Tencent 腾讯 | 腾讯云

云原生的Elasticsearch服务

腾讯云大数据及人工智能产品中心

张彬 ethanbzhang

ethanbzhang@tencent.com

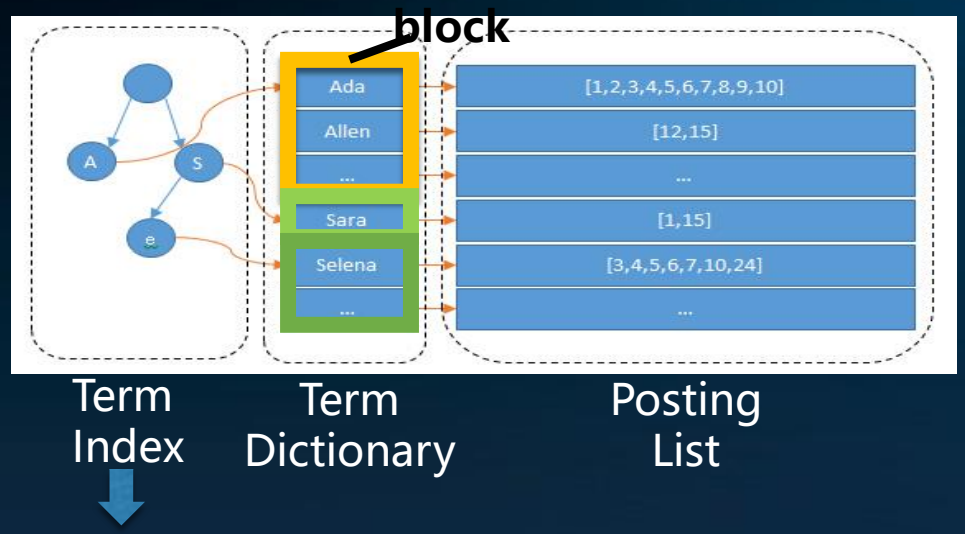
腾讯云Elasticsearch负责人

曾就职于新浪、乐视等公司，从事分布式redis，大数据平台、画像及推荐系统的开发

在日志分析及检索、Hadoop/Spark生态、分布式存储等方向有深入的研究

目录

全文索引优化



- FST：Finite State Transducer
- 同一个block中的terms，共用FST的同一路径
- 常驻JVM内存（Segment Memory），无法被回收，加速检索

优化

词典表改为Off-heap

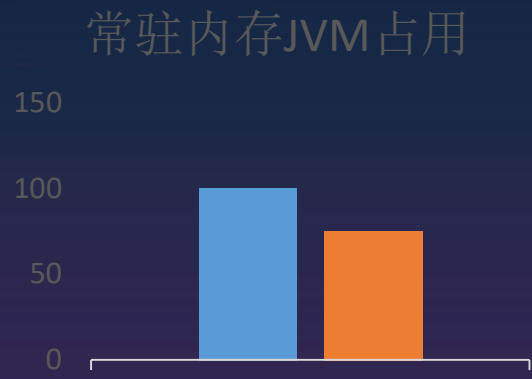
Elasticsearch 7.x

Lucene8.0 lazy loading

LUCENE-8635

腾讯云ES内核优化

改为LRU方式实现 支持 5.X 6.X



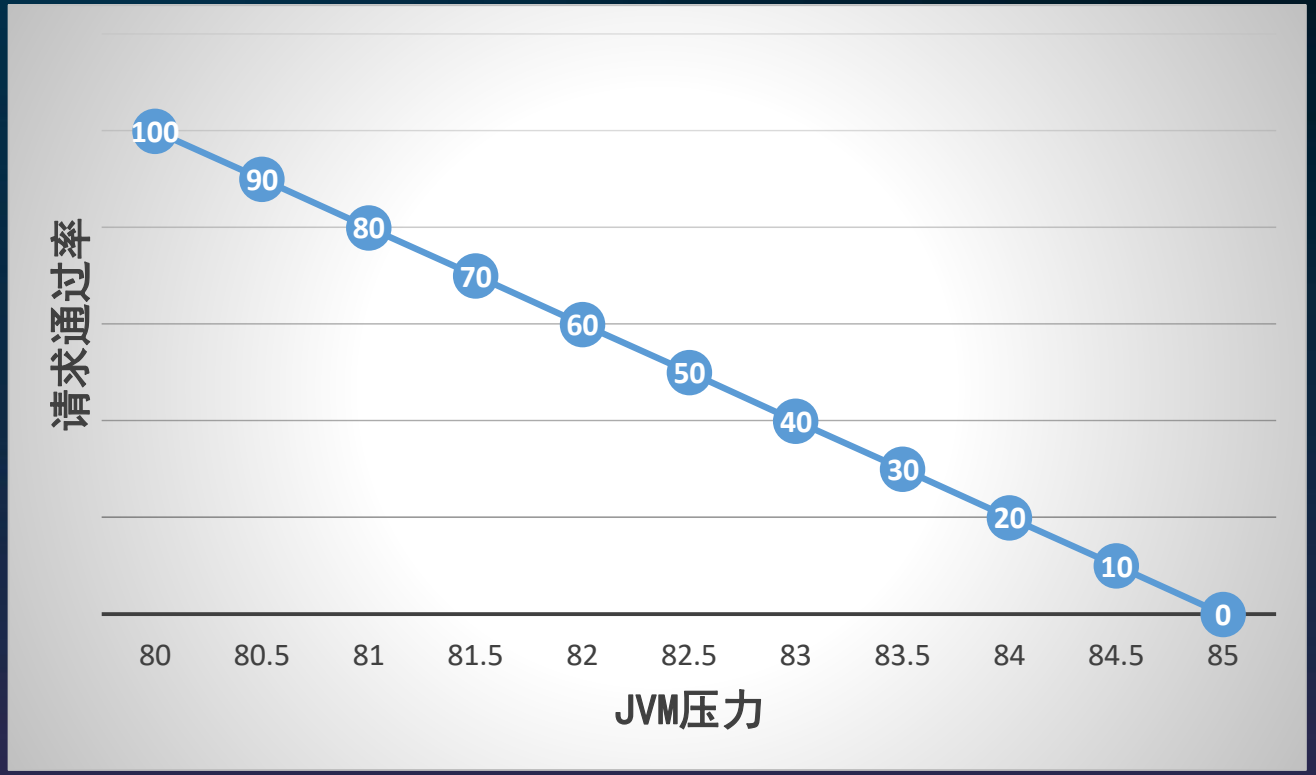
稳定性优化

原生熔断器

- 单查询、请求体控制
- 粗粒度查询缓存控制

优化熔断器

- 根据JVM老年代使用率控制请求
- 线性梯度熔断
- 内存高压力场景下始终保证集群稳定的写入吞吐



大时间跨度查询性能优化


```
GET host_monitor_60/_search?routing=b3528e3a
{
  "size": 65536,
  "stored_fields": "_none_",
  "docvalue_fields": ["cpu_usage"],
  "sort": {"timestamp": {"order": "asc"}},
  "query": {
    "bool": {
      "filter": [
        {
          "range": {
            "timestamp": {
              "gte": "2018-11-1 00:00:00",
              "lte": "2018-11-8 00:00:00",
              "format": "yyyy-MM-dd HH:mm:ss",
              "time_zone": "+08:00"
            }
          }
        },
        {
          "term": {
            "host_uuid": "84c0d3a9-dbdf-4e02-9460-2a86c05afdfc"
          }
        }
      ]
    }
  }
}
```

1. 7天超10亿条数据，缓存耗时间、空间
2. host数据分布分散，随机io高

- Query阶段根据查询子句的频次做缓存
- Fetch阶段抓取文档无序性，维护缓存队列

优化

Query阶段：根据查询cost过滤大范围的查询cache 降低查询毛刺

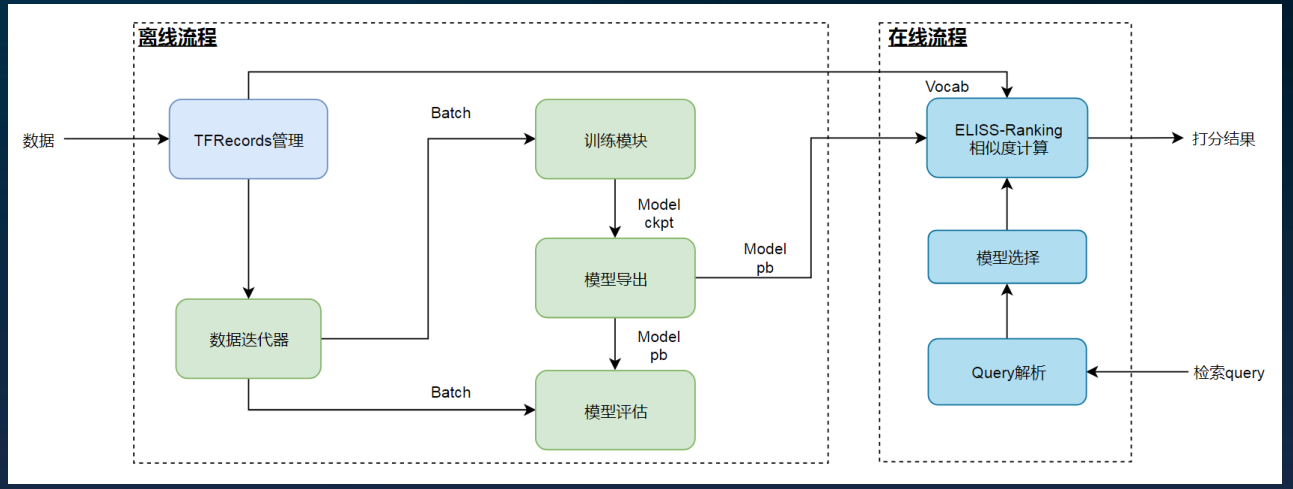
查询耗时 700ms -> 50ms

Fetch阶段：根据segment对文档id排序 提高cache利用率

查询性能 +20%

中文分词

- 支持IK中文分词插件、Pinyin分词插件
- 支持用户自定义词库上传，启用词库，停用词库，同义词库。自定义词库热加载



- 支持基于TensorFlow模型训练的NLP相似度打分插件
- ELISS-Ranking

目录

常见运维问题

- 磁盘总被写满，如何预估磁盘容量？
- 我的写入为什么很慢？
- JVM持续压满，最终导致OOM，如何避免？
- 怎样提高并发和吞吐？
- 怎样设置字段类型？

Disk Utilization

From JSON to indexed size on disk

- Replica num: at least one
- Index cost: 10%
- OS Reserved: 5%
- Cluster Reserved: 20%
- Safety Reserved: 15%

disk size = origin data * (1 + replica) * (1 + index cost) / (1– OS Reserved) / (1 – Cluster Reserved) / (1- Safety Reserved)

disk size = origin data * 3.38 (replica num==1)

1. _all
2. _source
3. keyword or text
4. best compression

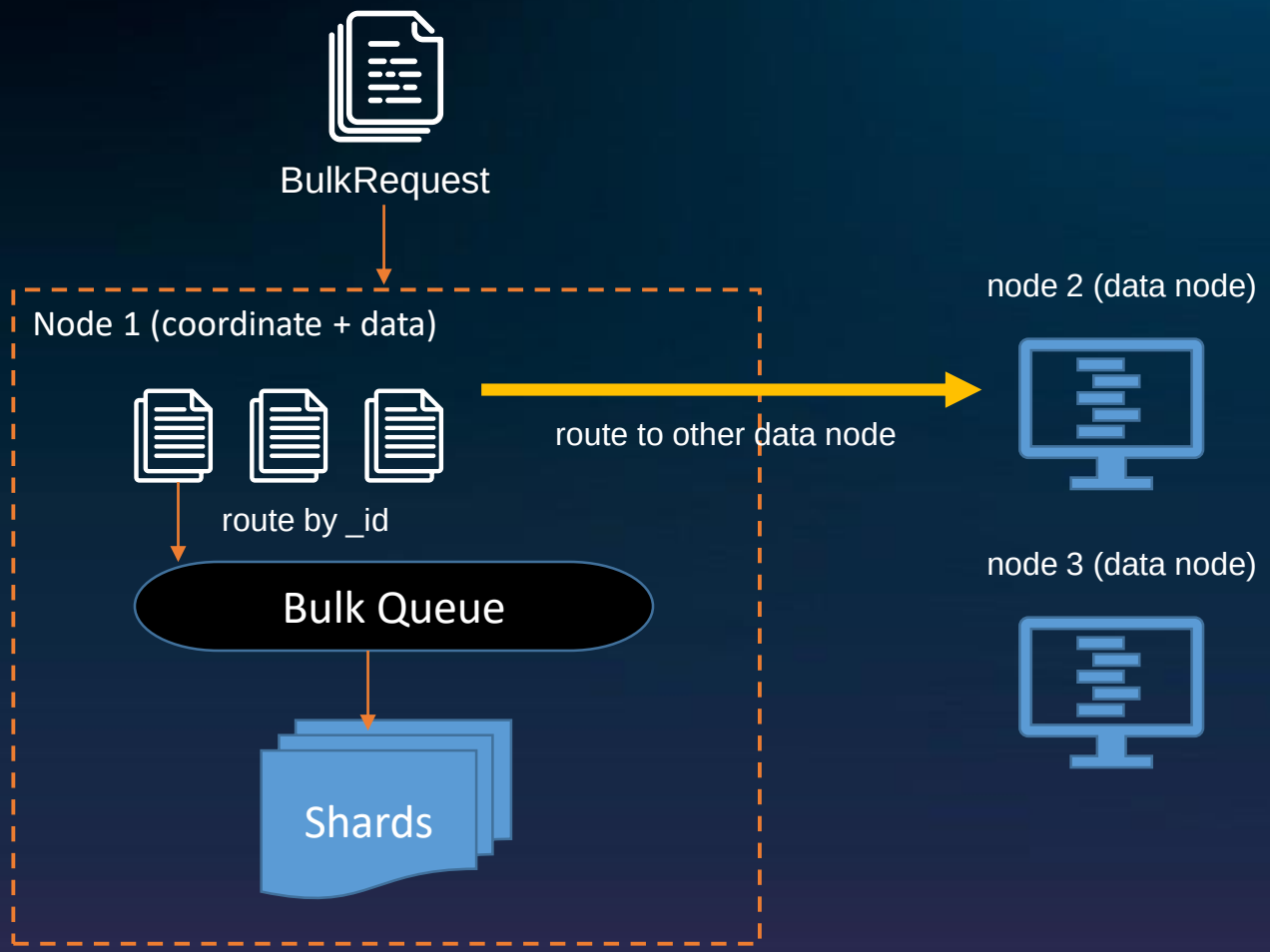
Memory Predicted

Memory Used

- Index Meta
- Shard Meta
- Segment Meta
- filter cache
- aggregations cache
- fielddata cache

	Ratio
Aggregation Search	1 : 8
Text Analyzed Search	1 : 40
Normal	1 : 24

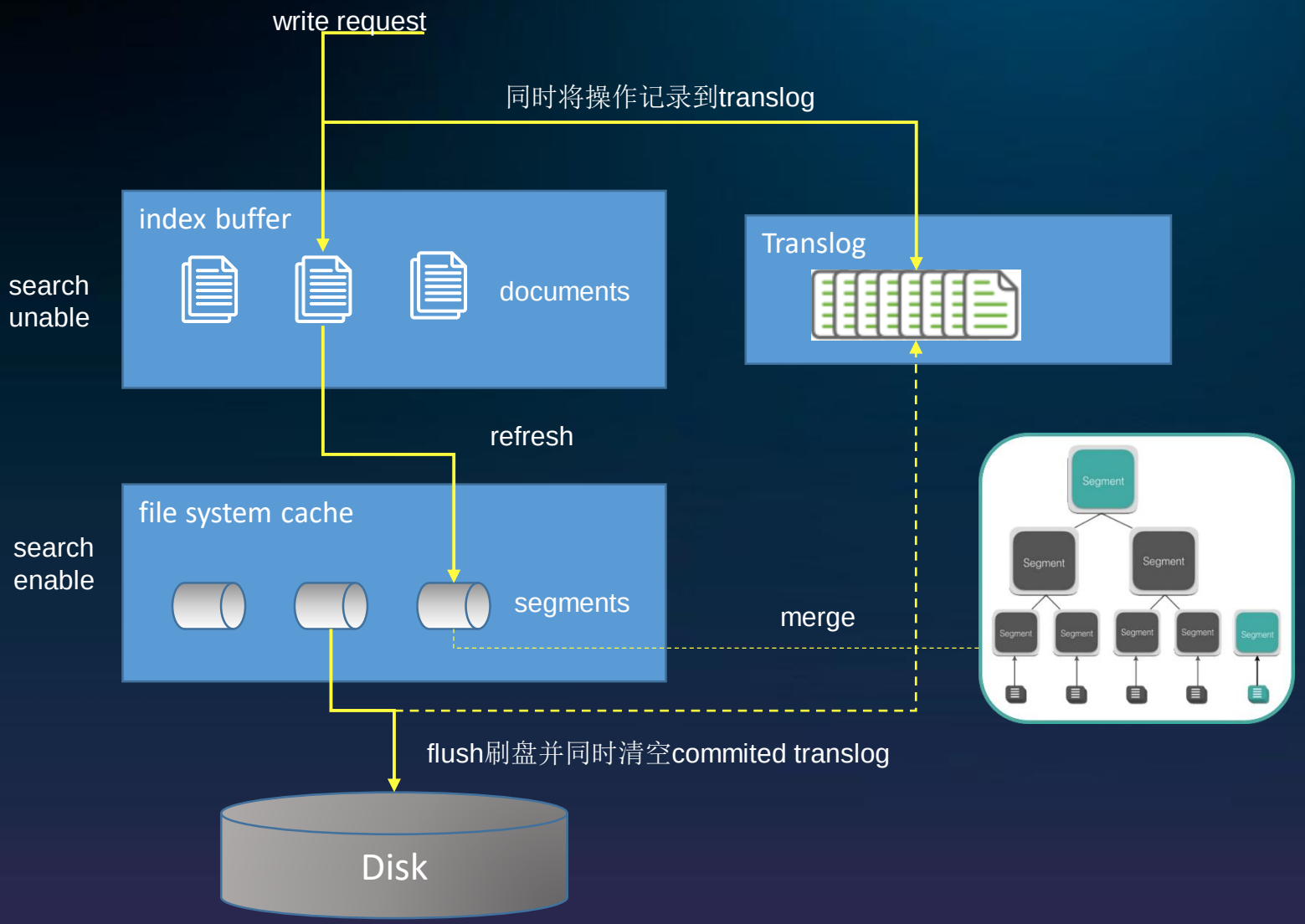
Shard Predicted



shard数过多
写入放大，bulk queue打满，导致拒绝率上升

shard数太少
不能很好的利用多核多节点的优势，并发不够，容易造成写入热点

写入性能优化



- **Translog**
  - Durability
  - sync_interval
- **Refresh**
  - make doc search enable
  - every 1 sec
- **Flush**
  - IndexWriter commit
  - Fsync index to segment file
- **Merge**
  - merge segments to big one

运维问题分析工具

Tencent 腾讯 | 腾讯云

- GET /_task?nodes=nodeId1,nodeId2
- GET /_nodes/nodeId1,nodeId2/hot_threads
- GET /_cat/indices?v&h=index,store.size,fielddata.memory_size&s=fielddata.memory_size:desc
- GET /_cluster/allocation/explain
- {"profile": true}

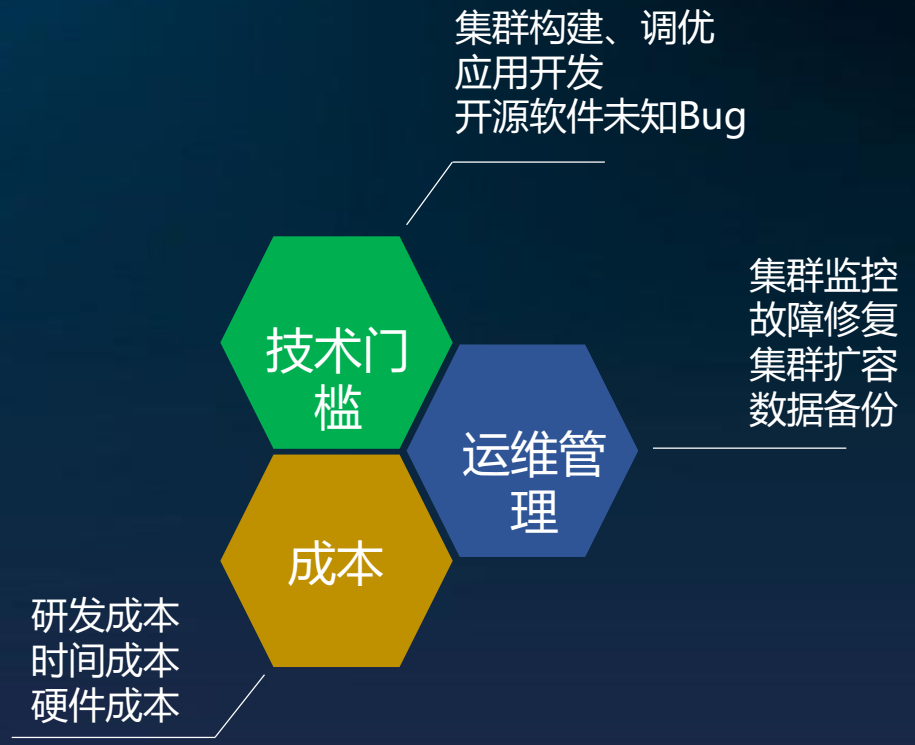
目录

Tencent 腾讯 | 腾讯云

- 腾讯云ES内核及功能优化
- 常见集群优化案例分析
- 弹性云原生功能介绍

Common customer issues

- 我的集群写入很慢，到底问题在哪儿？
- 怎样搭建一个多节点的集群，扩容和维护？
- 管理多个集群，监控和运维谁来管？
- 我的数据很多，如何在集群里面做冷热分离？
- 怎样做索引的管理？
- 怎样保证数据的安全和稳定，如何做数据备份？



智能运维系统

- 集群异常类
- 集群资源类
- 集群运维类
- 业务规范类

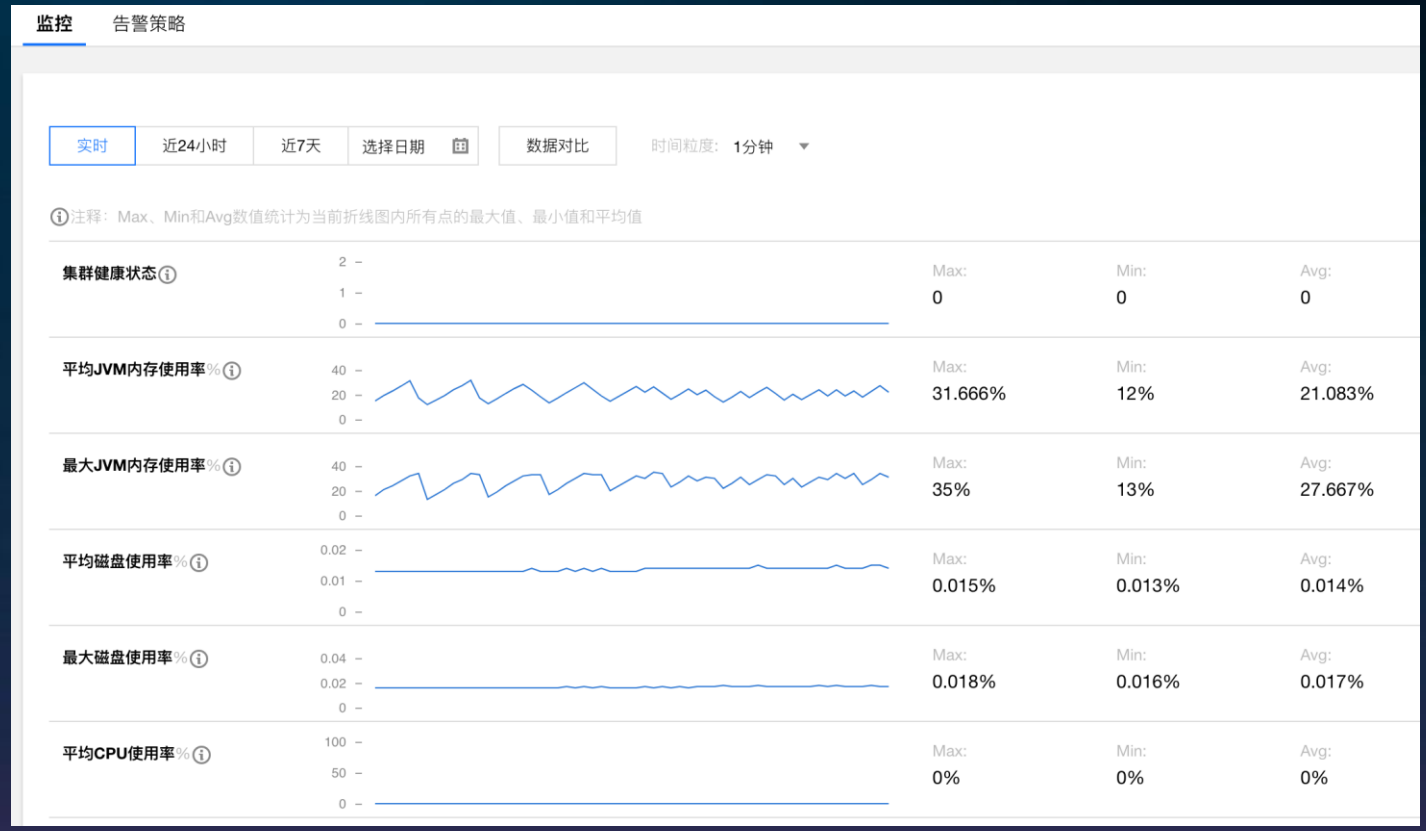
Scaling

- 支持按需对机型、存储或节点个数扩容及缩容
- 平滑扩缩容，业务无感知
- 升级re-allocation算法，避免新加节点负载过高

监控告警

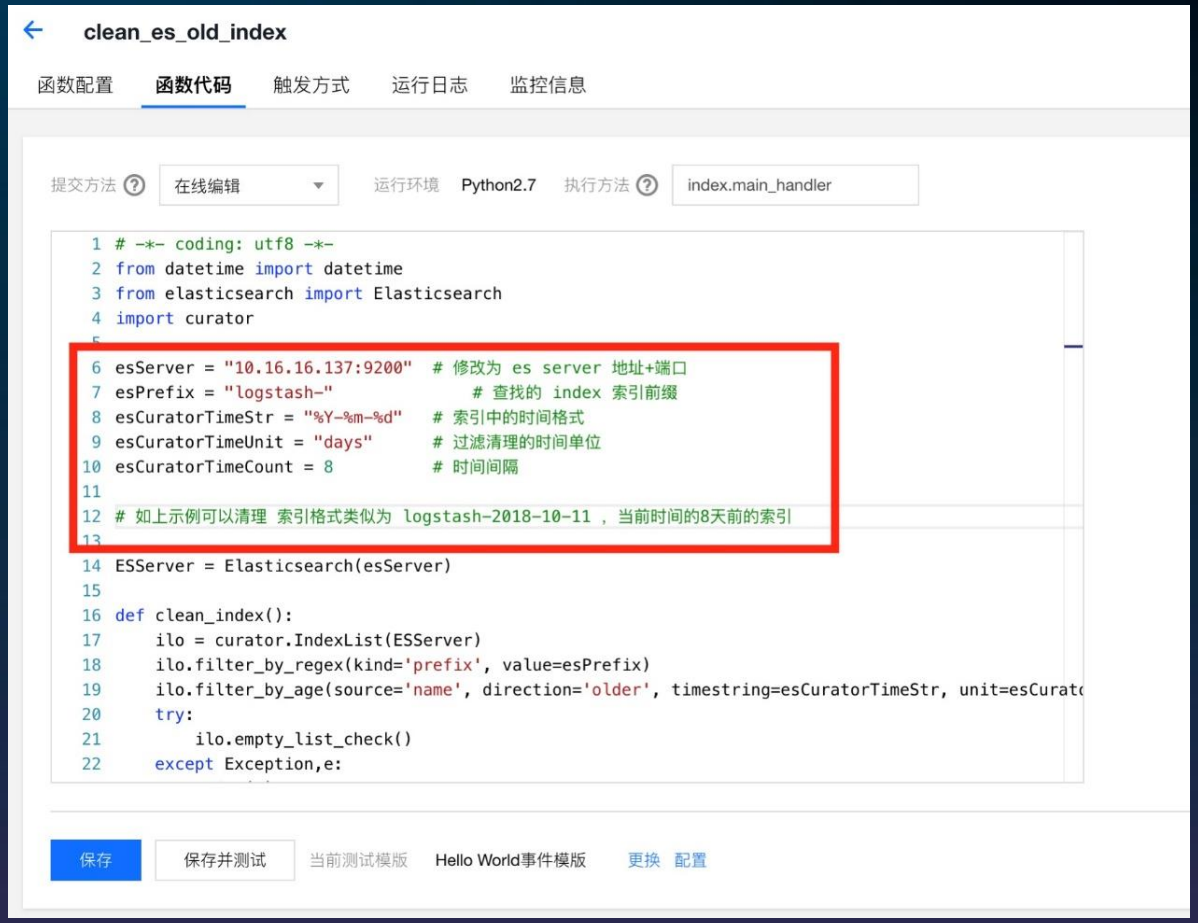
对集群各项指标的监控

短信、微信等告警策略，支持默认告警模板



结合无服务器函数做索引生命周期管理

- 支持索引冷热分离，索引关闭，索引清理，数据上卷，自动备份等等
- 操作方便，可简单套用模板
- 定时触发，条件触发，用户可配置通知机制，可查看执行日志



冷热分离

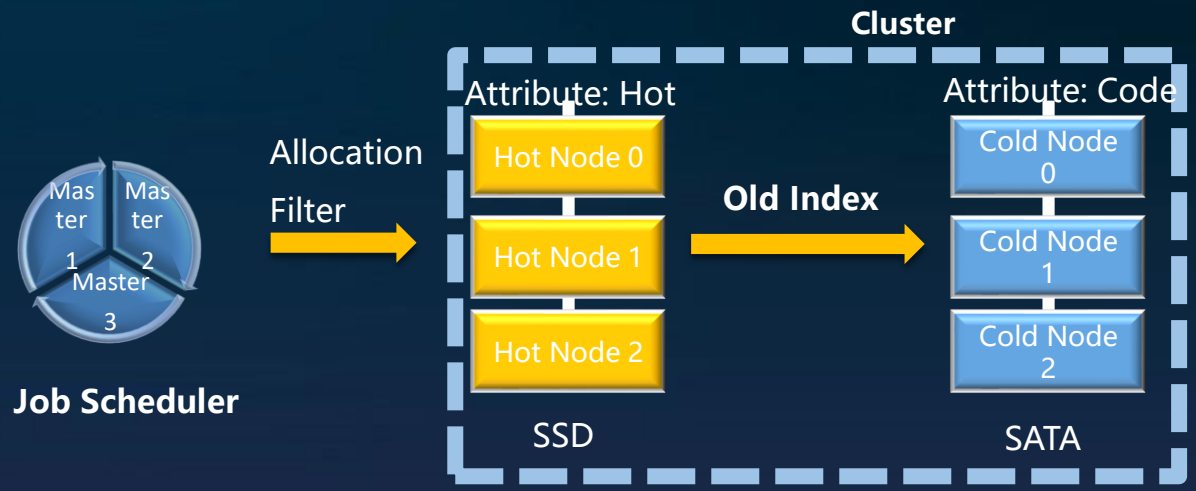
- **Node Attribute**

```
node.attr.temprature: hot

node.attr.temprature: cold
```

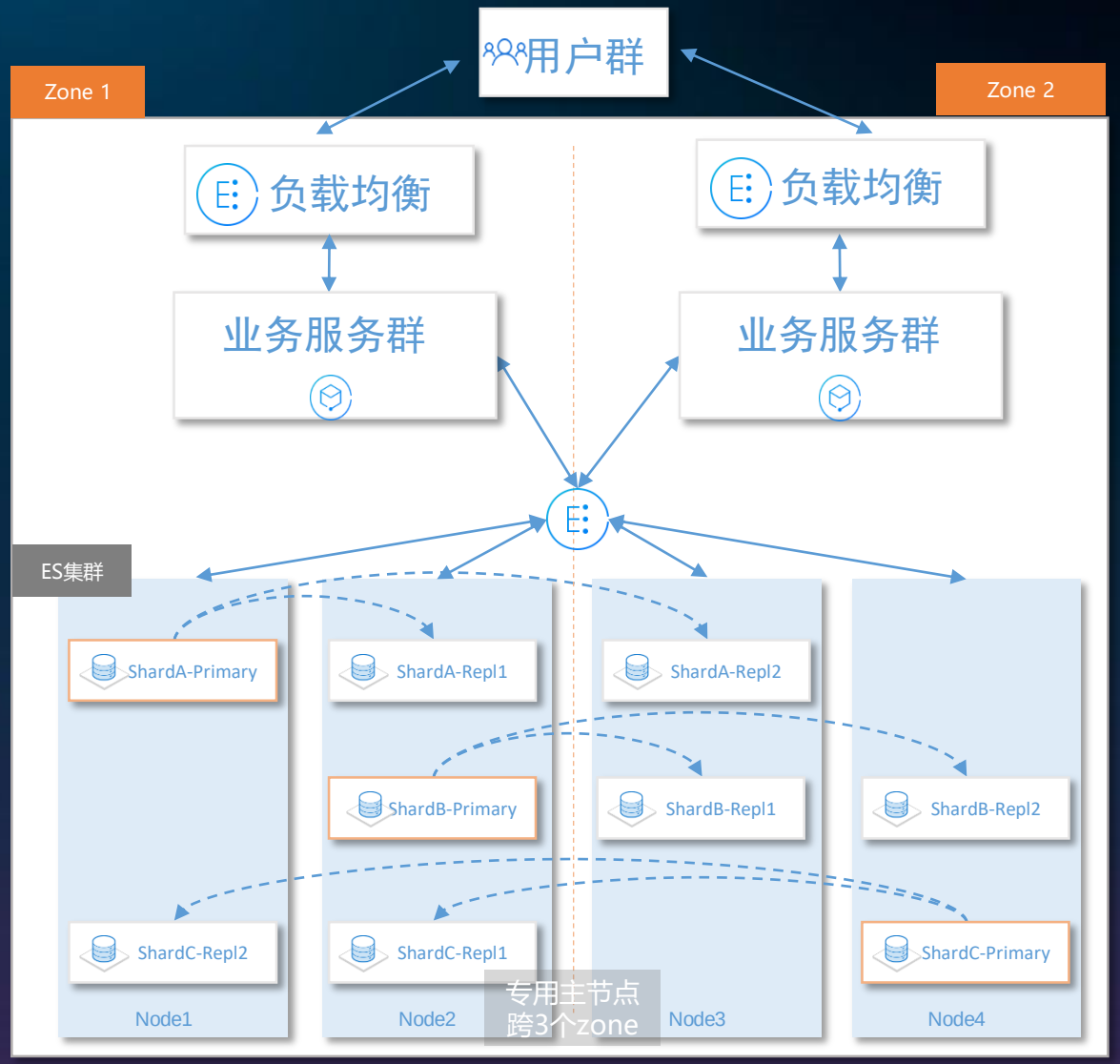
- **Index Allocation Filter**

```
PUT test/_settings
{
  "index.routing.allocation.include.temprature": "cold"
}
```



跨多可用区多活

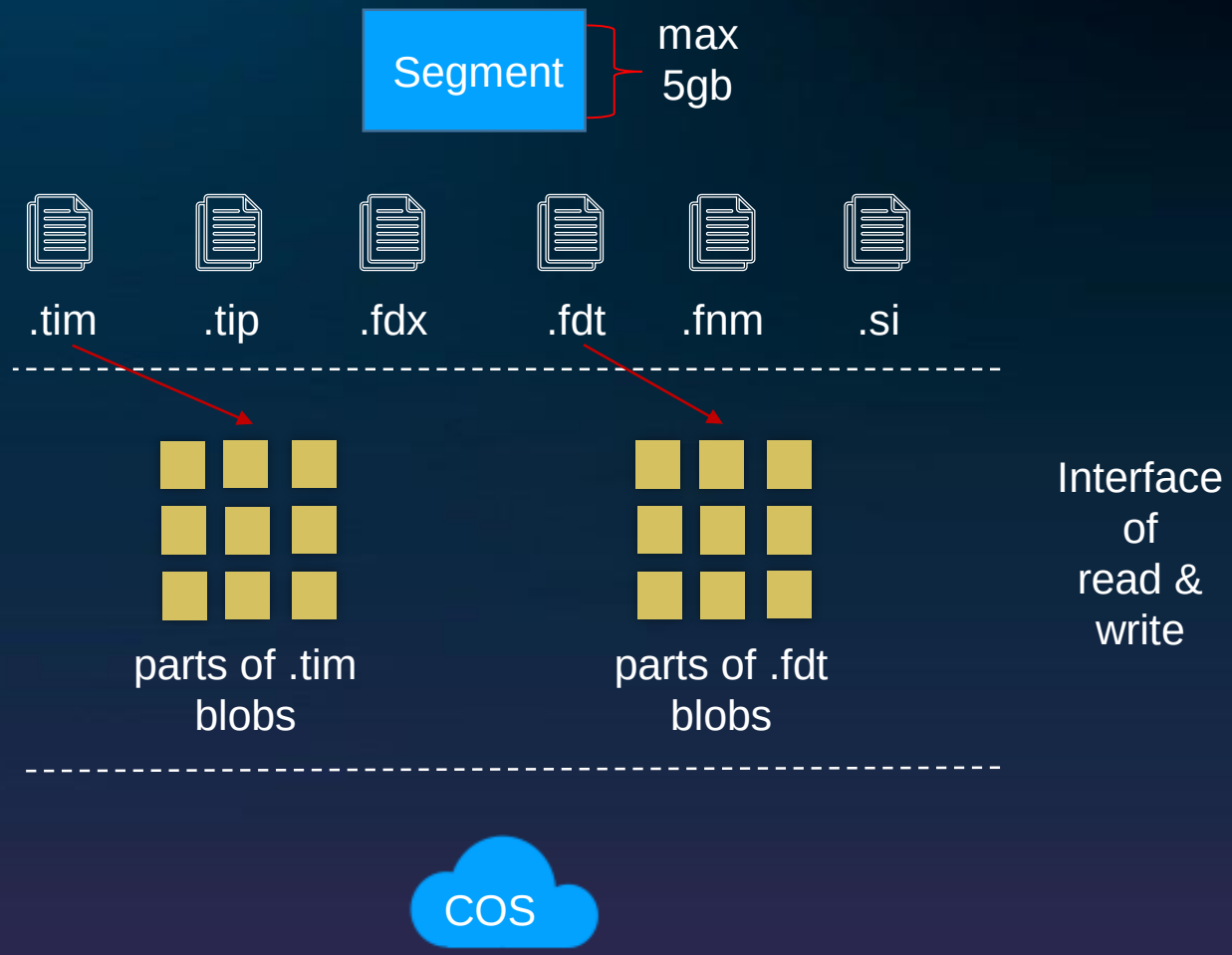
- 跨多个可用区部署集群，业务配置VIP进行访问
- Shard设置多副本，保证任意Shard有可用区副本
- 跨3可用区部署专用主节点，保证集群稳定性
- 单可用区故障，其它副本继续提供服务能力，切换速度快，业务无感知



数据备份及恢复

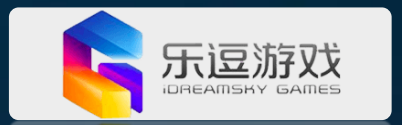
- 实现文件读写接口，直接将底层数据文件写入到COS对象存储中，速度快，成本低
- 使用ES API操作，可以备份到用户任意账号下的COS Bucket
- 每日一次自动备份，限时免费，数据保存7天
- 大版本升级前免费自动备份，防止升级失败造成的数据丢失

- 下半年针对部分用户开放基于Translog的流式备份，提高RPO



Tencent Cloud ElasticSearch Customers

Tencent 腾讯 | 腾讯云



IT Operations
Application Management
Security Analytics

Marketing insights
Business Analytics
Customer Sentiment

Website/App Search
URL Search

One More Thing

Tencent 腾讯 | 腾讯云

腾讯云

elastic

战略合作

将在X-Pack等深度能力领域展开合作，提供更有竞争力的服务

技术贡献

更紧密的参与 ES 社区、回馈社区，参与产品建设

开发者资源

举办主题展会、线下沙龙、线上活动、提供开发者试用/购买优惠等

Tencent 腾讯 | 腾讯云

Thank You

腾讯云，连接智能未来

腾讯云ES交流@广州