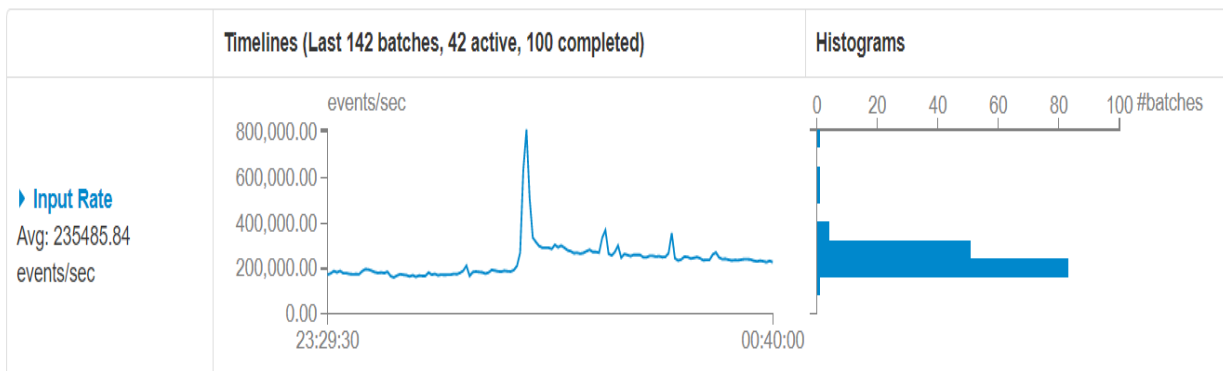
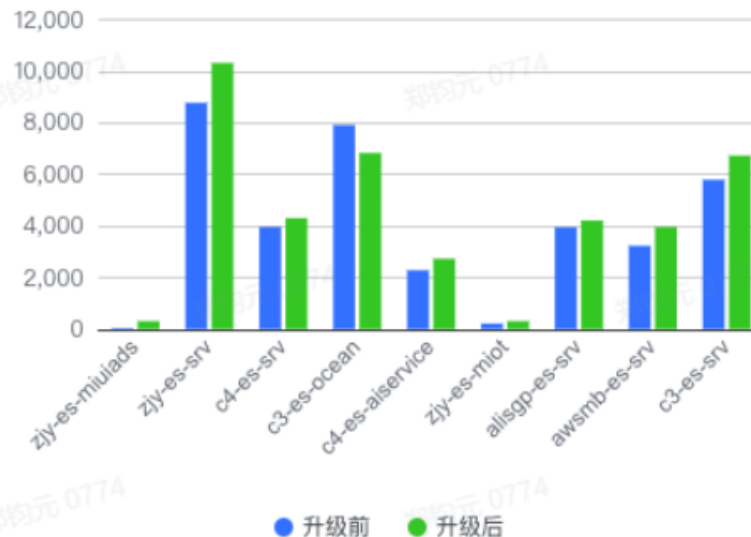
场景应用 – 日志场景



各集群load 1min指标统计



各主要集群存储方案落地前后索引数对比



整体写入性能**提升1倍**。双11零点日志数据流量**攀升4倍**。在**2000万 EPS**的数据写入流量下，提供**SLA 99.95%** 稳定服务。

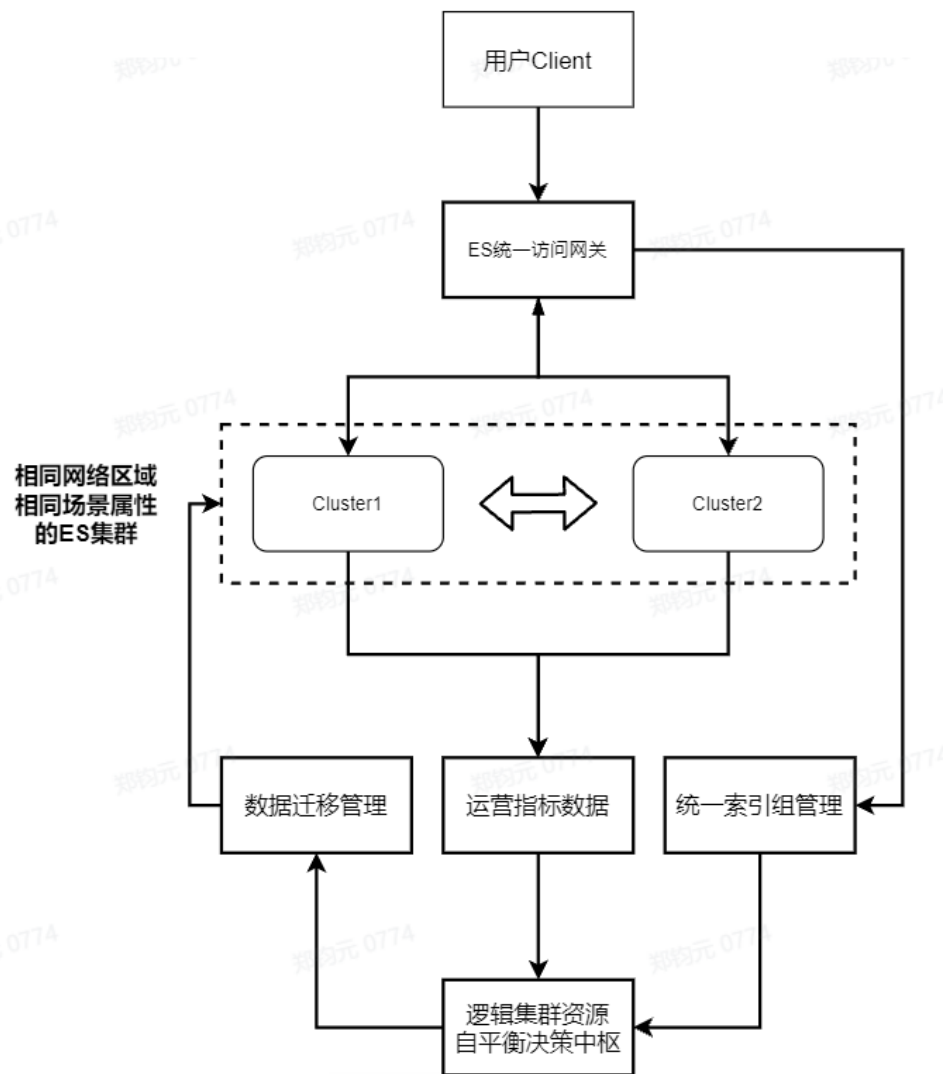


场景应用 – 资源自平衡

问题二：在解决了写入性能问题后，数据接入业务大量增加，内部公共日志集群索引个数以及分片个数不断增长，集群元数据压力陡增：

- **Cluster_state**变更后增量信息同步超时，导致部分节点离线，影响线上集群SLA
- 大量索引 ILM 生命周期控制请求同时触发，导致master节点pending_task堆积，非urgent级别的任务执行周期变长，影响集群管理请求响应效率

场景应用 – 资源自平衡



- 在相同网络区域与相同场景属性的物理集群间组合逻辑集群
- 由各个集群的**物理硬件性能指标和ES索引分片计数、分片大小、生命周期策略、存储总量及增量**等运营数据指标作为基础运算数据
- 对基础数据分集群计算**加权平均负载**，对指标数据进行向量化化简后完成奇偶装箱，最终达到逻辑集群内各物理集群的资源相对均衡
- 逻辑集群内新增索引基于当前各物理集群加权负载进行物理集群选择



场景应用 – 资源自平衡

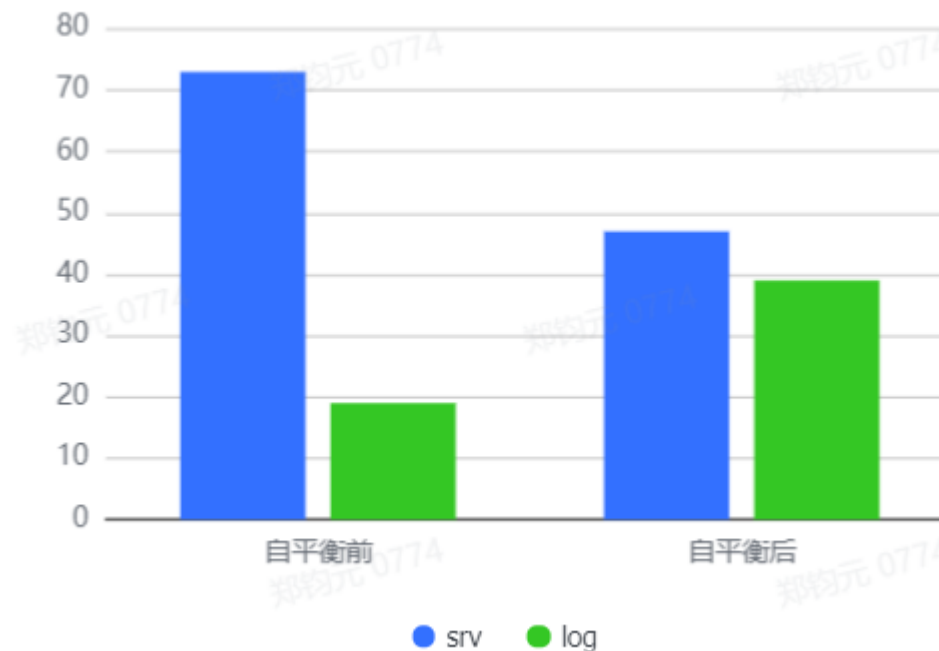
自平衡前:

集群名称	srv	log
存储量	1.1PB	450TB
日增量	67.2TB	19.4TB
索引个数	10276	2074
分片数量	65448	13090

自平衡后:

集群名称	srv	log
存储量	854TB	694TB
日增量	47.2TB	39.4TB
索引个数	7283	5067
分片数量	43172	35366

各集群加权负载对比



- 在实现自平衡后，各物理集群间负载差异从3.84 : 1下降至1.2 : 1。资源利用率更均衡。
- 实现索引平衡分布，单个集群元数据压力更平均。
- 解决单个集群扩展性上限

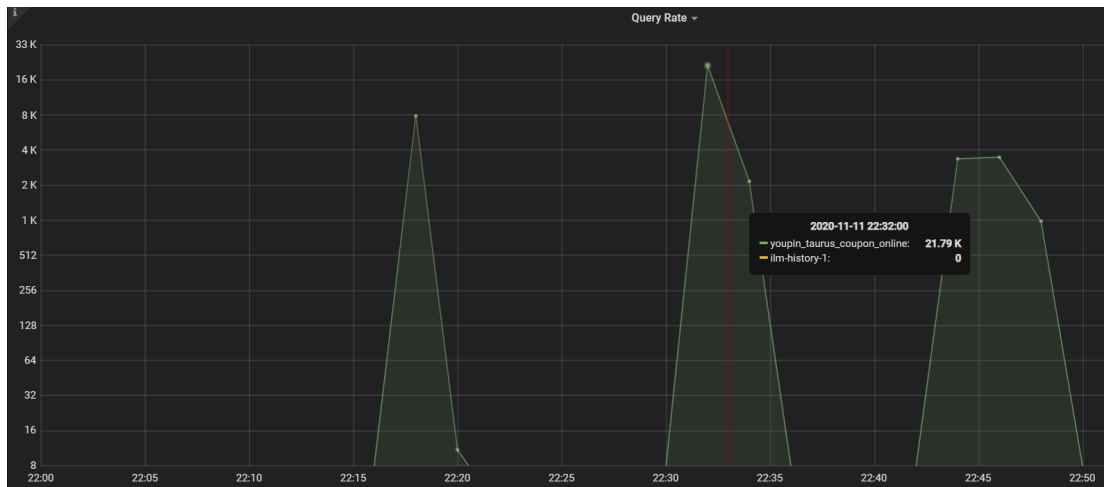


场景应用 – 跨集群HA

场景二： 实时数据检索，商品信息、视频信息、优惠券查询等

场景特点： 数据相对固定，数据量不大，没有频繁的写操作。支撑在线业务实现快速检索或实现数据库加速相关多级索引支撑

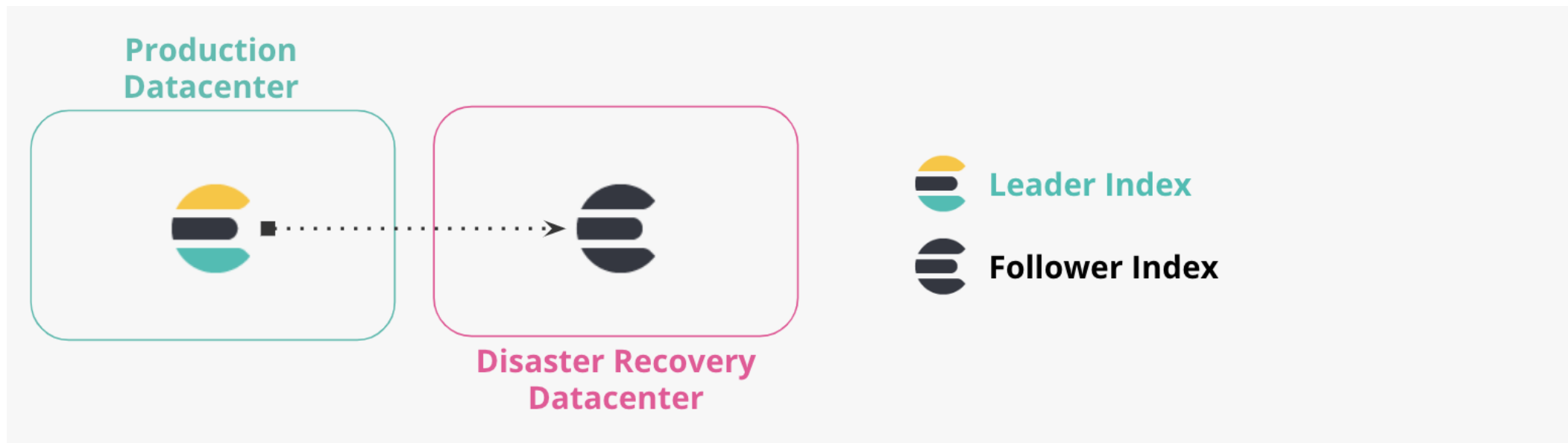
问题： 实际场景中，索引数据量不大，没有频繁的写操作。但是数据检索实时QPS高，对业务可用性要求高。如何在保证检索性能的同时提升索引HA的能力？



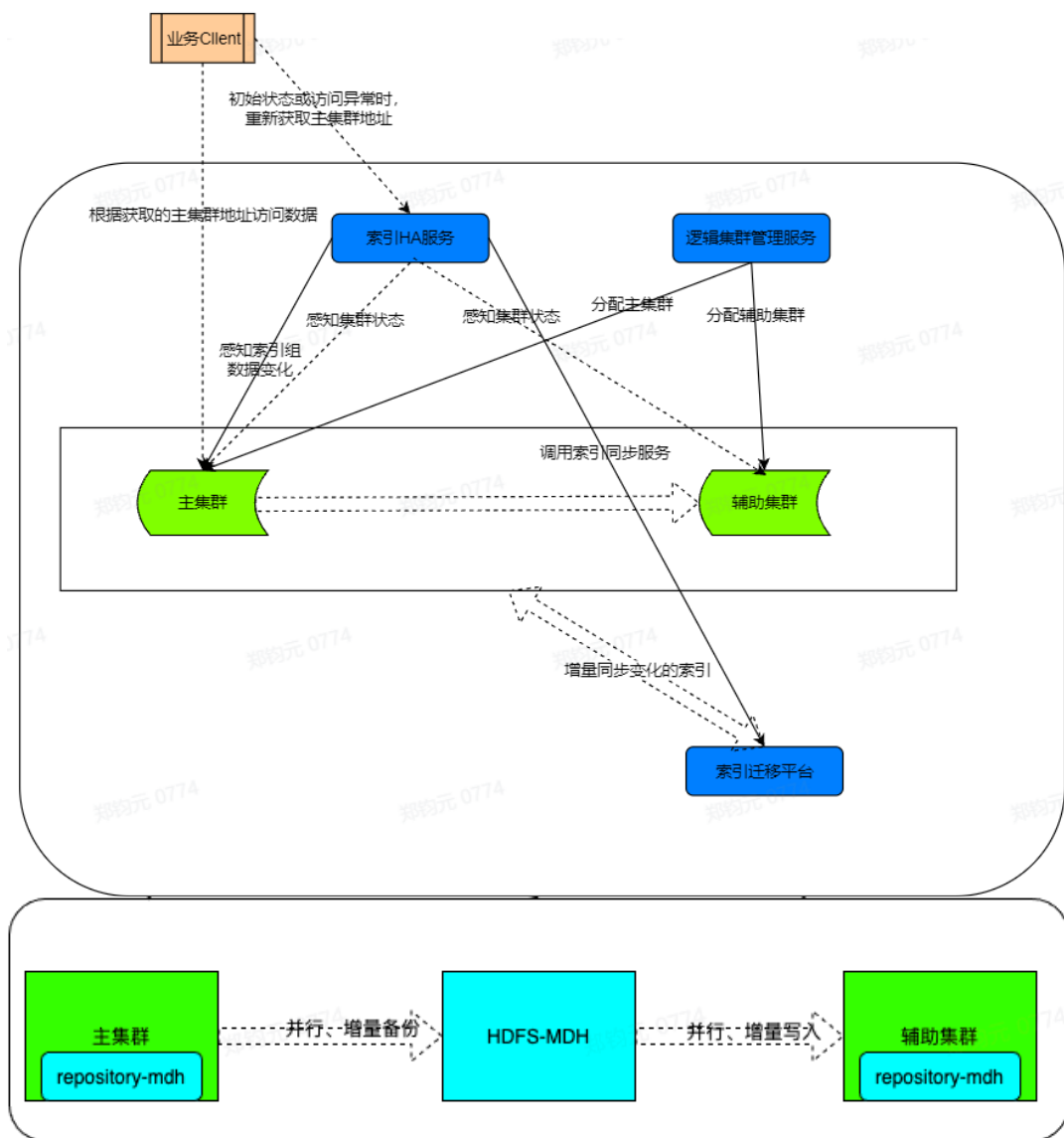
场景应用 – 跨集群HA



方案一：Elasticsearch跨集群复制功能CCR



场景应用 – 跨集群HA



- 索引在集群间的准实时数据同步
- 实现跨数据中心的灾难恢复和高可用需求，保障服务持续可用。
- 基于小米HDFS定制的ES插件 repository-mdh
- 实现HA索引增量数据同步
- 实现索引数据冷备策略

场景应用 – 跨集群HA



	自研索引HA服务	CCR服务
同步级别	索引组级别	索引级别
同步策略	增量同步	增量同步
同步方向	主动	被动
数据访问	主集群可读可写 从集群可读	Leader索引可读可写 Follower索引可读
异常处理	自动切换主从同步方向	Follower可读不可写， 需手动切换同步方向
连接切换	客户端接口自动实现集 群连接切换	需在客户端实现集群切 换逻辑



Contents

- 系统介绍
- 场景应用
- 运营升级
- 未来可期



运营升级 – 集群自动巡检

2021-10-28 Elasticsearch Inspection Report

Title	Cluster	Status	Time	MaxTime	MinTime
red	red	red	22	2021-10-29 07:17:05	2021-10-28 08:23:03
	red	red	8	2021-10-29 05:20:51	2021-10-28 10:27:51
	red	red	7	2021-10-29 07:13:51	2021-10-29 06:45:49
	red	red	4	2021-10-28 22:09:55	2021-10-28 18:09:56
	red	red	4	2021-10-28 15:12:42	2021-10-28 15:09:42
	red	red	4	2021-10-29 07:10:43	2021-10-28 08:18:43
	red	red	4	2021-10-29 05:37:53	2021-10-28 09:07:12
	red	red	3	2021-10-28 21:44:52	2021-10-28 21:42:52
	red	red	2	2021-10-29 04:31:42	2021-10-28 23:15:42
	red	red	1	2021-10-28 08:06:07	2021-10-28 08:06:07

Title	Cluster	IndexPrefix
error_bill	staging	miuiat-63
	staging	miuiad_dba_s_meltuan_257
	srv	cloud_container_staging_al-redis-on-k8s_694-to-696
	srv	cloud_
	srv	cloud_staging_al-redis-on-
	srv	cloud_container_staging_al-redis-on-k8s_696-to-694
	srv	cloud_cor_er_staging_al-redis-on-k8s_696-
	srv	cloud_omainer_staging_al-redis-on-k8s_696-to-694
	srv	cloud_omainer_staging_al-redis-on-k8s_696-to-694
	srv	cloud_omainer_staging_al-redis-on-k8s_696-to-694

Title	Day	Bills	Method	UploadTime	Message
upload_result	2021-10-28	50001	append	2021-10-29 00:30:38	upload success
upload_result	2021-10-28	4778	append	2021-10-29 00:30:44	upload success

Title	Cluster	Slo	Sli
goc	zly-es-srv	99.95	99.9923
	ali-es-srv	99.95	99.9972
	aws-es-srv	99.95	99.9938
	c3-es-srv	99.95	99.9937
	es-srv	99.95	99.9986
	es-srv	99.95	99.9972
	es-srv	99.95	99.9979
	es-srv	99.95	99.9979
	es-srv	99.95	99.9979
	es-srv	99.95	99.9757

Title	Cluster	Reason
warm_cluster	az-es-srv	cluster 75.00%hotnode disk low, duration beyond 8 hour
	zly-es-mitor	The number of cluster shards exceeds the expansion threshold
	es-srv	Cluster heap memory usage exceeds the expansion threshold, cluster 92.31%warmnode heap cost more than 70.00%, duration beyond 8 hour
	es-srv	cluster 100.00%warmnode heap cost more than 70.00%, duration beyond 8 hour
es-srv	cluster 72.73%warmnode heap cost more than 70.00%, duration beyond 8 hour	

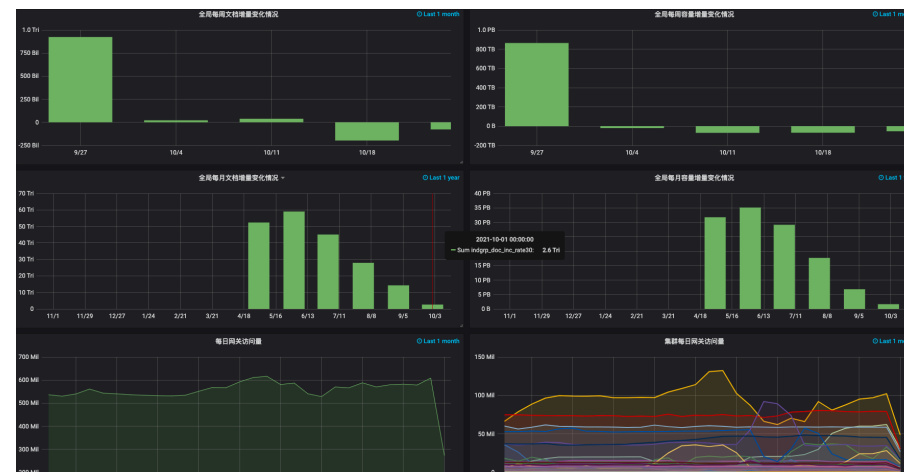
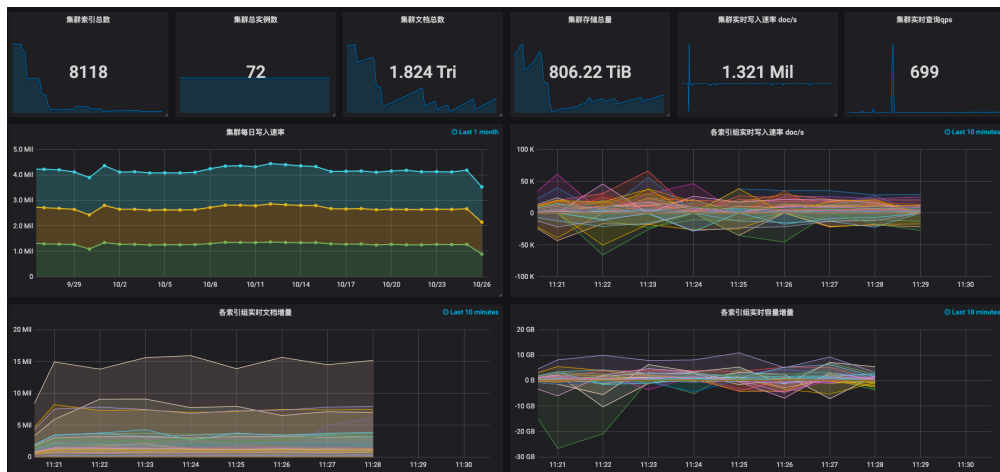
Title	HostList

- 各集群 CPU Load 时间阈值及资源阈值过载巡检
- 各集群堆内存占用资源阈值过载巡检
- 各集群存储占用资源阈值过载巡检
- Master 节点 Pending Task 阈值巡检
- 集群异常状态时间范围统计
- 维修或exclude节点列表
- 线上 P1 以上级别单日告警汇总
- 存储量过大分片、分片数过多索引阈值巡检
- 生命周期异常索引及空索引巡检
- 索引数过多索引组巡检

13项巡检策略充分保障集群稳定可用



运营升级 – 实时数据监控



- 全局及各集群实时写入流量监控
- 全局及各集群实时检索请求监控
- 全局及各集群数据存储变化趋势监控
- 全局及各集群存储文档数趋势监控
- 全局及各集群索引及分片个数变化趋势监控
- 各集群异常索引组监控

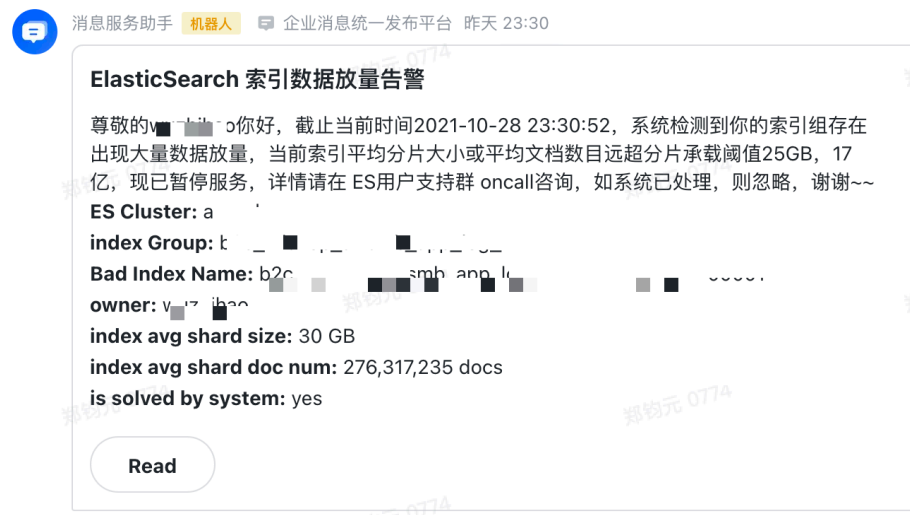
- 各索引组实时写入流量监控
- 各索引组实时访问读取流量监控
- 各索引组实时存储量变化监控



运营升级 – 告警推送



- 与小米办公和 Falcon 系统形成一体化异常告警体系
- 内核级别异常埋点检测
- 从字段解析异常、写入堆积异常、分片过大、分片 doc 数过多等多维度告警规则配置
- 实时精准推送给索引所属业务 owner，先于业务发现问题实现可用性风险预警





Contents

- 系统介绍
- 场景应用
- 运营升级
- 未来可期

未来可期



- 基于 searchable-snapshot 的成本优化策略
- 存储计算分离架构
- 逻辑集群在实时检索场景的扩展性及负载应用
- 智能化集群异常分析与自愈机制



谢谢!

