云搜索服务（ES）

用户操作指南

天翼云科技有限公司

目 录

1 产品介绍

1.1 什么是云搜索服务

什么是云搜索服务

云搜索服务（Cloud Search Service，简称 CSS）是一个基于 Elasticsearch、OpenSearch 且完全托管的在线分布式搜索服务，为用户提供结构化、非结构化文本、以及基于 AI 向量的多条件检索、统计、报表。云搜索服务是 ELK 生态的一系列软件集合，为您全方位提供托管的 ELK 生态云服务，兼容 Elasticsearch、Logstash、Kibana、Cerebro 等软件。

- Elasticsearch 和 OpenSearch

  Elasticsearch、OpenSearch 是开源搜索引擎，可以实现单机和集群部署，并提供托管的分布式搜索引擎服务。在 ELK 整个生态中，Elasticsearch 集群支持结构化、非结构化文本的多条件检索、统计、报表。Elasticsearch 搜索引擎相关内容的深入介绍可参见《Elasticsearch：权威指南》。OpenSearch 搜索引擎相关内容的深入介绍可参见《OpenSearch Documentation》。

  云搜索服务支持自动部署，快速创建 Elasticsearch 集群和 OpenSearch 集群，免运维，内置搜索调优实践；拥有完善的监控体系，提供一系列系统、集群以及查询性能等关键指标，让用户更专注于业务逻辑的实现。

- Logstash

  Logstash 是一个开源数据收集引擎，具有实时管道功能。在 ELK 整个生态中，Logstash 承担着数据接入的重要功能，可以动态地将来自不同数据源的数据统一起来，进行标准化的转换，然后将数据发送到指定的位置。

  云搜索服务支持快速创建 Logstash 集群，Logstash 是一款全托管的数据接入处理服务，100%兼容开源 Logstash 的能力。在生产系统中，数据往往以各种各样的形式，或分散或集中地存在于很多系统中。Logstash 的出现，能够很轻松的帮助您处理各种来源的数据并转储到 Elasticsearch 云服务中，从而更加方便的发现其中的价值。同时您也可以单独使用 Logstash 云服务处理数据发送到其他的系统中。

产品功能

- 开源兼容

兼容开源 Elasticsearch 和 OpenSearch 软件原生接口，并支持 Logstash、Beats、Kibana 等周边生态。

- 接入多种数据源

  无缝对接 FTP/OBS/Hbase/Kafka 等多种数据源，仅需简单配置，无需编程。

- 一键化操作

  一键申请集群、一键扩容、一键重启，从小规模测试到大规模上线，所有主要操作都是一键可达。

- 灵活词库管理

  支持自定义词库与拼音分词，支持词库热更新，无需重启，配置即生效。

- 自定义快照策略

  支持用户触发以及定时触发的快照备份能力，支持恢复到本集群以及其他集群的能力，随时恢复误删数据或者迁移数据到新的搜索集群。

1.2 产品优势

云搜索服务主要有以下特点与显著优势：

高效易用

TB 级数据毫秒级返回检索结果，提供可视化平台方便数据展示和分析。

弹性灵活

按需申请，在线扩容，零业务中断，快速应对业务增长。

无忧运维

全托管服务，开箱即用，主要操作一键可达，专业团队贴身看护。

内核增强

- **向量检索**

  云搜索服务的向量检索引擎支持对图像、视频、语料等非结构化数据提取的特征向量数据进行最近邻或近似近邻检索。

- **存算分离**

  云搜索服务提供冻结索引 API，支持将存储在 SSD 的热数据转储到 OBS 中以降低数据的存储成本，实现存算分离。

- **流量控制**

  云搜索服务支持流量控制，提供节点级别的流量控制功能，可提供单个节点基于黑白名单的访问限制、HTTPS 并发连接数限制、HTTP 最大连接数限制等。每个功能配置独立的控制开关。

- **大查询隔离**

  云搜索服务的大查询隔离特性支持对查询请求进行独立管理，将高内存、长耗时的查询请求进行隔离，保证节点内存安全。

- **索引监控**

  云搜索服务的索引监控特性提供了丰富的监控指标，用以监控集群索引的运行状况和变化趋势，衡量业务使用情况，同时可以针对可能存在的风险及时处理，保障集群的稳定运行。

- **集群监控增强**

  云搜索服务支持集群监控增强，支持对集群 Search 请求的 P99 时延进行监控、对集群 HTTP 状态码进行监控等。

高可靠性

支持用户手动触发以及定时触发的快照备份，支持恢复到本集群以及其他集群的能力，通过快照恢复支持集群的数据迁移。

- 自动备份（备份快照）

  云搜索服务提供备份功能，可以在控制台的备份恢复界面开启自动备份功能，并根据实际业务需要设置备份周期。

  自动备份是将集群的索引数据进行备份。索引的备份是通过创建集群快照实现，第一次备份时，建议将所有索引数据进行备份。

  云搜索服务支持将 ES 实例的快照数据保存到对象存储（OBS）服务中，借助 OBS 的跨 region 复制功能，可实现数据的跨 region 备份。

- 恢复数据（恢复快照）

  当数据发生丢失或者想找回某一时间段数据时，可以在“集群快照”界面上单击“恢复”功能，将已有的快照，通过恢复快照功能，将备份的索引数据恢复到指定的集群中，可以快速获得数据。

高安全性

云搜索服务主要从以下几个方面保障数据和业务运行安全：

- 网络隔离

  整个网络划分为 2 个平面，即业务平面和管理平面。两个平面采用物理隔离的方式进行部署，保证业务、管理各自网络的安全性。

  - 业务平面：主要是集群的网络平面，支持为用户提供业务通道，对外提供数据定义、索引、搜索能力。
  - 管理平面：主要是管理控制台，用于管理云搜索服务。
  - 通过 VPC 或安全组专有网络来确保主机的安全。

- 访问控制
  - 通过网络访问控制列表（ACL），可以允许或拒绝进入和退出各个子网的网络流量。
  - 内部安全基础设施（包括网络防火墙、入侵检测和防护系统）可以监视通过 IPsec VPN 连接进入或退出 VPC 的所有网络流量。
  - 支持用户认证与索引级别鉴权，支持对接第三方管理用户系统。

- 数据安全
  - 在云搜索服务中，通过多副本机制保证用户的数据安全。

  - 支持客户端与服务端通过 SSL 加密通信。
- 操作审计

  通过云审计服务支持对关键日志与操作进行审计。

高可用性

云搜索服务支持跨可用区部署方案。为了防止数据丢失并在服务中断时最大限度地减少集群停机时间，在创建集群时，可以选择部署在同一个区域中的两个或三个可用区，系统将在选择的可用区之间自动分配节点。当某一可用区出现故障时，剩余的可用区依然可以不间断地提供服务，显著增强了集群的可用性，提升了服务的稳定性。

1.3 产品组件

CSS 服务支持 Kibana 和 Cerebro 组件。

Kibana

Kibana 是一个开源的数据分析与可视化平台，与 Elasticsearch 搜索引擎一起使用。通过 Kibana 可以搜索、查看存放在 Elasticsearch 索引中的数据，也可以实现以图表、地图等方式展示数据。Kibana 的官方文档请参见：
https://www.elastic.co/guide/en/kibana/current/index.html

云搜索服务的 Elasticsearch 集群默认提供 Kibana，无需安装部署，即可一键访问 Kibana。云搜索服务兼容了开源 Kibana 可视化展现和 Elasticsearch 统计分析能力。

- 支持 10 余种数据呈现方式
- 支持近 20 种数据统计方式
- 支持时间、标签等各种维度分类

Cerebro

Cerebro 是使用 Scala、Play Framework、AngularJS 和 Bootstrap 构建的基于 Elasticsearch Web 的开源可视化管理工具。通过 Cerebro 可以对集群进行 Web 可视化管理，如执行 Rest 请求、修改 Elasticsearch 配置、监控实时的磁盘、集群负载、内存使用率等。

云搜索服务的 Elasticsearch 集群默认提供 Cerebro，无需安装部署，即可一键访问 Cerebro。云搜索服务完全兼容开源 Cerebro，适配最新 0.8.4 版本。

- 支持 Elasticsearch 可视化实时负载监控。
- 支持 Elasticsearch 可视化数据管理。

1.4 应用场景

云搜索服务可以帮助网站和 APP 搭建搜索框，提升用户的搜索体验；也可以用于搭建日志分析平台，助力企业实现数据驱动运维，数据驱动运营；它的向量检索能力可以帮助客户快速构建基于 AI 的图搜、推荐、语义搜索等丰富的应用。

站内搜索

云搜索服务可用于对网站内容进行关键字检索、对电商网站商品进行检索与推荐。

- 实时检索：站内资料或商品信息更新数秒至数分钟内即可被检索。
- 分类统计：检索同时可以将符合条件的商品进行分类统计。
- 高亮提示：提供高亮能力，页面可自定义高亮显示方式。

全场景日志分析

云搜索服务可用于全场景日志分析，包括 ELB 日志、服务器日志、容器和应用日志。其中 Kafka 作为消息缓冲队列，用于削峰填谷，Logstash 负责数据 ETL，Elasticsearch 负责数据检索与分析，最后由 Kibana 以可视化的方式呈现给用户。

- 性价比高：采用冷热分离、存算分离，成本同比降低 30%+。
- 易用性好：支持丰富的可视化查询语句与拖拽式报表。
- 强大的处理能力：支持每天百 TB 级数量入库，提供 PB 级以上数据处理能力。

数据库查询加速

云搜索服务可用于加速数据库查询。在电商、物流企业等有订单查询的业务场景，存在数据量大、查询并发高、吞吐大、查询延迟低的要求，关系型数据库具备较好的事务性与原子性，但其 TP 与 AP 处理能力较弱，通过将 CSS 作为备数据库，可提升整个系统的 TP 与 AP 处理能力。

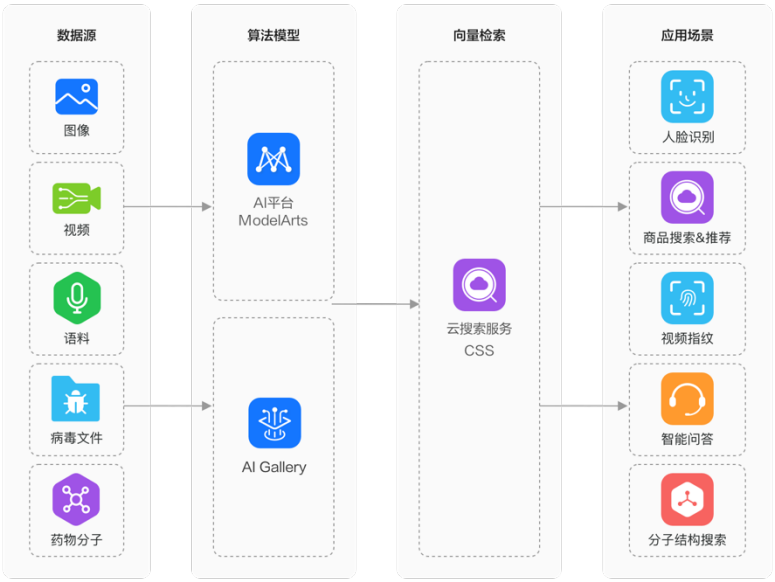
- 高性能：支持文本、时间、数字、空间等数据类型；亿级数据查询毫秒级响应。
- 高可扩展性：支持 200+数据节点，支持 1000+个数据字段。
- 业务“0”中断：规格变更、配置更新采用滚动重启，双副本场景下业务 0 中断。

向量检索

云搜索服务支持对图像、视频、语料等非结构化数据提取的特征向量数据进行最近邻或近似近邻检索。

- 高效可靠：向量检索引擎，提供优秀的搜索性能以及分布式容灾能力。
- 索引丰富：支持多种索引算法及相似度度量方式，满足各类应用场景及需求。
- “0”学习成本：完全兼容开源 ES 语法与生态。

图1-1 向量检索场景



1.5 Elasticsearch 集群版本特性差异

表1-1 Elasticsearch 集群版本特性

版本特性	5.x 版本	6.x 版本	7.x 版本
多 type 支持情况	一个 index 里面支持包含多个 type，每个 type 名称可以自定义。	一个 index 里面只支持有一个 type，type 名称可以自定义。	一个 index 里面只支持有一个 type，type 名称是固定的，_doc 不能自定义。
客户端接入	支持 TransportClient，可以同时使用 tcp 和 http 进行连接请求。	支持 TransportClient，可以同时使用 tcp 和 http 进行连接请求。建议使用 Java High Level REST Client。	只支持 RestClient，只支持使用 http 进行连接请求。建议使用 Java High Level REST Client。

	Elasticsearch 5.x 版本使用 TransportClient 接入 Elasticsearch 集群的样例：
	`// 初始化客户端，连接`
		Elasticsearch 6.x 版本、7.x 版本中使用 Java High Level REST Client 接入集群的样例：
		`// 初始化客户端，连接 9200 端口`
		`RestHighLevelClient client = new RestHighLevelClient(`
		`RestClient.builder(`
		`new HttpHost("localhost",`	

版本特性	5.x 版本	6.x 版本	7.x 版本

	9300 端口
	TransportClient client = new PreBuiltTransportClient(Settings.EMPTY)
	.addTransportAddress(new InetSocketTransportAddress(InetAddress.getByName("host1"), 9300))
	.addTransportAddress(new InetSocketTransportAddress(InetAddress.getByName("host2"), 9300));
	// 关闭客户端
	client.close();
		9200, "http"),
		new HttpHost("localhost", 9201, "http")));
		// 关闭客户端
		client.close();	

模板配置	Elasticsearch 5.x 版本中创建模板使用的是 template 字段。
	Elasticsearch 5.x 版本样例：
	PUT _template/template_1
	{
	"template": "te*",
	"settings": {
	"number_of_shards": 1
	}
	}
		Elasticsearch 6.x 及以上版本开始使用 index_pattern 字段。
		Elasticsearch 6.x 及以上版本样例：
		PUT _template/template_1
		{
		"index_patterns": ["te*"],
		"settings": {
		"number_of_shards": 1
		}
		}	

boolean 类型解析变化	Elasticsearch 5.x 版本中如下值都可以被解析成 boolean：true, false, on, off, yes, no, 0, 1。	Elasticsearch 6.x 及以上版本只接受 true/false，其他值会发生异常错误。
		以下语句在 Elasticsearch 5.x 版本不会报错，在 Elasticsearch 6.x/7.x 版本会直接报错：
		GET data1/_search
		{
		"profile": "noprofile",
		"query": {
		"match_all": {}
		}
		}	

<table>
<tr><th>版本特性</th><th>5.x 版本</th><th>6.x 版本</th><th>7.x 版本</th></tr>
<tr><td>JSON 格式校验</td><td>Elasticsearch 5.x 中允许 JSON 中存在重复的 key，后台会自动去掉。</td><td colspan="2">Elasticsearch 6.x 及以上版本不允许 JSON 存在重复的 key，会直接报解析错误。
以下语句在 Elasticsearch 5.x 版本不会报错，在 Elasticsearch 6.x/7.x 版本中会报错：
<pre>POST data1/doc
{
  "isl": 0,
  "isl": 1
}</pre></td></tr>
<tr><td>DELETE 文档变化</td><td>Elasticsearch 5.x 中，执行 DELETE index1/doc/1，如果 index1 不存在，会将 index1 创建出来。</td><td colspan="2">Elasticsearch 6.x 及以上版本，如果执行删除文档的索引不存在，会报错索引不存在。</td></tr>
<tr><td rowspan="2">_alias API 校验</td><td>Elasticsearch 5.x 中，_alias API 允许在 index 字段中指定为别名，能正常解析。
类似的，Elasticsearch 5.x 版本中允许使用别名删除一个索引。</td><td colspan="2">Elasticsearch 6.x 版本中，_alias API 中的 index 字段只能指定为索引名，不允许是别名。
Elasticsearch 6.x 版本中不再允许，必须使用索引名进行删除。</td></tr>
<tr><td colspan="3">如以下示例，在 Elasticsearch 5.x 版本中能正常工作，但是在 Elasticsearch 6.x 版本/7.x 版本中会报错：
<pre>PUT log-2023.11.11
POST _aliases
{
  "actions": [
    {
      "add": {
        "index": "log-2023.11.11",
        "alias": "log"
      }
    }
  ]
}
POST _aliases
{
  "actions": [
    {
      "remove": {
        "index": "log",
        "alias": "log"
      }
    }</pre></td></tr>
</table>

版本特性	5.x 版本	6.x 版本	7.x 版本

	]
	}
	报错信息：
	{
	"error" : {
	"root_cause" : [
	{
	"type" : "illegal_argument_exception",
	"reason" : "The provided expression [log] matches an alias, specify the corresponding concrete indices instead."
	}
	],
	"type" : "illegal_argument_exception",
	"reason" : "The provided expression [log] matches an alias, specify the corresponding concrete indices instead."
	},
	"status" : 400
	}		

默认配置变化	新建索引默认分片数为 5。		新建索引默认分片数为 1。
默认 routing 变化	Elasticsearch 5.x 版本/6.x 版本使用以下公式计算文档应该落在哪个 shard。 shard_num = hash(routing) % num_of_primary_shards		Elasticsearch 7.x 版本使用以下公式计算文档应该落在哪个 shard routing_factor = num_routing_shards / num_primary_shards shard_num = (hash(routing) % num_routing_shards) / routing_factor 其中 num_routing_shards 可以由以下配置指定。 index.number_of_routing_shards 如果不显式指定，则 Elasticsearch 会自动计算该值，以达到对索引进行 split 的能力。
Refresh 时机变化	默认定期每秒钟执行 refresh。		Elasticsearch 7.x 版本中如果没有显式的指定 index.refresh_interv

<table>
<tr><th>版本特性</th><th>5.x 版本</th><th>6.x 版本</th><th>7.x 版本</th></tr>
<tr><td></td><td colspan="2"></td><td>al，并且索引长时间没有 search 请求，这里的长时间是由配置 index.search.idle.after 指定，默认 30 秒，Elasticsearch 就不会再定期的进行 refresh，而是等到有新的 search 请求进来时再进行 refresh，这时候进行的 search 请求会等待，直到下一轮 refresh 完成才进行检索并返回，所以第一次 search 请求一般耗时会相对较长。</td></tr>
<tr><td>父熔断器变化</td><td colspan="2">父熔断器是在多个子熔断器中内存统计之和超限的情况下触发，超限阈值为 70%。</td><td>父熔断器会在堆内存超限的情况下触发，默认超限阈值为 95%。</td></tr>
<tr><td>Field Data 熔断器阈值变化</td><td colspan="2">Field Data 熔断器超限阈值 indices.breaker.fielddata.limit 默认为 60%。</td><td>Field Data 熔断器超限阈值 indices.breaker.fielddata.limit 默认为 40%。</td></tr>
<tr><td>_all 字段支持情况</td><td>支持_all 字段。</td><td>_all 字段被废弃。</td><td>已删除_all 字段，不再支持。</td></tr>
<tr><td>search API 返回中 hits.total</td><td colspan="2">Elasticsearch 5.x 版本/6.x 版本中，search API 返回中，hits.total 为数字，表示命中条数：<pre>{
  "took": 0,
  "timed_out": false,
  "_shards": {
    "total": 5,
    "successful": 5,
    "failed": 0
  },
  "hits": {
    "total": 4,
    "max score": 1,
  }
}</pre></td><td>Elasticsearch 7.x 版本中，hits.total 不再是数字：<pre>{
  "took" : 76,
  "timed_out" :
false,
  "_shards" : {
    "total" : 1,
    "successful" :
1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {</pre></td></tr>
</table>

版本特性	5.x 版本	6.x 版本	7.x 版本

			<pre>"value" : 4,
			"relation" : "eq"
			},
			"max_score" : 1.0
			}
			}</pre>其中：
			value 表示命中的条数。
			relation 表示 value 参数中的命中条数是否是准确值。
			eq 表示是准确值。
			gte 表示命中条数大于等于 value 参数。

_cache/clear API	_cache/clear API 支持 POST/GET 方式。		_cache/clear API 只支持 POST 方式，不再支持 GET 方式。

1.6 产品规格

当您在云搜索服务创建集群时，系统将为您提供多种规格以满足您按需选择的要求。具体规格说明和适用场景请见表 1-2。

表1-2 节点规格

CPU 架构	节点规格类型	CPU 内存比	适合场景
x86 计算	通用计算型	1:4	默认规格，使用较频繁，各种场景都能使用，如果没有特别要求，可选择该规格。
	内存优化型	1:8	内存特别大，优势较明显，在内存使用量较多并且对时延没有太大要求的场景可优先选择该规格，比如多聚合（filedata 堆内）、排序、列存 docvalue（系统堆外内存）等场景。
鲲鹏计算	鲲鹏通用计算型	1:2 和 1:4	同上 X86 计算型场景，相比于 X86 计算型，ARM 性价比较高。

1.7 约束与限制

集群和节点限制

下表显示了云搜索服务的集群和节点的限制。

表1-3 Elasticsearch 类型集群和节点限制

集群和节点	限制
每个集群的最大节点数（节点数量）	默认值 32，最大支持 200 个节点，如果需要更改默认值，请联系技术支持。
每个集群的最小节点数（节点数量）	1

浏览器限制

- 访问云搜索服务管理控制台，建议使用如下版本浏览器
  - Google Chrome：36.0 及更高版本
  - Mozilla FireFox：35.0 及更高版本
- 访问云搜索服务中 Kibana 和 Cerebro，建议使用如下版本浏览器
  - Google Chrome：36.0 及更高版本
  - Mozilla FireFox：35.0 及更高版本

1.8 与其他服务之间的关系

CSS 与其他服务的关系如图 1-2 所示。

图1-2 CSS 与其他服务的关系

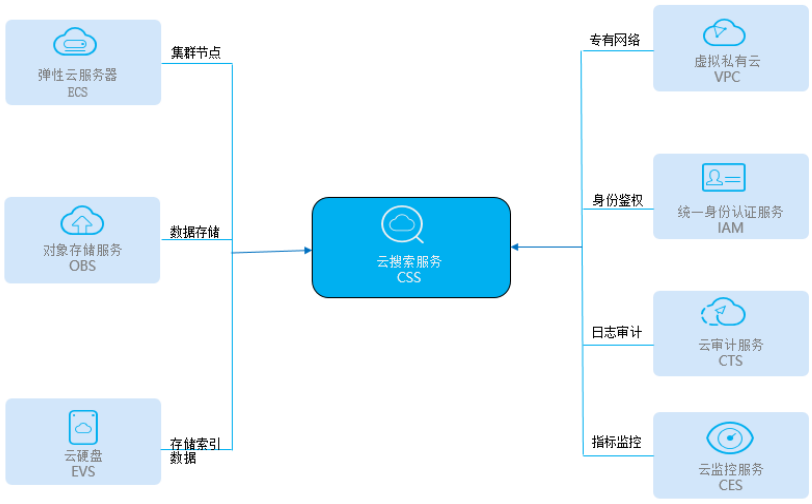


表1-4 CSS 服务与其他服务的关系

相关服务	交互功能
虚拟私有云（Virtual Private Cloud，简称 VPC）	云搜索服务 CSS 的集群创建在虚拟私有云（VPC）的子网内，VPC 通过逻辑方式进行网络隔离，为用户的集群提供安全、隔离的网络环境。
弹性云服务器（Elastic Cloud Server，简称 ECS）	云搜索服务 CSS 的集群中每个节点为一台弹性云服务器（ECS）。创建集群时将自动创建弹性云服务器作为节点。
云硬盘（Elastic Volume Service，简称 EVS）	云搜索服务 CSS 使用云硬盘（EVS）存储索引数据。创建集群时，将自动创建云硬盘用于集群存储。
对象存储服务（Object Storage Service，简称 OBS）	云搜索服务 CSS 的集群快照存储在对象存储服务（OBS）的桶中。
统一身份认证服务（Identity and Access Management，简称 IAM）	云搜索服务 CSS 使用统一身份认证服务（IAM）进行鉴权。
云监控服务（Cloud Eye）	云搜索服务使用云监控服务实时监测集群的指标信息，保障服务正常运行。云搜索服务当前支持的监控指标为磁盘使用率和集群健康状态。用户通过磁盘使用率指标可以及时了解集群的磁盘使用情况。通过集群健康状态指标，用户可以了解集群的健康状态。
云审计服务（Cloud Trace Service，简称	云审计服务（CTS）可以记录与 CSS 云搜索服务相关的操作事件，便于日后的查询、审计和回溯。

相关服务	交互功能
CTS）	

1.9 基本概念

集群

云搜索服务是以集群为单位进行组织，一个集群代表一个独立运行的搜索服务，由多个节点构成。

索引

用于存储 Elasticsearch 的数据，是一个或多个分片分组在一起的逻辑空间。

Shard

索引可以存储数据量超过 1 个节点硬件限制的数据。为满足这样的需求，Elasticsearch 提供了一个能力，将一个索引拆分为多个，称为 Shard。当您创建一个索引时，您可以根据实际情况指定 Shard 的数量。每个 Shard 托管在集群中的任意一个节点中，且每个 Shard 本身是一个独立的、全功能的“索引”。

Shard 的数量只能在创建索引前指定，且在索引创建成功后无法修改。

Replica（副本）

Shard 下的实际存储索引的一个副本。可以理解为备份 Shard。副本的存在可以预防单节点故障。使用过程中，您可以根据业务情况增加或减少 Replica 数量。

文档

Elasticsearch 存储的实体，是可以被索引的基本单位，相当于关系型数据库中的行。

文档类型

类似关系型数据库中的表，用于区分不同的数据。

Elasticsearch 7.x 以下版本中，1 个索引里面可以包含若干个文档类型，每个文档必须设定它的文档类型。

Elasticsearch 7.x 及以上版本中，文档类型只支持“_doc”。

映射

用来约束字段的类型，可以根据数据自动创建。相当于数据库中的 Schema。

字段

组成文档的最小单位。相当于数据库中的 Column。

2 快速入门

2.1 快速开始使用云搜索服务

本章节为您提供了一个简单的商品搜索示例，您可以参考此示例来了解云搜索服务，包括创建索引、导入数据、搜索数据等操作。

场景描述

某女装品牌在网上经营电商业务，以前使用传统数据库来为用户提供商品搜索功能，但随着用户数量和业务的增长，使用传统数据库的弊端愈来愈明显。主要问题表现为响应速度慢、准确性低。为了改善用户体验从而避免用户流失，该电商网站开始使用云搜索服务为用户提供商品搜索功能，不仅解决了之前使用传统数据库产生的问题，而且实现了用户数量的增长。

假设该电商网站经营商品的数据如下所示：

```
{
"products":[
{"productName":"2017 秋装新款文艺衬衫女装","size":"L"}
{"productName":"2017 秋装新款文艺衬衫女装","size":"M"}
{"productName":"2017 秋装新款文艺衬衫女装","size":"S"}
{"productName":"2018 春装新款牛仔裤女装","size":"M"}
{"productName":"2018 春装新款牛仔裤女装","size":"S"}
{"productName":"2017 春装新款休闲裤女装","size":"L"}
{"productName":"2017 春装新款休闲裤女装","size":"S"}
]
}
```

操作流程

- 步骤 1：创建集群
- 步骤 2：导入数据
- 步骤 3：搜索数据
- （可选）步骤 4：删除索引
- （可选）步骤 5：删除集群

步骤 1：创建集群

本示例您需要创建一个名称为“Sample-ESCluster”的 Elasticsearch 集群。此集群仅用于入门指导使用，建议选用“节点规格”为“ess.spec-4u8g”，“节点存储”为“高I/O”，“节点存储容量”为“40GB”。

1. 登录云搜索服务管理控制台。
2. 单击右上角的“创建集群”，进入“创建集群”页面。
3. 选择“当前区域”和“可用区”。

表2-1 区域和可用区参数说明

参数	说明
当前区域	在右侧下拉框中选择集群的所在区域。
可用区	选择集群工作区域下关联的可用区。最多支持配置 3 个“可用区”。 本场景示例选择 1 个可用区。

4. 配置集群基本信息。

图2-1 配置集群信息

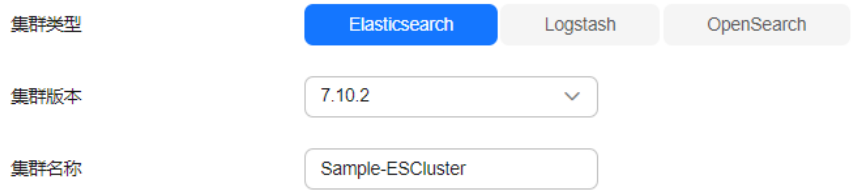


表2-2 基本参数说明

参数	说明
集群版本	选择所需的集群版本，支持的版本以界面可选项为准。
集群名称	自定义集群名称。 本场景示例您可以创建一个名称为“Sample-ESCluster”的集群。

5. 配置集群的规格信息。

表2-3 规格参数说明

参数	说明
节点数量	集群中的节点个数。可选节点数为 1~32。 本场景示例选择 1 个节点。
CPU 架构	具体支持的类型由实际区域环境决定。
节点规格	集群中的节点规格。 本场景示例节点规格建议选择“ess.spec-4u8g”。
节点存储	选择存储类型。 本示例节点存储建议选择高 I/O。
节点存储容量	存储空间大小，其取值范围与节点规格关联，不同的规格允许的取值范围不同。节点存储容量只支持配置为 20 的倍数。 本场景示例存储容量建议选择 40GB。
启用 Master 节点	Master 节点用于管理集群中的所有节点。当需要存储和分析的数据量大，所需节点数量大于 20 个节点时，建议启用 Master 节点，保证集群的稳定性。 本场景示例无需用 Master 节点。
启用 Client 节点	Client 节点用于提供客户端接入集群和分析数据的服务。当需要存储和分析的数据量大，所需节点数量大于 20 个节点时，建议启用 Client 节点，保证集群的稳定性。 本场景示例无需启用 Client 节点。
启用冷数据节点	冷数据节点用于存放对于历史数据要求分钟级别的返回。当用户对历史数据返回时间要求不是很高的话，可以将这部分数据存储在冷数据节点上，从而降低成本。 本场景示例无需启用冷数据节点。

6. 设置集群的企业项目。

   如果您开通了“企业项目”，在创建集群时，可以给集群绑定一个企业项目。您可以在右侧下拉框中选择当前用户下已创建的企业项目，也可以通过单击“查看项目管理”按钮，前往“企业项目管理”管理控制台，新建企业项目和查看已有的企业项目。

7. 单击“下一步：网络配置”，设置集群的网络配置。

图2-2 配置网络

表2-4 网络配置参数说明

参数	说明
虚拟私有云	VPC 即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。 选择创建集群需要的 VPC，单击“查看虚拟私有云”进入 VPC 服务查看已创建的 VPC 名称和 ID。如果没有 VPC，需要创建一个新的 VPC。 说明 此处您选择的 VPC 必须包含网段（CIDR），否则集群将无法创建成功。新建的 VPC 默认包含网段（CIDR）。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。 选择创建集群需要的子网，可进入 VPC 服务查看 VPC 下已创建的子网名称和 ID。
安全组	安全组是一个逻辑上的分组，为同一个 VPC 内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。单击“查看安全组”可了解安全组详情。
安全模式	集群支持选择是否开启安全模式，开启之后将对集群进行通讯加密和安全认证。 本场景示例选择关闭安全模式。

8. 单击“下一步：高级配置”，可以设置集群自动快照和其他高级功能。
   本集群仅做入门指导使用，无需开启集群快照和高级功能。
9. 单击“下一步：确认配置”，确认完成后单击“立即创建”开始创建集群。
10. 单击“返回集群列表”，系统将跳转到“集群管理”页面。您创建的集群将展现在集群列表中，且集群状态为“创建中”，创建成功后集群状态会变为“可用”。

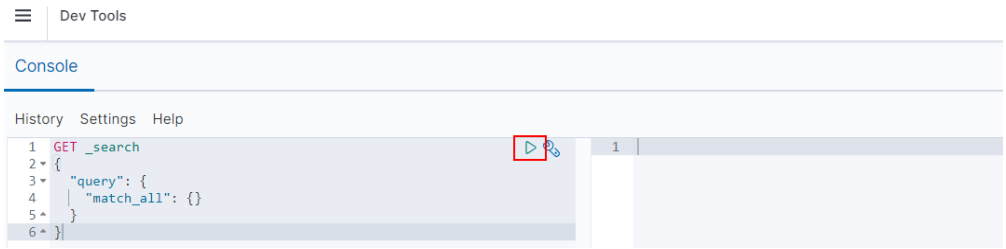
步骤 2：导入数据

云搜索服务支持通过云数据迁移（简称 CDM）、数据接入服务（简称 DIS）、Logstash、Kibana 或 API 将数据导入到 Elasticsearch。Kibana 是一个开源的数据分析与可视化平台，通过 Kibana 可以搜索、查看存放在 Elasticsearch 索引中的数据，也可以实现以图表、地图等方式展示数据。云搜索服务的 Elasticsearch 集群默认提供 Kibana，本示例将以 Kibana 为例介绍将数据导入到 Elasticsearch 的操作流程。

1. 在“集群管理”页面选择已创建的“Sample-ESCluster”集群，单击“操作”列中的“Kibana”进入 Kibana 控制台。
2. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。

   Console 左侧区域为输入框，输入框右侧的三角形为执行命令按钮，Console 右侧为结果输出区域。

   图2-3 Console 界面



3. 在 Console 界面，执行如下命令创建索引“my_store”。

   （低于 7.x 版本）

```
PUT /my_store
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "products": {
      "properties": {
        "productName": {
          "type": "text",
          "analyzer": "ik_smart"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
}
```

   （7.x 版本及高于 7.x 版本）

```
PUT /my_store
{
  "settings": {
    "number_of_shards": 1
```

```
  },
  "mappings": {
    "properties": {
      "productName": {
        "type": "text",
        "analyzer": "ik_smart"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
```

返回结果如下所示。

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "my_store"
}
```

4. 在 Console 界面，执行如下命令，将数据导入到“my_store”索引中。

（低于 7.x 版本）

```
POST /my_store/products/_bulk
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"L"}
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"M"}
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"S"}
{"index":{}}
{"productName":"2018 春装新款牛仔裤女装","size":"M"}
{"index":{}}
{"productName":"2018 春装新款牛仔裤女装","size":"S"}
{"index":{}}
{"productName":"2017 春装新款休闲裤女装","size":"L"}
{"index":{}}
{"productName":"2017 春装新款休闲裤女装","size":"S"}
```

（7.x 版本及高于 7.x 版本）

```
POST /my_store/_doc/_bulk
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"L"}
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"M"}
{"index":{}}
{"productName":"2017 秋装新款文艺衬衫女装","size":"S"}
{"index":{}}
{"productName":"2018 春装新款牛仔裤女装","size":"M"}
{"index":{}}
{"productName":"2018 春装新款牛仔裤女装","size":"S"}
{"index":{}}
{"productName":"2017 春装新款休闲裤女装","size":"L"}
{"index":{}}
{"productName":"2017 春装新款休闲裤女装","size":"S"}
```

当返回结果信息中“errors”字段的值为“false”时，表示导入数据成功。

步骤 3：搜索数据

- **全文检索**

  假设用户进入该电商网站，她想要查找名称包含“春装牛仔裤”的商品信息，可以搜索“春装牛仔裤”。这里使用 Kibana 演示用户搜索数据在后台的执行命令和返回结果。

  执行命令如下所示。

  （低于 7.x 版本）

```
GET /my_store/products/_search
{
  "query": {"match": {
    "productName": "春装牛仔裤"
  }}
}
```

  （7.x 版本及高于 7.x 版本）

```
GET /my_store/_search
{
  "query": {"match": {
    "productName": "春装牛仔裤"
  }}
}
```

  返回结果如下所示。

```
{
  "took" : 3,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 4,
      "relation" : "eq"
    },
    "max_score" : 1.7965372,
    "hits" : [
      {
        "_index" : "my_store",
        "_type" : "_doc",
        "_id" : "9xf6VHIBfClt6SDjw7H5",
        "_score" : 1.7965372,
        "_source" : {
          "productName" : "2018 春装新款牛仔裤女装",
          "size" : "M"
        }
      },
      {
        "_index" : "my_store",
```

```
      "_type" : "_doc",
      "_id" : "-Bf6VHIBfClt6SDjw7H5",
      "_score" : 1.7965372,
      "_source" : {
        "productName" : "2018 春装新款牛仔裤女装",
        "size" : "S"
      }
    },
    {
      "_index" : "my_store",
      "_type" : "_doc",
      "_id" : "-Rf6VHIBfClt6SDjw7H5",
      "_score" : 0.5945667,
      "_source" : {
        "productName" : "2017 春装新款休闲裤女装",
        "size" : "L"
      }
    },
    {
      "_index" : "my_store",
      "_type" : "_doc",
      "_id" : "-hf6VHIBfClt6SDjw7H5",
      "_score" : 0.5945667,
      "_source" : {
        "productName" : "2017 春装新款休闲裤女装",
        "size" : "S"
      }
    }
  ]
 }
}
```

- Elasticsearch 支持 IK 分词，上面执行命令会将“春装牛仔裤”分词为“春装”和“牛仔裤”。
- Elasticsearch 支持全文检索，上面执行命令会在所有商品信息中搜索包含“春装”或“牛仔裤”的商品信息。
- Elasticsearch 与传统数据库不同，它能借助倒排索引在毫秒级返回结果。
- Elasticsearch 支持评分排序，在上面返回结果中，前两条商品信息中同时出现了“春装”和“牛仔裤”，后两条商品信息中只出现了“春装”，所以前两条比后两条与检索关键词的匹配度更高，分数更高，排序也更靠前。

- **聚合结果显示**

  该电商网站可以提供聚合结果显示功能，例如： 对“春装”对应的产品按照尺码分类，统计不同尺码的数量。这里使用 Kibana 演示聚合结果显示功能在后台的执行命令和返回结果。

  执行命令如下所示。

  （低于 7.x 版本）

```
GET /my_store/products/_search
{
"query": {
"match": { "productName": "春装" }
},
```

```
"size": 0,
"aggs": {
"sizes": {
"terms": { "field": "size" }
}
}
}
```

（7.x 版本及高于 7.x 版本）

```
GET /my_store/_search
{
"query": {
"match": { "productName": "春装" }
},
"size": 0,
"aggs": {
"sizes": {
"terms": { "field": "size" }
}
}
}
```

返回结果如下所示。

（低于 7.x 版本）

```
{
  "took" : 31,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 4,
    "max_score" : 0.0,
    "hits" : [ ]
  },
  "aggregations" : {
    "sizes" : {
      "doc_count_error_upper_bound" : 0,
      "sum_other_doc_count" : 0,
      "buckets" : [
        {
          "key" : "S",
          "doc_count" : 2
        },
        {
          "key" : "L",
          "doc_count" : 1
        },
        {
          "key" : "M",
          "doc_count" : 1
        }
```

```
      ]
    }
  }
}
```

（7.x 版本及高于 7.x 版本 ）

```
{
  "took" : 3,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 4,
      "relation" : "eq"
    },
    "max_score" : null,
    "hits" : [ ]
  },
  "aggregations" : {
    "sizes" : {
      "doc_count_error_upper_bound" : 0,
      "sum_other_doc_count" : 0,
      "buckets" : [
        {
          "key" : "S",
          "doc_count" : 2
        },
        {
          "key" : "L",
          "doc_count" : 1
        },
        {
          "key" : "M",
          "doc_count" : 1
        }
      ]
    }
  }
}
```

（可选）步骤 4：删除索引

如果不再使用对应的索引数据，您可以执行如下命令，删除对应索引，避免造成浪费资源。

```
DELETE /my_store
```

返回结果如下所示。

```
{
  "acknowledged" : true
}
```

（可选）步骤 5：删除集群

已完成数据搜索业务，无需继续使用集群时，可删除集群释放资源。

说明

由于集群删除后，数据无法恢复，请谨慎操作。

1. 登录云搜索服务管理控制台。在左侧菜单栏选择“集群管理 > Elasticsearch”。
2. 进入集群管理页面，选中“Sample-ESCluster”集群所在行，在操作列单击“更多”>“删除”。
3. 在弹出的确认对话框中，确认要删除的集群名称，单击“确定”完成操作。

3 权限管理

3.1 创建用户并授权使用 CSS

本章节介绍创建 CSS 用户操作，将 CSS 服务的策略授予用户组，并将用户添加至用户组中（一个用户组下面的用户具有相同的权限），从而使用户拥有对应的 CSS 权限，操作流程如图 3-1 所示。

CSS 具有两种类型用户权限（CSS 管理员权限和只读权限），在权限规划的时候请规划这两种类型的用户组。

前提条件

给用户组授权之前，请您了解用户组可以添加的 CSS 系统策略，请参见权限管理。

示例流程

图3-1 给用户授权 CSS 权限流程

1. 参考 IAM 用户指南中创建用户组授权

   在 IAM 控制台创建用户组，并授予云搜索服务权限。
2. 参考 IAM 用户指南中创建用户并加入用户组

   在 IAM 控制台创建用户，并将其加入 1 中创建的用户组。
3. 用户登录并验证权限

   新创建的用户登录控制台，切换至授权区域，验证权限：
   - 在“服务列表”中选择云搜索服务 CSS，进入 CSS 主界面，单击右上角“创建集群”，尝试购买 CSS 集群，如果无法购买 CSS 集群（假设当前权限仅包含 CSS ReadOnlyAccess），表示“CSS ReadOnlyAccess”已生效。
   - 在“服务列表”中选择除云搜索服务外（假设当前策略仅包含 CSS ReadOnlyAccess）的任一服务，如果提示权限不足，表示“CSS ReadOnlyAccess”已生效。

3.2 CSS 自定义策略

如果系统预置的 CSS 权限，不满足您的授权要求，可以创建自定义策略。自定义策略中可以添加的授权项（Action）请参考云搜索服务 API 参考中的权限策略和授权项。

目前支持以下两种方式创建自定义策略：

- 可视化视图创建自定义策略：无需了解策略语法，按可视化视图导航栏选择云服务、操作、资源、条件等策略内容，可自动生成策略。

- JSON 视图创建自定义策略：可以在选择策略模板后，根据具体需求编辑策略内容；也可以直接在编辑框内编写 JSON 格式的策略内容。

具体创建步骤请参见：IAM 用户指南中的创建自定义策略。本章为您介绍常用的 CSS 自定义策略样例。

说明

IAM 权限和数据面集群权限分开控制，如果要开启数据面的安全能力，需要使用安全模式。

CSS 自定义策略样例

说明

如果是子账号，需要同时设置 GetBucketStoragePolicy、GetBucketLocation、ListBucket、ListAllMyBuckets 权限，才能看到 OBS 桶。

示例 1：授权用户创建集群。

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Action": [
            "css:cluster:create",
            "vpc:securityGroups:get",
            "vpc:securityGroups:create",
            "vpc:securityGroups:delete",
            "vpc:securityGroupRules:get",
            "vpc:securityGroupRules:create",
            "vpc:securityGroupRules:delete",
            "vpc:vpcs:list",
            "vpc:privateIps:list",
            "vpc:ports:get",
            "vpc:ports:create",
            "vpc:ports:update",
            "vpc:ports:delete",
            "vpc:quotas:list",
            "vpc:subnets:get",
            "ecs:cloudServerFlavors:get",
            "ecs:serverInterfaces:use",
            "ecs:cloudServers:addNics",
            "ecs:quotas:get",
            "evs:types:get",
            "evs:quotas:get"
         ],
         "Effect": "Allow"
      }
   ]
}
```

示例 2：拒绝用户删除集群。

拒绝策略需要同时配合其他策略使用，否则没有实际作用。用户被授予的策略中，一个授权项的作用如果同时存在 Allow 和 Deny，则遵循 **Deny 优先原则**。

如果您给用户授予 CSS Admin 的系统策略，但不希望用户拥有 CSS admin 中定义的删除云服务器权限，您可以创建一条拒绝删除云服务的自定义策略，然后同时将 CSS Admin 和拒绝策略授予用户，根据 Deny 优先原则，则用户可以对 CSS 执行除了删除集群外的所有操作。拒绝策略示例如下：

```
{
     "Version": "1.1",
     "Statement": [
          {
          "Effect": "Deny",
               "Action": [
                    "css:cluster:delete"
               ]
          }
     ]
}
```

示例 3：多个授权项策略。

一个自定义策略中可以包含多个授权项，且除了可以包含本服务的授权项外，还可以包含其他服务的授权项，可以包含的其他服务必须跟本服务同属性，即都是项目级服务或都是全局级服务。多个授权语句策略描述如下：

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Action": [
            "ecs:cloudServers:resize",
            "ecs:cloudServers:delete",
            "ecs:cloudServers:delete",
            "css:cluster:restart",
            "css:*:get*",
            "css:*:list*"
         ],
         "Effect": "Allow"
      }
   ]
}
```

4 Elasticsearch

4.1 创建集群

4.1.1 创建安全模式集群

Elasticsearch 集群支持开启安全模式。参考本章可以完成安全模式的 Elasticsearch 集群的创建。

说明

公网访问和 Kibana 公网访问需要开启安全模式才能使用。

背景信息

- 新建集群时，当设置不同节点类型时支持的节点数量区间会有区别，具体情况请参考表 4-1。

表4-1 不同节点类型的节点数量说明

集群包含的节点类型	节点数量的取值范围
ess	ess：1~32
ess、ess-master	ess：1~200 ess-master：3~9 的奇数
ess、ess-client	ess：1~32 ess-client：1~32
ess、ess-cold	ess：1~32 ess-cold：1~32
ess、ess-master、ess-client	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32
ess、ess-master、ess-cold	ess：1~200

集群包含的节点类型	节点数量的取值范围
	ess-master：3~9 的奇数 ess-cold：1~32
ess、ess-client、ess-cold	ess：1~32 ess-client：1~32 ess-cold：1~32
ess、ess-master、ess-client、ess-cold	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32 ess-cold：1~32
四种节点类型的说明： • ess：默认节点类型，即创建集群时必选的数据节点类型，其他 3 种节点类型都是基于业务需要可选的类型。 • ess-master：Master 节点 • ess-client：Client 节点 • ess-cold：冷数据节点	

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“总览”页面单击右上角的“创建集群”，进入“创建集群”页面。
   或者左侧导航栏单击“集群管理>Elasticsearch”，单击右上角的“创建集群”，进入“创建集群”页面。
3. 选择“当前区域”和“可用区”。

表4-2 区域和可用区参数说明

参数	说明
当前区域	集群工作区域在右侧下拉框中选择。
可用区	选择集群工作区域下关联的可用区。 最多支持配置 3 个“可用区”，多可用区的使用请参见 4.1.4 部署跨 AZ 集群。

4. 配置集群基本信息。

表4-3 基本参数说明

参数	说明
集群版本	选择所需的集群版本，支持的版本以界面可选项为准。
集群名称	自定义集群名称，可输入的字符范围为 4～32 个字符，只能包含数字、字母、中划线和下划线，且必须以字母开头。 说明 当集群创建成功后，您可以根据需求修改集群名称。单击需要修改的集群名称，进入集群基本信息页面，单击“集群名称”后面的，修改完成后，单击，进行保存。如果需要取消修改，可单击进行取消。

5. 配置集群的规格信息。

表4-4 规格参数说明

参数	说明
节点数量	集群中的节点个数。可选节点数为 1~32，建议节点数为 3 或 3 以上，以提升集群可用性。 • 如果未启用 Master 节点和 Client 节点时，此参数指定的节点将被作为 Master 节点和 Client 节点，同时具备集群管理、存储数据、提供接入集群和分析数据的服务。此时，为保证集群中数据的稳定性，建议设置节点数量大于等于 3 个。 • 如果启用 Master 节点，且未启用 Client 节点，此参数指定的节点将用于存储数据并提供 Client 节点功能。 • 如果已启用 Master 节点和 Client 节点，此参数指定的节点将仅用于存储数据。 • 如果启用 Client 节点，且未启用 Master 节点，此参数指定的节点将用于存储数据并提供 Master 节点功能。
CPU 架构	支持“X86 计算”和“鲲鹏计算”两种类型。具体支持的类型由实际区域环境决定。
节点规格	集群中的节点规格。您可以根据需求，选择对应的规格。每个集群只能选择一个规格。
节点存储	选择存储类型，支持普通 I/O、高 I/O、超高 I/O。
节点存储容量	存储空间大小，其取值范围与节点规格关联，不同的规格允许的取值范围不同。 节点存储容量只支持配置为 20 的倍数。
启用 Master 节点	Master 节点用于管理集群中的所有节点。当需要存储和分析的

参数	说明
	数据量大，所需节点数量大于 20 个节点时，建议启用 Master 节点，保证集群的稳定性。反之，建议仅设置集群的“节点数量”参数，同时作为 Master 和 Client 节点即可。 启用 Master 节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”必须是不小于 3 的奇数，最多设置 9 个节点。“节点存储”的存储容量为固定值，存储类型可以根据实际情况选择。
启用 Client 节点	Client 节点用于提供客户端接入集群和分析数据的服务。当需要存储和分析的数据量大，所需节点数量大于 20 个节点时，建议启用 Client 节点，保证集群的稳定性。反之，建议仅设置集群的“节点数量”参数，同时作为 Master 和 Client 节点即可。 启用 Client 节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”可设置为 1~32 任意数值。“节点存储”的存储容量为固定值，存储类型可以根据实际情况选择。
启用冷数据节点	当用户对历史数据返回时间要求不是很高的话，可以将这部分数据存储在冷数据节点上，从而降低成本。 启用冷数据节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”可设置为 1~32 任意数值。“节点存储”的存储类型和存储容量可以根据实际情况选择。 开启冷数据节点之后，云搜索服务将会自动的给相关节点打上冷热标签。

6. 设置集群的企业项目。

   如果您开通了“企业项目”，在创建集群时，可以给集群绑定一个企业项目。您可以在右侧下拉框中选择当前用户下已创建的企业项目，也可以通过单击“查看项目管理”按钮，前往“企业项目管理”管理控制台，新建企业项目和查看已有的企业项目。

7. 单击“下一步：网络配置”，设置集群的网络配置。

表4-5 网络配置参数说明

参数	说明
虚拟私有云	VPC 即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。 选择创建集群需要的 VPC，单击“查看虚拟私有云”进入 VPC 服务查看已创建的 VPC 名称和 ID。如果没有 VPC，需要创建一个新的 VPC。

参数	说明
	说明 此处您选择的 VPC 必须包含网段（CIDR），否则集群将无法创建成功。新建的 VPC 默认包含网段（CIDR）。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。 选择创建集群需要的子网，可进入 VPC 服务查看 VPC 下已创建的子网名称和 ID。
安全组	安全组是一个逻辑上的分组，为同一个 VPC 内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。单击“查看安全组”可了解安全组详情。 说明 • 为了确保您能够正常访问集群，需要放通安全组 9200 规则。 • 如果创建的集群为 7.6.2 及以上版本，则需要确保同安全组内节点之间的端口全放通。如果无法放通同安全组内节点之间的全部端口，请至少确保 9300 端口的通信。 • 放开 9300 端口通信后，如果集群磁盘使用率较高，可清理过期数据，释放磁盘存储空间。
安全模式	集群支持选择是否开启安全模式，开启之后将对集群进行通讯加密和安全认证。 • **管理员账户名**默认为 admin。 • 设置并确认**管理员密码**。要记住设置的密码，后续访问集群需要输入密码。
HTTPS 访问	只有开启集群的安全模式才可以启用 HTTPS 访问，开启 HTTPS 访问后，访问集群将进行通讯加密。 说明 安全集群使用 HTTPS 通信，相比非安全集群使用 HTTP 通信在读取性能上会降低，预期相对 HTTP 集群在大并发压力下有 20%的性能劣化。如果想要读取性能快，又想要使用安全集群所提供的用户权限隔离资源（索引、文档、字段等）的功能，则可以关闭 HTTPS 访问。关闭 HTTPS 访问后，会使用 HTTP 协议与集群通信，无法保证数据安全性，并且无法开启公网访问功能。
公网访问	开启“HTTPS 访问”后，可以选择是否配置“公网访问”，配置公网访问后，用户可以获得一个公网访问的 IP，通过这个 IP 可以在公网访问该安全集群，详细配置请参考 4.3.2 公网访问集群。

8. 单击“下一步：高级配置”，设置集群自动快照和其他高级功能。
   a. 设置集群自动快照开关、基础配置和快照配置。

系统默认打开集群快照开关，如果您不需要启用自动快照，可以在“集群快照开关”右侧关闭。自动快照会创建委托访问对象存储服务 OBS，快照存储在标准存储中需额外计费。

表4-6 集群快照基础配置的参数说明

参数	说明
OBS 桶	在下拉框中选择存储快照的 OBS 桶。也可以单击右侧的“创建桶”新建 OBS。 创建或者已存在的 OBS 桶需满足如下条件： • “存储类别”为“标准存储”。 • “区域”须与创建的集群所在区域相同。
备份路径	快照在 OBS 桶中的存放路径。 备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
IAM 委托	指当前账号授权云搜索服务访问或维护存储在 OBS 中数据。也可以单击右侧的“创建委托”新建 IAM 委托。 创建或者已存在的 IAM 委托需满足如下条件： • “委托类型”选择“云服务”。 • “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 设置当前委托具备“全局服务”中“对象存储服务”项目的“OBS Administrator”权限。

表4-7 集群快照自动创建快照的参数说明

参数	说明
快照名称前缀	快照名称前缀的长度为 1～32 个字符，只能包含小写字母、数字、中划线和下划线，且必须以小写字母开头。快照名称由快照名称前缀加上时间戳组成，例如自动生成的快照名称为“snapshot-1566921603720”。
时区	指备份时间对应的时区，不支持修改。基于此时区选择备份开始时间。
备份开始时间	指每天自动开始备份的时间，只能指定整点时间，如 00:00、01:00，取值范围为 00:00～23:00。请在下拉框中选择时间。

b. 配置集群高级功能。

- **默认配置**：默认关闭“终端节点服务”、“Kibana 公网访问”和“标签”功能，在集群创建完成后，如果有需要也可以人工启用这些功能。
- **自定义**：根据需要选择开启“终端节点服务”、“Kibana 公网访问”和“标签”功能。

表4-8 高级配置参数

参数	说明
Kibana 公网访问	只有开启“安全模式”的集群，才能配置 Kibana 公网访问。开启 Kibana 公网访问后，用户可以获得一个 Kibana 公网访问地址，通过这个地址，可以在公网上面访问该集群。详细配置请参考 4.13.2 Kibana 公网访问集群。
标签	为集群添加标签，可以方便用户识别和管理拥有的集群资源。此处您可以选择“标签管理服务”中已定义好的“预定义标签”，也可以自己定义标签。详细标签使用请参考 4.7.3 管理标签。 如您的组织已经设定云搜索服务的相关标签策略，则需按照标签策略规则为集群添加标签。标签如果不符合标签策略的规则，则可能会导致集群创建失败，请联系组织管理员了解标签策略详情。

9. 单击“下一步：确认配置”，确认完成后单击“立即创建”开始创建集群。
10. 单击“返回集群列表”，系统将跳转到“集群管理”页面。您创建的集群将展现在集群列表中，且集群状态为“创建中”，创建成功后集群状态会变为“可用”。

    如果集群创建失败，请根据界面提示，重新创建集群。

4.1.2 创建非安全模式集群

参考本章节可以完成非安全模式的 Elasticsearch 集群的创建。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“总览”页面单击右上角的“创建集群”，进入“创建集群”页面。

   或者左侧导航栏单击“集群管理>Elasticsearch”，单击右上角的“创建集群”，进入“创建集群”页面。
3. 选择“当前区域”和“可用区”。

表4-9 区域和可用区参数说明

参数	说明
当前区域	集群工作区域在右侧下拉框中选择。

参数	说明
可用区	选择集群工作区域下关联的可用区。 最多支持配置 3 个“可用区”，多可用区的使用请参见 4.1.4 部署跨 AZ 集群。

4. 配置集群基本信息。

表4-10 基本参数说明

参数	说明
集群版本	选择所需的集群版本，支持的版本以界面可选项为准。
集群名称	自定义集群名称，可输入的字符范围为 4～32 个字符，只能包含数字、字母、中划线和下划线，且必须以字母开头。 说明 当集群创建成功后，您可以根据需求修改集群名称。单击需要修改的集群名称，进入集群基本信息页面，单击“集群名称”后面的，修改完成后，单击，进行保存。如果需要取消修改，可单击进行取消。

5. 配置集群的规格信息。

表4-11 规格参数说明

参数	说明
节点数量	集群中的节点个数。可选节点数为 1~32，建议节点数为 3 或 3 以上，以提升集群可用性。 • 如果未启用 Master 节点和 Client 节点时，此参数指定的节点将被作为 Master 节点和 Client 节点，同时具备集群管理、存储数据、提供接入集群和分析数据的服务。此时，为保证集群中数据的稳定性，建议设置节点数量大于等于 3 个。 • 如果启用 Master 节点，且未启用 Client 节点，此参数指定的节点将用于存储数据并提供 Client 节点功能。 • 如果已启用 Master 节点和 Client 节点，此参数指定的节点将仅用于存储数据。 • 如果启用 Client 节点，且未启用 Master 节点，此参数指定的节点将用于存储数据并提供 Master 节点功能。
CPU 架构	支持“X86 计算”和“鲲鹏计算”两种类型。具体支持的类型由实际区域环境决定。

参数	说明
节点规格	集群中的节点规格。您可以根据需求，选择对应的规格。每个集群只能选择一个规格。
节点存储	选择存储类型，支持普通 I/O、高 I/O、超高 I/O。
节点存储容量	存储空间大小，其取值范围与节点规格关联，不同的规格允许的取值范围不同。 节点存储容量只支持配置为 20 的倍数。
启用 Master 节点	Master 节点用于管理集群中的所有节点。当需要存储和分析的数据量大，所需节点数量大于 20 个节点时，建议启用 Master 节点，保证集群的稳定性。反之，建议仅设置集群的“节点数量”参数，同时作为 Master 和 Client 节点即可。 启用 Master 节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”必须是不小于 3 的奇数，最多设置 9 个节点。“节点存储”的存储容量为固定值，存储类型可以根据实际情况选择。
启用 Client 节点	Client 节点用于提供客户端接入集群和分析数据的服务。当需要存储和分析的数据量大，所需节点数量大于 20 个节点时，建议启用 Client 节点，保证集群的稳定性。反之，建议仅设置集群的“节点数量”参数，同时作为 Master 和 Client 节点即可。 启用 Client 节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”可设置为 1~32 任意数值。“节点存储”的存储容量为固定值，存储类型可以根据实际情况选择。
启用冷数据节点	当用户对历史数据返回时间要求不是很高的话，可以将这部分数据存储在冷数据节点上，从而降低成本。 启用冷数据节点后，在下方选择对应的“节点规格”、“节点数量”和“节点存储”。“节点数量”可设置为 1~32 任意数值。“节点存储”的存储类型和存储容量可以根据实际情况选择。 开启冷数据节点之后，云搜索服务将会自动的给相关节点打上冷热标签。

6. 设置集群的企业项目。

   如果您开通了“企业项目”，在创建集群时，可以给集群绑定一个企业项目。您可以在右侧下拉框中选择当前用户下已创建的企业项目，也可以通过单击“查看项目管理”按钮，前往“企业项目管理”管理控制台，新建企业项目和查看已有的企业项目。

7. 指定集群的网络规格相关参数。

表4-12 参数说明

参数	说明
虚拟私有云	VPC 即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。 选择创建集群需要的 VPC，单击“查看虚拟私有云”进入 VPC 服务查看已创建的 VPC 名称和 ID。如果没有 VPC，需要创建一个新的 VPC。 说明 此处您选择的 VPC 必须包含网段（CIDR），否则集群将无法创建成功。新建的 VPC 默认包含网段（CIDR）。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。 选择创建集群需要的子网，可进入 VPC 服务查看 VPC 下已创建的子网名称和 ID。
安全组	安全组是一个逻辑上的分组，为同一个 VPC 内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。单击“查看安全组”可了解安全组详情。 说明 • 为了确保您能够正常访问集群，需要放通安全组 9200 规则。 • 如果创建的集群为 7.6.2 及以上版本，则需要确保同安全组内节点之间的端口全放通。如果无法放通同安全组内节点之间的全部端口，请至少确保 9300 端口的通信。 • 放开 9300 端口通信后，如果集群磁盘使用率较高，可清理过期数据，释放磁盘存储空间。
安全模式	关闭安全模式。

8. 单击“下一步：高级配置”，设置集群自动快照和其他高级功能。
   a. 设置集群自动快照开关、基础配置和快照配置。
      系统默认打开集群快照开关，如果您不需要启用自动快照，可以在“集群快照开关”右侧关闭。自动快照会创建委托访问对象存储服务 OBS，快照存储在标准存储中需额外计费。

表4-13 集群快照基础配置的参数说明

参数	说明
OBS 桶	在下拉框中选择存储快照的 OBS 桶。也可以单击右侧的“创建桶”新建 OBS。 创建或者已存在的 OBS 桶需满足如下条件： • “存储类别”为“标准存储”。

参数	说明
	• “区域”须与创建的集群所在区域相同。
备份路径	快照在 OBS 桶中的存放路径。 备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
IAM 委托	指当前账号授权云搜索服务访问或维护存储在 OBS 中数据。也可以单击右侧的“创建委托”新建 IAM 委托。 创建或者已存在的 IAM 委托需满足如下条件： • “委托类型”选择“云服务”。 • “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 设置当前委托具备“全局服务”中“对象存储服务”项目的“OBS Administrator”权限。

表4-14 集群快照自动创建快照的参数说明

参数	说明
快照名称前缀	快照名称前缀的长度为 1～32 个字符，只能包含小写字母、数字、中划线和下划线，且必须以小写字母开头。快照名称由快照名称前缀加上时间戳组成，例如自动生成的快照名称为“snapshot-1566921603720”。
时区	指备份时间对应的时区，不支持修改。基于此时区选择备份开始时间。
备份开始时间	指每天自动开始备份的时间，只能指定整点时间，如 00:00、01:00，取值范围为 00:00～23:00。请在下拉框中选择时间。

b. 配置集群高级配置功能。

- **默认配置**：默认关闭“终端节点服务”、“Kibana 公网访问”和“标签”功能，在集群创建完成后，如果有需要也可以人工启用这些功能。
- **自定义**：根据需要选择开启“终端节点服务”和“标签”功能。

表4-15 高级配置参数

参数	说明
Kibana 公网访问	非安全模式的集群，不支持启用“Kibana 公网访问”。

参数	说明
标签	为集群添加标签，可以方便用户识别和管理拥有的集群资源。此处您可以选择“标签管理服务”中已定义好的“预定义标签”，也可以自己定义标签。详细标签使用请参考 4.7.3 管理标签。 如您的组织已经设定云搜索服务的相关标签策略，则需按照标签策略规则为集群添加标签。标签如果不符合标签策略的规则，则可能会导致集群创建失败，请联系组织管理员了解标签策略详情。

9. 单击“下一步：确认配置”，确认完成后单击“立即创建”开始创建集群。
10. 单击“返回集群列表”，系统将跳转到“集群管理”页面。您创建的集群将展现在集群列表中，且集群状态为“创建中”，创建成功后集群状态会变为“可用”。

    如果集群创建失败，请根据界面提示，重新创建集群。

4.1.3 安全模式集群

CSS 服务在创建 Elasticsearch 集群时，支持创建安全模式的集群，当集群开启安全模式后，访问集群时需要进行安全认证，且支持对集群进行授权、加密等功能。

安全认证

当访问安全模式的集群时，需要输入用户名和密码才能访问。支持以下两类用户的安全认证：

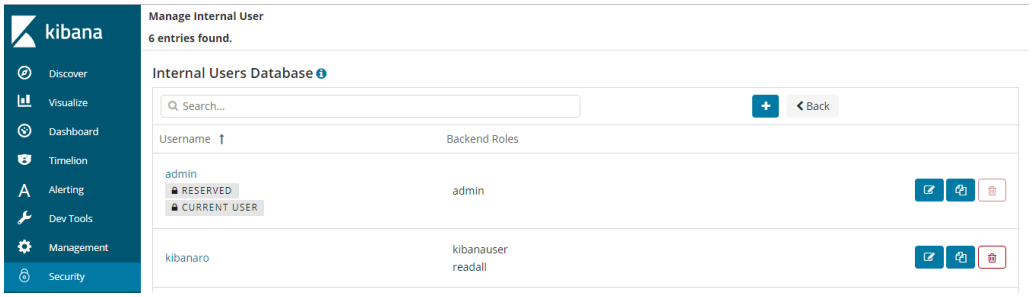
- 集群的管理员：管理员账户名默认为 **admin**，密码为创建集群时设置的管理员密码。
- 集群的 User：通过 kibana 创建的集群用户和密码。

授权

在 kibana 使用界面您可以在 Security 菜单中控制用户在 ES 集群中的权限，并且可以针对集群、索引、文档和字段四个级别进行分层权限设置。详细操作请参见 4.13.3 使用 Kibana 创建用户并授权。

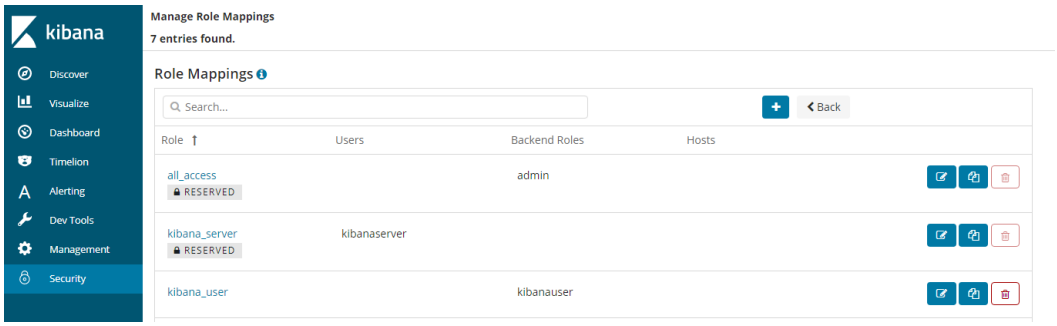
您可以增删用户，并将用户映射到角色类型设置权限。

图4-1 用户设置



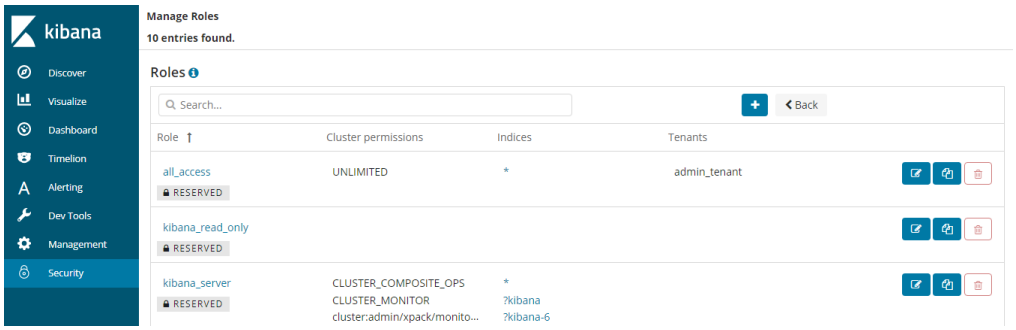
可以使用角色映射配置角色成员，可使用用户名、后端角色和主机名将用户分配给角色。

图4-2 角色映射



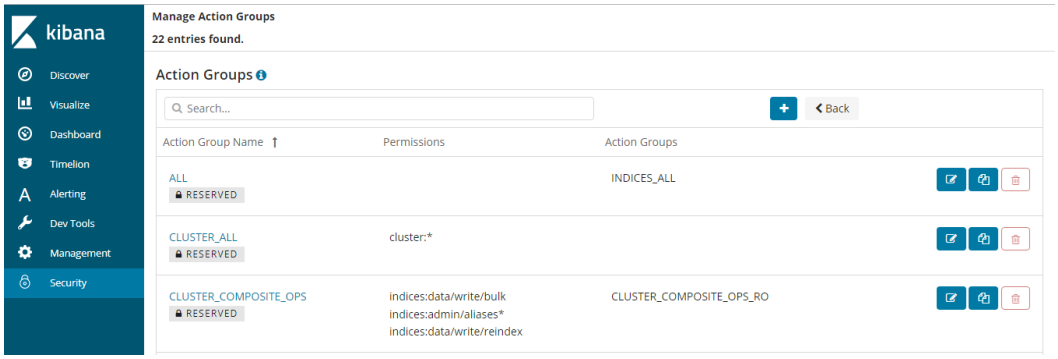
可以设置每种角色的集群访问权限、索引和文档访问权限以及 kibana 租户。

图4-3 角色权限设置



可以设置操作组，并将操作组分配给角色配置角色对索引和文档类型的访问权限。

图4-4 操作组设置

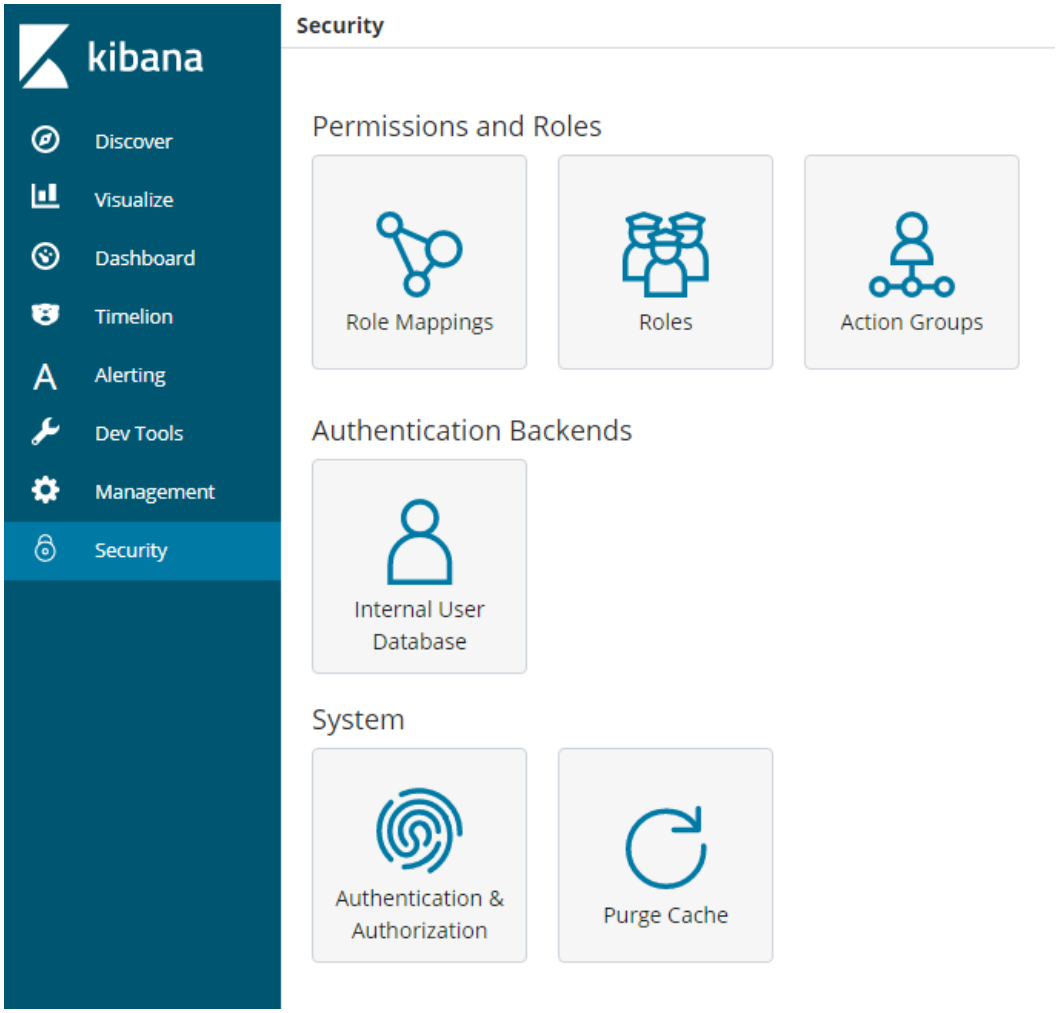


可以查询集群当前设置的身份验证及授权模块的参数。使用 securityadmin 命令行可修改相关配置。

图4-5 集群参数查看

最后，安全模块还为您提供了清除所有安全缓存的功能。

图4-6 安全缓存清除



加密

当您使用节点对节点传输或者 HTTP 传输方式传输关键数据时，可以借助 SSL/TLS 加密，对数据安全进行保护。

以上功能除了可以使用 Kibana 可视化界面操作，还可以使用.yml 文件（不推荐）和 REST API 操作，更多安全模式相关内容可以查看安全模式官方介绍。

重置管理员密码

当您想要更换安全模式集群的管理员密码，或者忘记管理员密码时，可以对密码进行重置。

1. 在集群管理列表，选择需要重置密码的集群，单击集群名称，进入集群基本信息页面。
2. 在“配置信息”区域，单击“重置密码”后的“重置”，设置并确认新的管理员密码。

说明

- 可输入的字符串长度为 8~32 个字符。
- 密码至少包含大写字母、小写字母、数字和特殊字符四类中的三类。其中支持的特殊字符有：~!@#$%^&*()-_=+\|[{}];:,<.>/?
- 不能与管理员账户名或倒序的管理员账户名相同。
- 建议定期修改密码。

4.1.4 部署跨 AZ 集群

为了防止数据丢失，或者在服务中断时能够最大限度地减少集群停机时间，CSS 服务支持集群选择跨 AZ 部署，提升集群的高可用能力。在创建集群时选择同一个区域中的两个或三个可用区，系统自动在选择的可用区之间分配节点。

选择节点数

当创建集群时，选择了两个或者三个可用区时，CSS 服务将自动为开启跨 AZ 高可用特性，节点将会被均衡的分布在不同的 AZ，不同节点数量的 AZ 分布情况可以参见表 4-16。

说明

- 创建集群时，选择的节点数量要大于等于 AZ 数量，否则不支持跨 AZ 部署。
- 部署跨 AZ 集群时，如果选择了“启用 Master 节点”，Master 节点也会被均匀的分布在不同的 AZ 上。
- 系统分配的节点，满足各个 AZ 之间节点数量差小于等于 1。

表4-16 节点数量和 AZ 分布

集群节点个数	单 AZ	两 AZ		三 AZ		
	AZ1	AZ1	AZ2	AZ1	AZ2	AZ3
1 个节点	1	不支持		不支持		
2 个节点	2	1	1	不支持		
3 个节点	3	2	1	1	1	1
4 个节点	4	2	2	2	1	1
…	…	…	…	…	…	…

设置副本

设置副本能有效的利用 AZ 的高可用能力。

- 在跨两个可用区的部署中，当其中一个 AZ 不可用时，剩下的 AZ 需要继续提供服务，因此索引的副本个数至少为 1 个。由于 Elasticsearch 默认副本数为 1 个，因此如果您对读性能没有特殊要求，可以直接使用默认值。
- 在跨三个可用区部署中，为了保证其中任意一个 AZ 不可用时，剩余的 AZ 可以继续提供服务，因此索引的副本数至少要为 1 个。为了提高集群的查询能力，也可以设置更多的副本。由于 Elasticsearch 默认的副本数为 1 个，因此需要用户修改 setting 配置来实现修改索引副本个数。

  可以通过如下命令修改索引的副本个数，如：

  curl -XPUT http://ip:9200/{index_name}/_settings -d '{"number_of_replicas":2}'

  也可以通过在模板中指定所有索引的副本个数，如：

  curl -XPUT http://ip:9200/ _template/templatename -d '{ "template": "*", "settings": {"number_of_replicas": 2}}'

说明

- ip：内网访问地址。
- index_name：索引名称。
- number_of_replicas：修改后的索引副本个数。命令中的取值表示修改为 2 个索引副本。

可用区中断的行为分析

当创建集群时，选择两个或三个 AZ，如果一个 AZ 故障，业务故障行为分析如表 4-17 所示。

表4-17 AZ 故障的业务故障行为分析

选择的 AZ 数量	开启主节点个数	业务中断行为
2	0	• 如果节点个数为 2 的倍数： – 一半的数据节点故障，需要替换故障可用区中的一个节点，才能继续选择主节点。 • 如果节点数为奇数： – 故障 AZ 含多一个节点，需要替换故障可用区中一个节点，才能继续选择主节点。相关替换请联系技术支持。 – 故障 AZ 含少一个节点，不中断业务，能够继续选择主节点。
2	3	有 50%机会的停机时间。当两个专用主节点分配到一个可用区中，一个主节点分配到另一个可用区中时： • 如果具有一个专用主节点的可用区遇到中断，则剩余可用区具有两个专用主节点，这两个专用主节点可以选择出主节点。 • 如果具有两个专用主节点的可用区遇到中断，剩余可用区只有一个专用主节点，无法选择出主节点，业务中断，需要联系技术支持。

选择的AZ数量	开启主节点个数	业务中断行为
3	0	当您选择3个可用区，节点个数为4，三个可用区的节点分布数为2，1，1，如果节点个数为2的可用区故障，那么此时业务中断，建议您选择三个可用区时避免选择4个节点。 一般不会出现业务中断时间。
3	3	无业务中断时间。

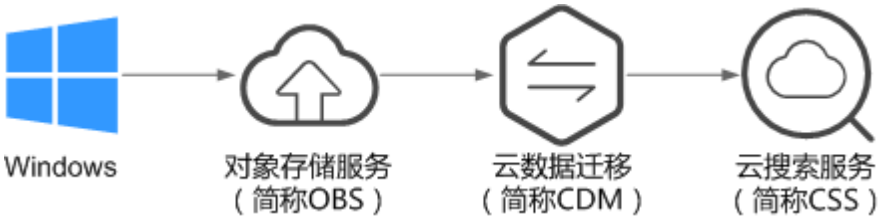
4.2 导入数据

4.2.1 使用 CDM 从 OBS 导入数据到 Elasticsearch

云搜索服务支持通过 CDM 的向导式界面，将存储在对象存储服务（简称 OBS）中的数据导入到 Elasticsearch 中。数据文件支持 JSON、CSV 等格式。

数据传输流程如图 4-7 所示。

图4-7 使用 CDM 从 OBS 导入数据到 Elasticsearch 时的数据传输流程



操作步骤

1. 登录 OBS 管理控制台。
2. 创建待存储数据的 OBS 桶。
   具体操作请参见《对象存储服务控制台指南》中的创建桶。
   创建的 OBS 桶需满足“区域”必须跟创建集群的区域保持一致。
3. 将数据文件上传到 OBS 桶中。
   具体操作请参见《对象存储服务控制台指南》中的上传文件。
   例如：将如下数据保存为 json 格式的文件，上传到创建的 OBS 桶中。

```
{"productName":"2017 秋装新款文艺衬衫女装","size":"L"}
{"productName":"2017 秋装新款文艺衬衫女装","size":"M"}
{"productName":"2017 秋装新款文艺衬衫女装","size":"S"}
{"productName":"2018 春装新款牛仔裤女装","size":"M"}
{"productName":"2018 春装新款牛仔裤女装","size":"S"}
```

```
{"productName":"2017 春装新款休闲裤女装","size":"L"}
{"productName":"2017 春装新款休闲裤女装","size":"S"}
```

4. 登录云搜索服务管理控制台。
5. 在左侧导航栏中，选择“集群管理 > Elasticsearch”，进入集群管理列表页面。
6. 在集群列表页面中，单击待导入数据的集群“操作”列的“Kibana”。
7. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。
8. 在 Console 界面，执行命令查看集群是否存在索引。

```
GET _cat/indices?v
```

   - 如果待导入数据的集群已存在可用的索引，则不需要再创建索引，请执行步骤 10。
   - 如果待导入数据的集群不存在可用的索引，则需要执行下一步创建索引。

9. 在 Console 界面，执行命令创建待存储数据的索引，并指定自定义映射来定义数据类型。

   例如：在 Console 界面，执行如下命令，创建索引“demo”，并指定自定义映射来定义数据类型。

   低于 7.x 版本

```
PUT /demo
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "products": {
      "properties": {
        "productName": {
          "type": "text",
          "analyzer": "ik_smart"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
}
```

   7.x 版本及高于 7.x 版本

```
PUT /demo
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
      "properties": {
        "productName": {
          "type": "text",
          "analyzer": "ik_smart"
        },
        "size": {
          "type": "keyword"
```

```
      }
     }
    }
  }
```

执行成功后显示如下：

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "demo"
}
```

10. 登录 CDM 管理控制台。
11. 购买云数据迁移集群。

    具体操作请参见《云数据迁移用户指南》中的创建集群。
12. 新建 CDM 和云搜索服务的连接。

    具体操作请参见《云数据迁移用户指南》中的新建连接。
13. 新建 CDM 和 OBS 的连接。

    具体操作请参见《云数据迁移用户指南》中的新建连接。
14. 在已购买的云数据迁移集群上新建作业，将 OBS 桶中的数据迁移到云搜索服务的待导入数据的集群中。

    具体操作请参见《云数据迁移用户指南》中的表/文件迁移。
15. 在已打开的 Kibana 的 Console 界面，通过搜索获取已导入的数据。

    在 Kibana 控制台，执行如下命令，搜索数据。查看搜索结果，如果数据与导入数据一致，表示数据文件的数据已导入成功。

```
GET demo/_search
```

执行成功后显示如下：

```
{
  "took": 18,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 7,
    "max_score": 1,
    "hits": [
      {
        "_index": "demo",
        "_type": "products",
        "_id": "g6UepnEBuvdFwWkRmn4V",
        "_score": 1,
        "_source": {
          "size": """size":"L"}""",
          "productName": """{"productName":"2017 秋装新款文艺衬衫女装"""
        }
      },
```

```
{
  "_index": "demo",
  "_type": "products",
  "_id": "hKUepnEBuvdFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """size":"M"}""",
    "productName": """{"productName":"2017 秋装新款文艺衬衫女装"""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "haUepnEBuvdFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """size":"S"}""",
    "productName": """{"productName":"2017 秋装新款文艺衬衫女装"""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "hqUepnEBuvdFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """size":"M"}""",
    "productName": """{"productName":"2018 春装新款牛仔裤女装"""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "h6UepnEBuvdFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """size":"S"}""",
    "productName": """{"productName":"2018 春装新款牛仔裤女装"""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "iKUepnEBuvdFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """size":"L"}""",
    "productName": """{"productName":"2017 春装新款休闲裤女装"""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "iaUepnEBuvdFwWkRmn4V",
  "_score": 1,
```

```
      "_source": {
        "size": """size":"S"}""",
        "productName": """{"productName":"2017 春装新款休闲裤女装"""
      }
    }
  ]
 }
}
```

说明

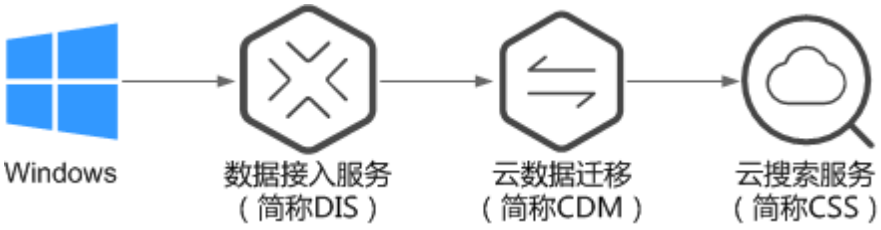
demo 为创建的索引名称，需根据实际情况填写。

4.2.2 使用 DIS 导入本地数据到 Elasticsearch

通过 DIS 可以将本地 windows 系统上的日志数据上传到 DIS 队列中，然后通过 CDM 可以将 DIS 队列中的数据迁移到云搜索服务的 Elasticsearch 中，从而方便用户使用 Elasticsearch 搜索引擎高效管理和获取日志。数据文件支持 JSON、CSV 等格式。

数据传输流程如图 4-8 所示。

图4-8 使用 DIS 导入本地数据到 Elasticsearch 时的数据传输流程



操作步骤

1. 登录 DIS 管理控制台。
2. 购买接入通道。
   具体操作请参见《数据接入服务用户指南》中的开通 DIS 通道。
3. 安装并配置 DIS Agent。
   具体操作请参见《数据接入服务用户指南》中的安装 DIS Agent 和。
4. 启动 DIS Agent，将采集的本地数据上传到 DIS 队列中。
   具体操作请参见《数据接入服务用户指南》中的启动 DIS Agent。
   例如：将如下数据通过 DIS Agent 上传到 DIS 队列中。

```
{"logName":"aaa","date":"bbb"}
{"logName":"ccc","date":"ddd"}
{"logName":"eee","date":"fff"}
{"logName":"ggg","date":"hhh"}
{"logName":"mmm","date":"nnn"}
```

5. 登录云搜索服务管理控制台。
6. 在左侧导航栏中，选择“集群管理 > Elasticsearch”，进入集群管理列表页面。
7. 在集群列表页面中，单击待导入数据的集群“操作”列的“Kibana”。

8. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。
9. 在 Console 界面，执行命令查看集群是否存在索引。

```
GET _cat/indices?v
```

如果待导入数据的集群已存在可用的索引，则不需要再创建索引，请执行步骤11。

如果待导入数据的集群不存在可用的索引，则需要执行下一步创建索引。

10. 在 Console 界面，执行命令创建待存储数据的索引，并指定自定义映射来定义数据类型。

例如：在 Console 界面，执行如下命令，创建索引“apache”，并指定自定义映射来定义数据类型。

低于 7.x 版本

```
PUT /apache
{
   "settings": {
      "number_of_shards": 1
   },
   "mappings": {
      "logs": {
         "properties": {
            "logName": {
               "type": "text",
               "analyzer": "ik_smart"
            },
            "date": {
               "type": "keyword"
            }
         }
      }
   }
}
```

7.x 版本及高于 7.x 版本

```
PUT /apache
{
   "settings": {
      "number_of_shards": 1
   },
   "mappings": {
              "properties": {
            "logName": {
               "type": "text",
               "analyzer": "ik_smart"
            },
            "date": {
               "type": "keyword"
            }
         }
      }
}
```

执行成功后显示如下：

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "apache"
}
```

11. 登录 CDM 管理控制台。
12. 购买云数据迁移集群。

    具体操作请参见《云数据迁移用户指南》中的创建集群。
13. 新建 CDM 和云搜索服务的连接。

    具体操作请参见《云数据迁移用户指南》中的新建连接。
14. 新建 CDM 和 DIS 的连接。

    具体操作请参见《云数据迁移用户指南》中的新建连接。
15. 在已购买的云数据迁移集群上新建作业，将 DIS 队列中的数据迁移到云搜索服务的待导入数据的集群中。

    具体操作请参见《云数据迁移用户指南》中的表/文件迁移。
16. 在已打开的 Kibana 的 Console 界面，通过搜索获取已导入的数据。

    在 Kibana 控制台，输入如下命令，搜索数据。查看搜索结果，如果数据与导入数据一致，表示数据文件的数据已导入成功。

    ```
    GET apache/_search
    ```

    执行成功后显示如下：

    ```
    {
      "took": 81,
      "timed out": false,
      " shards": {
        "total": 1,
        "successful": 1,
        "skipped": 0,
        "failed": 0
      },
      "hits": {
        "total": 5,
        "max_score": 1,
        "hits": [
          {
            "_index": "apache",
            "_type": "logs",
            "_id": "txfbqnEBPuwwWJWL-qvP",
            "_score": 1,
            "_source": {
              "date": """{"logName":"aaa""",
              "logName": """date":"bbb"}"""
            }
          },
          {
            "_index": "apache",
            "_type": "logs",
            "_id": "uBfbqnEBPuwwWJWL-qvP",
            "_score": 1,
            "_source": {
    ```

```
        "date": """{"logName":"ccc""",
        "logName": """date":"ddd"}"""
      }
    },
    {
      "_index": "apache",
      "_type": "logs",
      "_id": "uRfbqnEBPuwwWJWL-qvP",
      "_score": 1,
      "_source": {
        "date": """{"logName":"eee""",
        "logName": """date":"fff"}"""
      }
    },
    {
      "_index": "apache",
      "_type": "logs",
      "_id": "uhfbqnEBPuwwWJWL-qvP",
      "_score": 1,
      "_source": {
        "date": """{"logName":"ggg""",
        "logName": """date":"hhh"}"""
      }
    },
    {
      " _index": "apache",
      " _type": "logs",
      "_id": "uxfbqnEBPuwwWJWL-qvP",
      "_score": 1,
      "_source": {
        "date": """{"logName":"mmm""",
        "logName": """date":"nnn"}"""
      }
    }
  ]
 }
}
```

说明

apache 为创建的索引名称，需根据实际情况填写。

4.2.3 使用 Logstash 导入数据到 Elasticsearch

云搜索服务支持使用 Logstash 将其收集的数据迁移到 Elasticsearch 中，方便用户通过 Elasticsearch 搜索引擎高效管理和获取数据。数据文件支持 JSON、CSV 等格式。

Logstash 是开源的服务器端数据处理管道，能够同时从多个来源采集数据、转换数据，然后将数据发送到 Elasticsearch 中。Logstash 的官方文档请参见：https://www.elastic.co/guide/en/logstash/current/getting-started-with-logstash.html。

数据导入分为如下 2 种场景：

- Logstash 部署在外网时导入数据
- Logstash 部署在弹性云服务器上时导入数据

前提条件

- 为方便操作，建议采用 Linux 操作系统的机器部署 Logstash。
- Logstash 的下载路径为：https://www.elastic.co/cn/downloads/logstash-oss

说明

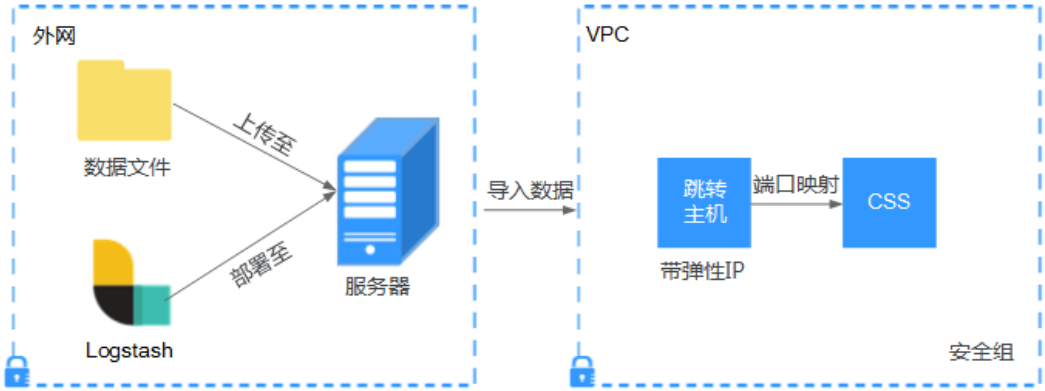
Logstash 要求使用 OSS 版本，选择和 CSS 一致版本。

- 安装 Logstash 之前，需要先安装 JDK。在 Linux 操作系统中，您可以执行 **yum -y install java-1.8.0** 命令直接安装 1.8.0 版本 JDK。在 Windows 操作系统中，您可以访问 JDK 官网，下载符合操作系统版本的 JDK，并根据指导安装。
- 安装完 Logstash 后，再根据如下步骤导入数据。安装 Logstash 的操作指导，请参见：https://www.elastic.co/guide/en/logstash/current/installing-logstash.html
- 在“Logstash 部署在弹性云服务器上时导入数据”场景中，请确保此弹性云服务器与接入的 Elasticsearch 集群在同一个 VPC 下。

Logstash 部署在外网时导入数据

当 Logstash 部署在外网时，导入数据的流程说明如图 4-9 所示。

图4-9 Logstash 部署在外网时导入数据示意图



1. 创建一个跳转主机，并按如下要求进行配置。
   - 跳转主机为一台 Linux 操作系统的弹性云服务器，且已绑定弹性 IP。
   - 跳转主机与 CSS 集群在同一虚拟私有云下。
   - 已开放跳转主机的本地端口，用于 SSH 转发，能够从本地端口转发至 CSS 集群某一节点的 9200 端口。
   - 关于跳转主机的本地端口转发配置，请参见 SSH 官方文档。
2. 使用 PuTTY，通过弹性 IP 登录已创建的跳转主机。
3. 执行如下命令进行端口映射，将发往跳转主机对外开放端口的请求转发到待导入数据的集群中。

```
ssh -g -L <跳转主机的本地端口:节点的内网访问地址和端口号> -N -f root@<跳转主机的私网IP地址>
```

说明

- <*跳转主机的本地端口*>：为步骤 1 中的端口。
- <*节点的内网访问地址和端口号*>：为集群中某一节点的内网访问地址和端口号。当该节点出现故障时，将导致命令执行失败。如果集群包含多个节点，可以将<*节点的内网访问地址和端口号*>替换为集群中另一节点的内网访问地址和端口号；如果集群只包含一个节点，则需要将该节点修复之后再次执行命令进行端口映射。
- <*跳转主机的私网 IP 地址*>：打开弹性云服务器管理控制台，从“IP 地址”列中获取标有“私网”对应的 IP 地址。

例如：跳转主机对外开放的端口号为 9200，节点的内网访问地址和端口号为 192.168.0.81:9200，跳转主机的私网 IP 地址为 192.168.0.227，需要执行如下命令进行端口映射。

```
ssh -g -L 9200:192.168.0.81:9200 -N -f root@192.168.0.227
```

4. 登录部署了 Logstash 的服务器，将需要进行操作的数据文件存储至此服务器中。

   例如，需要导入的数据文件“access_20181029_log”，文件存储路径为“/tmp/access_log/”，此数据文件中包含的数据如下所示：

   说明

   文件存储路径中的 access_log 文件夹如果不存在，用户可以自建。

```
|   All |             Heap used for segments |                       |    18.6403 |
MB |
|   All |           Heap used for doc values |                       |   0.119289
|     MB |
|   All |               Heap used for terms |                       |    17.4095 |
MB |
|   All |               Heap used for norms |                       |  0.0767822 |
MB |
|   All |              Heap used for points |                       |   0.225246 |
MB |
|   All |        Heap used for stored fields |                       |   0.809448
|     MB |
|   All |                     Segment count |                       |        101 |
|
|   All |                    Min Throughput |         index-append |    66232.6
|  docs/s |
|   All |                 Median Throughput |         index-append |    66735.3
|  docs/s |
|   All |                    Max Throughput |         index-append |    67745.6
|  docs/s |
|   All |            50th percentile latency |         index-append |
510.261 |     ms |
```

5. 在部署 Logstash 的服务器中，执行如下命令在 Logstash 的安装目录下新建配置文件 logstash-simple.conf。

```
cd /<Logstash 的安装目录>/
vi logstash-simple.conf
```

6. 在配置文件 logstash-simple.conf 中输入如下内容。

```
input {
数据所在的位置
}
```

```
filter {
数据的相关处理
}
output {
   elasticsearch {
      hosts => "<跳转主机的公网 IP 地址>:<跳转主机对外开放的端口号>"
   }
}
```

- input：指定数据的来源。实际请根据用户的具体情况来设置。input 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/input-plugins.html。
- filter：指定对数据进行处理的方式。例如，对日志进行了提取和处理，将非结构化信息转换为结构化信息。filter 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/filter-plugins.html。
- output：指定数据的目的地址。output 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/output-plugins.html。*<跳转主机的公网 IP 地址>*请从弹性云服务器管理控制台的“IP 地址”列中获取标有“弹性公网”对应的 IP 地址。*<跳转主机对外开放的端口号>*即为步骤 1 中的端口，例如：9200。

以步骤 4 中“/tmp/access_log/”的数据文件为例，输入数据文件从首行开始，且过滤条件保持为空，即不做任何数据处理操作。跳转主机的公网 IP 和端口号为“192.168.0.227:9200”。导入数据的索引名称为“myindex”。配置文件的示例如下所示，配置文件按实际数据情况修改完成后，输入“:wq”保存。

```
input {
   file{
     path => "/tmp/access_log/*"
     start_position => "beginning"
   }
}
filter {
}
output {
   elasticsearch {
     hosts => "192.168.0.227:9200"
     index => "myindex"

   }
}
```

说明

如果在使用中出现 license 相关的报错，可以尝试设置 ilm_enabled => false。

如果集群开启了安全模式，则需要先下载证书。

a. 在云搜索服务管理控制台。
b. 在“集群管理”页面，单击需要下载证书的集群名称，进入集群基本信息页面。
c. 在集群基本信息页面下载证书。

图4-10 下载证书

d. 将下载的证书存放到部署 logstash 服务器中。

e. 修改配置文件 logstash-simple.conf。

以步骤 4 中“/tmp/access_log/”的数据文件为例，输入数据文件从首行开始，且过滤条件保持为空，即不做任何数据处理操作。跳转主机的公网 IP 和端口号为“192.168.0.227:9200”。导入数据的索引名称为“myindex”，证书存放路径为“/logstash/logstash6.8/config/CloudSearchService.cer”。配置文件的示例如下所示，配置文件按实际数据情况修改完成后，输入“:wq”保存。

```
input{
    file {
        path => "/tmp/access_log/*"
        start_position => "beginning"
    }
}
filter {
    }
output{
    elasticsearch{
```

```
        hosts => ["https://192.168.0.227:9200"]
        index => "myindex"
        user => "admin"
        password => "******"
        cacert => "/logstash/logstash6.8/config/CloudSearchService.cer"
        manager_template => false
        ilm_enabled => false
        ssl => true
        ssl_certificate_verification => false
    }
}
```

说明

password：登录安全集群的密码。

7. 执行如下命令将 Logstash 收集的数据导入到集群中。

```
./bin/logstash -f logstash-simple.conf
```

说明

此命令需要在存放 logstash-simple.conf 文件的目录下执行。例如，logstash-simple.conf 文件存放在/root/logstash-7.1.1/，则需要先进入该路径，再执行此命令。

8. 登录云搜索服务管理控制台。
9. 在左侧导航栏中，选择“集群管理 > Elasticsearch”，进入集群管理列表页面。
10. 在集群列表页面中，单击待导入数据的集群“操作”列的“Kibana”。
11. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。
12. 在已打开的 Kibana 的 Console 界面，通过搜索获取已导入的数据。

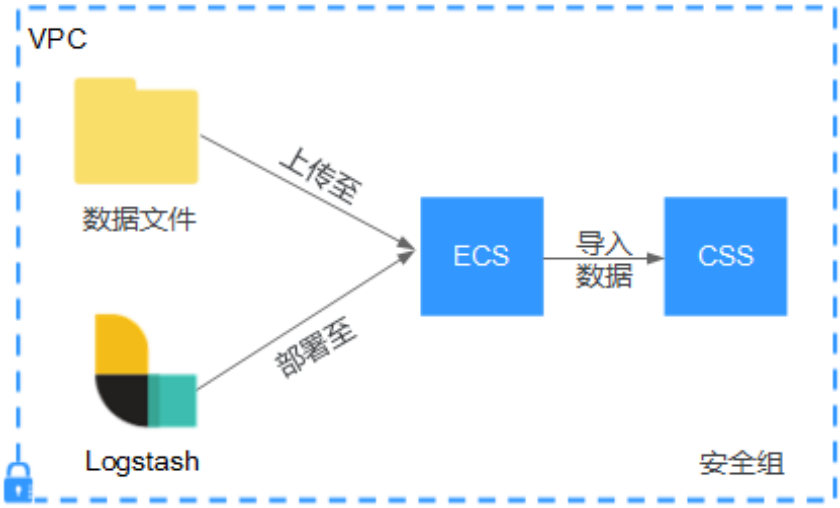
    在 Kibana 控制台，输入如下命令，搜索数据。查看搜索结果，如果数据与导入数据一致，表示数据文件的数据已导入成功。

```
GET myindex/_search
```

Logstash 部署在弹性云服务器上时导入数据

当 Logstash 部署在同一 VPC 的弹性云服务时，导入数据的流程说明如图 4-11 所示。

图4-11 Logstash 部署在弹性云服务器上时导入数据示意图



1. 确保已部署 Logstash 的弹性云服务器与待导入数据的集群在同一虚拟私有云下，已开放安全组的 9200 端口的外网访问权限，且弹性云服务器已绑定弹性 IP。

说明

- 如果同一个 VPC 内有多台服务器，只要其中一台绑定了弹性 IP，其他的服务器可以不需要绑定弹性 IP。通过绑定弹性 IP 的节点跳转到部署 Logstash 的节点即可。
- 如果有专线或者 VPN，也不需要绑定弹性 IP。

2. 使用 PuTTY 登录弹性云服务器。

   例如此服务器中存储了需要导入的数据文件“access_20181029_log”，文件存储路径为“/tmp/access_log/”，此数据文件中包含的数据如下所示：

```
|   All |             Heap used for segments |                    |    18.6403 |
MB |
|   All |           Heap used for doc values |                    |    0.119289
|     MB |
|   All |               Heap used for terms |                    |    17.4095 |
MB |
|   All |               Heap used for norms |                    |   0.0767822 |
MB |
|   All |              Heap used for points |                    |    0.225246 |
MB |
|   All |        Heap used for stored fields |                    |    0.809448
|     MB |
|   All |                    Segment count |                    |       101 |
|
|   All |                   Min Throughput |         index-append |    66232.6
|  docs/s |
|   All |                Median Throughput |         index-append |    66735.3
|  docs/s |
|   All |                   Max Throughput |         index-append |    67745.6
|  docs/s |
|   All |            50th percentile latency |         index-append |
510.261 |     ms |
```

3. 执行如下命令在 Logstash 的安装目录下新建配置文件 logstash-simple.conf。

```
cd /<Logstash 的安装目录>/
vi logstash-simple.conf
```

在配置文件 logstash-simple.conf 中输入如下内容。

```
input {
数据所在的位置
}
filter {
数据的相关处理
}
output {
   elasticsearch{
      hosts => "<节点的内网访问地址和端口号>"}
}
```

- input：指定数据的来源。实际请根据用户的具体情况来设置。input 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/input-plugins.html。
- filter：对日志进行了提取和处理，将非结构化信息转换为结构化信息。filter 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/filter-plugins.html。
- output：指定数据的目的地址。output 参数的详细解释和使用介绍请参见 https://www.elastic.co/guide/en/logstash/current/output-plugins.html。<*节点的内网访问地址和端口号*>为集群中节点的内网访问地址和端口号。

  当集群包含多个节点时，为了避免节点故障，建议将上述命令中<*节点的内网访问地址和端口号*>替换为该集群中多个节点的内网访问地址和端口号，多个节点的内网访问地址和端口号之间用英文逗号隔开，填写格式请参见如下示例。

  ```
  hosts => ["192.168.0.81:9200","192.168.0.24:9200"]
  ```

  当集群只包含一个节点时，填写格式请参见如下示例。

  ```
  hosts => "192.168.0.81:9200"
  ```

以步骤 2 中“/tmp/access_log/”的数据文件为例，输入数据文件从首行开始，且过滤条件保持为空，即不做任何数据处理操作。需导入数据的集群，其节点内网访问地址和端口号为“192.168.0.81:9200”。导入数据的索引名称为“myindex”。配置文件的示例如下所示，配置文件按实际数据情况修改完成后，输入“:wq”保存。

```
input {
   file{
     path => "/tmp/access_log/*"
     start_position => "beginning"
   }
}
filter {
}
output {
   elasticsearch {
     hosts => "192.168.0.81:9200"
     index => "myindex"
```

```
    }
}
```

如果集群开启了安全模式，则需要先下载证书。

a. 在集群基本信息页面下载证书。

图4-12 下载证书

b. 将下载的证书存放到部署 logstash 服务器中。

c. 修改配置文件 logstash-simple.conf。

以步骤 2 中“/tmp/access_log/”的数据文件为例，输入数据文件从首行开始，且过滤条件保持为空，即不做任何数据处理操作。跳转主机的公网 IP 和端口号为“192.168.0.227:9200”。导入数据的索引名称为“myindex”，证书存放路径为“/logstash/logstash6.8/config/CloudSearchService.cer”。配置文件的示例如下所示，配置文件按实际数据情况修改完成后，输入“:wq”保存。

```
input{
   file {
      path => "/tmp/access_log/*"
      start_position => "beginning"
```

```
    }
}
filter {
    }
output{
    elasticsearch{
        hosts => ["https://192.168.0.227:9200"]
        index => "myindex"
        user => "admin"
        password => "******"
        cacert => "/logstash/logstash6.8/config/CloudSearchService.cer"
        manager_template => false
        ilm_enabled => false
        ssl => true
        ssl_certificate_verification => false
    }
}
```

说明

password：登录安全集群的密码。

4. 执行如下命令将 Logstash 收集的弹性云服务器的数据导入到集群中。

```
./bin/logstash -f logstash-simple.conf
```

5. 登录云搜索服务管理控制台。
6. 在左侧导航栏中，选择“集群管理 > Elasticsearch”，进入集群管理列表页面。
7. 在集群列表页面中，单击待导入数据的集群“操作”列的“Kibana”。
8. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。
9. 在已打开的 Kibana 的 Console 界面，通过搜索获取已导入的数据。

   在 Kibana 控制台，输入如下命令，搜索数据。查看搜索结果，如果数据与导入数据一致，表示数据文件的数据已导入成功。

```
GET myindex/_search
```

4.2.4 使用 Kibana 或 API 导入数据到 Elasticsearch

云搜索服务支持使用 Kibana 或者 API 将数据导入到 Elasticsearch 中，数据文件支持 JSON 等格式。

使用 Kibana 导入数据

在导入数据之前，您可以使用 Kibana 接入集群。如下操作步骤介绍如何使用 POST 命令导入数据。

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏中，选择“集群管理 > Elasticsearch”，进入集群管理列表页面。
3. 选择已创建的集群，单击操作列“Kibana”，登录 Kibana。
4. 单击左侧导航栏的“Dev Tools”进入操作页面。
5. 在 Kibana 操作界面，执行命令查看集群是否存在索引。

```
GET _cat/indices?v
```

如果待导入数据的集群已存在可用的索引，则不需要再创建索引，请执行步骤 7。

如果待导入数据的集群不存在可用的索引，则需要执行下一步创建索引。

6. 执行命令创建待存储数据的索引，并指定自定义映射来定义数据类型。

   例如：在 Console 界面，执行如下命令，创建索引“my_store”，并指定自定义映射来定义数据类型。

   低于 7.x 版本

```
PUT /my_store
{
   "settings": {
      "number_of_shards": 1
   },
   "mappings": {
      "products": {
         "properties": {
            "productName": {
               "type": "text"
            },
            "size": {
               "type": "keyword"
            }
         }
      }
   }
}
```

   7.x 版本及高于 7.x 版本

```
PUT /my_store
{
   "settings": {
      "number_of_shards": 1
   },
   "mappings": {
      "properties": {
         "productName": {
            "type": "text"
         },
         "size": {
            "type": "keyword"
         }
      }
   }
}
```

7. 执行命令导入数据，以导入一条数据为例，执行如下命令。

   低于 7.x 版本

```
POST /my_store/products/_bulk
{"index":{}}
{"productName":"Latest art shirts for women in 2017 autumn","size":"L"}
```

   7.x 版本及高于 7.x 版本

```
POST /my_store/_bulk
{"index":{}}
{"productName":"Latest art shirts for women in 2017 autumn","size":"L"}
```

返回结果如图 4-13 所示，当返回结果信息中“errors”字段的值为“false”时，表示导入数据成功。

图4-13 返回消息


```
{
  "took": 42,
  "errors": false,
  "items": [
    {
      "index": {
        "_index": "my_store",
        "_type": "products",
        "_id": "AWTGbHt7BwpN-hb3LKau",
        "_version": 1,
        "result": "created",
        "_shards": {
          "total": 2,
          "successful": 2,
          "failed": 0
        },
        "created": true,
        "status": 201
      }
    }
  ]
}
```


使用 API 导入数据

使用 bulk API 通过 curl 命令导入数据文件，如下操作以 JSON 数据文件为例。

说明

- 使用 API 导入数据文件时，建议导入的数据文件大小不能超过 50MB。
- 本案例以非安全模式的集群为例介绍 curl 命令导入数据。

1. 登录即将接入集群的弹性云服务器。
2. 将 JSON 数据文件上传至 ECS。
3. 在 ECS 中存放 JSON 数据文件的路径下执行如下命令，将 JSON 数据导入到 Elasticsearch 集群。

   其中，***{Private network address and port number of the node}***需替换为集群中节点的内网访问地址和端口号，当该节点出现故障时，将导致命令执行失败。如果集群包含多个节点，可以将***{Private network address and port number of the node}***替换为集群中另一节点的内网访问地址和端口号；如果集群只包含一个节点，则需要将该节点修复之后再次执行命令进行导入数据。**test.json** 为导入数据的 json 文件。

```
curl -X PUT "http://{Private network address and port number of the node}
/_bulk" -H 'Content-Type: application/json' --data-binary @test.json
```

说明

其中，-X 参数的参数值为命令，如“-X PUT”，-H 参数的参数值为消息头，如“-H 'Content-Type: application/json' --data-binary @test.json”。添加的-k 参数时，请勿将-k 参数放置在参数与参数值之间。

示例：将“test.json”数据文件中的数据导入至 Elasticsearch 集群，此集群未进行通信加密，其中一个节点内网访问地址为“192.168.0.90”，端口号为“9200”。其中 test.json 文件中的数据如下所示：

低于 7.x 版本

```
{"index": {"_index":"my_store","_type":"products"}}
{"productName": "2019 秋装新款文艺衬衫女装","size": "M"}
{"index": {"_index":"my_store","_type":"products"}}
{"productName": "2019 秋装新款文艺衬衫女装","size": "L"}
```

7.x 版本及高于 7.x 版本

```
{"index": {" index":"my store"}}
{"productName": "2019 秋装新款文艺衬衫女装","size": "M"}
{"index": {" index":"my store"}}
{"productName": "2019 秋装新款文艺衬衫女装","size": "L"}
```

导入数据的操作步骤如下所示：

a. 可执行以下命令，创建 my_store 索引。

低于 7.x 版本

```
curl -X PUT http://192.168.0.90:9200/my_store -H 'Content-Type:
application/json' -d '
 {
   "settings": {
     "number_of_shards": 1
   },
   "mappings": {
     "products": {
       "properties": {
         "productName": {
           "type": "text"
           },
         "size": {
           "type": "keyword"
         }
       }
     }
   }
 }'
```

7.x 版本及高于 7.x 版本

```
curl -X PUT http://192.168.0.90:9200/my_store -H 'Content-Type:
application/json' -d '
{
   "settings": {
      "number_of_shards": 1
   },
   "mappings": {
      "properties": {
         "productName": {
```

```
            "type": "text"
         },
         "size": {
            "type": "keyword"
         }
      }
   }
}'
```

b. 执行以下命令，导入 test.json 文件中的数据。

```
curl -X PUT "http://192.168.0.90:9200/_bulk" -H 'Content-Type:
application/json' --data-binary @test.json
```

本案例回显如下信息，表示数据导入成功。

```
{"took":204,"errors":false,"items":[{"index":{"_index":"my_store","_type":
"_doc","_id":"DJQkBIwBbJvUd2769Wi-
","_version":1,"result":"created","_shards":{"total":2,"successful":1,"fai
led":0},"_seq_no":0,"_primary_term":1,"status":201}},{"index":{"_index":"m
y_store","_type":"_doc","_id":"DZQkBIwBbJvUd2769Wi_","_version":1,"result"
:"created","_shards":{"total":2,"successful":1,"failed":0},"_seq_no":1,"_p
rimary_term":1,"status":201}}]}
```

4.3 访问 Elasticsearch 集群

4.3.1 快速访问 Elasticsearch 集群

CSS 服务创建的 Elasticsearch 集群自带 Kibana 和 Cerebro 组件，支持一键打开 Kibana 和 Cerebro，快速访问 Elasticsearch 集群。

通过 Kibana 访问集群

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择需要登录的集群，单击“操作”列中的“Kibana”进入 Kibana 登录界面。
   - 非安全模式的集群：将直接进入 Kibana 操作界面。
   - 安全模式的集群：需要在登录页面输入用户名和密码，单击“Log In”进入 Kibana 操作界面。用户名默认为 admin，密码为创建集群时设置的管理员密码。
3. 登录成功后，可在 Kibana 界面进行相关操作访问集群。

通过 Cerebro 访问集群

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择需要登录的集群，单击“操作”列中的“更多 > Cerebro”进入 Cerebro 登录页面。
   - 非安全模式的集群：单击 Cerebro 登录页面的集群名称即可进入 Cerebro 操作界面。

- 安全模式的集群：单击 Cerebro 登录页面的集群名称，再输入用户名和密码，单击“Authenticate”进入 Cerebro 操作界面。用户名默认为 admin，密码为创建集群时设置的管理员密码。

3. 登录成功后，可在 Cerebro 界面进行相关操作访问集群。

4.3.2 公网访问集群

针对启用 HTTPS 访问的安全集群（6.5.4 及之后版本的 Elasticsearch 集群支持开启“安全模式”），云搜索服务的集群支持配置公网访问，配置完成后，通过提供的公网 IP，您可以在外网接入安全集群。

CSS 服务默认通过共享型负载均衡器实现公网访问，您也可以通过性能更优的独享型负载均衡器接入集群实现公网访问。

说明

CSS 开启公网访问后，会使用到 EIP 和带宽资源，涉及相关资源费用。

创建集群时配置公网访问

1. 登录云搜索服务管理控制台。
2. 在创建集群页面，开启“安全模式”。设置管理员密码，并启用 HTTPS 访问。
3. “公网访问”选择“自动绑定”，配置公网访问相关参数。

表4-18 公网访问参数说明

参数	说明
带宽	设置公网访问的带宽。
访问控制开关	如果关闭访问控制开关，则允许任何 IP 通过公网 IP 访问集群。如果开启访问控制开关，则只允许白名单列表中的 IP 通过公网 IP 访问集群。
白名单	设置允许访问的 IP 地址或网段，中间用英文逗号隔开。仅当打开“访问控制开关”时才需要配置。

已有集群公网访问管理

您可以对已经创建集群的公网访问进行修改，查看，解绑，也可以配置公网访问。

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击需要配置公网访问的集群名称，进入集群基本信息页面，管理公网访问相关配置。
   - 配置公网访问

     如果创建安全集群时，开启了 HTTPS 访问但未配置公网访问，集群创建成功后，可以在集群基本信息页面配置公网访问。

     单击“公网访问”参数右侧的“绑定”，设置访问带宽后，单击“确定”。

如果绑定失败，用户可以等待几分钟后，再次尝试重新绑定公网访问。

- 修改

  对已经配置了公网访问的集群，可以通过单击“带宽”参数右侧的“修改”，修改带宽大小，也可以通过单击“访问控制”右侧的“设置”，设置访问控制开关和访问白名单。

- 查看

  在“基本信息”页面，可以查看当前集群绑定的公网 IP 地址。

- 解绑

  对于已经绑定的公网 IP，可以通过单击“公网访问”参数右侧的“解绑”，解绑公网 IP。

通过公网 IP 接入集群

公网访问配置完成后，集群将会获得一个“公网访问”的 IP 地址，用户可以通过公网 IP 地址和端口接入集群。

例如，查看集群中的索引信息，集群中某一个节点的公网访问地址为“10.62.179.32”，端口为“9200”，使用 curl 执行如下命令。

- 如果接入集群未启用安全模式，接入方式为：

  ```
  curl 'http://10.62.179.32:9200/_cat/indices'
  ```

- 如果接入集群已启用安全模式，则需要使用 https 方式访问，并附加用户名和密码，在 curl 命令中添加-u 选项。

  ```
  curl -u username:password -k 'https://10.62.179.32:9200/_cat/indices'
  ```

4.3.3（可选）对接独享型负载均衡器

4.3.3.1 场景描述

云搜索服务集成了共享型负载均衡器的能力，支持绑定公网访问和开启终端节点服务。相比共享型负载均衡器，独享型负载均衡器功能更丰富、性能更强，本章主要介绍如何使用独享型负载均衡器接入集群。

使用独享型负载均衡器接入集群具有如下优点：

- 非安全模式集群也支持集成弹性负载均衡服务的能力。
- 支持用户使用自定义证书进行 HTTPS 双向认证。
- 支持 7 层流量监控及告警配置，方便用户随时查看监控情况。

不同安全模式的集群对接独享型负载均衡器共有 8 种不同的服务形态，不同服务形态对应的 ELB 能力如表 4-19 所示，8 种组合的配置概览如表 4-20 所示。

须知

如果 ELB 绑定公网，则不推荐接入非安全模式的集群。非安全模式的集群使用 HTTP 通信且不需要安全认证即可访问，如果对外公网访问存在安全风险。

表4-19 不同集群对应的 ELB 能力

集群安全模式	ELB 最终对外提供的服务形态	ELB 负载均衡	ELB 流量监控	ELB 双向认证
非安全	无认证	支持	支持	不支持
	单向认证 双向认证	支持	支持	支持
安全+HTTP	密码认证	支持	支持	不支持
	单向认证+密码认证 双向认证+密码认证	支持	支持	支持
安全+HTTPS	单向认证+密码认证 双向认证+密码认证	支持	支持	支持

表4-20 不同集群对接 ELB 的配置概览

集群安全模式	ELB 最终对外提供的服务形态	ELB 监听器			后端服务器组		
		前端协议	前端端口	SSL 解析方式	后端协议	健康检查端口	健康检查路径
非安全	无认证	HTTP	9200	无认证	HTTP	9200	/
	单向认证	HTTPS	9200	单向认证	HTTP	9200	
	双向认证	HTTPS	9200	双向认证	HTTP	9200	
安全+HTTP	密码认证	HTTP	9200	无认证	HTTP	9200	/_opendistro/_security/health
	单向认证+密码认证	HTTPS	9200	单向认证	HTTP	9200	
	双向认证+密码认证	HTTPS	9200	双向认证	HTTP	9200	
安全+HTTPS	单向认证+密码认证	HTTPS	9200	单向认证	HTTPS	9200	
	双向认证+密码认证	HTTPS	9200	双向认证	HTTPS	9200	

4.3.3.2 对接独享型负载均衡器

本文介绍 CSS 集群对接独享型负载均衡器的操作步骤。

（可选）准备自签名证书

如果规划的 ELB 监听器的协议为 HTTP 则跳过此步骤。

准备并上传自签名证书。

说明

建议使用云证书管理服务 CCM 购买的证书，或者其他权威机构颁发的证书。

1. 登录到任意一台安装有 OpenSSL 工具和 JDK 的 Linux 客户端。
2. 执行如下命令制作自签名证书。

```
mkdir ca
mkdir server
mkdir client

#使用 OpenSSL 制作 CA 证书
cd ca
#创建 CA 证书的 openssl 配置文件 ca_cert.conf
cat >ca_cert.conf <<EOF
[ req ]
distinguished_name     = req_distinguished_name
prompt                 = no

[ req_distinguished_name ]
 O                     = ELB
EOF
#创建 CA 证书私钥文件 ca.key
openssl genrsa -out ca.key 2048
#创建 CA 证书的 csr 请求文件 ca.csr
openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf
#创建自签名的 CA 证书 ca.crt
openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key
#把 ca 证书格式转为 p12 格式
openssl pkcs12 -export -clcerts -in ca.crt -inkey ca.key -out ca.p12
#把 ca 证书格式转为 jks 格式
keytool -importkeystore -srckeystore ca.p12 -srcstoretype PKCS12 -deststoretype
JKS -destkeystore ca.jks


#使用 CA 证书签发服务器证书
cd ../server
#创建服务器证书的 openssl 配置文件 server_cert.conf，CN 字段根据需求改为服务器对应的域名、
IP 地址
cat >server_cert.conf <<EOF
[ req ]
distinguished_name     = req_distinguished_name
prompt                 = no

[ req_distinguished_name ]
 O                     = ELB
 CN                    = 127.0.0.1
EOF
#创建服务器证书私钥文件 server.key
openssl genrsa -out server.key 2048
#创建服务器证书的 csr 请求文件 server.csr
```

```
openssl req -out server.csr -key server.key -new -config ./server_cert.conf
#使用 CA 证书签发服务器证书 server.crt
openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days
5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key
#把服务端证书格式转为 p12 格式
openssl pkcs12 -export -clcerts -in server.crt -inkey server.key -out
server.p12
#把服务证书格式转为 jks 格式
keytool -importkeystore -srckeystore server.p12 -srcstoretype PKCS12 -
deststoretype JKS -destkeystore server.jks


#使用 CA 证书签发客户端证书
cd ../client
#创建客户端证书的 openssl 配置文件 client_cert.conf，CN 字段根据需求改为服务器对应的域名、
IP 地址
cat >client_cert.conf <<EOF
[ req ]
distinguished_name    = req_distinguished_name
prompt              = no

[ req distinguished name ]
O                   = ELB
CN                  = 127.0.0.1
EOF
#创建客户端证书私钥文件 client.key
openssl genrsa -out client.key 2048
#创建客户端证书的 csr 请求文件 client.csr
openssl req -out client.csr -key client.key -new -config ./client_cert.conf
#使用 CA 证书签发客户端证书 client.crt
openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days
5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key
#把客户端证书格式转为浏览器可识别的 p12 格式
openssl pkcs12 -export -clcerts -in client.crt -inkey client.key -out
client.p12
#把客户端证书格式转为 jks 格式
keytool -importkeystore -srckeystore client.p12 -srcstoretype PKCS12 -
deststoretype JKS -destkeystore client.jks
```

3. 上传自签名证书，具体操作请参见。

创建独享型负载均衡器

1. 登录弹性负载均衡管理控制台。
2. 参考，创建独享型负载均衡器。CSS 集群对接独享型负载均衡器所需要关注的参数如表 4-21 所示，其他参数请根据实际需要填写。

表4-21 独享型负载均衡器的配置说明

参数	配置说明	取值样例
实例类型	选择“独享型”。	独享型
计费模式	性能独享型负载均衡器的收费类型。	按需计费

参数	配置说明	取值样例
区域	选择 CSS 集群所在的区域。	-
跨 VPC 后端	开启跨 VPC 后端才能连接 CSS 集群。	开启
网络类型	负载均衡器对外提供服务所使用的网络类型。	IPv4 私网
所属 VPC	所属虚拟私有云。无论选择哪种网络类型，均需配置此项。 需要选择和 CSS 集群同一 VPC。	-
子网	选择创建负载均衡实例的子网。无论选择哪种网络类型，均需配置此项。 需要选择和 CSS 集群同一子网。	-
规格	建议选择功能和性能更优的应用型规格。	应用型(HTTP/HTTPS) 小型 I

通过 Curl 命令接入集群

通过执行如下命令，测试独享型负载均衡器是否能够正常接入集群。

表4-22 不同集群的接入命令

集群安全模式	ELB 最终对外提供的服务形态	接入集群的 Curl 命令
非安全	无认证	`curl http://IP:9200`
	单向认证	`curl -k --cert ./client.crt --key ./client.key https://IP:9200`
	双向认证	`curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200`
安全+HTTP	密码认证	`curl http://IP:9200 -u user:pwd`
	单向认证+密码认证	`curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd`
	双向认证+密	`curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd`

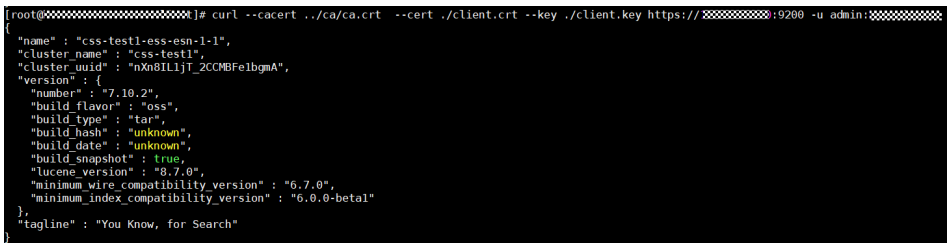
集群安全模式	ELB 最终对外提供的服务形态	接入集群的 Curl 命令
	码认证	
安全+HTTPS	单向认证+密码认证	`curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd`
	双向认证+密码认证	`curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd`

表4-23 变量说明

变量名	说明
IP	弹性负载均衡的 IP 地址。
user	访问 CSS 集群的用户名。
pwd	用户名对应的密码。

当正常返回集群信息时，表示连接成功，例如安全模式+HTTPS 协议的集群对接 ELB 双向认证模式的返回信息如图 4-14 所示。

图4-14 接入集群



4.3.3.3 接入集群进行双向认证的代码示例

本文介绍通过 Java 客户端连接集群进行双向认证的代码示例。

ESSecuredClientWithCerDemo 代码

```
import org.apache.commons.io.IOUtils;
import org.apache.http.auth.AuthScope;
import org.apache.http.auth.UsernamePasswordCredentials;
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.client.BasicCredentialsProvider;
import org.apache.http.HttpHost;
import org.apache.http.nio.conn.ssl.SSLIOSessionStrategy;
import org.elasticsearch.action.search.SearchRequest;
```

```
import org.elasticsearch.action.search.SearchResponse;
import org.elasticsearch.client.RequestOptions;
import org.elasticsearch.client.RestClient;
import org.elasticsearch.client.RestClientBuilder;
import org.elasticsearch.client.RestHighLevelClient;
import org.elasticsearch.index.query.QueryBuilders;
import org.elasticsearch.search.SearchHit;
import org.elasticsearch.search.SearchHits;
import org.elasticsearch.search.builder.SearchSourceBuilder;
import java.io.FileInputStream;
import java.io.IOException;
import java.security.KeyStore;
import java.security.SecureRandom;
import javax.net.ssl.HostnameVerifier;
import javax.net.ssl.KeyManagerFactory;
import javax.net.ssl.SSLContext;
import javax.net.ssl.SSLSession;
import javax.net.ssl.TrustManagerFactory;
public class ESSecuredClientWithCerDemo {
   private static final String KEY_STORE_PWD = "";
   private static final String TRUST_KEY_STORE_PWD = "";
   private static final String CA_JKS_PATH = "ca.jks";
   private static final String CLIENT_JKS_PATH = "client.jks";
   private static final String ELB_ADDRESS = "127.0.0.1";
   private static final int ELB_PORT = 9200;
   private static final String CSS_USERNAME = "user";
   private static final String CSS_PWD = "";
   public static void main(String[] args) {
      // 建立客户端
      RestHighLevelClient client = initESClient(ELB_ADDRESS, CSS_USERNAME,
CSS_PWD);
      try {
         // 查询 match_all，相当于 {\"query\": {\"match_all\": {}}}
         SearchRequest searchRequest = new SearchRequest();
         SearchSourceBuilder searchSourceBuilder = new SearchSourceBuilder();
         searchSourceBuilder.query(QueryBuilders.matchAllQuery());
         searchRequest.source(searchSourceBuilder);
         // query
         SearchResponse searchResponse = client.search(searchRequest,
RequestOptions.DEFAULT);
         System.out.println("query result: " + searchResponse.toString());
         SearchHits hits = searchResponse.getHits();
         for (SearchHit hit : hits) {
            System.out.println(hit.getSourceAsString());
         }
         System.out.println("query success");
         Thread.sleep(2000L);
      } catch (InterruptedException | IOException e) {
         e.printStackTrace();
      } finally {
         IOUtils.closeQuietly(client);
      }
   }
   private static RestHighLevelClient initESClient(String clusterAddress, String
userName, String password) {
```

```
      final CredentialsProvider credentialsProvider = new
BasicCredentialsProvider();
      credentialsProvider.setCredentials(AuthScope.ANY, new
UsernamePasswordCredentials(userName, password));
      SSLContext ctx = null;
      try {
         KeyStore ks = getKeyStore(CLIENT_JKS_PATH, KEY_STORE_PWD, "JKS");
         KeyManagerFactory kmf = KeyManagerFactory.getInstance("SunX509");
         kmf.init(ks, KEY_STORE_PWD.toCharArray());
         KeyStore tks = getKeyStore(CA_JKS_PATH, TRUST_KEY_STORE_PWD, "JKS");
         TrustManagerFactory tmf = TrustManagerFactory.getInstance("SunX509");
         tmf.init(tks);
         ctx = SSLContext.getInstance("SSL", "SunJSSE");
         ctx.init(kmf.getKeyManagers(), tmf.getTrustManagers(), new
SecureRandom());
      } catch (Exception e) {
         e.printStackTrace();
      }
      SSLIOSessionStrategy sessionStrategy = new SSLIOSessionStrategy(ctx, new
HostnameVerifier() {
         @Override
         public boolean verify(String arg0, SSLSession arg1) {
            return true;
         }
      });
      SecuredHttpClientConfigCallback httpClientConfigCallback = new
SecuredHttpClientConfigCallback(sessionStrategy,
         credentialsProvider);
      RestClientBuilder builder = RestClient.builder(new HttpHost(clusterAddress,
ELB_PORT, "https"))
         .setHttpClientConfigCallback(httpClientConfigCallback);
      RestHighLevelClient client = new RestHighLevelClient(builder);
      return client;
   }
   private static KeyStore getKeyStore(String path, String pwd, String type) {
      KeyStore keyStore = null;
      FileInputStream is = null;
      try {
         is = new FileInputStream(path);
         keyStore = KeyStore.getInstance(type);
         keyStore.load(is, pwd.toCharArray());
      } catch (Exception e) {
         e.printStackTrace();
      } finally {
         IOUtils.closeQuietly(is);
      }
      return keyStore;
   }
}
```

SecuredHttpClientConfigCallback 代码

```
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.nio.client.HttpAsyncClientBuilder;
import org.apache.http.nio.conn.ssl.SSLIOSessionStrategy;
import org.elasticsearch.client.RestClientBuilder;
```

```
import org.elasticsearch.common.Nullable;
import java.util.Objects;
class SecuredHttpClientConfigCallback implements
RestClientBuilder.HttpClientConfigCallback {
   @Nullable
   private final CredentialsProvider credentialsProvider;
   /**
    * The {@link SSLIOSessionStrategy} for all requests to enable SSL / TLS
encryption.
    */
   private final SSLIOSessionStrategy sslStrategy;
   /**
    * Create a new {@link SecuredHttpClientConfigCallback}.
    *
    * @param credentialsProvider The credential provider, if a username/password
have been supplied
    * @param sslStrategy       The SSL strategy, if SSL / TLS have been supplied
    * @throws NullPointerException if {@code sslStrategy} is {@code null}
    */
   SecuredHttpClientConfigCallback(final SSLIOSessionStrategy sslStrategy,
      @Nullable final CredentialsProvider credentialsProvider) {
      this.sslStrategy = Objects.requireNonNull(sslStrategy);
      this.credentialsProvider = credentialsProvider;
   }
   /**
    * Get the {@link CredentialsProvider} that will be added to the HTTP client.
    *
    * @return Can be {@code null}.
    */
   @Nullable
   CredentialsProvider getCredentialsProvider() {
      return credentialsProvider;
   }
   /**
    * Get the {@link SSLIOSessionStrategy} that will be added to the HTTP client.
    *
    * @return Never {@code null}.
    */
   SSLIOSessionStrategy getSSLStrategy() {
      return sslStrategy;
   }
   /**
    * Sets the {@linkplain
HttpAsyncClientBuilder#setDefaultCredentialsProvider(CredentialsProvider)
credential provider},
    *
    * @param httpClientBuilder The client to configure.
    * @return Always {@code httpClientBuilder}.
    */
   @Override
   public HttpAsyncClientBuilder customizeHttpClient(final HttpAsyncClientBuilder
httpClientBuilder) {
      // enable SSL / TLS
      httpClientBuilder.setSSLStrategy(sslStrategy);
      // enable user authentication
```

```
        if (credentialsProvider != null) {
            httpClientBuilder.setDefaultCredentialsProvider(credentialsProvider);
        }
        return httpClientBuilder;
    }
}
```

pom.xml 代码

```
<?xml version="1.0" encoding="UTF-8"?>
<project xmlns="http://maven.apache.org/POM/4.0.0"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="http://maven.apache.org/POM/4.0.0
http://maven.apache.org/xsd/maven-4.0.0.xsd">
    <modelVersion>4.0.0</modelVersion>
    <groupId>1</groupId>
    <artifactId>ESClient</artifactId>
    <version>1.0-SNAPSHOT</version>
    <name>ESClient</name>

    <properties>
        <maven.compiler.source>8</maven.compiler.source>
        <maven.compiler.target>8</maven.compiler.target>
        <project.build.sourceEncoding>UTF-8</project.build.sourceEncoding>
        <elasticsearch.version>7.10.2</elasticsearch.version>
    </properties>
    <dependencies>
        <dependency>
            <groupId>org.elasticsearch.client</groupId>
            <artifactId>transport</artifactId>
            <version>${elasticsearch.version}</version>
        </dependency>
        <dependency>
            <groupId>org.elasticsearch</groupId>
            <artifactId>elasticsearch</artifactId>
            <version>${elasticsearch.version}</version>
        </dependency>
        <dependency>
            <groupId>org.elasticsearch.client</groupId>
            <artifactId>elasticsearch-rest-high-level-client</artifactId>
            <version>${elasticsearch.version}</version>
        </dependency>
        <dependency>
            <groupId>commons-io</groupId>
            <artifactId>commons-io</artifactId>
            <version>2.11.0</version>
        </dependency>
    </dependencies>
</project>
```

4.4 备份与恢复索引

4.4.1 备份与恢复概述

为避免数据丢失，您可以将集群的索引数据进行备份，当数据发生丢失或者想找回某一时间段数据时，您可以通过恢复索引操作快速获得数据。索引的备份是通过创建集群快照实现。第一次备份时，建议将所有索引数据进行备份。

说明

备份与恢复索引功能上线之前（即 2018 年 3 月 10 日之前）创建的 Elasticsearch 集群，无法创建快照。

- 4.4.2 管理自动创建快照：自动创建快照指按照设置的规则，每天在指定时间自动创建快照。您可以开启自动创建功能、设置自动创建的策略、和关闭自动创建功能。
- 4.4.3 手动创建快照：在任意时间，您通过手动创建快照的方式，针对当时的数据或某几个索引创建快照进行备份。
- 4.4.4 恢复数据：将已有的快照，通过恢复快照功能，将备份的索引数据恢复到指定的集群中。
- 4.4.5 删除快照：对于已失效的快照，建议删除以释放存储资源。

4.4.2 管理自动创建快照

自动创建快照指按照设置的规则，在指定时间自动创建快照。您可以开启自动创建功能、设置自动创建的策略、和关闭自动创建功能。

说明

备份与恢复索引功能上线之前（即 2018 年 3 月 10 日之前）创建的集群，无法创建快照。

前提条件

登录云搜索服务管理控制台的账号或 IAM 用户必须同时具备如下权限才能使用创建或恢复快照功能。

- “全局服务”中“对象存储服务”项目的“OBS Administrator”权限。
- 当前所属区域的“Elasticsearch Administrator”权限。

注意事项

- 第一次备份时，建议将所有索引数据进行备份。
- 集群快照会导致 CPU、磁盘 IO 上升等影响，建议在业务低峰期进行操作。
- 创建快照之前，您需要进行基础配置，包含存储快照的 OBS 桶、快照的备份路径及安全认证使用的 IAM 委托。
- 集群快照存储的 OBS 桶，在首次设置后，不管自动创建快照还是手动创建快照，如果快照列表中已有可用的快照，则 OBS 桶将无法再变更，请谨慎选择存储 OBS 桶。

- 如果 OBS 桶已经存储了快照，OBS 无法变更，您可以使用这个方法修改：首先关闭快照功能，然后再开启快照功能，指定新的 OBS 桶。一旦关闭快照功能，之前创建的快照将无法用于恢复集群。
- 当集群处于“不可用”状态时，快照功能中，除了恢复快照功能外，其他快照信息或功能只能查看，无法进行编辑。
- 备份与恢复过程中，支持集群扩容、访问 Kibana、查看监控、删除其他快照的操作。不支持重启此集群、删除此集群、删除正在创建或恢复的快照、再次创建或恢复快照的操作。补充说明，当此集群正在进行创建快照或者恢复快照时，此时，自动创建快照任务将被取消。
- CSS 集群第一次快照是全量，后面再备份快照是在之前的快照基础上增量，CSS 是增量快照逻辑，快照之间的文件会相互依赖。

管理自动创建快照

1. 在云搜索服务管理控制台，单击左侧导航栏的“集群管理”。
2. 在“集群管理”页面，单击需要进行备份的集群名称，进入集群基本信息页面。在左侧导航栏选择“集群快照”，进入“集群快照”管理页面。
3. 在“集群快照”管理页面，在“集群快照开关”右侧单击开关，打开集群快照功能。
4. 打开集群快照功能后，云搜索服务会自动为客户创建 OBS 桶和 IAM 委托，用于存储快照。自动创建的 OBS 桶和 IAM 委托将直接展示在界面中。如果您不希望使用自动创建的 OBS 桶和 IAM 委托，您可以在“基础配置”右侧单击进行配置。

表4-24 集群快照基础配置的参数说明

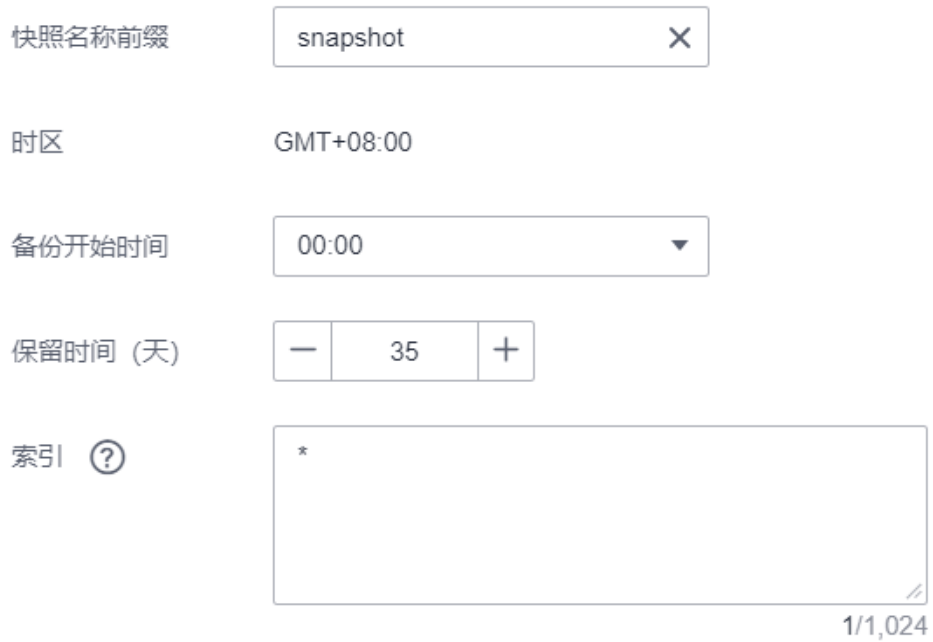
参数	说明
OBS 桶	在下拉框中选择存储快照的 OBS 桶。也可以单击右侧的“创建桶”新建 OBS。 创建或者已存在的 OBS 桶需满足如下条件： • “存储类别”为“标准存储”。 • “区域”须与创建的集群所在区域相同。
备份路径	快照在 OBS 桶中的存放路径。 备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
IAM 委托	指当前账号授权云搜索服务访问或维护存储在 OBS 中数据。也可以单击右侧的“创建委托”新建 IAM 委托。 创建或者已存在的 IAM 委托需满足如下条件： • “委托类型”选择“云服务”。

参数	说明
	• “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 设置当前委托具备“全局服务”中“对象存储服务”项目的“OBS Administrator”权限。

5. 在“自动创建快照”右侧，单击开关开启自动创建快照功能，弹出“创建快照策略”页面。如果已启用自动创建快照功能，也可以在开关右侧单击进行快照策略修改。
   - “快照名称前缀”：快照名称前缀的长度为 1～32 个字符，只能包含小写字母、数字、中划线和下划线，且必须以小写字母开头。快照名称由快照名称前缀加上时间组成，例如自动生成的快照名称 snapshot-2018022405925。
   - “时区”：指备份时间对应的时区。请基于此时区选择“备份开始时间”。
   - “备份开始时间”：指每小时、每天或者每周指定一天自动开始备份的时间，当选择每天或每周指定一天时，可以指定整点时间，如 00:00、01:00，取值范围为 00:00～23:00。请在下拉框中选择备份时间。
   - “索引”：填写索引名称，支持选择索引进行备份。索引名称不能包含空格和大写字母，且不能包含"\<|>/?特殊字符，多个索引之间使用英文逗号隔开。如果不填写，则默认备份集群中所有索引。支持使用“*”匹配多个索引，例如：index*，表示备份名称前缀是 index 的所有索引的数据。

     在 Kibana 中使用 **GET /_cat/indices** 命令，可以查询集群中所有索引的名称。

   图4-15 自动创建快照

6. 设置完成后，单击“确定”保存快照策略。

按照策略自动创建的快照将呈现在快照管理列表中。快照列表同时展示自动创建和手动创建的快照，您可以通过快照类型参数进行区分。在快照列表右上角，您可以输入快照名称或快照 ID 的关键字进行查找。

7. （可选）关闭自动创建快照功能。

   关闭自动创建快照功能后，系统将停止继续自动创建快照。如果系统正在根据策略自动创建快照，而快照列表还未呈现正在创建的快照时，无法关闭自动创建快照功能。如果您单击了关闭按钮，系统将提示您无法关闭。建议等快照自动创建成功后，即快照列表已出现最新创建的快照时，再单击关闭按钮，关闭自动创建快照功能。

   关闭自动创建快照功能时，您可以在弹出窗口中通过“删除自动创建的快照”选项，选择是否立即删除之前已自动创建的快照，默认不勾选。

   - 不勾选：表示不会删除关闭此功能前已自动创建的快照。如果不删除，后续还可以在快照列表中通过删除按钮手动删除，详细操作指导请参见 4.4.5 删除快照。如果未手动删除，且之后用户又重新开启了自动创建快照功能，那么此集群中所有“快照类型”为自动创建的快照（包含开启自动创建快照功能前已存在的自动创建的快照）都无法手动删除，只会被系统自动删除。系统会基于重新开启自动创建快照功能时的配置策略进行自动删除，例如此策略中定义的保留快照个数为 10 个，那么当快照个数超过 10 个，系统在半点时刻会自动删除超过保留个数的快照。
   - 勾选：表示删除此集群快照列表中所有“快照类型”为自动创建的快照。

说明

快照功能关闭时，不会主动清理快照数据。如果需要清理快照数据，可前往 OBS 控制台存储快照的 OBS 桶对快照数据进行主动管理。

4.4.3 手动创建快照

在任意时间，您通过手动创建快照的方式，针对当时的数据或某几个索引创建快照进行备份。

说明

备份与恢复索引功能上线之前（即 2018 年 3 月 10 日之前）创建的集群，无法创建快照。

前提条件

登录云搜索服务管理控制台的账号或 IAM 用户必须同时具备如下权限才能使用创建或恢复快照功能。

- “全局服务”中“对象存储服务”项目的“OBS Administrator”权限。
- 当前所属区域的“Elasticsearch Administrator”权限。

注意事项

- 第一次备份时，建议将所有索引数据进行备份。
- 集群快照会导致 CPU、磁盘 IO 上升等影响，建议在业务低峰期进行操作。
- 创建快照之前，您需要进行基础配置，包含存储快照的 OBS 桶、快照的备份路径及安全认证使用的 IAM 委托。

- 集群快照存储的 OBS 桶，在首次设置后，不管自动创建快照还是手动创建快照，如果快照列表中已有可用的快照，则 OBS 桶将无法再变更，请谨慎选择存储 OBS 桶。
- 如果 OBS 桶已经存储了快照，OBS 无法变更，您可以使用这个方法修改：首先关闭快照功能，然后再开启快照功能，指定新的 OBS 桶。一旦关闭快照功能，之前创建的快照将无法用于恢复集群。
- 当集群处于“不可用”状态时，快照功能中，除了恢复快照功能外，其他快照信息或功能只能查看，无法进行编辑。
- 备份与恢复过程中，支持集群扩容、访问 Kibana、查看监控、删除其他快照的操作。不支持重启此集群、删除此集群、删除正在创建或恢复的快照、再次创建或恢复快照的操作。补充说明，当此集群正在进行创建快照或者恢复快照时，此时，自动创建快照任务将被取消。
- CSS 集群第一次快照是全量，后面再备份快照是在之前的快照基础上增量，CSS 是增量快照逻辑，快照之间的文件会相互依赖。

手动创建快照

1. 在云搜索服务管理控制台，单击左侧导航栏的“集群管理”。
2. 在“集群管理”页面，单击需要进行备份的集群名称，进入集群基本信息页面。在左侧导航栏选择“集群快照”，进入“集群快照”管理页面。
3. 在“集群快照”管理页面，单击“集群快照开关”右侧的开关，打开集群快照功能。
4. 打开集群快照功能后，云搜索服务会自动为客户创建 OBS 桶和 IAM 委托，用于存储快照。自动创建的 OBS 桶和 IAM 委托将直接展示在界面中。如果您不希望使用自动创建的 OBS 桶和 IAM 委托，您可以在“基础配置”右侧单击 进行配置。

表4-25 集群快照基础配置的参数说明

参数	说明
OBS 桶	在下拉框中选择存储快照的 OBS 桶。也可以单击右侧的“创建桶”新建 OBS。 创建或者已存在的 OBS 桶需满足如下条件： • “存储类别”为“标准存储”。 • “区域”须与创建的集群所在区域相同。
备份路径	快照在 OBS 桶中的存放路径。 备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
IAM 委托	指当前账号授权云搜索服务访问或维护存储在 OBS 中数据。也可以

参数	说明
	单击右侧的“创建委托”新建 IAM 委托。 创建或者已存在的 IAM 委托需满足如下条件： • “委托类型”选择“云服务”。 • “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 设置当前委托具备“全局服务”中“对象存储服务”项目的“OBS Administrator”权限。

5. 完成基础配置后，单击“创建快照”可手动创建快照。
   - “快照名称”：手动创建的快照名称，4～64 个字符，只能包含小写字母、数字、中划线和下划线，且必须以字母开头。与自动创建不同，手动创建的快照名称按照用户设置的名称，不会自动加上时间信息。
   - “索引”：填写索引名称，支持选择索引进行备份。索引名称不能包含空格和大写字母，且不能包含"\<|>/?特殊字符，多个索引之间使用英文逗号隔开。如果不填写，则默认备份集群中所有索引。支持使用“*”匹配多个索引，例如：index*，表示备份名称前缀是 index 的所有索引的数据。

     在 Kibana 中使用 **GET /_cat/indices** 命令，可以查询集群中所有索引的名称。
   - “快照描述”：创建的快照描述信息。0～256 个字符，不能包含“<>”字符。

   图4-16 手动创建快照

6. 单击“确定”开始创建快照。

   快照创建完成后，将直接呈现在快照管理列表中，快照状态为“可用”表示快照创建成功。快照列表同时展示自动创建和手动创建的快照，您可以通过快照类型

参数进行区分。在快照列表右上角，您可以输入快照名称或快照 ID 的关键字进行查找。

4.4.4 恢复数据

将已有的快照，通过恢复快照功能，将备份的索引数据恢复到指定的集群中。

说明

备份与恢复索引功能上线之前（即 2018 年 3 月 10 日之前）创建的集群，无法创建快照。

前提条件

登录云搜索服务管理控制台的账号或 IAM 用户必须同时具备如下权限才能使用创建或恢复快照功能。

- “全局服务”中“对象存储服务”项目的“OBS Administrator”权限。
- 当前所属区域的“Elasticsearch Administrator”权限。

注意事项

- 集群快照会导致 CPU、磁盘 IO 上升等影响，建议在业务低峰期进行操作。
- 如果 OBS 桶已经存储了快照，OBS 无法变更，您可以使用这个方法修改：首先关闭快照功能，然后再开启快照功能，指定新的 OBS 桶。一旦关闭快照功能，之前创建的快照将无法用于恢复集群。
- 当集群处于“不可用”状态时，快照功能中，除了恢复快照功能外，其他快照信息或功能只能查看，无法进行编辑。
- 备份与恢复过程中，支持集群扩容、访问 Kibana、查看监控、删除其他快照的操作。不支持重启此集群、删除此集群、删除正在创建或恢复的快照、再次创建或恢复快照的操作。补充说明，当此集群正在进行创建快照或者恢复快照时，此时，自动创建快照任务将被取消。
- 快照在恢复过程中不可查询集群数据。
- CSS 集群快照恢复到另一个集群会覆盖目标集群中的同名索引，不同名的索引不会覆盖。如果两个集群的 shard 不一样，则同名的索引不会被覆盖。
- 用于恢复的目的集群版本不得低于源端集群，一般保持版本一致即可。

恢复数据

快照管理列表中“快照状态”为“可用”的快照，可以恢复集群中的数据。已存储的快照数据可恢复至其他集群。

恢复数据将覆盖集群中当前的数据，请谨慎操作。

1. 在快照管理列表中，选择需要恢复的快照，单击“操作”列的“恢复”。
2. 在“恢复”页面配置集群的恢复参数。

   “索引”：指定需要进行恢复的索引名称，默认为空。如保持默认值，即不指定索引名称，则表示恢复所有的索引数据。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?特殊字符。支持使用“*”匹配多个索引，比如 index*，表示恢复快照中名称前缀是 index 的所有索引。

“索引名称匹配模式”：在恢复时，可以根据文本框中定义的过滤条件去恢复符合条件的索引，过滤条件请使用正则表达式。默认值“index_(.+)”表示所有的索引。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?,特殊字符。

“索引名称替换模式”：索引重命名的规则。默认值“restored_index_$1”表示在所有恢复的索引名称前面加上“restored_”。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?,特殊字符。

说明

“索引名称匹配模式”和“索引名称替换模式”需要同时设置才会生效。

“集群”：选择需要进行恢复的集群名称，可选择当前集群或者其他集群。只能选择处于“可用”状态的集群，如果快照所属的集群处于“不可用”状态，那么也无法将快照恢复到本集群。恢复到其他集群时，目标集群中的版本不低于本集群。如果已选择其他集群，且该集群中存在同名的索引，则恢复完成后，该同名的索引中的数据将会被覆盖，请谨慎操作。

“覆盖目标集群同名同 shard 结构索引”：默认不覆盖，通过快照恢复数据，是以快照文件覆盖的形式进行数据恢复。覆盖目标集群同名的索引后，可能导致目标集群同名索引数据丢失，请谨慎勾选操作。

图4-17 恢复快照

3. 单击“确定”开始恢复。恢复成功，快照列表中“任务状态”将变更为“恢复成功”，索引数据将根据快照信息重新生成。

图4-18 恢复成功

4.4.5 删除快照

当快照信息不需要使用时，您可以删除快照释放存储资源。当自动创建快照功能开启时，自动创建的快照无法手动删除，系统会按照设置的策略在半点时刻自动删除超过“保留时间”的快照。当自动创建快照功能关闭，且之前已自动创建的快照并未同步删除时，快照列表中自动创建的快照，可通过删除按钮手动删除。如果未手动删除，且之后用户又重新开启了自动创建快照功能，那么此集群中所有“快照类型”为自动创建的快照（包含开启自动创建快照功能前已存在的自动创建的快照）都无法手动删除，只会被系统自动删除。

说明

快照信息删除后，数据将无法恢复，请谨慎操作。

1. 在快照管理页面中，选择需要删除的快照。
2. 单击“操作”列的“删除”，在弹窗中确认要删除的快照信息后，单击“确定”删除快照。

4.5 形态变更

4.5.1 形态变更概述

CSS 集群支持形态变更，包括集群扩容、集群规格变更、集群缩容等。当创建的集群规格不能满足业务需求时，可以通过形态变更，提高集群的使用效率，降低运维成本。

4.5.2 扩容

- 当**集群数据节点**（ess）的写入与查询压力大、响应时间过长时，可以通过扩容数据节点的“节点存储容量”保证数据的持久性。如果因数据量过大或操作不当导致数据节点状态异常时，可以扩容“节点数量”保证集群的可用性。
- **冷数据节点**（ess-cold）主要用于分担 ess 数据节点的压力，当发现冷数据有丢失的风险时，可以扩容冷数据节点的“节点存储容量”保证冷数据的持久性，同时也支持扩容节点个数保证集群的可用性。

4.5.3 变更规格

- 当新增索引或分片分配的处理时间过长，或管理集群各个节点的协调、调度不足时，可以变更 **Master 节点**（ess-master）的“节点规格”保证集群的正常使用。
- 当数据节点任务分发量、结果汇聚量过大时，需要变更 **Client 节点**（ess-client）的“节点规格”。
- 当数据的写入与查询突然变得缓慢时，可以变更**数据节点**（ess）的“节点规格”提高数据节点的查询与写入效率。
- 当存在冷数据查询缓慢时，可以变更**冷数据节点**（ess-cold）的“节点规格”提高对数据查询的效率。

4.5.4 缩容集群节点数量

- 当集群有充足的能力处理当前数据时，为节省资源可以随机减小集群占用的资源。

4.5.5 缩容指定节点

- 当集群有充足的能力处理当前数据时，为节省资源可以指定一个或多个节点进行缩容。

4.5.6 替换指定节点

- 当集群中的节点发生故障，可以通过删除故障节点，重建一个一样的新节点，实现指定节点替换。

4.5.7 添加 Master/Client 节点

- 当集群数据面业务增长，需要动态调整集群形态时，可以增加 Master/Client 节点。

4.5.8 更改安全模式

集群创建完成后，可以通过更改安全模式进行集群模式变更。CSS 服务支持如下几种安全模式更改：

- 非安全模式切换为安全模式：“非安全模式”切换为“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”
- 安全模式切换为非安全模式：“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”切换为“非安全模式”
- 切换安全模式下的协议：“安全模式+HTTP 协议”切换为“安全模式+HTTPS 协议”、“安全模式+HTTPS 协议”切换为“安全模式+HTTP 协议”

4.5.9 切换可用区

切换可用区包含两大场景：可用区高可用改造和可用区平移切换。

- 可用区高可用改造：适用于单 AZ 改造成两 AZ、单 AZ 改造成三 AZ 或两 AZ 改造成三 AZ 的场景，目的是为了提升集群的高可用性。
- 可用区平移切换：适用于从一个 AZ 完全迁移到另一个 AZ 的场景，目的是为了解决当前可用区资源不足的问题。

4.5.2 扩容

当集群数据面业务变化，需要动态调整集群节点的数量和容量时，可以执行“扩容”任务。扩容集群时，业务不会中断。

前提条件

- 集群处于“可用”状态，且无正在进行的任务。
- 有足够的配额支持集群扩容。

约束限制

- 扩容操作不支持修改“节点规格”。修改“节点规格”请执行 4.5.3 变更规格操作。

- 扩容什么节点类型的“节点数量”和“节点存储容量”，扩容完成后只生效该节点类型的“节点数量”和“节点存储容量”，其他节点类型的“节点数量”和“节点存储容量”保持不变。
- 当集群包含的节点类型不同时，扩容的节点数量区间会有区别，具体情况请参考表 4-26。

表4-26 不同节点类型的节点数量说明

集群包含的节点类型	节点数量的取值范围
ess	ess：1~32
ess、ess-master	ess：1~200 ess-master：3~9 的奇数
ess、ess-client	ess：1~32 ess-client：1~32
ess、ess-cold	ess：1~32 ess-cold：1~32
ess、ess-master、ess-client	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32
ess、ess-master、ess-cold	ess：1~200 ess-master：3~9 的奇数 ess-cold：1~32
ess、ess-client、ess-cold	ess：1~32 ess-client：1~32 ess-cold：1~32
ess、ess-master、ess-client、ess-cold	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32 ess-cold：1~32
四种节点类型的说明： • ess：默认节点类型，即创建集群时必选的数据节点类型，其他 3 种节点类型都是基于业务需要可选的类型。 • ess-master：Master 节点 • ess-client：Client 节点 • ess-cold：冷数据节点	

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置扩容参数。
   - “变更类型”：“扩容”。
   - “变更的资源”：资源的变化量。
   - “变更的角色”：此处修改的是默认数据节点类型的节点数量和节点存储容量。
     - “节点数量”支持修改的取值范围请参考表 4-26。
     - “节点存储”的容量取值范围由“节点规格”决定。只支持配置为 20 的倍数。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列显示为“扩容”，表示集群正在扩容。当集群状态变为“可用”，则表示扩容成功。

4.5.3 变更规格

当集群数据面业务变化，需要动态调整集群节点的规格时，可以执行“变更规格”任务。

前提条件

- 集群处于“可用”状态，且无正在进行的任务。
- 有足够的配额支持集群变更规格。
- 变更规格时，为了不中断业务，请确认业务数据都有副本。

  在 Kibana 中执行命令 **GET _cat/indices?v**，如果回显的“rep”值大于“0”，则表示有副本；如果“rep”值等于“0”，则表示没有副本，请先为集群 4.4.3 手动创建快照再变更规格。
- 如果数据量比较大的情况下，更改节点规格耗时会比较长，因此，建议在业务低峰期更改节点规格，利于更快完成规格更改。

约束限制

- 变更规格操作不支持修改“节点数量”和“节点存储容量”。增加“节点数量”和“节点存储容量”请执行 4.5.2 扩容操作。减少“节点数量”请执行 4.5.4 缩容集群节点数量操作。
- 如果将大规格更改为小规格，集群的处理性能将会降低，将会影响业务能力，请谨慎操作。
- 当集群包含多种节点类型时，一次只支持变更一种类型的节点规格，且变更完成后只生效所选类型的节点规格。
- 变更规格过程中，Kibana 不可用。

- 变更规格过程中，会依次对节点进行关机，完成更改后依次开机。是一个滚动的变更过程。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置变更规格的参数。
   - “变更类型”：“变更规格”。
   - “变更的资源”：显示资源的变化量。
   - “变更的角色”：此处修改的是默认数据节点类型的节点规格，在对应节点规格下拉框中选择所需的规格，然后勾选需要变更的节点。
   - 如果集群启用了 Master 节点、Client 节点或冷数据节点，还可以更改 Master 节点、Client 节点与冷数据节点的“节点规格”。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 在弹出的窗口确认是否勾选“进行索引副本校验”和“检测集群状态”，单击“确认”启动集群规格变更。

   索引副本校验：

   索引副本校验用于保证索引正常提供服务。如果跳过校验，变更规格操作将不会要求每一个索引都有副本，变更过程中可能会影响业务。
   - 没有 Master 节点的集群更改节点规格时，如果选择进行索引副本校验，则要求所有索引至少有 1 个副本，且“节点数量”总和不小于 3。
   - 有 Master 节点的集群更改节点规格时，如果选择进行索引副本校验，则要求所有索引至少有 1 个副本。

   检测集群状态：

   默认检测集群状态，变更规格为滚动变更，变更过程中，为了保证变更成功率以及数据安全，会确保每个节点进程拉起来后继续后续节点操作。当集群负载过高业务故障，无法正常下发变更请求，依赖更多的资源才能恢复的紧急情况下可忽略检测集群状态，忽略后，变更过程中因为忽略集群状态检测可能会导致集群故障并中断业务，请谨慎跳过。
8. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“规格修改”，表示集群正在更改规格。当集群状态变为“可用”，则表示规格变更成功。

4.5.4 缩容集群节点数量

当集群有充足的能力处理当前数据时，为节省资源可以执行“缩容”任务，减小集群占用的资源。建议在业务低峰期进行缩容操作。

前提条件

集群处于“可用”状态，且无正在进行的任务。

约束限制

- 缩容操作仅支持修改“节点数量”，不支持修改“节点规格”和“节点存储容量”。修改“节点规格”请执行 4.5.3 变更规格操作。修改“节点存储容量”请执行 4.5.2 扩容操作。
- 缩容什么节点类型的“节点数量”，缩容完成后只生效新该节点类型的“节点数量”，其他节点类型的“节点数量”保持不变。
- 要确保缩容之后的磁盘使用量小于 80%，且集群每个节点类型中每个 AZ 的节点数至少为 1。
- 缩容过程会涉及数据迁移，将要下线的节点数据迁移到其他节点上，数据迁移的超时阈值为 5 小时。当超过 5 小时数据还未迁移完成，那么缩容会失败。建议在集群数据量较大的情况下，分多次进行缩容。
- 如果集群没有启用 Master 节点，缩容后剩余的数据节点个数（包含冷数据节点和其他类型节点）须大于之前的一半，并大于索引的最大副本个数。
- 关于有 Master 节点的集群，每次缩容的 Master 节点个数要小于当前 Master 节点总数的一半，缩容后的 Master 节点个数必须是奇数且不小于 3。
- 两个节点的集群暂不支持缩容，可使用单节点重新创建集群。
- 当集群包含的节点类型不同时，缩容的节点数量区间会有区别，具体情况请参考表 4-27。

表4-27 不同节点类型的节点数量说明

集群包含的节点类型	节点数量的取值范围
ess	ess：1~32
ess、ess-master	ess：1~200 ess-master：3~9 的奇数
ess、ess-client	ess：1~32 ess-client：1~32
ess、ess-cold	ess：1~32 ess-cold：1~32
ess、ess-master、ess-client	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32
ess、ess-master、ess-cold	ess：1~200 ess-master：3~9 的奇数 ess-cold：1~32
ess、ess-client、ess-cold	ess：1~32 ess-client：1~32 ess-cold：1~32

集群包含的节点类型	节点数量的取值范围
ess、ess-master、ess-client、ess-cold	ess：1~200 ess-master：3~9 的奇数 ess-client：1~32 ess-cold：1~32
四种节点类型的说明： • ess：默认节点类型，即创建集群时必选的数据节点类型，其他 3 种节点类型都是基于业务需要可选的类型。 • ess-master：Master 节点 • ess-client：Client 节点 • ess-cold：冷数据节点	

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置缩容参数。
   - “变更类型”：“缩容”。
   - “变更的资源”：显示资源的变化量。
   - “变更的角色”：设置缩容后的节点数量，支持修改的取值范围请参考表 4-27。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“缩容中”，表示集群正在缩容。当集群状态变为“可用”，则表示缩容成功。

4.5.5 缩容指定节点

当集群有充足的能力处理当前数据时，为节省资源可以执行“指定节点缩容”任务，指定一个或多个节点进行缩容。建议在业务低峰期进行缩容操作。

前提条件

集群处于“可用”状态，且无正在进行的任务。

约束限制

- 要确保缩容之后的磁盘使用量小于 80%，且集群每个节点类型中每个 AZ 的节点数至少为 1。
- 关于跨 AZ 的集群，在不同 AZ 中同类型节点个数的差值要小于等于 1。

- 关于没有 Master 节点的集群，每次缩容的数据节点和冷数据节点个数之和要小于缩容前数据节点和冷数据节点个数之和的一半，缩容后的数据节点和冷数据节点个数之和要大于索引的最大副本个数。
- 关于有 Master 节点的集群，每次缩容的 Master 节点个数要小于当前 Master 节点总数的一半，缩容后的 Master 节点个数必须是奇数且不小于 3。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 选择“缩容指定节点”页签。
5. 在缩容指定节点页面，设置如下参数：
   - 在数据节点表格中勾选待缩容的节点。
6. 单击“下一步”。
7. 确认变更信息后，单击“提交申请”，在数据迁移弹窗中确认是否进行数据迁移（建议勾选“进行数据迁移”避免数据丢失），单击“确认”提交申请。

   数据迁移过程中，系统会把待替换节点中的所有数据分片移动到剩余节点，待数据迁移完成后再进行节点替换操作。当待替换节点上的数据在剩余节点存在副本时，可跳过数据迁移，缩短集群变更时间，减少数据迁移带来的负载压力。
8. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“缩容中”，表示集群正在缩容。当集群状态变为“可用”，则表示缩容成功。

4.5.6 替换指定节点

当集群中的节点发生故障时，可以执行“指定节点替换”任务。通过删除故障节点，重建一个一样的新节点，实现节点替换。替换指定节点过程中会将替换节点的数据提前转移，不会造成数据丢失。

前提条件

集群处于“可用”状态，且无正在进行的任务。

约束限制

- 一次只能选择一个节点进行替换。
- 节点替换过程会按照原节点的 ID、IP 地址、规格、AZ 等信息重建节点。
- 节点替换过程中不会保留手动操作。例如被替换的节点手动加过回程路由，那么节点替换完成后，需要重新添加回程路由。
- 如果替换的是数据节点（ess 或 ess-cold），需要确认集群/节点是否满足以下条件：
  a. 替换数据节点或冷数据节点的过程中，会先将被替换节点的数据迁移到其他数据节点，因此集群中每个索引的副本数和主分片数之和的最大值要小于集群的数据节点个数之和（含 ess 和 ess-cold）。替换过程耗时跟数据迁移到其他节点的耗时强相关。

b. 版本号在 7.6.2 以前的集群，不能有 close 的索引，否则不支持替换数据节点或冷数据节点。
c. 被替换的数据节点或冷数据节点所在的 AZ 需要有两个及以上的数据节点（含 ess 和 ess-cold）。
d. 如果替换的数据节点或冷数据节点所在集群不存在 Master 节点（ess-master），则集群中可用的数据节点（含 ess 和 ess-cold）个数要大于等于 3。
e. 如果替换的是 Master 节点（ess-master）或 Client 节点（ess-client），则不受以上四条约束。
f. 如果替换的是故障节点，不管什么类型都不受以上四条约束。因为故障节点不包含在“_cat/nodes”中。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 选择“替换指定节点”页签。
5. 在替换指定节点页面，设置如下参数：
   - 在数据节点表格中勾选待替换的节点。
6. 单击“提交申请”。在数据迁移弹窗中确认是否进行数据迁移（建议勾选“进行数据迁移”避免数据丢失），单击“确认”提交申请。

   数据迁移过程中，系统会把待替换节点中的所有数据分片移动到剩余节点，待数据迁移完成后再进行节点替换操作。当待替换节点上的数据在剩余节点存在副本时，可跳过数据迁移，缩短集群变更时间，减少数据迁移带来的负载压力。
7. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“节点替换中”，表示集群正在替换节点。当集群状态变为“可用”，则表示节点替换成功。

4.5.7 添加 Master/Client 节点

当集群数据面业务增长，需要动态调整集群形态，增加 Master/Client 节点时，可以执行“添加 Master/Client 节点”任务。添加 Master/Client 节点时，业务不会中断。

前提条件

集群处于“可用”状态，且无正在进行的任务。

约束限制

- 如果集群已经具有 Master 节点和 Client 节点，则“形态变更”页面将不存在“添加 Master/Client 节点”页签。如此时需要添加 Master 节点和 Client 节点，可以执行 4.5.2 扩容任务。
- 添加 Master 节点或 Client 节点时，不同节点类型的可配置节点数量不同，具体情况请参考表 4-28。

表4-28 不同节点类型的节点数量说明

节点类型	节点数量的取值范围
Master 节点	3~9 的奇数
Client 节点	1~32

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 选择“添加 Master/Client 节点”页签。
5. 在添加 Master/Client 节点页面，勾选需要添加的节点，并配置“节点规格”、“节点数量”和“节点存储”。
   - 一次任务只能添加一种节点类型；如需添加 2 种，需要执行 2 次“添加 Master/Client 节点”任务。
   - 如果集群本身已经有 Master 节点或 Client 节点时，则只能添加另一种类型的节点。
6. 单击“下一步”。
7. 确认变更信息后，单击“提交申请”。

   返回集群列表页面，集群的“任务状态”列中显示为“扩容中”。
   - 如果添加 Master 节点，当“集群状态”变为“可用”时表示节点添加成功。

须知

如果集群版本小于 7.x，当“集群状态”变为“可用”时，还需要重启集群的所有数据节点和冷数据节点，使新增节点的配置生效。如果不重启使新增节点配置生效，使用集群时因为上报机制问题可能显示不可用状态（集群业务并未不可用）。重启操作请参见 4.7.8 重启集群。

   - 如果添加 Client 节点，当“集群状态”变为“可用”时表示节点添加成功。可选择是否重启数据节点和冷数据节点，下线数据节点上的 Cerebro 和 Kibana 的进程。

4.5.8 更改安全模式

集群创建完成后，可以通过更改安全模式进行集群模式变更。CSS 服务支持如下几种安全模式更改：

- 非安全模式切换为安全模式：“非安全模式”切换为“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”

- 安全模式切换为非安全模式：“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”切换为“非安全模式”
- 切换安全模式下的协议：“安全模式+HTTP 协议”切换为“安全模式+HTTPS 协议”、“安全模式+HTTPS 协议”切换为“安全模式+HTTP 协议”

背景信息

CSS 服务支持创建多种安全模式的集群，不同安全模式的差异请参见表 4-29。

表4-29 集群安全模式对比

集群安全模式	适用场景	优点	缺点
非安全模式	适合内网业务，用于测试场景。	简单，接入集群容易。	安全性差，谁都可以访问集群。
安全模式+HTTP 协议	可以实现用户权限隔离，适用于对集群性能敏感的场景。	访问集群需要安全认证，提升了集群安全性，且通过 HTTP 协议访问集群能保留集群的高性能。	无法公网访问集群。
安全模式+HTTPS 协议	有非常高的安全要求，且需要公网访问集群的场景。	访问集群需要安全认证，提升了集群安全性，且 HTTPS 协议的通讯加密可以实现集群公网访问功能。	通过 HTTPS 协议访问集群，集群性能相对 HTTP 协议来说，会下降 20%左右。

前提条件

- 建议更改集群安全模式前，先完成数据备份。
- 集群必须处于“可用”状态，且无正在进行的任务。

约束限制

- 2022 年 11 月之后创建的集群（且集群版本不小于 6.5.4）才支持切换安全模式。
- 在更改集群安全模式的过程中，集群会自动重启。重启会导致业务中断，并且重启后调用集群的认证方式也会发生改变，客户端需要做相应调整。
- 对于已打开过 Kibana 会话框的集群，在更改集群安全模式后，Kibana 会提示 Session 错误，此时需要清理缓存再打开 Kibana 才能正常访问。

非安全模式切换为安全模式

介绍集群从“非安全模式”切换为“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”的操作。当集群从非安全模式变更为安全模式后，访问集群将需要进行安全认证。

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择“集群管理>Elasticsearch”，进入 Elasticsearch 集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 选择“更改安全模式”页签。
5. 在更改安全模式页面，打开“安全模式”的开关，输入并确认集群的管理员密码。

   图4-19 非安全模式切换为安全模式

6. 选择集群是否启用“HTTPS 访问”。
   - 打开开关：实现“非安全模式”切换为“安全模式+HTTPS 协议”。集群通讯使用 HTTPS 协议，通讯数据将进行加密，且可以启用集群的公网访问功能。
   - 关闭开关：实现“非安全模式”切换为“安全模式+HTTP 协议”。集群通讯使用 HTTP 协议，无法启用集群的公网访问功能。
7. 单击“提交申请”，确认后将返回集群列表页面。

   集群的“任务状态”列中显示为“安全模式更改中”，当“集群状态”变为“可用”表示变更成功。

安全模式切换为非安全模式

介绍集群从“安全模式+HTTP 协议”或“安全模式+HTTPS 协议”切换为“非安全模式”的操作。当集群从安全模式变更为非安全模式后，访问集群将不再需要进行安全认证。

须知

- 非安全模式的集群将无需安全认证即可访问，且采用的是 HTTP 协议传输数据，所以请确保集群访问环境的安全性，勿将访问接口暴露到公网环境上。
- 在安全模式切换为非安全模式的过程中，变更任务会删除原安全模式的集群用到的索引。切换前请做好数据备份，以免数据丢失。
- 如果集群已绑定公网 IP，则无法将安全模式切换为非安全模式，需要先解绑公网 IP，才能切换。
- 如果集群已启用 Kibana 公网访问，则无法将安全模式切换为非安全模式，需要先关闭 Kibana 公网访问功能，才能切换。

1. 登录云搜索服务管理控制台。
2. 左侧导航栏选择“集群管理”，进入集群列表页面，选择需要更改安全模式的集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
3. 选择“更改安全模式”页签。
4. 在更改安全模式页面，关闭“安全模式”的开关。

   图4-20 安全模式切换为非安全模式

5. 单击“提交申请”，确认后将返回集群列表页面。

   集群的“任务状态”列中显示为“安全模式更改中”，当“集群状态”变为“可用”表示变更成功。

切换安全模式下的协议

介绍安全模式的集群切换访问协议的操作，包括“安全模式+HTTP 协议”切换为“安全模式+HTTPS 协议”、“安全模式+HTTPS 协议”切换为“安全模式+HTTP 协议”。

须知

如果集群已绑定公网 IP，则无法将协议从 HTTPS 切换到 HTTP，需要先解绑公网 IP，才能切换安全模式下的协议。

1. 登录云搜索服务管理控制台。
2. 左侧导航栏选择“集群管理”，进入集群列表页面，选择需要更改安全模式的集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
3. 选择“更改安全模式”页签。
4. 在更改安全模式页面，修改“HTTPS 访问”的开关切换安全模式下的集群访问协议。

图4-21 切换协议

- 打开开关：实现“安全模式+HTTP 协议”切换为“安全模式+HTTPS 协议”。
  切换为 HTTPS 协议后，集群通讯将进行加密，且可以启用集群的公网访问功能。
- 关闭开关：实现“安全模式+HTTPS 协议”切换为“安全模式+HTTP 协议”。关闭开关会弹出告警提示，单击“确定”后关闭开关。
  切换为 HTTP 协议后，集群通讯将不再加密，且无法启用集群的公网访问功能。

5. 单击“提交申请”，确认后将返回集群列表页面。
   集群的“任务状态”列中显示为“安全模式更改中”，当“集群状态”变为“可用”表示变更成功。

4.5.9 切换可用区

CSS 服务支持跨可用区的能力，在集群使用过程中遇到可用区资源不足或需要提升集群的高可用性时，可以通过切换可用区进行升配。本文介绍如何切换可用区。

场景描述

切换可用区包含两大场景：可用区高可用改造和可用区平移切换。

- 可用区高可用改造：适用于单 AZ 改造成两 AZ、单 AZ 改造成三 AZ 或两 AZ 改造成三 AZ 的场景，目的是为了提升集群的高可用性。
- 可用区平移切换：适用于从一个 AZ 完全迁移到另一个 AZ 的场景，目的是为了解决当前可用区资源不足的问题。

前提条件

- 确保存在资源充足的可用区。
- 集群处于“可用”状态，且无正在进行的任务。
- 确认集群不存在非标操作。当集群进行过修改回程路由、系统参数、Kibana 配置等这些非标操作时，这些操作无法被记录到系统中，因此在切换过程中将无法继承，切换后可能会影响业务。

约束限制

- 为了保证业务的连续性，集群中数据节点和冷数据节点的个数之和要不小于 3。
- 切换过程中会逐一下线节点再新建节点，需要保证单节点下线后，其余节点的磁盘容量能够接纳该节点的所有数据。
- 集群中索引的最大主备分片数能够被分配到剩余数据节点和冷数据节点中，防止切换过程中出现节点下线后副本无法分配的情况。即“索引的主备分片数的最大值+1 ≤ 切换前的数据节点数和冷数据节点之和”。
- 切换前建议做好数据备份，防止升级故障数据丢失。
- 当集群处于切换可用区过程中，部分节点的 AZ 属性可能已经更改，此时节点的可用区可能会显示出中间状态，等待任务结束后，集群的可用区显示将恢复正常状态。
- **可用区高可用改造**是基于原可用区进行扩展，切换时必须保留原可用区。如果是单 AZ 改造成两 AZ 或单 AZ 改造成三 AZ 的场景，必须所有节点一起改造；如果是两 AZ 改造成三 AZ 的场景，则既支持同时操作集群中所有类型的节点，也支持单独操作集群中某一类型的节点（例如，在已有两 AZ 的集群中，可以只对 Master 节点进行切换可用区，将 Master 节点形变为三 AZ，而其他节点依旧是两 AZ）。高可用改造会尽可能挪动最小的节点重建集群，改造完成后，因未挪动节点的 yml 配置更新，需要重启集群才能生效。
- **可用区平移切换**一次只能迁移一个可用区，切换时选择的可用区只能有一个不同。平移切换支持同时操作集群中所有类型的节点，或者单独操作集群中某一类型的节点（例如，在已有两 AZ 的集群中，可以只对 Master 节点进行平移切换，将 Master 节点从当前 AZ 迁移到另一个 AZ，而其他节点依旧在当前 AZ）。除了单 AZ 平移切换场景，其他的多 AZ 平移切换场景均需要重启集群生效。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 选择“切换可用区”页签。
5. 在切换可用区页面，配置切换参数。

表4-30 切换可用区的参数说明

参数	描述
切换类型	• 可用区高可用改造：支持单 AZ 改造成两 AZ、单 AZ 改造成三 AZ 或两 AZ 改造成三 AZ。 高可用改造会尽可能挪动最小的节点重建集群，改造完成后，因未挪动节点的 yml 配置更新，需要重启集群才能生效。 • 可用区平移切换：支持从一个 AZ 迁移到另一个 AZ。 除了单 AZ 平移切换场景，其他的多 AZ 平移切换场景均需要重启集群生效。
当前节点类	选择进行 AZ 切换的节点类型。每次可选择一种节点类型进行操作，

参数	描述
型	当集群有多种节点类型时，可选择“所有节点”进行全部切换。 说明 如果是单 AZ 改造成两 AZ 或单 AZ 改造成三 AZ 的场景，只支持所有节点一起改造，因此只能选择“所有节点”。
当前 AZ 集合	显示当前集群所在的可用区。
目标 AZ 集合	选择需要切换的目标可用区。 • 可用区高可用改造：高可用区改造是基于原可用区进行扩展，因此“目标 AZ 集合”必须包含“当前 AZ 集合”且至少再选择一个可用区，最多选择 3 个可用区。 • 可用区平移切换：一次切换任务只能迁移一个 AZ，因此“目标 AZ 集合”只能有一个可用区和“当前 AZ 集合”不同。
配置委托	选择 IAM 委托，授权当前账号切换可用区的操作权限。 如果没有合适的委托，可以单击“创建委托”跳转到 IAM 控制台新建委托。 说明 选择的委托必须授权了“Tenant Administrator”或“VPC Administrator”策略。

6. 配置完成后，单击“提交申请”。确认是否进行全量索引快照备份检测，单击“确定”启动切换任务。
7. 在“任务记录”列表，显示当前切换任务。当“任务状态”为“运行中”时，可以展开任务列表，单击“查看进度”查看详细的切换进度。

   当“任务状态”为“失败”时，可以重试任务或者直接终止任务。

   - 重试切换任务：在任务列表的操作列，单击“重试”，重新切换可用区。
   - 终止切换任务：在任务列表的操作列，单击“终止”，结束切换可用区。

     当任务终止后，如果原节点还未切换成功，可以通过 4.5.6 替换指定节点尝试恢复节点。

说明

如果存在部分节点已经完成切换，此时集群的 AZ 形态已经发生变化，此时终止切换任务，可能无法再次按照上一次的切换请求进行命令下发，请谨慎执行终止切换任务。

4.6 升级集群版本

版本升级功能支持同版本升级、跨版本升级和跨引擎升级，同版本升级是升级集群的内核补丁，用于修复问题或优化性能；跨版本升级是升级集群的版本，用于功能加强或版本收编；跨引擎升级是支持 Elasticsearch 集群升级至 OpenSearch 集群。

场景描述

升级原理

升级过程采用的是 one-by-one 的方式，不会中断业务。升级时，先下线一个节点，将该节点的数据迁移到其余节点，再创建一个目标版本的新节点，将已下线节点的网卡 port 挂载到新节点，以此保留节点 IP 地址，待新节点加入到集群后，再依次将其余节点进行替换。当集群的数据量很大时，升级耗时将依赖数据迁移耗时。

升级流程

1. 升级前检查
2. 4.4.3 手动创建快照
3. 创建升级任务

版本限制

待升级的集群版本不同，支持升级的目标版本也不同，具体请参见表 4-31。

表4-31 版本升级的版本限制

当前版本	目标版本
Elasticsearch：6.2.3	Elasticsearch：6.5.4，6.8.23
Elasticsearch：6.5.4	Elasticsearch：6.8.23
Elasticsearch：6.8.23	Elasticsearch：7.6.2，7.10.2
Elasticsearch：7.1.1	Elasticsearch：7.6.2，7.10.2
Elasticsearch：7.6.2	Elasticsearch：7.10.2
Elasticsearch：7.9.3	Elasticsearch：7.10.2
说明： • ELasticsearch 集群以 7.6.2 和 7.10.2 为主力版本，建议其他版本的集群收编到该版本。实际支持的目标版本请以升级页面中“目标镜像”的可选值为准。 • 5.x 版本的 ELasticsearch 集群不支持跨版本升级，6.2.3 和 6.5.4 版本 Elasticsearch 集群可先升级到 6.8.23 版本，再升级到 7.x.x 版本。	

约束限制

- 最多同时支持 20 个集群升级，建议在业务低峰期进行升级操作。
- 待升级的集群不能存在正在进行中的任务。
- 升级任务一旦启动就无法中止，直到升级任务的“任务状态”显示“失败”或“成功”才结束。
- 升级过程中，存在节点替换的过程，此时如果有请求发送到替换中的节点，请求可能会失败。为防止该情况出现，建议通过终端节点服务或者独享型负载均衡器接入集群进行访问。

- 升级过程中会重建 Kibana 和 Cerebro 组件，重建过程中，Kibana 和 Cerebro 可能会访问失败。而且，由于不同 Kibana 版本相互不兼容，升级过程中 Kibana 还可能因为版本不兼容而无法访问。访问失败的问题，待集群升级成功后会自动恢复。

升级前检查

为了保证升级成功，需要做升级前检查，升级前检查主要包括如下事项：

表4-32 升级前检查项

检查项	检查方式	描述	正常状态
集群状态	系统检查	升级任务启动后，系统会自动检查集群状态。集群状态为 green 或者 yellow，表示集群可以正常提供服务，不存在未分配的主分片。	“集群状态”为“可用”。
节点个数	系统检查	升级任务启动后，系统会自动检查节点个数。为了保证业务的连续性，集群中数据节点和冷数据节点的个数之和要不小于 3。	集群的数据节点和冷数据节点个数之和 ≥ 3
磁盘容量	系统检查	升级任务启动后，系统会自动检查磁盘容量。升级过程中会逐一下线节点再新建节点，需要保证单节点下线后，其余节点的磁盘容量能够接纳该节点的所有数据。	下线单节点后，剩余节点能够包含集群的所有数据。
数据副本	系统检查	检测集群中索引的最大主备分片数是否能够分配到剩余数据节点和冷数据节点中，防止升级过程中出现节点下线后副本无法分配的情况。	索引的主备分片数的最大值 +1 ≤ 升级前的数据节点数和冷数据节点之和
数据备份	系统检查	升级前要做好数据备份，防止升级故障数据丢失。在提交升级任务时，可以选择是否需要系统检查全量索引的快照备份。	确认是否存在数据备份。
资源充足	系统检查	升级任务启动后，系统会自动检查资源。升级过程中会新建资源，需要保证有资源可用。	资源可用且配额充足。
自定义插件	系统检	历史版本的集群装有自定义插件才需要进行该项检查。如果集群装有自定义插件，则需要在插件	装有自定义插件的集群已经把对应版本的插件包上传到

检查项	检查方式	描述	正常状态
	查+人工检查	管理界面上传目标版本的所有插件包，升级过程中才能将自定义插件转入新节点中，否则集群升级成功后，自定义插件将丢失。升级任务启动后，系统会自动检查是否已经上传自定义插件包，但是上传的插件包是否正确需要人工检查。 说明 如果上传的插件包不正确或者不兼容将导致升级过程中无法自动安装插件包，升级任务会失败。如果需要恢复集群，可以终止升级任务，执行4.5.6 替换指定节点操作，修复当前升级失败的节点。 升级完成后，自定义插件对应的“插件状态”将会重置为“已上传”状态。	升级插件列表中。
自定义配置	系统检查	升级过程中，系统会自动同步集群配置文件“elasticsearch.yml”中的内容。	如果集群有自定义的参数配置，升级完成后，配置都未丢失。
非标操作	人工检查	确认是否存在非标操作。非标操作指的是没有被记录下来的手动操作，这些操作在升级过程中无法自动传递，比如“Kibana.yml”配置文件修改、系统配置、回程路由等。	有些非标操作是兼容的，例如安全插件的修改，可以通过元数据保留下来；系统配置修改，可以通过镜像的操作保留下来。但是有些非标操作例如“kibana.yml”文件修改就无法保留，需要提前备份。
兼容性检测	系统检查+人工检查	跨版本升级的升级任务启动后，系统会自动检测集群升级前后两个版本是否有不兼容配置。如果集群装有自定义插件，则自定义插件的版本兼容性需要人工检测。	跨版本升级时，升级前后没有不兼容的配置。

创建升级任务

1. 登录云搜索服务管理控制台。
2. 左侧导航栏选择“集群管理”，进入集群列表页面，单击目标集群名称，进入集群基本信息页面。
3. 选择“版本升级”。
4. 在升级页面，配置升级参数。

表4-33 升级参数说明

参数	描述
升级类型	• 同版本升级：升级集群的内核补丁，集群版本号不变。 • 跨版本升级：升级集群的版本，集群版本号升级。 • 跨引擎升级：Elasticsearch 集群升级至 OpenSearch 集群，当前仅支持 7.10.2 版本 Elasticsearch 集群升级至 1.3.6 版本 OpenSearch 集群。
目标镜像	选择目标版本的镜像。选中镜像后，下方会显示镜像名称和目标版本的详细说明。 实际支持的目标版本请以升级页面中“目标镜像”的可选值为准。如果无法选择目标镜像，有如下几个原因： • 当前集群已是最新版本集群。 • 当前集群是 23 年之前创建的旧集群，且存在向量索引。 • 当前局点暂未录入新版本镜像。
配置委托	选择 IAM 委托，授权当前账号升级操作权限。 如果没有合适的委托，可以单击“创建委托”跳转到 IAM 控制台新建委托。 **说明** **选择的委托必须授权了**“Tenant Administrator”**或**“VPC Administrator”**策略。**

5. 配置完成后，单击“确认提交”。确认是否进行“全量索引快照备份检测”和“集群负载检测”，单击“确定”启动升级任务。

   当集群负载过高时，升级任务大概率会卡住或失败。增加“集群负载检测”可以有效避免失败。检测中如果出现以下四种情况，请等待或者主动降低负载。如果需求紧急且已了解升级失败风险，也可以不做“集群负载检测”。“集群负载检测”检测项包含：

   - 最大 search 队列排队数小于 1000（nodes.thread_pool.search.queue < 1000）
   - 最大 write 队列排队数小于 200（nodes.thread_pool.write.queue < 200）
   - 最大 cpu 使用率小于 90%（nodes.process.cpu.percent < 90）
   - 最大 load 相对核数占比小于 80%（ nodes.os.cpu.load_average / cpu 核数 < 80%）

6. 在“任务记录”列表，显示当前升级任务。当“任务状态”为“运行中”时，可以展开任务列表，单击“查看进度”查看详细的升级进度。

   当“任务状态”为“失败”时，可以重试任务或者直接终止任务。

   - 重试升级：在任务列表的操作列，单击“重试”，重新升级。
   - 终止升级：在任务列表的操作列，单击“终止”，结束升级。

须知

- 同版本升级：当升级的“任务状态”为“失败”即可终止升级任务。
- 跨版本升级：当升级的“任务状态”为“失败”时，且没有任何节点升级成功的，才可以终止升级任务。

当升级任务终止后，集群的“任务状态”将清空“升级失败”的状态，相当于集群回退到升级前状态，不影响集群进行其他任务。

4.7 管理集群

4.7.1 集群列表概览

集群列表显示云搜索服务所有的集群，集群数量较多时，可采用翻页显示，您可以查看任何状态下的集群。

集群列表默认按时间顺序排列，时间最近的集群显示在最前端，集群列表参数说明如表 4-34 所示。

在集群列表单击“导出”可以下载全量的集群列表信息。

表4-34 集群列表说明

参数	描述
名称/ID	表示集群的名称和 ID。单击集群名称可进入集群“基本信息”页面，展现了集群的基本信息。集群 ID 是系统自动生成的，是集群在服务中的唯一标示。
集群状态	展示集群当前的状态。集群状态说明请参见 6 查看集群运行状态和存储容量状态。
任务状态	展示重启集群、扩容集群、备份集群、恢复集群等任务的状态。
版本	表示此集群中 Elasticsearch 的版本号。
创建时间	表示集群的创建时间。
企业项目	表示集群所归属的企业项目。

参数	描述
内网访问地址	集群的内网访问地址和端口号，您可以使用此参数接入集群。集群有多个节点时，此处显示多个节点的内网访问地址和端口号。
计费模式	呈现集群的计费模式。
操作	展示集群可执行的操作入口，包含 Kibana、监控信息、重启、删除等其他更多操作。当某一操作无法执行时，显示为灰色链接。

4.7.2 查看集群基本信息

在集群的基本信息页面，可以获取集群的内网访问地址、公网访问地址、版本、节点等信息。

1. 登录云搜索服务管理控制台。
2. 选择“集群管理 > Elasticsearch”，进入集群列表页面。
3. 单击集群名称进入集群“基本信息”页面，查看集群的基本信息。

表4-35 基本信息的参数说明

类别	参数	描述
基本信息	集群名称	集群名称。支持自定义名称。 单击右侧  可以修改集群名称。
	ID	集群的唯一标识，是系统自动生成的。 同一个区域下，集群 ID 是唯一的。
	集群版本	集群的版本信息。
	集群状态	集群当前的状态。
	任务状态	集群当前的任务状态，如果没有进行中的任务则显示“--”。
	创建时间	集群创建的时间。
	集群存储容量（GB）	集群设置的存储容量。
	集群存储使用量（GB）	集群已使用的存储容量。
配置信息	区域	集群所在区域。
	可用区	集群所在的可用区。
	虚拟私有云	集群所属的虚拟私有云。

类别	参数	描述
	子网	集群所属的子网。
	安全组	集群所属的安全组。 单击右侧的“更改安全组”可以修改集群的安全组信息。 须知 进行安全组切换前，请确保业务接入时需要的 9200 端口已经放开，错误的安全组配置可能会导致业务无法访问，请谨慎操作。
	安全模式	集群的安全模式。 • 启用：表示当前集群是安全模式的集群。 • 未开启：表示当前集群是非安全集群。
	重置密码	仅安全模式的集群显示。 单击“重置”可以修改安全集群的管理员账户 **admin** 的密码。 说明 管理员密码的规则： • 可输入的字符串长度为 8~32 个字符。 • 密码至少包含大写字母、小写字母、数字和特殊字符四类中的三类。其中支持的特殊字符有：~!@#$%^&*()-_=+|[{}];:,<.>/? • 不能与管理员账户名或倒序的管理员账户名相同。 • 建议定期修改密码。
	企业项目	集群所属的企业项目。 单击项目名称可以跳转到项目管理页面查看企业项目的基本信息。
	公网访问	公网访问信息，仅安全模式的集群显示。 • 启用公网访问的安全集群，此处显示公网访问地址，通过该地址可以在公网访问安全集群。 • 未启用公网访问的安全集群，此处显示“--”。 使用公网地址访问集群时，建议开启访问控制并配置访问白名单，提高集群安全性。

类别	参数	描述
	访问控制	启用公网访问的集群是否设置访问控制，仅启用公网访问的集群显示。 • 已开启：启用访问控制开关，只允许白名单列表中的 IP 地址通过公网访问集群。 • 未开启：未启用访问控制开关，允许任何 IP 地址通过公网访问集群。 单击“设置”，可以更新访问控制开关和白名单。
	带宽	公网访问的带宽，仅启用公网访问的集群显示。 单击“修改”，可以更新带宽大小。
	HTTPS 访问	集群是否启用 HTTPS 访问协议。 • 关闭：表示集群未启用 HTTPS 访问，集群使用 HTTP 访问协议。 • 开启：表示集群启用了 HTTPS 访问协议，仅安全模式的集群支持开启 HTTPS 访问。启用 HTTPS 访问的安全集群可以单击“下载证书”获取 CER 安全证书，用于接入安全模式的集群。安全证书暂不支持在公网环境下使用。
	内网访问 IPv4 地址	集群的内网 IP 地址和端口号，使用此参数可以接入集群。如果集群只有一个节点，此处仅显示 1 个节点的 IP 地址和端口号，例如“10.62.179.32:9200”；如果集群有多个节点，此处显示所有节点的 IP 地址和端口号，例如“10.62.179.32:9200,10.62.179.33:9200”。
节点信息	节点规格	集群中节点的规格信息。
	节点存储	集群中节点的存储容量和存储类型。
	节点数量	集群中节点的个数。

4.7.3 管理标签

标签是集群的标识。为集群添加标签，可以方便用户识别和管理拥有的集群资源。

您可以在创建集群时添加标签，也可以在集群创建完成后，在集群的详情页添加标签。

如您的组织已经设定云搜索服务的相关标签策略，则需按照标签策略规则为集群添加标签。标签如果不符合标签策略的规则，则可能会导致集群创建失败，请联系组织管理员了解标签策略详情。

新建集群的标签管理

1. 登录云搜索服务管理控制台。
2. 单击右上角的“创建集群”，进入创建集群页面。
3. 在创建集群页面，“高级配置”选择“自定义”后，为集群添加标签。

   您可以选择预定义标签，并为此标签设置“标签值”。您可以单击“查看预定义标签”，进入“标签管理服务”，了解此用户下已有的标签。

   您也可以自定义“标签键”和“标签值”。

   云搜索服务的每个集群最多可以设置 10 个标签。当设置不正确时，可单击标签右侧的“删除”按钮，删除此标签。当不设置标签时，可保持为空。

表4-36 标签命名规则

参数	说明
标签键	• 对于同一个集群，标签键值唯一。 • 长度不超过 64 个字符。 • 只能包含数字、英文字母、下划线、中划线以及特殊字符_ . : = + - @。不能以空格开头和结尾。 • 不能为空。
标签值	• 长度不超过 64 个字符。 • 只能包含数字、英文字母、下划线、中划线以及特殊字符_ . : = + - @/。不能以空格开头和结尾。 • 不能为空。

已有集群的标签管理

您可以对已经创建的集群的标签进行修改，删除，也可以添加标签。

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击待管理标签的集群名称。

   系统跳转至该集群“基本信息”页面。
3. 左侧菜单栏选择“标签”，在此可以对集群标签进行添加，修改，删除操作。
   - 查看

     在“标签”页，可以查看当前集群的标签详情，包括标签个数，以及每个标签的键和值。
   - 添加

     单击左上角的“添加标签”，在弹出的“添加标签”窗口，输入新添加标签的键和值，并单击“确定”。

- 修改

  只能修改已有标签的标签值。

  单击标签所在行“操作”列下的“编辑”，在弹出的“编辑标签”窗口，输入修改后标签值，并单击“确定”。

- 删除

  单击标签所在行“操作”列下的“删除”，如果确认删除，在弹出的“删除标签”窗口，单击“确定”。

通过标签搜索集群

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击集群列表右上角的“标签搜索”。
3. 选择或输入需要搜索的标签键和标签值，单击“添加”将标签加入搜索输入框中。

   标签键和标签值仅支持从下拉列表中选择，当标签键和标签值全匹配时，系统可以自动查询到目标集群。当有多个标签条件时，会取各个标签的交集，进行集群查询。

   系统最多支持 10 个不同标签的组合搜索。
4. 单击“搜索”。

   系统根据标签键和标签值搜索目标集群。

4.7.4 管理日志

为了方便用户使用日志定位问题，云搜索服务提供了日志备份和日志查询功能。日志备份可以定期将集群的日志存储在 OBS 桶中，通过 OBS 可以直接下载需要的日志文件，进行问题分析定位。

日志查询

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，单击需要查询日志的集群名称，进入集群基本信息页面。
3. 左侧导航栏，选择“日志管理”。
4. 在日志管理页面进行日志查询。
5. 选择需要查询的节点、日志类型和日志级别信息后，单击  ，显示查询结果。

   查询日志时，是从最近时刻的 1 万条日志中进行匹配，查询结果最多显示 100 条。

备份日志

日志备份可以定期将 CSS 集群的日志存储到 OBS 桶中。

1. 登录云搜索服务管理控制台。

2. 在“集群管理”页面，单击需要配置日志备份的集群名称，进入集群基本信息页面。
3. 左侧导航栏，选择“日志管理”，在“日志备份开关”右侧单击开关，打开集群的日志管理功能。
4. 在“编辑日志备份配置”弹窗中，完成参数配置。

   弹窗中默认填写了云搜索服务自动为用户创建的“OBS 桶”、“备份路径”和“IAM 委托”，用于日志备份。支持用户参考表 4-37 修改默认值。

   如果集群已经启用了日志备份功能，也可以单击“日志备份配置”右侧的编辑图标，在“编辑日志备份配置”窗口，参考表 4-37 更新日志备份的配置参数。

表4-37 日志备份配置的参数说明

参数	说明	注意事项
“OBS 桶”	选择日志存储的 OBS 桶。单击右侧的“创建桶”支持新建 OBS 桶。	OBS 桶的所在区域必须跟集群的所在区域保持一致。 **说明** **如果是子账号，需要同时设置** GetBucketStoragePolicy、GetBucketLocation、ListBucket、ListAllMyBuckets **权限，才能看到** OBS **桶。**
“备份路径”	填写日志在 OBS 桶中的存放路径。	备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
“IAM 委托”	选择 IAM 委托，指当前账号授权云搜索服务访问或维护存储在 OBS 中数据。单击右侧的“创建委托”支持新建委托。	IAM 委托需满足如下条件： • “委托类型”选择“云服务”。 • “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 必选策略：“OBS Administrator”

5. 备份日志。
   - 自动备份日志。

     在“自动备份开关”右侧，单击开关，开启自动备份日志功能。

     开启“自动备份开关”后，在“修改日志备份策略”弹窗中设置“备份开始时间”。设置成功后，系统会按照设置的时间进行自动备份。

     打开“自动备份开关”后，单击开关右侧的编辑图标，可以修改“备份开始时间”。
   - 手动备份日志。

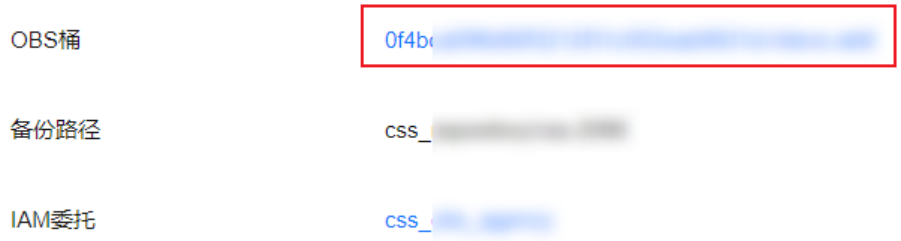
选择“日志备份”页签，单击“开始备份”，在弹出的确认提示框中单击“确定”，开始备份日志。

日志备份列表中的“任务状态”为“Successful”时，表示日志备份成功。

6. 查看日志文件。

日志备份成功后，用户可以单击“OBS 桶”跳转到桶列表，找到存放备份日志的桶查看日志文件。

图4-22 进入 OBS



日志文件介绍

Elasticsearch 和 OpenSearch 集群备份的日志文件主要包括废弃操作日志、运行日志、慢索引日志、慢查询日志。

表4-38 日志文件类型

日志名称	描述
clustername_deprecation.log	弃用操作的日志记录。
clustername_index_indexing_slowlog.log	慢索引日志。
clustername_index_search_slowlog.log	慢索引查询日志。
clustername.log	Elasticsearch 运行日志。
clustername_access.log	接入日志。

4.7.5 配置 YML 参数

云搜索服务支持用户修改 elasticsearch.yml 文件。

修改参数配置

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，单击需要修改参数配置的集群名称，进入集群基本信息页面。

3. 选择“参数配置”，单击“编辑”，根据需求修改对应模块的参数值。

表4-39 模块参数信息说明

模块名称	参数名称	说明
跨域访问	http.cors.allow-credentials	跨域访问是否返回头部的 Access-Control-Allow-Credentials。 取值范围：true、false。 默认值：false。
	http.cors.allow-origin	允许跨域访问的 IP，配置样例如 122.122.122.122:9200。
	http.cors.max-age	浏览器默认缓存时间。如果超过设置的时间后，缓存将自动清除。 单位：秒。 默认值：1728000。
	http.cors.allow-headers	跨域访问允许的 headers，包括 X-Requested-With, Content-Type, Content-Length，中间用英文逗号和空格分开。
	http.cors.enabled	是否允许跨域访问。 取值范围：true、false。 默认值：false。
	http.cors.allow-methods	跨域访问允许的方法，包括 OPTIONS, HEAD, GET, POST, PUT, DELETE，中间用英文逗号和空格分开。
集群索引重建	reindex.remote.whitelist	配置该参数可以将本集群数据通过 reindex 接口迁移到配置的集群，配置样例如 122.122.122.122:9200。
自定义缓存	indices.queries.cache.size	查询阶段的缓存大小。 取值范围：1-100。 单位：%。 默认值：10%。
线程池队列大小	thread_pool.force_merge.size	用来做 forcemerge 的队列大小。输入的参数值为整数类型。 默认值：1。
自定义	用户可以根据实际情况，添加相关参数名称。	自定义参数的取值。 **说明** • 如果自定义参数有多个取值，则取值的输入格式为[value1, value1, value1...]。

模块名称	参数名称	说明
		• 取值之间用英文逗号和空格隔开。 • 自定义参数值中不能包含冒号。

4. 修改完成后，单击上方的“提交”弹出“提交配置”窗口，确认参数无误后勾选“参数修改后需要手动重启才能生效”，单击“确定”。

   当下方的参数修改列表显示“作业状态”为“成功”时，表示修改保存成功。系统最多显示 20 条修改记录。

5. 返回集群列表，单击集群操作列的“更多 > 重启”重启集群，使修改的配置生效。
   - 如果修改了参数配置，未重启集群，则在“集群管理”页面的“任务状态”栏显示为“配置未更新”。
   - 如果修改后重启集群，“任务状态”显示“配置错误”，则表示修改参数配置文件失败。

4.7.6 查看系统默认插件列表

云搜索服务的 Elasticsearch 集群自带系统默认插件。可以通过控制台查看或在 Kibana 查询系统默认插件信息。

通过控制台查看

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击需要查看插件的集群名称，跳转至该集群基本信息页面。
3. 选择“插件管理”。
4. 在“系统默认插件列表”页查看当前版本支持的系统默认插件信息。

在 Kibana 查询

1. 登录云搜索服务控制台。
2. 在集群管理列表，选择需要查看插件的集群，单击操作列的“Kibana”登录 Kibana 界面。
   - 非安全模式的集群：将直接进入 Kibana 操作界面。
   - 安全模式的集群：需要在登录页面输入用户名和密码，单击“Log In”进入 Kibana 操作界面。用户名默认为 admin，密码为创建集群时设置的管理员密码。
3. 进入 Dev Tools，执行如下命令查看集群插件信息：

```
GET _cat/plugins?v
```

响应体示例如下：

```
name              component                    version
css-test-ess-esn-1-1 analysis-dynamic-synonym      7.6.2-xxxx-ei-css-v1.0.1
css-test-ess-esn-1-1 analysis-icu                7.6.2-xxxx-ei-css-v1.1.6
css-test-ess-esn-1-1 analysis-ik                 7.6.2-xxxx-ei-css-v1.0.1
......
```

“name”是集群的节点名称，“component”是插件名称，“version”是插件版本。

4.7.7 绑定企业项目

企业可以根据组织架构规划企业项目，将企业分布在不同区域的资源按照企业项目进行统一管理，同时可以为每个企业项目设置拥有不同权限的用户组和用户。本章节为您介绍 CSS 集群如何绑定、修改企业项目。

前提条件

在绑定企业项目前，您已在“企业项目管理控制台”创建企业项目。

绑定企业项目

在创建集群时，可以在“企业项目”绑定已创建的企业项目，也可以单击“查看项目管理”，前往企业项目管理控制台，新建企业项目和查看已有的企业项目。

修改企业项目

针对之前已创建的集群，其绑定的企业项目可根据实际情况进行修改。

1. 登录在云搜索服务管理控制台，
2. 在左侧导航栏，选择对应的集群类型，进入集群管理页面。
3. 在集群列表中，单击集群名称进入集群“基本信息”页面。
4. 在集群“基本信息”页面，单击“企业项目”右侧的企业项目名称，进入项目管理页面。
5. 在“资源”页签下，“区域”选项中选择当前集群所在的区域，“服务”选项中选“云搜索服务 CSS”。此时，资源列表将筛选出对应的 CSS 集群。

图4-23 筛选 CSS 集群

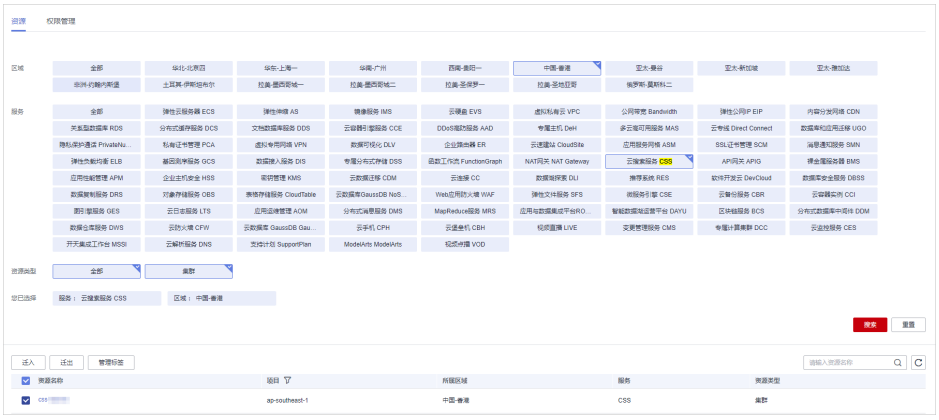
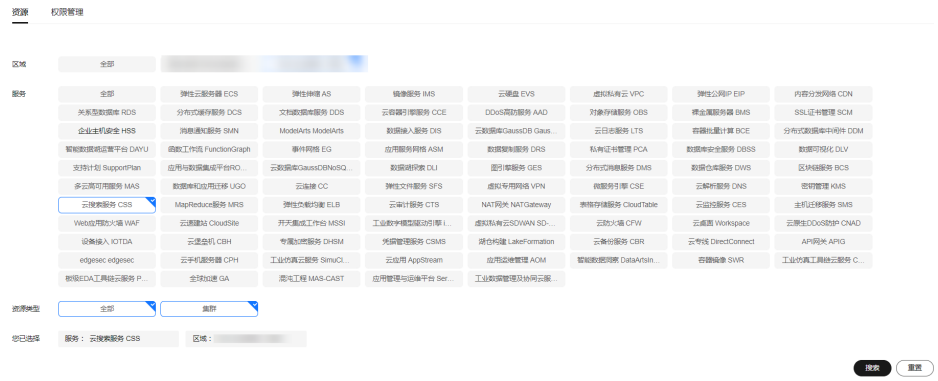


图4-24 筛选 CSS 集群



6. 勾选需要修改企业项目的集群，然后单击“迁出”。
7. 在“迁出资源”页面，选择“迁出方式”，再选择“请选择要迁入的企业项目”，然后单击“确定”。

图4-25 迁出资源

8. 迁出完成后，可以在云搜索服务管理控制台集群管理页面，查看修改后的集群企业项目信息。

4.7.8 重启集群

集群停止工作时，您可通过重启集群恢复运行。

前提条件

- 确认集群的“任务状态”没有正在执行中的任务，且集群未被冻结。
- 当集群处于可用状态时，确认集群已停止处理业务数据（如导入数据、搜索数据），否则重启集群时可能导致数据丢失等。建议在业务空闲时操作。

背景信息

重启集群支持快速重启和滚动重启。

快速重启

- 所有集群都支持。
- 当选择“节点类型”快速重启时，所选类型的所有节点会一起重启。
- 当选择“节点名称”快速重启时，一次只能重启一个节点。
- 快速重启过程中，集群不可用。

滚动重启

- 仅当集群的节点数量（含 Master 节点、Client 节点和冷数据节点）大于等于 3 时，才支持滚动重启。
- 滚动重启只支持根据“节点类型”进行重启。选择节点类型滚动重启时，所选类型的节点会依次重启。
- 滚动重启过程中，只有正在重启的节点不可用，不在重启过程中的节点可以正常提供服务。
- 当数据量比较大时，滚动重启耗时较长。

快速重启

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏，选择对应的集群类型，进入集群管理列表界面。
3. 在对应集群的“操作”列中单击“更多>重启”。
4. 在“重启集群”页面，选择“快速重启”。

   快速重启支持根据“节点类型”或者“节点名称”重启。如果选择“节点类型”，则支持选择多种节点类型同时进行快速重启。如果选择“节点名称”，则一次只能快速重启一个节点。
5. 重启集群后，请刷新页面，观察集群状态。重启过程中，集群状态为“处理中”，任务状态为“重启中”。如果集群状态变更为“可用”，表示集群已重启成功。

滚动重启

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏，选择对应的集群类型，进入集群管理列表界面。
3. 在对应集群的“操作”列中单击“更多>重启”。
4. 在“重启集群”页面，选择“滚动重启”。

   滚动重启支持根据“节点类型”进行重启。如果只需要重启集群中的某些类型的节点时，可以选择需要重启的节点类型。
5. 重启集群后，请刷新页面，观察集群状态。重启过程中，集群状态为“处理中”，任务状态为“重启中”。如果集群状态变更为“可用”，表示集群已重启成功。

4.7.9 删除集群

当用户已完成数据搜索业务，无需继续使用某一集群时，可删除集群释放资源。

说明

- 删除集群时，会清理集群业务数据，请谨慎操作。
- 如果集群启用过快照功能，且 OBS 桶中创建的快照并未被删除，删除集群时，并不会释放这部分备份数据。如果有需要，可以通过 OBS 桶中存储的快照信息恢复集群。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏，选择对应的集群类型，进入集群列表界面。
3. 在对应集群的“操作”列中单击“更多>删除”。
4. 在弹出的确认提示框中，输入需要删除的集群名称，单击“确定”完成集群删除。

4.8 自定义词库

4.8.1 配置自定义词库

云搜索服务的自定义词库功能，支持对文本进行分词，使得一些特殊词语在分词的时候能够被识别出来，便于根据关键词搜索文本数据。例如，根据公司名称来查询；或者根据网络流行词来查询，如“喜大普奔”。

说明

- 自定义词库功能上线之前创建的集群，无法使用自定义词库功能。
- 自定义词库一般用于中文分词，如果用于英文分词，会按照除#&+-.@_外的特殊符号进行切分。

背景信息

自定义词库使用的分词器包括 IK 分词器和同义词分词器。

IK 分词器配备主词词库和停词词库；同义词分词器配备同义词词库。配置自定义词库需要提前准备词库文件并上传至 OBS，详情请见上传词库文件至 OBS。

其中，IK 分词器包含 ik_max_word 和 ik_smart 分词策略。同义词分词器使用的是 ik_synonym 分词策略。

- ik_max_word：会将文本做最细粒度的拆分，比如会将“昨夜西风吹折千林梢”拆分为“昨夜西风,昨夜,西风,吹折千林梢,吹折,千林梢,千,林,折千林,千林,吹”，会穷尽各种可能的分词组合。
- ik_smart：会做最粗粒度的拆分，比如会将“昨夜西风吹折千林梢”拆分为“昨夜西风,吹折千林梢”。

前提条件

登录云搜索服务管理控制台的账号或 IAM 用户必须同时具备如下两个权限才能使用自定义词库功能。

- “全局服务”中“对象存储服务”项目的“OBS Administrator”权限。
- 当前所属区域的“Elasticsearch Administrator”权限。

上传词库文件至 OBS

配置自定义词库需要提前将词库上传至 OBS 桶。

1. 按表 4-40 要求准备词库文件。

表4-40 词库说明

词库类型	说明	文件要求
主词词库	主词为用户希望进行分词的特殊词语，例如“智能手机”和“喜大普奔”。主词词库则是用户自定义的特殊词语的集合。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，主词文件最大支持 100M。如果涉及单词，必须改成小写字母。
停词词库	停词为用户不希望进行分词或者关注的词语，例如“的”、“什么”、“怎么”等。停词词库是用户自定义的停词词语的集合。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，停词文件最大支持 100M。
同义词词库	同义词为意义相同的一组词语，例如“开心”和“高兴”。同义词词库是用户自定义的同义词词语的集合。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一组同义词对，同义词对中的同义词间逗号隔开，同义词文件最大支持 100MB。
静态主词词库	静态主词词库是 CSS 服务预置的常用主词集合，当需要查看静态主词词库时，可以单击地址 https://github.com/infinilabs/analysis-ik/blob/master/config/main.dic 获取词库。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，静态主词文件最大支持 100M。
静态停词词库	静态停词词库是 CSS 服务预置的常用停词集合，当需要查看静态停词词库时，可以单击地址 https://github.com/infinilabs/analysis-ik/blob/master/config/stopword.dic 获取词库。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，静态停词文件最大支持 100M。
Extra 主词词库	Extra 主词词库是 CSS 服务预置的生僻主词集合，当需要查看 Extra 主词词库时，可以单击地址	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，Extra 主词文件最大支持

词库类型	说明	文件要求
	https://github.com/infinilabs/analysis-ik/blob/master/config/extra_main.dic 获取词库。	100M。
Extra 停词词库	Extra 停词词库是 CSS 服务预置的生僻停词集合，当需要查看 Extra 停词词库时，可以单击地址 https://github.com/infinilabs/analysis-ik/blob/master/config/extra_stopword.dic 获取词库。	词库文件必须是 UTF-8 无 BOM 格式编码的文本文件，一行一个分词，Extra 停词文件最大支持 100M。

2. 上传词库文件至 OBS 桶，详细操作步骤请参见《对象存储服务用户指南》。上传的 OBS 桶必须和集群在相同“区域”。

管理自定义词库

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏，选择“集群管理 > Elasticsearch”，进入集群列表页面。
3. 在“集群管理”页面，单击需要配置自定义词库的集群名称，进入集群基本信息页面。
4. 选择“自定义词库”。
5. 在“自定义词库”页面，配置集群的自定义词库或修改预置词库。
   a. 当需要配置用户自定义的词库时，参考表 4-41 修改对应的词库。

表4-41 配置自定义词库

参数	说明
OBS 桶	选择词库文件存储的 OBS 位置。 单击“创建桶”可以跳转到创建桶页面新建 OBS 桶，新建 OBS 桶必须和集群在相同“区域”。
主词词库	主词词库是用户自定义的词库，初始状态为空。默认选择“不更新”表示不配置该词库。 • 当需要添加自定义的主词词库时，单击“更新”，选择 txt 格式的词库文件。 • 当无需添加自定义的主词词库时，单击“不使用此词库”，删除词库。
停词词库	停词词库是用户自定义的词库，初始状态为空。默认选择“不更新”表示不配置该词库。 • 当需要添加自定义的停词词库时，单击“更新”，选择 txt 格式的词库文件。

参数	说明
	• 当无需添加自定义的停词词库时，单击“不使用此词库”，删除词库。
同义词词库	同义词词库是用户自定义的词库，初始状态为空。默认选择“不更新”表示不配置该词库。 • 当需要添加自定义的同义词词库时，单击“更新”，选择txt 格式的词库文件。 • 当无需添加自定义的同义词词库时，单击“不使用此词库”，删除词库。

b. 当需要修改预置词库时，单击打开“修改预置词库”右侧的开关，参考表 4-42 修改对应的预置词库。

说明

如果界面不存在静态词库和 Extra 词库这 4 个词库时，表示该集群版本不支持删除或修改这 4 个预置词库。如果想要使用该功能，建议升级集群版本，或者新建集群并进行数据迁移。

表4-42 配置预置词库

参数	说明
静态主词词库	静态主词词库是预置的常用词语的主词库。默认选择“不更新”表示使用该预置词库。 • 当需要修改预置的静态主词词库时，单击“更新”，选择txt 格式的词库文件。 • 当不使用静态主词库时，单击“不使用此词库”，删除词库。
静态停词词库	静态停词词库是预置的常用词语的停词库。默认选择“不更新”表示使用该预置词库。 • 当需要修改预置的静态停词词库时，单击“更新”，选择txt 格式的词库文件。 • 当不使用静态停词库时，单击“不使用此词库”，删除词库。
Extra 主词词库	Extra 主词词库是预置的生僻词语的主词库。默认选择“不更新”表示使用该预置词库。 • 当需要修改预置的 Extra 主词词库时，单击“更新”，选择txt 格式的词库文件。 • 当不使用 Extra 主词库时，单击“不使用此词库”，删除词库。
Extra 停词词库	Extra 停词词库是预置的生僻词语的 Extra 停词库。默认选择“不更新”表示使用该预置词库。 • 当需要修改预置的 Extra 停词词库时，单击“更新”，选择

参数	说明
	txt 格式的词库文件。 • 当不使用 Extra 停词库时，单击“不使用此词库”，删除词库。

6. 单击“保存”，在弹窗中单击“确定”。词库信息在下方呈现，此时词库状态为“更新中”。请耐心等待 1 分钟左右，当词库配置完成后，词库状态变更为“成功”。
7. 当删除或更新静态词库和 Extra 词库这 4 个词库时，需要重启集群才能使配置的词库生效；其他词库的更新为动态更新，无需重启集群。重启集群的操作指导请参见 4.7.8 重启集群。

4.8.2 使用示例

场景说明

通过给集群配置自定义词库，将“智能手机”设置为主词，“是”设置为停词，“开心”和“高兴”设置为同义词。使用配置好的集群，对文本内容“智能手机是很好用”进行关键词搜索，查看关键词查询效果；对文本内容“我今天获奖了我很开心”进行同义词搜索，查看同义词查询效果。

步骤一：配置自定义词库

1. 准备词库文件（UTF-8 无 BOM 格式编码的文本文件），上传到对应 OBS 路径下。

   主词词库文件中包含词语“智能手机”；停词词库文件中包含词语“是”；同义词词库文件中包含一组同义词“开心”和“高兴”。

   说明

   由于系统预置的静态停词词库包含了“是”、“的”等常用词，如果集群未删除或更新预置词库，则此类停用词可以不用上传。

2. 在云搜索服务管理控制台，单击左侧导航栏的“集群管理”。
3. 在“集群管理”页面，单击需要配置自定义词库的集群名称，进入集群基本信息页面。
4. 在左侧导航栏，选择“自定义词库”，参考管理自定义词库为集群配置 1 准备好的词库文件。
5. 待词库配置信息生效后，返回集群列表。单击集群操作列的“Kibana”接入集群。
6. 在 Kibana 界面，单击左侧导航栏的“Dev Tools”，进入操作页面。
7. 执行如下命令，查看自定义词库的不同分词策略的分词效果。
   - 使用 ik_smart 分词策略对文本内容“智能手机是很好用”进行分词。

     示例代码：

     ```
     POST /_analyze
     {
     ```

```
  "analyzer":"ik_smart",
  "text":"智能手机是很好用"
}
```

运行结束后，查看分词效果：

```
{
  "tokens": [
    {
      "token": "智能手机",
      "start_offset": 0,
      "end_offset": 4,
      "type": "CN_WORD",
      "position": 0
    },
    {
      "token": "很好用",
      "start_offset": 5,
      "end_offset": 8,
      "type": "CN_WORD",
      "position": 1
    }
  ]
}
```

- 使用 ik_max_word 分词策略对文本内容“智能手机是很好用”进行分词。

  示例代码：

```
POST /_analyze
{
  "analyzer":"ik_max_word",
  "text":"智能手机是很好用"
}
```

运行结束后，查看分词效果：

```
{
  "tokens" : [
    {
      "token" : "智能手机",
      "start_offset" : 0,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 0
    },
    {
      "token" : "智能",
      "start_offset" : 0,
      "end_offset" : 2,
      "type" : "CN_WORD",
      "position" : 1
    },
    {
      "token" : "智",
      "start_offset" : 0,
      "end_offset" : 1,
      "type" : "CN_WORD",
      "position" : 2
    },
```

```
    {
      "token" : "能手",
      "start_offset" : 1,
      "end_offset" : 3,
      "type" : "CN_WORD",
      "position" : 3
    },
    {
      "token" : "手机",
      "start_offset" : 2,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 4
    },
    {
      "token" : "机",
      "start_offset" : 3,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 5
    },
    {
      "token" : "很好用",
      "start_offset" : 5,
      "end_offset" : 8,
      "type" : "CN_WORD",
      "position" : 6
    },
    {
      "token" : "很好",
      "start_offset" : 5,
      "end_offset" : 7,
      "type" : "CN_WORD",
      "position" : 7
    },
    {
      "token" : "好用",
      "start_offset" : 6,
      "end_offset" : 8,
      "type" : "CN_WORD",
      "position" : 8
    },
    {
      "token" : "用",
      "start_offset" : 7,
      "end_offset" : 8,
      "type" : "CN_WORD",
      "position" : 9
    }
  ]
}
```


步骤二：使用关键词搜索

Elasticsearch 7.x 之前的版本和之后的版本，命令有差别，所以分开举例。

- **7.x 之前的版本**

  a. 创建索引“book”，配置分词策略。

  示例中“analyzer”和“search_analyzer”可以根据实际需要“ik_max_word”或“ik_smart”分词策略，此处以“ik_max_word”为例。

```
PUT /book
{
   "settings": {
      "number_of_shards": 2,
      "number_of_replicas": 1
   },
   "mappings": {
      "type1": {
         "properties": {
            "content": {
               "type": "text",
               "analyzer": "ik_max_word",
               "search_analyzer": "ik_max_word"
            }
         }
      }
   }
}
```

  b. 导入数据，将文本信息导入“book”索引中。

```
PUT /book/type1/1
{
  "content":"智能手机是很好用"
}
```

  c. 使用关键词“智能手机”进行文本搜索，并查看搜索结果。

```
GET /book/type1/_search
{
  "query": {
    "match": {
      "content": "智能手机"
    }
  }
}
```

  搜索结果：

```
{
  "took" : 20,
  "timed_out" : false,
  " shards" : {
    "total" : 2,
    "successful" : 2,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 1.1507283,
    "hits" : [
      {
        "_index" : "book",
```

```
      "_type" : "type1",
      "_id" : "1",
      "_score" : 1.1507283,
      "_source" : {
        "content" : "智能手机是很好用"
      }
    }
  ]
 }
}
```

- **7.x 及之后的版本**

  a. 创建索引“book”，配置分词策略。

  示例中“analyzer”和“search_analyzer”可以根据实际需要“ik_max_word”或“ik_smart”分词策略，此处以“ik_max_word”为例。

```
PUT /book
{
   "settings": {
      "number_of_shards": 2,
      "number_of_replicas": 1
   },
   "mappings": {
      "properties": {
         "content": {
            "type": "text",
            "analyzer": "ik_max_word",
            "search_analyzer": "ik_max_word"
         }
      }
   }
}
```

  b. 导入数据，将文本信息导入“book”索引中。

```
PUT /book/_doc/1
{
 "content":"智能手机是很好用"
}
```

  c. 使用关键词“智能手机”进行文本搜索，并查看搜索结果。

```
GET /book/_doc/_search
{
 "query": {
   "match": {
     "content": "智能手机"
   }
 }
}
```

  搜索结果：

```
{
 "took" : 16,
 "timed_out" : false,
 "_shards" : {
   "total" : 2,
   "successful" : 2,
```

```
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 1.7260926,
    "hits" : [
      {
        "_index" : "book",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 1.7260926,
        "_source" : {
          "content" : "智能手机是很好用"
        }
      }
    ]
  }
}
```

步骤三：使用同义词搜索

Elasticsearch 7.x 之前的版本和之后的版本，命令有差别，所以分开举例。

- **7.x 之前的版本**

  a. 创建索引“myindex”，配置分词策略。

```
PUT myindex
{
  "settings": {
    "analysis": {
      "filter": {
        "my_synonym": {
          "type": "dynamic_synonym"
        }
      },
      "analyzer": {
        "ik_synonym": {
          "filter": [
            "my_synonym"
          ],
          "type": "custom",
          "tokenizer": "ik_smart"
        }
      }
    }
  },
  "mappings": {
    "mytype" :{
      "properties": {
        "desc": {
          "type": "text",
```

```
        "analyzer": "ik_synonym"
       }
      }
    }
  }
}
```

b. 导入数据，将文本信息导入“myindex”索引中。

```
PUT /myindex/mytype/1
{
   "desc": "我今天获奖了我很开心"
}
```

c. 使用同义词“高兴”进行文本搜索，并查看搜索结果。

```
GET /myindex/_search
{
  "query": {
    "match": {
      "desc": "高兴"
    }
  }
}
```

搜索结果：

```
{
  "took" : 2,
  "timed_out" : false,
  "_shards" : {
    "total" : 5,
    "successful" : 5,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 0.49445358,
    "hits" : [
      {
        "_index" : "myindex",
        "_type" : "mytype",
        "_id" : "1",
        "_score" : 0.49445358,
        "_source" : {
          "desc" : "我今天获奖了我很开心"
        }
      }
    ]
  }
}
```

- **7.x 及之后的版本**

a. 创建索引“myindex”，配置分词策略。

```
PUT myindex
{
   "settings": {
      "analysis": {
```

```
            "filter": {
                "my_synonym": {
                    "type": "dynamic_synonym"
                }
            },
            "analyzer": {
                "ik_synonym": {
                    "filter": [
                        "my_synonym"
                    ],
                    "type": "custom",
                    "tokenizer": "ik_smart"
                }
            }
        }
    },
    "mappings": {
        "properties": {
            "desc": {
                "type": "text",
                "analyzer": "ik_synonym"
            }
        }
    }
}
```

b. 导入数据，将文本信息导入“myindex”索引中。

```
PUT /myindex/_doc/1
{
    "desc": "我今天获奖了我很开心"
}
```

c. 使用同义词“高兴”进行文本搜索，并查看搜索结果。

```
GET /myindex/ search
{
  "query": {
    "match": {
      "desc": "高兴"
    }
  }
}
```

搜索结果：

```
{
  "took" : 1,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
```

```
    },
    "max_score" : 0.1519955,
    "hits" : [
      {
        "_index" : "myindex",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 0.1519955,
        "_source" : {
          "desc" : "我今天获奖了我很开心"
        }
      }
    ]
  }
}
```

4.9 简繁体转换搜索（使用简繁分析插件）

默认情况下云搜索服务安装了简繁体转换插件，用户无需自行安装。简繁体转换插件是一款可以使中文简体和中文繁体相互转换的插件。通过该插件的转换，用户可以使用中文繁体关键字搜索出包含对应中文简体的索引数据，也可以使用中文简体关键字搜索出包含对应中文繁体的索引数据。

简繁体转换插件通常可以当做 analyzer、tokenizer、token-filter 或 char-filter 来使用。

简繁体转换插件的转换类型包含如下两种：

- s2t：将中文简体转换为中文繁体。
- t2s：将中文繁体转换为中文简体。

示例指导

1. 登录云搜索服务管理控制台。
2. 在左侧导航栏中，选择“集群管理”，进入集群列表页面。
3. 在集群列表中，单击需要使用的集群对应“操作”列的“Kibana”。
   如果开启了安全模式，需要输入创建集群时设置的管理员账户名和密码。
4. 在 Kibana 的左侧导航中选择“Dev Tools”，进入 Console 界面。
5. 在 Console 界面，执行如下命令，创建索引“stconvert”，并指定自定义映射来定义数据类型。
   低于 7.x 版本

```
PUT /stconvert
{
   "settings": {
      "number_of_shards": 1,
      "number_of_replicas": 0,
      "analysis": {
         "analyzer": {
            "ts_ik": {
               "tokenizer": "ik_smart",
               "char_filter": [
```

```
                    "tsconvert",
                    "stconvert"
                ]
            }
        },
        "char_filter": {
            "tsconvert": {
                "type": "stconvert",
                "convert_type": "t2s"
            },
            "stconvert": {
                "type": "stconvert",
                "convert_type": "s2t"
            }
        }
    },
    "mappings": {
        "type": {
            "properties": {
                "desc": {
                    "type": "text",
                    "analyzer": "ts ik"
                }
            }
        }
    }
}
```

7.x 版本及高于 7.x 版本

```
PUT /stconvert
{
    "settings": {
        "number_of_shards": 1,
        "number_of_replicas": 0,
        "analysis": {
            "analyzer": {
                "ts_ik": {
                    "tokenizer": "ik_smart",
                    "char_filter": [
                        "tsconvert",
                        "stconvert"
                    ]
                }
            },
            "char_filter": {
                "tsconvert": {
                    "type": "stconvert",
                    "convert_type": "t2s"
                },
                "stconvert": {
                    "type": "stconvert",
                    "convert_type": "s2t"
                }
            }
        }
```

```
    },
    "mappings": {
               "properties": {
              "desc": {
                  "type": "text",
                  "analyzer": "ts_ik"
              }
           }
         }
}
```

返回结果如下所示。

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "stconvert"
}
```

6. 在 Console 界面，执行如下命令，导入数据到“stconvert”索引中。

   低于 7.x 版本

   ```
   POST /stconvert/type/1
   {
      "desc": "國際電視臺"
   }
   ```

   7.x 版本及高于 7.x 版本

   ```
   POST /stconvert/_doc/1
   {
      "desc": "國際電視臺"
   }
   ```

   当返回结果信息中“failed”字段的值为“0”时，表示数据导入成功。

7. 在 Console 界面，执行如下命令，搜索关键字“国际”，并查看搜索结果。

   ```
   GET /stconvert/_search
   {
      "query": {
         "match": {
            "desc": "国际"
         }
      }
   }
   ```

   搜索结果如下所示。

   ```
   {
     "took" : 15,
     "timed_out" : false,
     "_shards" : {
       "total" : 1,
       "successful" : 1,
       "skipped" : 0,
       "failed" : 0
     },
     "hits" : {
       "total" : 1,
       "max_score" : 0.5753642,
       "hits" : [
   ```

```
    {
      "_index" : "stconvert",
      "_type" : "type",
      "_id" : "1",
      "_score" : 0.5753642,
      "_source" : {
        "desc" : "國際電視臺"
      }
    }
  ]
 }
}
```

4.10 使用 SQL 编写查询（使用 Open Distro sql 插件）

在 6.5.4 及之后版本中提供 Open Distro for Elasticsearch SQL 插件允许您使用 SQL 而不是 Elasticsearch 查询域特定语言（DSL）编写查询。

如果您已经熟悉 SQL 并且不想学习 DSL 查询，那么此功能是一个很好的选择。

基本操作

- Kibana（推荐）
  - 登录 Kibana，在 DevTools 中将请求发送到_opendistro/_sql，可以使用请求参数或请求正文。

    ```
    POST _opendistro/_sql
    {
      "query": "SELECT * FROM my-index LIMIT 50"
    }
    ```

  - 默认情况下，查询返回 JSON。您也可以选择 CSV 格式返回数据，选择 CSV 格式需要对 format 参数进行如下设置：

    ```
    POST _opendistro/_sql?format=csv
    {
      "query": "SELECT * FROM my-index LIMIT 50"
    }
    ```

    CSV 格式返回数据时，每行对应一个文档，每列对应一个字段。
- curl 命令

  您也可以在 ECS 中使用 curl 命令，来执行该 SQL 操作。

  ```
  curl -XPOST https://localhost:9200/_opendistro/_sql -u username:password -k -d '{"query": "SELECT * FROM kibana_sample_data_flights LIMIT 10"}' -H 'Content-Type: application/json'
  ```

支持操作

支持的 SQL 操作包括声明、条件、聚合函数、Include 和 Exclude、常用函数、连接 join 和展示等操作。

- 声明 statements

表4-43 声明 statements

Statement	Example
Select	SELECT * FROM my-index
Delete	DELETE FROM my-index WHERE _id=1
Where	SELECT * FROM my-index WHERE ['field']='value'
Order by	SELECT * FROM my-index ORDER BY _id asc
Group by	SELECT * FROM my-index GROUP BY range(age, 20,30,39)
Limit	SELECT * FROM my-index LIMIT 50 (default is 200)
Union	SELECT * FROM my-index1 UNION SELECT * FROM my-index2
Minus	SELECT * FROM my-index1 MINUS SELECT * FROM my-index2

说明

与任何复杂查询一样，大型 UNION 和 MINUS 语句可能会使集群资源紧张甚至崩溃。

- 条件 Conditions

表4-44 条件 Conditions

Condition	Example
Like	SELECT * FROM my-index WHERE name LIKE 'j%'
And	SELECT * FROM my-index WHERE name LIKE 'j%' AND age > 21
Or	SELECT * FROM my-index WHERE name LIKE 'j%' OR age > 21
Count distinct	SELECT count(distinct age) FROM my-index
In	SELECT * FROM my-index WHERE name IN ('alejandro', 'carolina')
Not	SELECT * FROM my-index WHERE name NOT IN ('jane')
Between	SELECT * FROM my-index WHERE age BETWEEN 20 AND 30
Aliases	SELECT avg(age) AS Average_Age FROM my-index
Date	SELECT * FROM my-index WHERE birthday='1990-11-15'
Null	SELECT * FROM my-index WHERE name IS NULL

- 聚合函数 Aggregation

表4-45 聚合函数 Aggregation

Aggregation	Example
avg()	SELECT avg(age) FROM my-index
count()	SELECT count(age) FROM my-index
max()	SELECT max(age) AS Highest_Age FROM my-index
min()	SELECT min(age) AS Lowest_Age FROM my-index
sum()	SELECT sum(age) AS Age_Sum FROM my-index

- Include 和 Exclude 字段

表4-46 Include 和 Exclude

Pattern	Example
include()	SELECT include('a*'), exclude('age') FROM my-index
exclude()	SELECT exclude('*name') FROM my-index

- 函数 Functions

表4-47 函数 Functions

Function	Example
floor	SELECT floor(number) AS Rounded_Down FROM my-index
trim	SELECT trim(name) FROM my-index
log	SELECT log(number) FROM my-index
log10	SELECT log10(number) FROM my-index
substring	SELECT substring(name, 2,5) FROM my-index
round	SELECT round(number) FROM my-index
sqrt	SELECT sqrt(number) FROM my-index
concat_ws	SELECT concat_ws(' ', age, height) AS combined FROM my-index
/	SELECT number / 100 FROM my-index
%	SELECT number % 100 FROM my-index

Function	Example
date_format	SELECT date_format(date, 'Y') FROM my-index

说明

必须在文档映射中启用 fielddata 才能使大多数字符串函数正常工作。

- 连接操作 Joins

表4-48 连接操作 Joins

Join	Example
Inner join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s JOIN school sc ON sc.name = s.school_name WHERE s.age > 20
Left outer join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s LEFT JOIN school sc ON sc.name = s.school_name
Cross join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s CROSS JOIN school sc

相关约束和限制，参考连接操作 Joins。

- 展示 Show

展示 show 操作与索引模式匹配的索引和映射。您可以使用*或%使用通配符。

表4-49 展示 show

Show	Example
Show tables like	SHOW TABLES LIKE logs-*

连接操作 Joins

Open Distro for Elasticsearch SQL 支持 inner joins, left outer joins,和 cross joins。Join 操作有许多约束：

- 您只能加入两个参数。
- 您必须为索引使用别名（例如 people p）。
- 在 ON 子句中，您只能使用 AND 条件。
- 在 WHERE 语句中，不要将包含多个索引的树组合在一起。例如，以下语句有效：

```
WHERE (a.type1 > 3 OR a.type1 < 0) AND (b.type2 > 4 OR b.type2 < -1)
```

以下声明无效：

```
WHERE (a.type1 > 3 OR b.type2 < 0) AND (a.type1 > 4 OR b.type2 < -1)
```

- 您不能使用 GROUP BY 或 ORDER BY 来获得结果。
- LIMIT 和 OFFSET 不支持一起使用（例如 LIMIT 25 OFFSET 25）。

JDBC 驱动

Java 数据库连接（JDBC）驱动程序允许您将 Open Distro for Elasticsearch 与您的商业智能（BI）应用程序集成。

有关下载和使用 JAR 文件的信息，请参阅 GitHub 仓库。

4.11 切换冷热数据

云搜索服务提供了冷数据节点供企业选择，企业可以将部分现查要求秒级返回的数据放在高性能机器上面，对于历史数据要求分钟级别返回的数据放在大容量低规格节点。

说明

- 创建集群时，数据节点为必选，只有当选择了冷数据节点后，数据节点才会变成热节点。
- 选择冷数据节点的同时，支持独立选择 Master 和 Client 节点。
- 冷数据节点支持节点和磁盘扩容，前提是冷节点规格支持（本地盘不支持磁盘扩容）。

冷热数据切换

用户在创建集群的时候如果选择启用冷数据节点，冷数据节点将会打上“cold”标签，用来表示冷节点，同时其他数据节点将会上升为热节点，会被打上“hot”标签。用户可以通过配置指定索引，将数据分配到冷热节点。

通过设置 template，可以通过模板将相应的 index 存储到指定冷热节点。

如下，登录集群的 Kibana Console 页面，配置 myindex 开头的索引，储存在冷节点上面。 这样可以通过模板在创建的时候把 myindex*的数据存储在冷数据节点上面。

- 5.x 版本使用以下命令创建模板：

```
PUT _template/test
{
   "order": 1,
   "template": "myindex*",
   "settings": {
      "index": {
         "refresh_interval": "30s",
         "number_of_shards": "3",
         "number_of_replicas": "1",
         "routing.allocation.require.box_type": "cold"
      }
```

```
    }
}
```

- 6.x 及以上版本使用以下命令创建模板：

```
PUT _template/test
{
  "order": 1,
  "index_patterns": "myindex*",
  "settings": {
    "refresh_interval": "30s",
    "number_of_shards": "3",
    "number_of_replicas": "1",
    "routing.allocation.require.box_type": "cold"
  }
}
```

同时也可以单独对已经创建好的索引进行操作。

```
PUT myindex/_settings
 {
       "index.routing.allocation.require.box_type": "cold"
   }
```

也可以去掉冷热数据配置，不受冷热数据标签影响。

```
PUT myindex/_settings
{
       "index.routing.allocation.require.box_type": null
   }
```

4.12 管理索引

4.12.1 创建及管理索引

Elasticsearch 7.6.2 及以上版本的集群支持索引状态管理。索引状态管理（ISM）是一个插件，通过该插件，您可以根据索引使用期限，索引大小或文档数的变化触发这些定期的管理操作，从而使它们自动化。使用 ISM 插件时，您可以根据需要定义自动处理索引滚动或删除的策略。

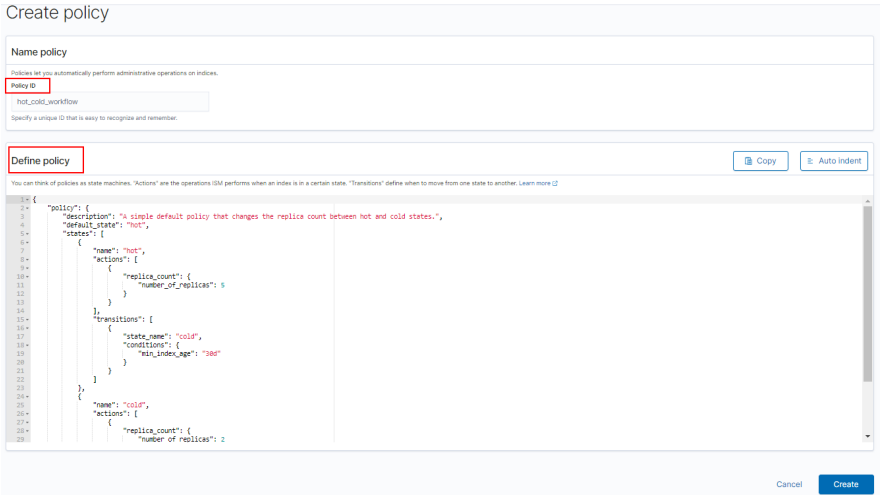
说明

如下操作步骤，以 7.6.2Elasticsearch 版本为例，不同版本的 Kibana 界面有些差别，但是操作类似。

创建索引策略

1. 登录 Kibana，在左侧选择“IM”或“Index Management”，进入索引管理页面 。
2. 右侧单击 **Create policy**，创建索引策略。
3. 在 **Policy ID** 部分输入策略 ID，**Define policy** 部分输入您的策略。

图4-26 配置策略



4. 单击 **Create** 完成索引策略的创建。

将策略附加到索引

创建索引策略后，可以将此策略附加到一个或多个索引，匹配该索引模板创建出的索引都将被附加该策略。

- **方式 1：Kibana 命令行**

  在 Kibana 的“Dev Tools”页面，执行如下命令在索引模板中关联策略 ID。

```
PUT _template/<template_name>
{
    "index_patterns": ["index_name-*"],
    "settings": {
        "opendistro.index_state_management.policy_id": "policy_id"
    }
}
```

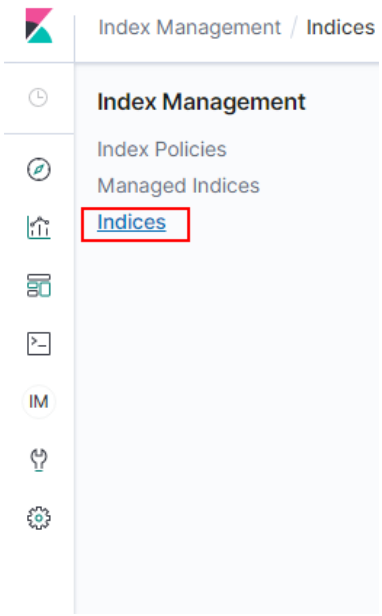
  - <template_name>：需要替换为创建的索引模板名。
  - policy_id：需要替换为自定义的策略 ID，即 Policy ID。

  更多创建索引模板的说明可参考索引模板。

- **方式 2：Kibana 控制台**

  a. 在 Kibana“Index Management”页面，选择 **Indices**。

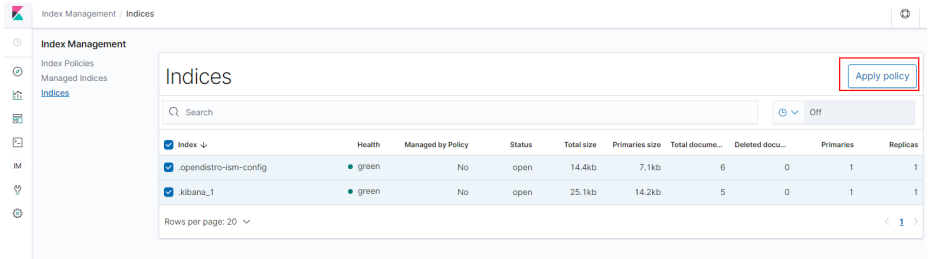
图4-27 选择 Indices



b. 在 **Indices** 列表中选择您要附加策略的一个或多个索引。

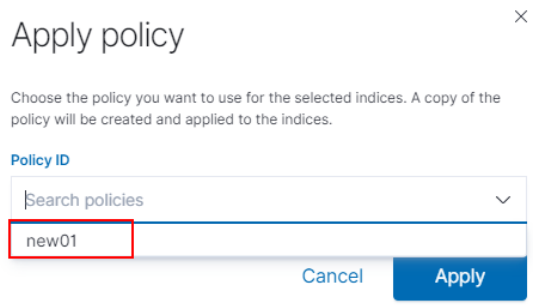
c. 单击右上角的 **Apply policy**，添加应用策略。

图4-28 添加引用策略



d. 从 **Policy ID** 菜单中，选择您创建的策略。

图4-29 选择



e. 单击 **Apply**。

将策略附加到索引后，ISM 会默认创建每 5 分钟运行一次的作业，以执行策略操作，检查条件并将索引转换为不同的状态。

管理索引策略

1. 选择 **Managed Indices**。
2. 如果您要更改策略，可以选择 **Change policy**，详情请参考 **4.12.2** 变更策略。
3. 如果您要删除策略，请选择您的策略，然后选择 **Remove policy**。
4. 如果您要重试策略，请选择您的策略，然后选择 **Retry policy**。

具体使用可参考索引管理官方介绍。

4.12.2 变更策略

您可以更改任何托管索引策略，但是 ISM 有一些约束条件可以确保策略更改不会破坏索引。

如果索引卡在其当前状态，永不进行，并且您想立即更新其策略，请确保新策略包括与旧策略相同的状态（名称，操作，顺序相同）。在这种情况下，即使策略处于执行操作中，ISM 也会应用新策略。

如果在不包含相同状态的情况下更新策略，则 ISM 仅在当前状态下的所有操作执行完成后才更新策略。或者，您可以在旧策略中选择特定状态，然后让新策略生效。

在 Kibana 中更改策略，操作步骤如下：

1. 在 **Managed indices** 下，选择需要更换新策略的索引。
2. 单击右上角的 **Change policy**，进入 **Choose managed indices** 页面，选择更换新策略的相关信息 。

表4-50 更换索引策略参数信息

参数	说明
Managed indices	选择需要更换策略的索引名称。支持选择多个索引。
State filters	选择索引状态。选择后，会将新策略附加到处于特定状态的索引。
New policy	选择新策略。

3. 选择完成后，单击 **Change**。

4.13 Kibana 可视化平台

4.13.1 登录 Kibana

创建 CSS 集群后，可通过控制台和公网访问两种方式登录 Kibana。

Kibana 使用限制

- Kibana 中可以自定义用户名、角色名、租户名等，不能包含中文字符。
- Kibana 不支持中文。

操作步骤

- 通过控制台访问方式登录
  a. 登录云搜索服务管理控制台。
  b. 在“集群管理”页面选择需要登录的集群，单击“操作”列中的“Kibana”进入 Kibana 登录界面。
     - 非安全模式的集群：将直接进入 Kibana 操作界面。
     - 安全模式的集群：需要在登录页面输入用户名和密码，单击“Log In”进入 Kibana 操作界面。用户名默认为 admin，密码为创建集群时设置的管理员密码。
  c. 登录成功后，可在 Kibana 界面进行相关操作访问 Elasticsearch 集群。

4.13.2 Kibana 公网访问集群

针对安全模式集群，云搜索服务支持配置 Kibana 开启公网访问，配置完成后，对应集群将会获得一个 Kibana 公网访问地址，通过这个地址可以在公网上面访问集群的 Kibana。

对于安全模式集群来说，支持在创建的时候配置 Kibana 公网访问，同时也支持安全模式集群创建完之后再开启 Kibana 公网访问。

说明

- 6.5.4 及之后版本的 Elasticsearch 集群支持开启“安全模式”。
- 在该特性上线之前（即 2020 年 6 月前）创建的 Elasticsearch 安全模式的集群，不支持此功能。
- Kibana 公网访问配置的白名单依赖 ELB 的白名单能力。更新白名单后，白名单对新建的连接是实时生效的，但对于已存在的长连接，可能会出现去掉的白名单 IP 地址还能访问 Kibana 的场景，这是因为要等长连接断开后才生效，预计 1 分钟左右。

创建集群时配置 Kibana 公网访问

1. 登录云搜索服务管理控制台。
2. 单击右上角的“创建集群”，进入创建集群页面。
3. 在创建集群页面，开启“安全模式”。
4. “高级配置”选择“自定义”后，开启 Kibana 公网访问，配置相关参数。

表4-51 Kibana 公网访问参数说明

参数	说明
带宽	设置公网访问的带宽。 取值范围：1-100。 单位：Mbit/s。
访问控制开关	如果关闭访问控制开关，则允许任何 IP 通过公网 IP 访问集群 Kibana。如果开启访问控制开关，则只允许白名单列表中的 IP 通过公网 IP 访问集群 Kibana。
白名单	设置允许访问的 IP 地址或网段，中间用英文逗号隔开。仅当打开“访问控制开关”时才需要配置。 建议开启白名单。

集群创建成功后，单击集群名称，进入集群基本信息页面，在“Kibana 公网访问”页签，可以查看 kibana 公网访问地址。

已有集群开启 Kibana 公网访问

您可以对已经创建的安全模式集群的 Kibana 公网访问进行开启、关闭、修改、查看等操作。

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击需要配置 Kibana 公网访问的集群名称，进入集群基本信息页面。
3. 选择“Kibana 公网访问”，在“Kibana 公网访问”右侧单击开关，打开 Kibana 公网访问功能。
4. 在开启 Kibana 公网访问页面，配置相关参数。

表4-52 Kibana 公网访问参数说明

参数	说明
带宽	设置公网访问的带宽。 取值范围：1-100。 单位：Mbit/s。
访问控制开关	如果关闭访问控制开关，则允许任何 IP 通过公网 IP 访问集群 Kibana。如果开启访问控制开关，则只允许白名单列表中的 IP 通过公网 IP 访问集群 Kibana。
白名单	设置允许访问的 IP 地址或网段，中间用英文逗号隔开。仅当打开“访问控制开关”时才需要配置。 建议开启白名单。

5. 配置完成后，单击“确定”。

修改 Kibana 公网访问

对已经配置了 Kibana 公网访问的集群，云搜索服务支持修改带宽、修改访问控制和关闭 Kibana 公网访问。

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击需要修改 Kibana 公网访问的集群名称，进入集群基本信息页面。
3. 选择“Kibana 公网访问”，修改 Kibana 公网访问。
   - 修改带宽

     单击“带宽”参数右侧的“修改”，在“修改 Kibana 公网访问带宽”页面修改带宽大小，修改完成后，单击“确定”。
   - 修改访问控制

     单击“访问控制开关”右侧的“修改”，在“修改 Kibana 公网访问控制”页面设置“访问控制开关”和访问“白名单”，修改完成后，单击“确定”。
   - 关闭 Kibana 公网访问

     在“Kibana 公网访问”右侧单击开关，确认关闭 Kibana 公网访问功能。

通过公网 IP 访问 Kibana

Kibana 公网访问配置完成后，将会获得一个 Kibana 公网访问地址，用户可以通过此 IP 地址访问集群的 Kibana。

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，单击需要配置 Kibana 公网访问的集群名称，进入集群基本信息页面。
3. 选择“Kibana 公网访问”，获取 kibana 公网访问地址。
4. 通过该地址，就可以在公网上面访问云搜索服务集群的 Kibana。

4.13.3 使用 Kibana 创建用户并授权

云搜索服务（Cloud Search Service，简称 CSS）用 opendistro_security 安全插件对外提供安全集群能力，opendistro_security 安全插件是基于 RBAC（Role-Based Access Control）模型构建。RBAC 包括三个重要核心概念：用户（User）、权限（Action）、角色（Role）。RBAC 简化了用户和权限的关系，降低了权限管理的难度，方便权限扩展易于维护。三者之前的关系如下图所示：

图4-30 用户、权限和角色

表4-53 参数说明

参数	描述
用户 User	用户可以向 Elasticsearch 集群发出操作请求。用户具有凭证（例如，用户名和密码）、零个或多个后端角色以及零个或多个自定义属性。
角色 Role	角色定义为权限或操作组的组合，包括对集群，索引，文档或字段的操作权限。
权限 Permission	单个动作，例如创建索引（例如 indices:admin/create）。
角色映射 Role mapping	用户在成功进行身份验证后会担任角色。角色映射，就是将角色映射到用户（或后端角色）。例如，kibana_user（角色）到 jdoe（用户）的映射意味着 John Doe 在获得 kibana_user 身份验证后获得了所有权限。同样，all_access（角色）到 admin（后端角色）的映射意味着具有后端角色 admin（来自 LDAP / Active Directory 服务器）的任何用户都获得了 all_access 身份验证后的所有权限。您可以将每个角色映射到许多用户和/或后端角色。
操作组 Action group	一组权限。例如，预定义的 SEARCH 操作组授权角色使用 _search 和_msearchAPI。

除了 RBAC 模型之外，Elasticsearch 还有一个重要的概念，叫做 Tenant。RBAC 能解决各个用户本身授权的问题，Tenant 则能解决了不同租户之间的共享信息，通过配置 Tenant 空间，各个 IAM 用户（子用户）可以在 Tenant 空间中共享 Dashboard、index_pattern 等信息。

本章节将介绍如何使用 Kibana 创建用户，并为用户授权。集群必须开启安全模式才支持使用 Kibana 创建用户并授权。

说明

- 不同的版本之间 Kibana 界面有差异，本章节以 7.6.2 版本为例。
- Kibana 中可以自定义用户名、角色名、租户名等，但不能包含中文字符。

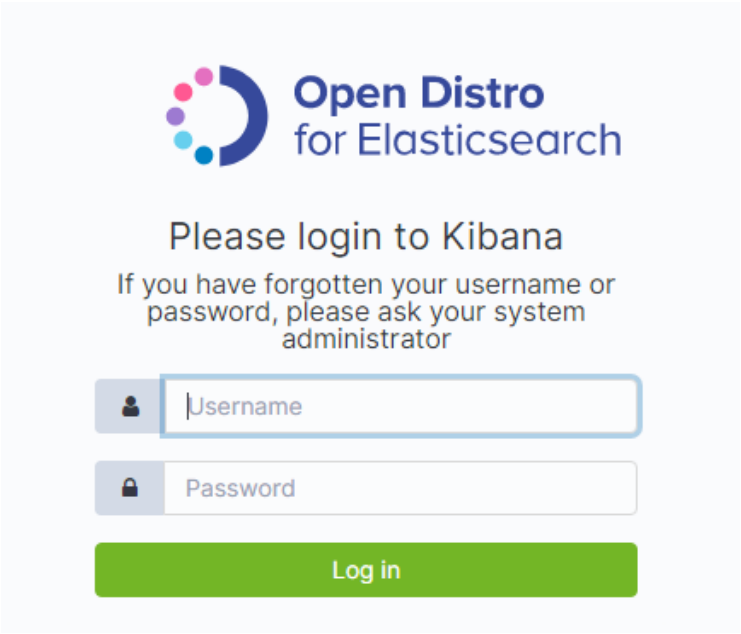
- 步骤一：登录 Kibana 界面
- 步骤二：创建用户
- 步骤三：创建角色并授权
- 步骤四：为用户配置角色

登录 Kibana 界面

1. 登录云搜索服务控制台。
2. 在集群管理列表，选择对应集群，单击操作列的“Kibana”。
   输入管理员账户名和密码登录 Kibana。
   - 账户名：admin（默认管理员账户名）

- 密码：创建安全模式的集群时，设置的管理员密码。

图4-31 登录页面

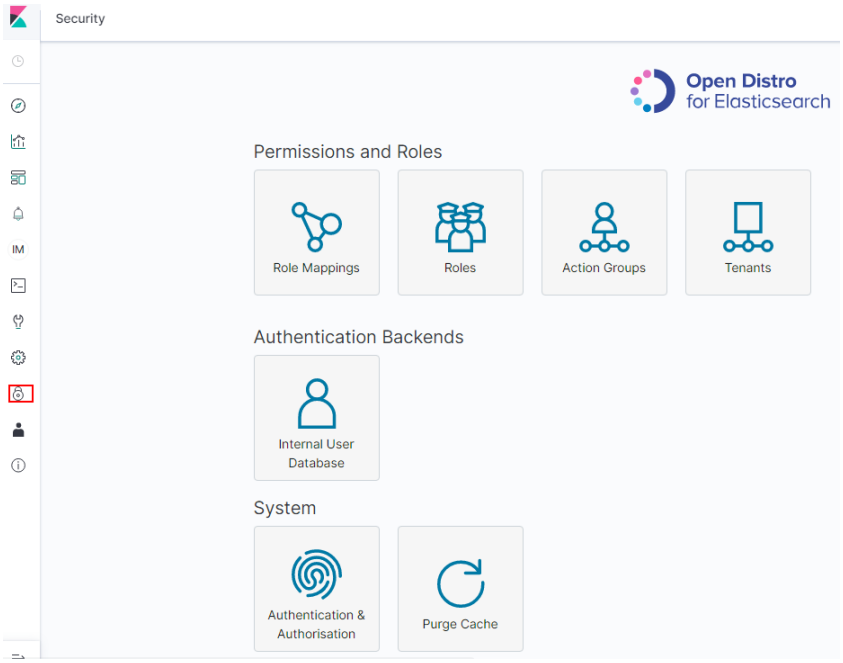


创建用户

登录 Kibana 操作界面后，在 Security 页面创建用户。

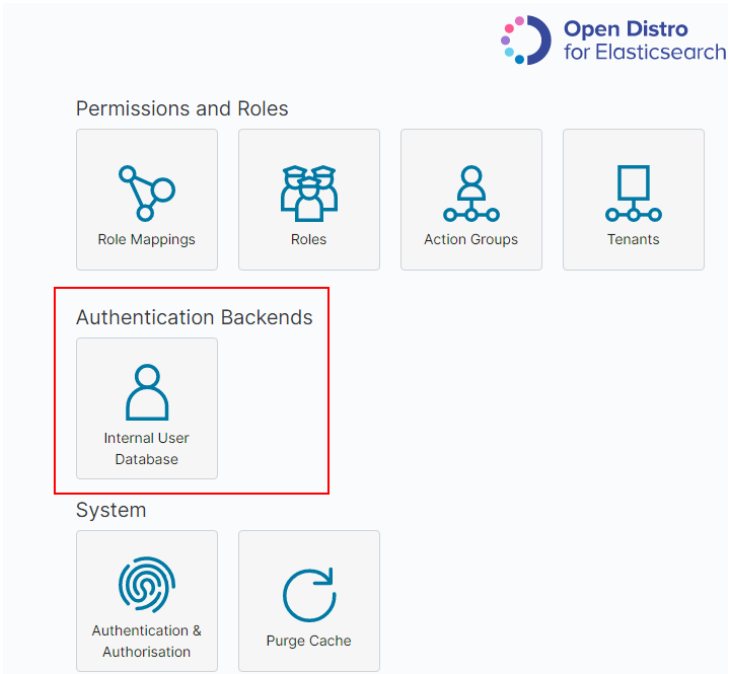
1. 登录成功后，在 Kibana 操作界面的左侧导航栏选择“Security”，进入“Security”页面。

图4-32 进入 Security 页面



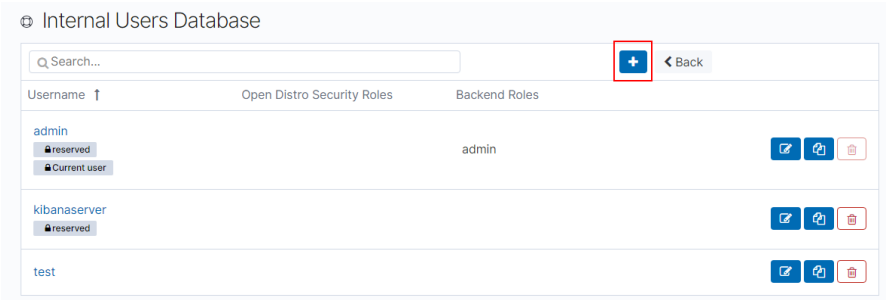
2. 选择“Authentication Backends”>“Internal Users Database”，进入创建用户页面。

图4-33 添加用户（1）



3. 在“Internal Users Database”页面，选择  ，进入添加用户信息页面。

图4-34 添加用户（2）



4. 在创建用户页面，输入“Username”、“Password”和“Repeatpassword”，单击“Submit”。

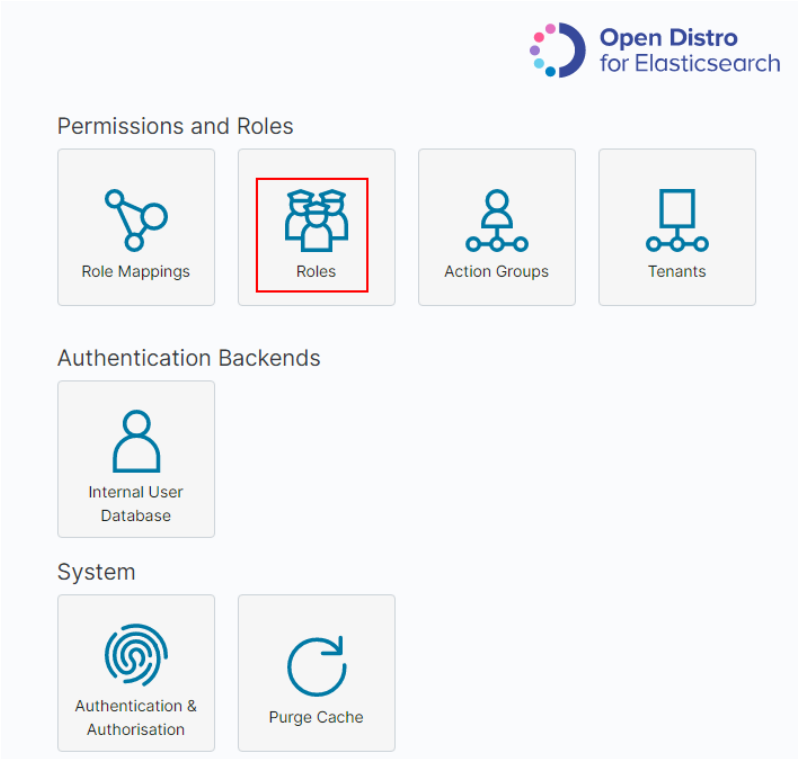
创建成功后，可以在列表中看到新创建的用户。

创建角色并授权

在 Security 页面创建角色，并为角色授权。

1. 在“Security”中选择“Roles”，进入 Open Distro Security Roles 页面。

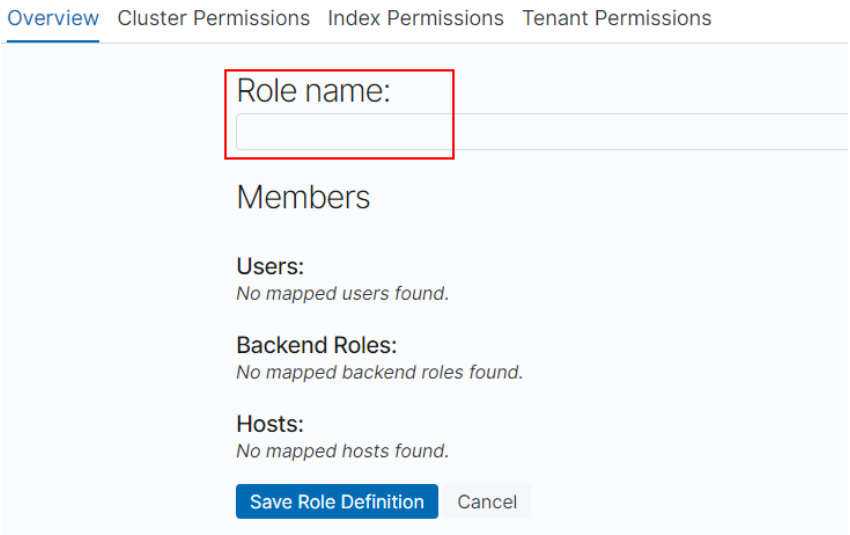
图4-35 添加角色



2. 在 Open Distro Security Roles 页面，单击  添加角色权限。

a. 在 Overview 页签设置角色名“Role name”。

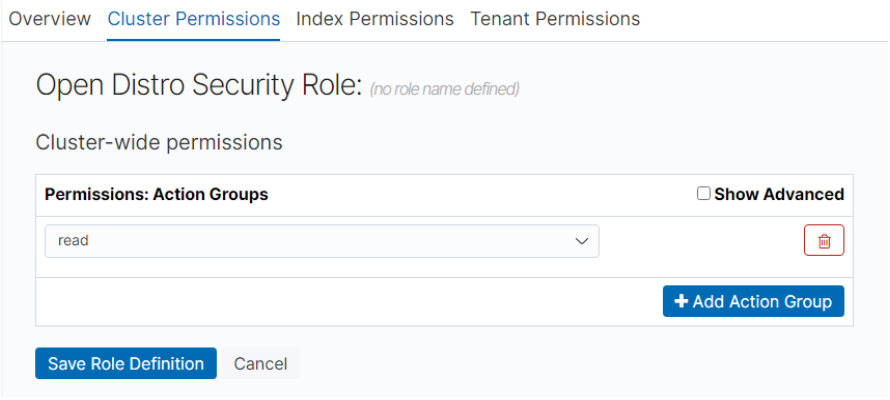
图4-36 添加角色名称



b. 在“Cluster Permissions”页签设置 CSS 集群权限。根据业务需要选择相应的集群权限，不配置时表示角色不具有任何集群级别的权限。

- “Permissions: Action Groups”：单击“Add Action Group”，可以设置集群权限。例如，集群只读权限选择 read，表示仅可查看集群状态、集群节点等信息。
- “Permissions: Single Permissions”：勾选 Show Advanced 后，单击“Add Single Permission”可以针对集群设置更精细的权限。例如设置为 indices:data/read，表示仅指定索引的读取权限。

图4-37 Cluster Permissions 页面



c. 在“Index Permissions”页面设置索引权限。

- “Index patterns”：单击“Add index permissions”配置为需要设置权限的索引名称，例如，索引模板名称为 my_store。

说明

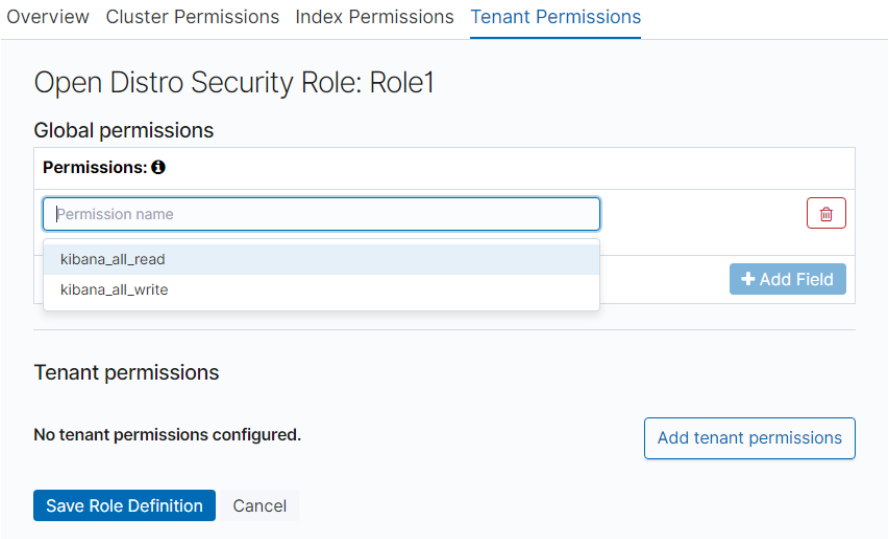
建议索引名称和创建的用户名不要相同。

- “Permissions: Action Groups”：单击“Add Action Group”，根据需要开通的权限设置。例如，只读权限选择 Search。

d. “Tenant Permissions”页面设置角色权限，根据业务需要配置。

- “Global permissions”：单击“Add Field”，设置角色的 kibana 读写权限，例如 kibana_all_read 或 kibana_all_write 权限。
- “Tenant permissions”：单击“Add tenant pattern”，自行添加租户模式，并且为新增的租户模式设置 kibana_all_read 或 kibana_all_write 权限。

图4-38 Tenant Permissions 页面



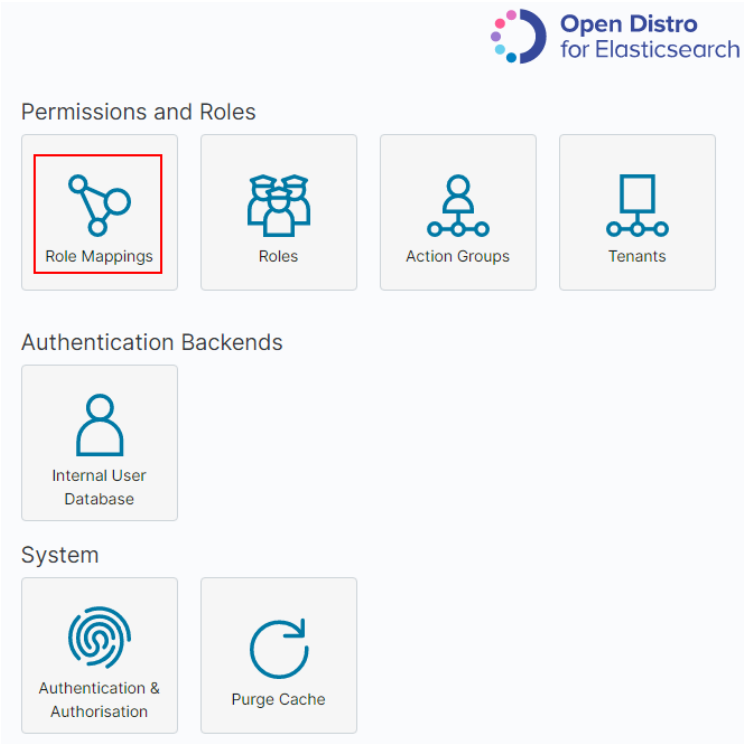
3. 单击“Save Role Definition”，保存角色设置，即可看到设置的角色。

为用户配置角色

创建好角色并为角色授予权限后，需要将角色映射到用户，用户才会获得被映射的角色所拥有的权限。

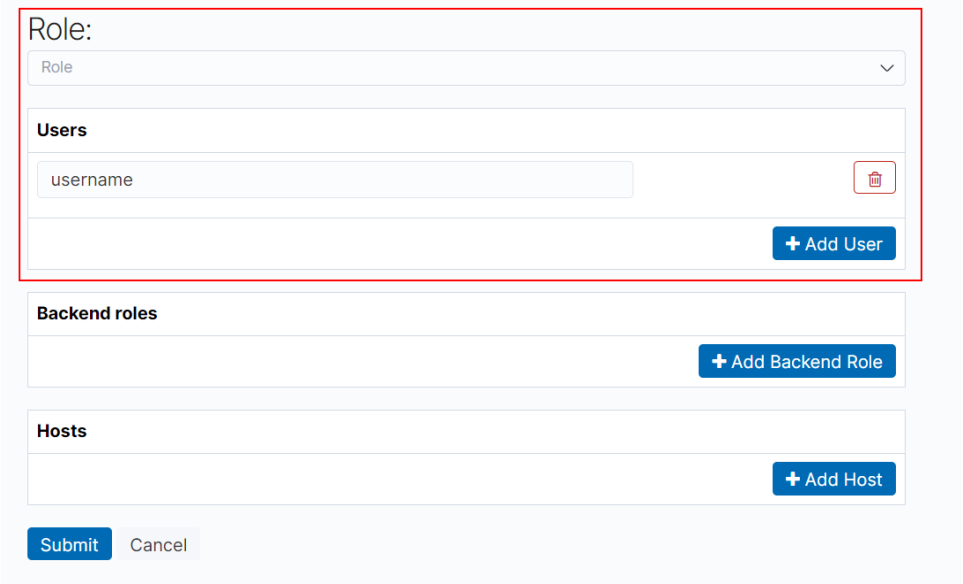
1. 在“Security”中选择“Role Mappings”，进入 Role Mappings 页面，进行角色映射。

图4-39 角色映射



2. 在 Role Mappings 页面，单击  ，选择角色“Role”，添加用户“Users”。
   - “Role”：选择待映射的角色名称。
   - “Users”：单击“Add User”，输入被映射角色的用户名称。

图4-40 用户和角色映射



3. 添加完成后，单击“Submit”。
4. 配置完成后，可以在 Kibana 中进行验证是否生效。

4.13.4 自建 Kibana 接入 Elasticsearch 集群

自建 Kibana 对接云搜索服务 Elasticsearch 集群，需满足如下条件：

- 本地环境需要支持外网访问。
- 通过同 vpc 下 ECS 服务搭建 Kibana，本地公网访问 Kibana 即可。
- 只支持 OSS 版本的 Kibana 镜像连接到云搜索服务的 Elasticsearch。

Kibana 配置文件参考：

- 安全模式：

```
elasticsearch.username: "***"
elasticsearch.password: "***"
elasticsearch.ssl.verificationMode: none
server.ssl.enabled: false
server.rewriteBasePath: false
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: https://10.0.0.xxx:9200
elasticsearch.requestHeadersWhitelist: ["securitytenant","Authorization"]
opendistro_security.multitenancy.enabled: true
opendistro_security.multitenancy.tenants.enable_global: true
opendistro_security.multitenancy.tenants.enable_private: true
opendistro_security.multitenancy.tenants.preferred: ["Private", "Global"]
opendistro_security.multitenancy.enable_filter: false
```

说明

- 安全模式需要安装插件 opendistro_security_kibana，详细请参考 https://github.com/opendistro-for-elasticsearch/security-kibana-plugin/tags?after=v1.3.0.0。
- 安装的插件版本需要和集群版本保持一致，可通过 **GET _cat/plugins** 获取到集群安全插件的版本号。

- 非安全模式：

```
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: http://10.0.0.xxx:9200
```

5 Logstash

5.1 创建集群

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“总览”页面单击右上角的“创建集群”，进入“创建集群”页面。
   或者左侧导航栏单击“集群管理>Logstash”，单击右上角的“创建集群”，进入“创建集群”页面。
3. 单击“创建集群”，进入“创建集群”页面。
4. 选择“计费模式”和“订购周期”。

表5-1 计费模式参数说明

参数	说明
计费模式	集群支持包年/包月和按需计费两种模式。 • 包年/包月：根据集群购买时长，一次性支付集群费用。最短时长为 1 个月，最长时长为 3 年。如果购买时长超过 9 个月，建议包年购买，价格更优惠。一年计费为购买 10 个月得 12 个月。 • 按需计费：按实际使用时长计费，计费周期为一小时，不足一小时按一小时计费。
订购周期	选择包年/包月模式后，需要选择购买时长。 您可以根据需求，选择是否需要自动续费。

5. 选择“当前区域”和“可用区”。

表5-2 区域和可用区参数说明

参数	说明
当前区域	集群工作区域在右侧下拉框中选择。
可用区	选择集群工作区域下关联的可用区。 最多支持配置 3 个“可用区”，多可用区的使用请参见 4.1.4 部署跨 AZ 集群。

6. 指定集群基本信息，选择“集群版本”，并输入“集群名称”。

表5-3 基本参数说明

参数	说明
集群类型	此处选择 Logstash 类型。 • Elasticsearch 类型用于创建 Elasticsearch 集群。 • OpenSearch 类型用于创建 OpenSearch 集群。
集群版本	当前支持 7.10.0。
集群名称	自定义的集群名称，可输入的字符范围为 4～32 个字符，只能包含数字、字母、中划线和下划线，且必须以字母开头。 说明 当集群创建成功后，您可以根据需求修改集群名称。单击需要修改的集群名称，进入集群基本信息页面，单击“集群名称”后面的，修改完成后，单击，进行保存。如果需要取消修改，可单击进行取消。

7. 指定集群的主机规格相关参数。

表5-4 参数说明

参数	说明
节点数量	集群中的节点个数。
CPU 架构	目前 Logstash 集群支持“x86 计算”类型。具体支持类型由实际区域环境决定。
节点规格	集群中的节点规格。 您可以选择任一系列，然后从对应系列中根据需要选择一个规格。
节点存储	目前只支持普通 I/O、高 I/O、超高 I/O。

参数	说明
	说明 如果当前局点的存储类型不支持，则界面不显示。
节点存储容量	存储空间大小。

8. 设置集群的企业项目。

   如果您开通了“企业项目”，在创建集群时，可以给集群绑定一个企业项目。您可以在右侧下拉框中选择当前用户下已创建的企业项目，也可以通过单击“查看项目管理”按钮，前往“企业项目管理”管理控制台，新建企业项目和查看已有的企业项目。

9. 单击“下一步，网络配置”，设置集群的网络配置。

表5-5 参数说明

参数	说明
虚拟私有云	VPC 即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。 选择创建集群需要的 VPC，单击“查看虚拟私有云”进入 VPC 服务查看已创建的 VPC 名称和 ID。如果没有 VPC，需要创建一个新的 VPC。 说明 此处您选择的 VPC 必须包含网段（CIDR），否则集群将无法创建成功。新建的 VPC 默认包含网段（CIDR）。
子网	通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。 选择创建集群需要的子网，可进入 VPC 服务查看 VPC 下已创建的子网名称和 ID。
安全组	安全组是一个逻辑上的分组，为同一个 VPC 内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。单击“查看安全组”可了解安全组详情。 说明 请确保安全组的“端口范围/ICMP 类型”为“Any”或者包含端口 9200 的端口范围。

10. 单击下一步：确认配置，确认完成后单击“立即创建”开始创建集群。
11. 单击“返回集群列表”，系统将跳转到“集群管理”页面。您创建的集群将展现在集群列表中，且集群状态为“创建中”，创建成功后集群状态会变为“可用”。

    如果集群创建失败，请根据界面提示，重新创建集群。

5.2 配置集群

5.2.1 配置中心

Logstash 类型的集群支持通过配置中心，修改 logstash 的配置文件，从不同的数据源（input）迁移数据到不同的目的端（output）。

连通性测试

在使用 Logstash 集群迁移数据时，可以先测试下数据源和 Logstash 集群的网络是否连通。

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择 Logstash 类型集群，单击需要配置数据导入导出文件的集群名称，进入集群基本信息页面，选择“配置中心”，或者直接单击目标集群操作列的“配置中心”，进入配置中心页面。
3. 在配置中心页面，单击“连通性测试”。
4. 在连通性测试弹窗中输入数据来源的 IP 地址和端口号，单击“测试”。

说明

连通性测试最多可一次性测试 10 个 IP 地址。您可以单击“继续添加”，添加多个 IP 地址，然后单击“批量测试”，进行一次性测试多个 IP 地址的连通性。

创建配置文件

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择 Logstash 类型集群，单击需要配置数据导入导出文件的集群名称，进入集群基本信息页面，选择“配置中心”页签，进入配置中心页面；或者直接单击目标集群操作列的“配置中心”，进入配置中心页面。
3. 单击右上角“创建”，进入创建配置文件页面。

   您可以选择系统模板或者自定义模板方式创建，也可以直接进行创建配置文件。

   - 如果选择模板方式，可以直接单击对应的模板操作列的“应用”，然后在“名称”、“配置文件内容”和“隐藏内容列表”中进行命名和修改。

     目前支持的系统模板类型有：

     elasticsearch：从 Elasticsearch 类型集群导入数据到 Elasticsearch 类型集群。

     各个模板的参数配置请参考 5.2.3 系统模板配置参数说明。
   - 如果直接创建配置文件，在“名称”和“配置文件内容”参数中直接输入对应内容即可。创建的配置文件内容大小不能超过 100k。支持创建配置文件个数不超过 50 个。
   - 隐藏内容列表：输入需要隐藏的敏感字串列表，按 Enter 创建；配置隐藏字符串列表后，在返回的配置内容中，会将所有在列表中的字串隐藏为***（列表最大支持 20 条，单个字串最大长度 512 字节）。
4. 配置完成后，单击“下一页”，配置参数。

   配置文件在迁移数据时管道中的配置。

表5-6 参数说明

参数	说明
pipeline.workers	并行执行管道的 Filters+Outputs 阶段的工作线程数，默认值为 CPU 核数，建议取值为 1-20 之间。
pipeline.batch.size	单个工作线程在尝试执行其 Filters 和 Outputs 之前将从 inputs 收集的最大事件数，该值较大通常更有效，但会增加内存开销，默认为 125。
pipeline.batch.delay	创建管道事件批时，在将过小的批调度到管道工作线程之前，等待每个事件的时间（以毫秒为单位），默认值为 50。
queue.type	用于事件缓冲的内部队列模型。memory 为基于内存的传统队列，persisted 为基于磁盘的 ACKed 持久化队列，默认值为 memory。
queue.checkpoint.writes	如果使用持久化队列，则表示强制执行检查点之前写入的最大事件数，默认值为 1024。
queue.max_bytes	如果使用持久化队列，则表示持久化队列的总容量，确保磁盘的容量大于该值，默认值为 1024。 单位：MB。

5. 配置完成后，单击“创建”。

   在配置中心页面可以看到创建的配置文件，状态为“可用”，表示创建成功。您还可以在操作列对创建的配置文件进行编辑、添加到自定义模板、删除等操作。

   - 编辑：单击操作列的“编辑”，可以修改配置文件的内容及配置参数。
   - 添加到自定义模板：可以将当前创建的配置文件，作为模板添加到自定义模板中，方便下次创建配置文件时使用。
   - 删除：如果不需要此配置文件，可以通过操作列进行删除。

说明

您也可以单击“操作记录”或“运行日志”，查看配置文件的相关操作记录和运行日志信息。

启动配置文件

配置文件创建完成后，在配置中心页面可以看到创建成功的配置文件。

1. 选择需要启动的配置文件，单击左上角的“启动”。

   配置文件可以同时选择多个进行启动，一次不能超过 50 个。

2. 在“启动 Logstash 服务”对话框中，根据业务需要，选择“是否保持常驻”。

   开启“保持常驻”适用于需要长期运行的业务，开启“保持常驻”以后，将会在每个节点上面配置一个守护进程，当 logstash 服务出现故障的时候，会主动拉起并修复。“保持常驻”不适用于短期运行的业务，短期业务开启保持常驻，如果源端无数据，会导致任务失败。

3. 单击“确定”，开始启动配置文件。
   可以在管道列表看到启动的配置文件。

说明

您也可以单击“操作记录”或“运行日志”，查看配置文件的相关操作记录和运行日志信息。

热启动配置文件

Logstash 服务已在运行时，可以选择热启动功能增加管道。

说明

- 使用 logstash stdin 插件的配置文件禁止使用热启动功能。
- 使用热启动功能时，如果配置文件热启动失败且导致 logstash 进程异常退出，会进行恢复机制重新启动原 logstash 进程，请谨慎操作。
- 只能选择一个配置文件进行热启动，且在管道列表中状态为“工作中”的配置数量小于 20 个。

1. 选择一个需要热启动的配置文件，单击左上角的“热启动”。
   对话框中“是否保持常驻”的值默认与管道列表中的“是否保持常驻”值保持一致。
2. 单击“确定”，开始热启动配置文件。
   可以在管道列表看到热启动的配置文件。

热停止配置文件

Logstash 服务已在运行时，可以选择热停止功能移除某个管道。

1. 在管道列表选择一个需要热停止的配置，单击管道列表上方的“热停止”。
2. 根据对话框提示，单击“确定”。
   热停止成功则可以看到管道列表中目标配置已被移除且该管道数据迁移中断。

停止全部配置文件

如果需要停止管道列表中全部配置文件的数据迁移，单击管道列表上方的“全部停止”。

根据对话框提示，单击“确定”，选择停止的所有管道会造成数据迁移中断。

全部停止成功则可以看到全部管道停止数据迁移。

5.2.2 Logstash 配置文件样例

说明

以下样例以源端、目的端的 elasticsearch 集群访问类型一样为例。访问类型一样指的是同样是非安全集群或者同样是安全集群没有开启 HTTPS 访问。

如果源端、目的端的 elasticsearch 集群访问类型不一样，可以由下面的 3 个样例文件的 input 和 output 部分自由组合出您需要的配置文件。

非安全集群

当创建的 Elasticsearch 类型集群未开启安全模式时，接入样例可参考如下。

```
input {
    elasticsearch {
        # 源端 ES 地址
        hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
        # 需要迁移的索引列表，以逗号分隔
        index => "xxx,xxx,xxx"
        # 以下保持默认即可
        docinfo => true
    }
}

filter {
    # 去掉一些 logstash 自己加的字段
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}

output {
    elasticsearch {
        # 目的端 ES 地址
        hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
        # 目的端索引名称，以下配置为和源端保持一致
        index => "%{[@metadata][_index]}"
        # 目的数据的_id，如果不需要保留原_id，可以删除以下这行，删除后性能会更好
        document_id => "%{[@metadata][_id]}"
        # 以下保持默认即可
        manage_template => false
        ilm_enabled => false
    }
}
```

安全集群（没有开启 HTTPS 访问）

当创建的集群开启了安全模式，但是关闭了 HTTPS 访问，接入样例可参考如下。

```
input {
    elasticsearch {
        # 源端用户名
        user => "xxx"
        # 源端密码
        password => "xxx"
        # 源端 ES 的地址
        hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
        # 需要迁移的索引列表，以逗号分隔
        index => "xxx,xxx,xxx"
        # 以下保持默认即可
        docinfo => true
    }
}
```

```
filter {
    # 去掉一些 logstash 自己加的字段
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}

output {
    elasticsearch {
        # 目的端用户名
        user => "xxx"
        # 目的端密码
        password => "xxx"
        # 目的端 ES 地址
        hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
        # 目的端索引名称，以下配置为和源端保持一致
        index => "%{[@metadata][_index]}"
        # 目的数据的_id，如果不需要保留原_id，可以删除以下这行，删除后性能会更好
        document_id => "%{[@metadata][_id]}"
        # 以下保持默认即可
        manage_template => false
        ilm_enabled => false
    }
}
```

安全集群（开启 HTTPS 访问）

当创建的集群开启了安全模式，并且开启了 HTTPS 访问，接入样例可参考如下。

```
input {
    elasticsearch {
        # 源端用户名
        user => "xxx"
        # 源端密码
        password => "xxx"
        # 源端 ES 的地址
        hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
        # 需要迁移的索引列表，以逗号分隔
        index => "xxx,xxx,xxx"
        # 源端 ES 证书，云上的集群保持以下不变；自建 logstash 集群可以在集群详情页面直接下载。这里输入对应的证书名称+证书路径
        ca_file => "/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"
        # 以下保持默认即可
        docinfo => true
        ssl => true
    }
}

filter {
    # 去掉一些 logstash 自己加的字段
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}
```

```
output {
   elasticsearch {
      # 目的端用户名
      user => "xxx"
      # 目的端密码
      password => "xxx"
      # 目的端 ES 地址
      hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
      # 目的端索引名称，以下配置为和源端保持一致
      index => "%{[@metadata][_index]}"
      # 目的数据的_id，如果不需要保留原_id，可以删除以下这行，删除后性能会更好
      document_id => "%{[@metadata][_id]}"
      # 目的端 ES 证书，云上的集群保持以下不变；自建 logstash 集群，则需要在集群详情页面下载到节点里，这里输入对应的证书名称+证书路径
      cacert => "/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"
      # 以下保持默认即可
      manage_template => false
      ilm_enabled => false
      ssl => true
      ssl certificate verification => false
   }
}
```

5.2.3 系统模板配置参数说明

- elasticsearch：从 Elasticsearch 类型集群导入数据到 Elasticsearch 类型集群。

  详细请参考 https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-elasticsearch.html

表5-7 配置项说明

配置项	是否必填	说明
hosts	是	导入数据的 Elasticsearch 集群的节点 IP。
user	否	登录 Elasticsearch 集群的用户名，一般为 admin。 如果是安全集群，需要输入此参数。
password	否	登录 Elasticsearch 集群的密码。此密码为创建集群时设置的密码。 如果是安全集群，需要输入此参数。
index	是	导入数据的索引，即用户需要从哪个索引迁移出数据。
docinfo	否	文档信息。 取值范围：true、false。 如果设置，请在事件中包括 Elasticsearch 文档信息，例如索引，类型和 ID。

配置项	是否必填	说明
ca_file	否	默认值"/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"，云上 logstash 集群保持默认值不变或使用自定义证书时输入相应的自定义证书路径；自建 Logstash 集群，可以在开启 SSL 的 ES 集群详情页面下载证书文件，在此处输入相应的路径。
ssl	否	源端 ES 集群启用 SSL 时，设置为 true。
hosts	是	输出数据的 Elasticsearch 集群的节点 IP。
user	否	登录 Elasticsearch 集群的用户名，一般为 admin。 如果是安全集群，需要输入此参数。
password	否	登录 Elasticsearch 集群的密码。此密码为创建集群时设置的密码。 如果是安全集群，需要输入此参数。
index	是	配置为需要迁移数据的索引，即用户需要迁移数据到哪个索引，就配置哪个索引。 不支持配置多条索引。
document_type	否	当 docinfo 参数配置为 true 时，此参数生效。 当 docinfo 参数配置为 false 时，需要从配置文件中删除此参数。
document_id	否	当 docinfo 参数配置为 true 时，此参数生效。 当 docinfo 参数配置为 false 时，需要从配置文件中删除此参数
cacert	否	默认值"/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"，云上 logstash 集群保持默认值不变或使用自定义证书时输入相应的自定义证书路径；自建 Logstash 集群，可以在开启 SSL 的 ES 集群详情页面下载证书文件，在此处输入相应的路径。
ssl	否	目的端 ES 集群启用 SSL 时，设置为 true。
ssl_certificate_verification	否	设置为 false 时，启用 SSL 时将忽略验证服务器证书。

5.3 形态变更

5.3.1 扩容

当集群数据面业务变化，需要动态调整集群节点的数量时，可以执行“扩容”任务。扩容集群时，业务不会中断。

前提条件

- 集群处于“可用”状态，且无正在进行的任务。
- 有足够的配额支持集群扩容。

约束限制

- 扩容操作不支持修改“节点规格”。修改“节点规格”请执行 5.3.3 变更规格操作。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置扩容参数。
   - “变更类型”：“扩容”。
   - “变更的资源”：资源的变化量。
   - “变更的角色”：此处修改的是默认数据节点类型的节点数量。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列显示为“扩容”，表示集群正在扩容。当集群状态变为“可用”，则表示扩容成功。

5.3.2 缩容

当集群有充足的能力处理当前数据时，为节省资源可以执行“缩容”任务，减小集群占用的资源。建议在业务低峰期进行缩容操作。

前提条件

集群处于“可用”状态，且无正在进行的任务。

约束限制

- 缩容操作仅支持修改“节点数量”，不支持修改“节点规格”和“节点存储容量”。
- 缩容什么节点类型的“节点数量”，缩容完成后只生效新该节点类型的“节点数量”，其他节点类型的“节点数量”保持不变。

- 要确保缩容之后的磁盘使用量小于 80%，且集群每个节点类型中每个 AZ 的节点数至少为 1。
- 缩容过程会涉及数据迁移，将要下线的节点数据迁移到其他节点上，数据迁移的超时阈值为 5 小时。当超过 5 小时数据还未迁移完成，那么缩容会失败。建议在集群数据量较大的情况下，分多次进行缩容。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择“集群管理>Logstash”中对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置缩容参数。
   - “变更类型”：“缩容”。
   - “变更的资源”：显示资源的变化量。
   - “变更的角色”：此处修改的是默认数据节点类型的节点数量。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“缩容中”，表示集群正在缩容。当集群状态变为“可用”，则表示缩容成功。

5.3.3 变更规格

当集群数据面业务变化，需要动态调整集群节点的规格时，可以执行“变更规格”任务。

前提条件

- 集群处于“可用”状态。
- 在集群的“配置中心”，配置文件列表不存在校验中的任务，管道列表不存在启动中、停止中或恢复中的任务，且管道列表工作中的任务需要保持常驻。
- 有足够的配额支持集群变更规格。
- 建议在业务低峰期更改节点规格，利于更快完成规格更改。

约束限制

- 变更规格操作不支持修改“节点数量”和“节点存储容量”。增加“节点数量”请执行 5.3.1 扩容操作。
- 如果将大规格更改为小规格，集群的处理性能将会降低，将会影响业务能力，请谨慎操作。
- 集群的“节点数量”大于等于 2 的集群才支持变更规格。
- 变更规格过程中，会依次对节点进行关机，完成更改后依次开机。是一个滚动的变更过程。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在左侧菜单栏，选择对应的集群类型，进入集群管理页面。
3. 选择目标集群，单击操作列的“更多>形态变更”进入更改集群规格页面。
4. 在更改集群规格页面，设置变更规格的参数。
   - “变更类型”：“变更规格”。
   - “变更的资源”：显示资源的变化量。
   - “变更的角色”：此处修改的是默认数据节点类型的节点规格，在对应节点规格下拉框中选择所需的规格，然后勾选需要变更的节点。
5. 单击“下一步”。
6. 确认变更信息后，单击“提交申请”。
7. 在弹出的窗口确认是否勾选“检测集群状态”，单击“确认”启动集群规格变更。

   默认检测集群状态，变更规格为滚动变更，变更过程中，为了保证变更成功率以及数据安全，会确保每个节点进程拉起来后继续后续节点操作。当集群负载过高业务故障，无法正常下发变更请求，依赖更多的资源才能恢复的紧急情况下可忽略检测集群状态，忽略后，变更过程中因为忽略集群状态检测可能会导致集群故障并中断业务，请谨慎跳过。
8. 单击“返回集群列表”跳转到集群管理页面。集群的“任务状态”列中显示为“规格修改”，表示集群正在更改规格。当集群状态变为“可用”，则表示规格变更成功。

5.4 查看集群的基本信息

在 Logstash 集群的基本信息页面，可以获取 Logstash 集群的内网访问地址、版本、节点等信息。

1. 登录云搜索服务管理控制台。
2. 选择“集群管理 ＞Logstash”，进入集群列表页面。
3. 单击集群名称进入集群“基本信息”页面，查看集群的基本信息。

表5-8 基本信息的参数说明

类别	参数	描述
基本信息	集群名称	集群名称。支持自定义名称。 单击右侧  可以修改集群名称。
	ID	集群的唯一标识，是系统自动生成的。 同一个区域下，集群 ID 是唯一的。
	集群版本	集群的版本信息。
	集群状态	集群当前的状态。

类别	参数	描述
	任务状态	集群当前的任务状态，如果没有进行中的任务则显示“--”。
	创建时间	集群创建的时间。
	集群存储容量（GB）	集群设置的存储容量。
	集群存储使用量（GB）	集群已使用的存储容量。
配置信息	区域	集群所在区域。
	可用区	集群所在的可用区。
	虚拟私有云	集群所属的虚拟私有云。
	子网	集群所属的子网。
	安全组	集群所属的安全组。 单击右侧的“更改安全组”可以修改集群的安全组信息。 **须知** **进行安全组切换前，请确保业务接入时需要的 9200 端口已经放开，错误的安全组配置可能会导致业务无法访问，请谨慎操作。**
	企业项目	集群所属的企业项目。 单击项目名称可以跳转到项目管理页面查看企业项目的基本信息。
	内网访问地址	集群的内网 IP 地址和端口号，使用此参数可以接入集群。如果集群只有一个节点，此处仅显示 1 个节点的 IP 地址和端口号，例如“10.62.179.32:9200”；如果集群有多个节点，此处显示所有节点的 IP 地址和端口号，例如“10.62.179.32:9200,10.62.179.33:9200”。
节点信息	节点规格	集群中节点的规格信息。
	节点存储	集群中节点的存储容量和存储类型。
	节点数量	集群中节点的个数。

5.5 绑定企业项目

企业可以根据组织架构规划企业项目，将企业分布在不同区域的资源按照企业项目进行统一管理，同时可以为每个企业项目设置拥有不同权限的用户组和用户。本章节为您介绍 CSS 中 Logstash 集群如何绑定、修改企业项目。

前提条件

在绑定企业项目前，您已在“企业项目管理控制台”创建企业项目。

绑定企业项目

在创建集群时，可以在“企业项目”绑定已创建的企业项目，也可以单击“查看项目管理”，前往企业项目管理控制台，新建企业项目和查看已有的企业项目。

修改企业项目

针对之前已创建的集群，其绑定的企业项目可根据实际情况进行修改。

1. 登录在云搜索服务管理控制台，
2. 在左侧导航栏，选择“集群管理 ＞Logstash”，进入集群管理页面。
3. 在集群列表中，单击集群名称进入集群“基本信息”页面。
4. 在集群“基本信息”页面，单击“企业项目”右侧的企业项目名称，进入项目管理页面。
5. 在“资源”页签下，“区域”选项中选择当前集群所在的区域，“服务”选项中选“云搜索服务 CSS”。此时，资源列表将筛选出对应的 CSS 集群。
6. 勾选需要修改企业项目的集群，然后单击“迁出”。
7. 在“迁出资源”页面，选择“迁出方式”，再选择“请选择要迁入的企业项目”，然后单击“确定”。

图5-1 迁出资源

8. 迁出完成后，可以在云搜索服务管理控制台集群管理页面，查看修改后的集群企业项目信息。

5.6 强制重启集群

当 Logstash 集群由于长时间运行或者其他未知原因导致 Logstash 故障不可用时，您可通过强制重启集群恢复集群运行。重启过程中集群不可用，请谨慎操作。

您可以通过以下步骤强制重启集群：

1. 登录云搜索服务管理控制台。
2. 单击“集群管理”>“Logstash”，进入 Logstash 类型的集群列表界面，在对应集群的“操作”列中单击“更多>强制重启”。
3. 在“强制重启集群 VMs”对话框中，仔细阅读注意事项后，单击“确定”。

图5-2 强制重启集群

集群重启过程中，集群状态为“处理中”，任务状态为“重启中”。如果集群状态变更为“可用”，表示集群已重启成功。

5.7 管理日志

为了方便用户使用日志定位问题，云搜索服务提供了日志备份和日志查询功能。日志备份可以定期将集群的日志存储在 OBS 桶中，通过 OBS 可以直接下载需要的日志文件，进行问题分析定位。

日志查询

1. 登录云搜索服务管理控制台。

2. 在“集群管理”页面，单击需要查询日志的集群名称，进入集群基本信息页面。
3. 左侧导航栏，选择“日志管理”。
4. 在日志管理页面进行日志查询。
5. 选择需要查询的节点、日志类型和日志级别信息后，单击  ，显示查询结果。

   查询日志时，是从最近时刻的 1 万条日志中进行匹配，查询结果最多显示 100 条。

备份日志

日志备份可以定期将 CSS 集群的日志存储到 OBS 桶中。

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，单击需要配置日志备份的集群名称，进入集群基本信息页面。
3. 左侧导航栏，选择“日志管理”，在“日志备份开关”右侧单击开关，打开集群的日志管理功能。
4. 在“编辑日志备份配置”弹窗中，完成参数配置。

   弹窗中默认填写了云搜索服务自动为用户创建的“OBS 桶”、“备份路径”和“IAM 委托”，用于日志备份。支持用户参考表 5-9 修改默认值。

   如果集群已经启用了日志备份功能，也可以单击“日志备份配置”右侧的编辑图标，在“编辑日志备份配置”窗口，参考表 5-9 更新日志备份的配置参数。

表5-9 日志备份配置的参数说明

参数	说明	注意事项
“OBS 桶”	选择日志存储的 OBS 桶。单击右侧的“创建桶”支持新建 OBS 桶。	OBS 桶的所在区域必须跟集群的所在区域保持一致。 **说明** **如果是子账号，需要同时设置 GetBucketStoragePolicy、GetBucketLocation、ListBucket、ListAllMyBuckets 权限，才能看到 OBS 桶。**
“备份路径”	填写日志在 OBS 桶中的存放路径。	备份路径配置规则： • 备份路径不能包括下列符号：\:*?"<>| • 备份路径不能以“/”开头。 • 备份路径不能以“.”开头或结尾。 • 备份路径的总长度不能超过 1023 个字符。
“IAM 委托”	选择 IAM 委托，指当前账号授权云搜索服	IAM 委托需满足如下条件： • “委托类型”选择“云服务”。

参数	说明	注意事项
	务访问或维护存储在OBS 中数据。单击右侧的“创建委托”支持新建委托。	• “云服务”选择“Elasticsearch”或者“云搜索服务 CSS”。 • 必选策略：“OBS Administrator”

5. 备份日志。
   - 自动备份日志。

     在“自动备份开关”右侧，单击开关，开启自动备份日志功能。

     开启“自动备份开关”后，在“修改日志备份策略”弹窗中设置“备份开始时间”。设置成功后，系统会按照设置的时间进行自动备份。

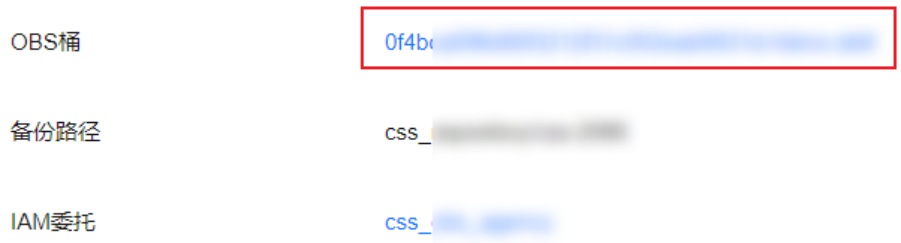
     打开“自动备份开关”后，单击开关右侧的编辑图标，可以修改“备份开始时间”。
   - 手动备份日志。

     选择“日志备份”页签，单击“开始备份”，在弹出的确认提示框中单击“确定”，开始备份日志。

     日志备份列表中的“任务状态”为“Successful”时，表示日志备份成功。
6. 查看日志文件。

   日志备份成功后，用户可以单击“OBS 桶”跳转到桶列表，找到存放备份日志的桶查看日志文件。

图5-3 进入 OBS



日志文件介绍

Logstash 集群备份的日志文件主要包括废弃操作日志、运行日志。

表5-10 日志类型信息

日志名称	描述
logstash-deprecation.log	弃用操作的日志记录。
logstash-plain.log	Logstash 运行日志。

6 查看集群运行状态和存储容量状态

在云搜索服务管理控制台总览页，直接展现当前云搜索服务中已有集群的状态以及集群存储容量状态。

表6-1 集群状态说明

状态	说明
可用	表示集群服务正常运行中，并为用户提供服务。
异常	表示集群创建失败或不可用。 如果此集群处于“不可用”状态，支持删除集群操作或将集群正常状态时创建的快照恢复至其他集群。无法执行扩容集群、访问 Kibana、创建快照或将快照恢复至此集群的操作。建议不要执行导入数据的操作，避免数据丢失。您可以查看监控或重启集群，但根据集群故障情况不同这些操作可能执行失败，当执行失败时，请及时联系技术支持。
处理中	表示集群正处在重启中、扩容中、备份中或恢复中。
创建中	表示集群正处在创建过程中。

表6-2 集群存储容量状态

状态	说明
正常	表示集群中所有节点存储容量使用率小于 50%。
警告	表示集群中任一节点存储容量使用率大于等于 50%，小于 80%。
危险	表示集群中任一节点存储容量使用率大于等于 80%。建议增加集群的存储容量，以便能够正常使用集群进行数据搜索或分析。
异常	表示未能查询到集群的存储容量信息。例如，集群运行故障，状态为“异常”时，此集群的存储容量状态为“异

状态	说明
	常”。

7 集群增强特性

7.1 向量检索

7.1.1 场景描述

图像检索、视频搜索、推荐等场景下日益增长的数据规模，对高维空间向量检索的时延和准确率提出了更高的要求。云搜索服务针对大规模的向量检索场景提供了具体的解决方案，基于自研的向量搜索引擎，结合 Elasticsearch 的插件机制，高效集成了向量检索能力。

原理

向量检索从本质上讲，其思维框架和传统的检索方法没有区别。为了提升向量检索的性能，通常需要解决以下两个问题：

- **减少候选向量集**

  和传统的文本检索类似，向量检索也需要某种索引结构来避免在全量的数据上做匹配，传统文本检索是通过倒排索引来过滤掉无关文档，而向量检索是通过对向量建立索引结构来绕过不相关的向量，减小需要考察的范围。

- **降低单个向量计算的复杂度**

  向量检索支持漏斗模型，先对所有向量进行量化和近似计算，筛选出一定量接近检索目标的数据集，然后基于筛选的数据集进行精细的计算和排序。本方法不需要对所有向量都进行复杂的计算，可以有效提高检索效率。

向量检索即在一个给定的向量数据集中，按照某种度量方式，检索出与查询向量相近的 K 个向量（K-Nearest Neighbor，KNN），但由于 KNN 计算量过大，通常只关注近似近邻（Approximate Nearest Neighbor，ANN）问题。

功能

自研向量检索引擎集成了暴力检索、图索引（HNSW）、乘积量化、IVF-HNSW 等多种向量索引，支持欧式、内积、余弦、汉明等多种相似度计算方式，召回率和检索性能均优于开源引擎。能够满足高性能、高精度、低成本、多模态等多种应用场景及需求。

向量检索支持原生 Elasticsearch 的所有能力，包括分布式、多副本、错误恢复、快照、权限控制等；兼容所有原生 Elasticsearch 生态，包括集群监测工具 Cerebro，可视化工具 Kibana，实时数据采集工具 Logstash 等；提供 Python/Java/Go/C++等多种客户端语言支持。

约束限制

- 仅 Elasticsearch 7.6.2、Elasticsearch 7.10.2、OpenSearch 1.3.6 集群支持 CSS 服务的向量检索引擎。
- 向量检索插件涉及较高的内存计算，内存要求比普通索引高，建议集群选择“内存优化型”的计算规格。
- 集群数据节点或冷数据节点的内存规格要大于 16G，否则无法使用 CSS 服务的向量检索插件，如果需要开启则请联系技术支持。

7.1.2 向量检索的集群规划

向量检索的索引构建与查询均使用堆外内存，所以集群容量与索引类型、总堆外内存大小等因素相关。通过预估全量索引所需的堆外内存大小，可以选择合适的集群规格。由于向量索引内存占用较高，CSS 对于内存规格为 8GB 及以下的集群默认禁用了向量检索插件。

不同类型的索引所需堆外内存大小的预估方式不同，计算公式如下：

- **GRAPH 类索引**

$$mem\_needs = (dim \times dim\_size + neighbros \times 4) \times num + delta$$

说明

如果有实时更新索引的需求，还需要考虑向量索引构建和自动 merge 所需的堆外内存开销，保守估计需要 1.5~2 倍 *mem_needs*。

- **PQ 类索引**

$$mem\_needs = frag\_num \times frag\_size \times num + delta$$

- **FALT、IVF 索引**

$$mem\_needs = dim \times dim\_size \times num + delta$$

表7-1 参数说明

参数	说明
dim	向量维度。
neighbors	图节点邻居数，默认值为 64。
dim_size	每一维度值所需的字节数，默认为 float 类型，需要 4 字节。
num	向量总条数。

参数	说明
delta	元数据大小，该项通常可以忽略。

frag_num	量化编码时的向量分段数，创建索引时如果未配置该值，则由向量维度“dim”决定。
	`if dim <= 256:`
	`  frag_num = dim / 4`
	`elif dim <= 512:`
	`  frag_num = dim / 8`
	`else :`
	`  frag_num = 64`

frag_size	量化编码时中心点编码的 size，默认为 1，当“frag_num”大于 256 时，该值等于 2。

基于上述计算方法，可预估出完整向量索引所需堆外内存的大小。选择集群规格时，还需考虑每个节点的堆内存开销。

节点的堆内存分配策略：每个节点的堆内存大小为节点物理内存的一半，且最大不超过 31GB。

例如，基于 SIFT10M 数据集创建 GRAPH 索引，其“dim”为“128”，“dim_size”为“4”，“neighbors”采用默认值“64”，“num”为“1000 万”，将各值代入上述公式得到 GRAPH 索引所需堆外内存大小约为：

$$mem\_needs = (128 \times 4 + 64 \times 4) \times 10000000 \approx 7.5GB$$

同时考虑到堆内存的开销，单台“8U 16G”规格的机器可以满足该场景的需求。如果实际场景还有实时写入或更新的需求，则需要考虑申请更大的内存规格。

7.1.3 创建向量索引

前提条件

- 已经参考 7.1.2 向量检索的集群规划完成集群创建，集群必须是 7.6.2 或 7.10.2 版本 Elasticsearch 集群、或者 1.3.6 版本 OpenSearch 集群。
- 根据实际需要参考集群高级配置完成集群高级设置。

创建向量索引

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择需要启用向量检索的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行如下命令创建向量索引。

   创建一个名为“my_index”的索引，该索引包含一个名为“my_vector”的向量字段和一个名为“my_label”的文本字段。其中，向量字段创建了 GRAPH 图索引，并使用欧式距离作为相似度度量。

```
PUT my_index
{
  "settings": {
    "index": {
      "vector": true
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "dimension": 2,
        "indexing": true,
        "algorithm": "GRAPH",
        "metric": "euclidean"
      },
      "my_label": {
        "type": "text"
      }
    }
  }
}
```

表7-2 创建索引参数说明

类型	参数	说明
Index settings 参数	vector	当需要使用向量索引加速时，需要设置该值为true。
Field mappings 参数	type	字段类型，“vector”表示该字段为向量字段。
	dimension	向量数据维度。取值范围：[1, 4096]。
	indexing	是否开启向量索引加速。 可选值： • false：表示关闭向量索引加速，向量数据仅写入docvalues，只支持使用 ScriptScore 以及 Rescore 进行向量查询。 • true：表示开启向量索引加速，系统将创建额外的向量索引，索引算法由"algorithm"字段指定，写入数据后可以使用 VectorQuery 进行查询。 默认值：false。
	algorithm	索引算法。仅当“indexing”为“true”时生效。 可选值： • FLAT：暴力计算，目标向量依次和所有向量进行距离计算，此方法计算量大，召回率 100%。适用于对召回准确率要求极高的场景。 • GRAPH：图索引，内嵌深度优化的 HNSW 算法，主要应用在对性能和精度均有较高要求且

类型	参数	说明
		单 shard 中文档数量在千万个以内的场景。 • GRAPH_PQ：将 HNSW 算法与 PQ 算法进行了结合，通过 PQ 降低原始向量的存储开销，能够使 HNSW 轻松支撑上亿规模的检索场景。 • IVF_GRAPH：算法将 IVF 与 HNSW 结合，对全量空间进行划分，每一个聚类中心向量代表了一个子空间，极大地提升检索效率，同时会带来微小的检索精度损失。适用于数据量在上亿以上同时对检索性能要求较高的场景。 • IVF_GRAPH_PQ：PQ 算法与 IVF-HNSW 的结合，PQ 可以通过配置选择与 HNSW 结合和 IVF 结合，进一步提升系统的容量并降低系统开销，适用于 shard 中文档数量在十亿级别以上同时对检索性能要求较高的场景。 • PV_GRAPH：图索引，同属内嵌优化的 HNSW 算法，主要应用在性能和精度均有较高要求，且单 shard 中文档数量在千万个以内，可用于内存资源充足的场景。同时该算法支持联合标量过滤，目前仅限关联 keyword 类型字段，相较于后置过滤和布尔查询，可大幅提升返回结果的填充率和查询性能。仅 Elasticsearch 集群 7.10.2 版本支持 PV_GRAPH 索引。 默认值：GRAPH。 **说明** **当选择 IVF_GRAPH 或者 IVF_GRAPH_PQ 索引时，需要额外进行预构建中心点索引以及注册等步骤，具体内容请参考 7.1.6 （可选）预构建与注册。**
	表 7-3	当使用向量索引加速时（即“indexing”为“true”时），为了获得更高的查询性能以及查询精度，CSS 提供了与向量索引相关的可选参数配置。
	metric	计算向量之间距离的度量方式。 可选值： • euclidean：欧式距离。 • inner_product：内积距离。 • cosine：余弦距离。 • hamming：汉明距离，仅支持设置"dim_type"为"binary"时使用。 默认值：euclidean。
	dim_type	向量维度值的类型。 可选值：binary、float（默认）。

类型	参数	说明
	sub_fields	新增参数字段 sub_fields 用于定义向量附属标量字段，仅支持 keyword 类型，对应新增支持联合过滤功能，仅支持 PV_GRPAH 类型。

表7-3 可选参数说明

类型	参数	说明
GRAPH 类索引配置参数	neighbors	图索引中每个向量的邻居数，默认值为 64，值越大查询精度越高。索引越大，构建速度以及后续的查询速度也会变慢。 取值范围：[10, 255]。
	shrink	构建 hnsw 时的裁边系数，默认值 1.0f。 取值范围：(0.1, 10)。
	scaling	构建 hnsw 时上层图节点数的缩放比例，默认值 50。 取值范围：(0, 128]。
	efc	构建 hnsw 时考察邻居节点的队列大小，默认值为 200，值越大精度越高，构建速度将会变慢。 取值范围：(0, 100000]。
	max_scan_num	扫描节点上限，默认值为 10000，值越大精度越高，索引速度变慢。 取值范围：(0, 1000000]。
PQ 类索引配置参数	centroid_num	每一段的聚类中心点数目，默认值为 255。 取值范围：(0, 65535]。
	fragment_num	段数，默认值为 0，插件自动根据向量长度设置合适的段数。 取值范围：[0, 4096]。

导入向量数据

执行如下命令，导入向量数据。向“my_index”索引中写入向量数据时，需要指定向量字段名称和向量数据。

- 向量数据输入格式为逗号分隔的浮点型数组时：

```
POST my_index/_doc
{
```

```
  "my_vector": [1.0, 2.0]
}
```

- 向量数据输入格式为小端字节序编码的 Base64 字符串时：

  在写入二值向量，或向量维度较高、数值有效位较多时，使用 Base64 编码格式传输、解析更加高效。

```
POST my_index/_doc
{
  "my_vector": "AACAPwAAAEA="
}
```

- 当写入大规模数据时，建议使用 Bulk 操作：

```
POST my_index/_bulk
{"index": {}}
{"my_vector": [1.0, 2.0], "my_label": "red"}
{"index": {}}
{"my_vector": [2.0, 2.0], "my_label": "green"}
{"index": {}}
{"my_vector": [2.0, 3.0], "my_label": "red"}
```

集群高级配置

- 在离线导入数据场景下，为了提高批量写入性能，建议将索引的 refresh_interval 参数设置为-1，即关闭自动刷新索引。
- 建议将备份数 number_of_replicas 设置为 0，当离线数据导入完成后，再设置为需要的值。
- 其他高级功能的参数配置说明：

表7-4 集群配置参数

参数	说明
native.cache.circuit_breaker.enabled	是否开启堆外内存熔断。 默认值：true。
native.cache.circuit_breaker.cpu.limit	向量索引堆外内存使用上限。 假设使用 128GB 内存的机器且堆内存大小为 31GB，默认堆外内存使用上限为(128 - 31) * 45% = 43.65GB，堆外内存使用量超过该值将会触发写入熔断。 默认值：45%。
native.cache.expire.enabled	是否开启缓存超时设置。开启时，如果某些缓存项长时间没有被访问过将会被清除。 取值范围：true、false。 默认值：false。
native.cache.expire.time	超时时长。 默认值：24h。
native.vector.index_threads	创建底层索引时所使用的线程数，每个 shard 均会使用多个构建线程。该值建议不要设置过大，避免产生过多的构建线

参数	说明
	程抢占查询资源。 默认值：4。

7.1.4 向量查询

标准查询

针对创建了向量索引的向量字段，提供了标准向量查询语法。下述查询命令将会返回所有数据中与查询向量最近的 size（topk）条数据。

```
POST my_index/_search
{
  "size":2,
  "_source": false,
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1, 1],
        "topk":2
      }
    }
  }
}
```

表7-5 标准查询的参数说明

参数	说明
vector（第一个）	表示该查询类型为 VectorQuery。
my_vector	指定了需要查询的向量字段名称。
vector（第二个）	指定查询向量的具体值，支持数组形式以及 Base64 编码形式的输入。
topk	topk 的值通常与 size 保持一致。
表 7-6	通过调整不同索引的查询参数，可以获得更高的查询性能或者查询精度。

表7-6 可选的查询参数说明

参数	子参数	说明
GRAPH 类索引配置参数	ef	查询时考察邻居节点的队列大小。值越大查询精度越高，查询速度会变慢。默认值为 200。 取值范围：(0, 100000]。

参数	子参数	说明
	max_scan_num	扫描节点上限。值越大精度越高，查询速度变慢。默认值为 10000。 取值范围：(0, 1000000]。
IVF 类索引配置参数	nprobe	查询考察中心点的数目。值越大精度越高，查询速度变慢。默认值为 100。 取值范围：(0, 100000]。

复合查询

向量检索支持与其他 ES 子查询组合进行复合查询，比如布尔查询、后置过滤等。

以下两个示例的查询结果：首先查询 top10 条与查询向量距离最近的结果，filter 作为过滤条件将仅保留 my_label 字段为“red”的结果。

- 布尔查询示例

```
POST my_index/_search
{
  "size": 10,
  "query": {
    "bool": {
      "must": {
        "vector": {
          "my_vector": {
            "vector": [1, 2],
            "topk": 10
          }
        }
      },
      "filter": {
        "term": { "my_label": "red" }
      }
    }
  }
}
```

- 后置过滤示例

```
GET my_index/_search
{
  "size": 10,
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1, 2],
        "topk": 10
      }
    }
  },
  "post_filter": {
    "term": { "my_label": "red" }
```

```
  }
}
```

ScriptScore 查询

写入向量数据后，针对向量字段可以使用 ScriptScore 进行最近邻查询，查询语法如下所示。

前置过滤条件可以为任意查询，script_score 仅针对前置过滤的结果进行遍历，计算向量相似度并排序返回。此种查询方式的性能取决于前置过滤后中间结果集的大小，当前置过滤条件为"match_all"时，相当于全局暴力检索。

```
POST my_index/_search
 {
   "size":2,
   "query": {
   "script_score": {
      "query": {
        "match_all": {}
      },
      "script": {
        "source": "vector_score",
        "lang": "vector",
        "params": {
          "field": "my_vector",
          "vector": [1.0, 2.0],
          "metric": "euclidean"
        }
      }
    }
   }
 }
```

表7-7 script_score 参数说明

参数	说明
source	script 脚本描述，使用向量相似度打分时为固定值"vector_score"。
lang	script 语法描述，使用固定值"vector"。
field	向量字段名称。
vector	查询向量数据。
metric	度量方式，可选值为：euclidean、inner_product、cosine、hamming。 默认值：euclidean。

重打分查询

当使用 GRAPH_PQ 索引或者 IVF_GRAPH_PQ 索引时，查询结果是根据 PQ 计算的非对称距离进行排序。CSS 支持 Rescore 的方式对查询结果进行重打分精排，提升召回率。

假设 my_index 是 PQ 类型的索引，Rescore 示例如下：

```
GET my_index/_search
 {
   "size": 10,
   "query": {
     "vector": {
       "my_vector": {
         "vector": [1.0, 2.0],
         "topk": 100
       }
     }
   },
   "rescore": {
     "window_size": 100,
     "vector_rescore": {
       "field": "my_vector",
       "vector": [1.0, 2.0],
       "metric": "euclidean"
     }
   }
 }
```

表7-8 Rescore 参数说明

参数	说明
window_size	向量检索将会返回 topk 条结果，仅取前 window_size 条结果精排。
field	向量字段名称。
vector	查询向量数据。
metric	度量方式，可选值为：euclidean、inner_product、cosine、hamming。 默认值：euclidean。

Painless 语法扩展支持

CSS 扩展实现了多种向量距离计算函数，可在自定义的 painless 脚本中直接使用，用以构建灵活的重打分公式。

示例如下：

```
POST my_index/_search
{
```

```
  "size": 10,
  "query": {
    "script_score": {
      "query": {
        "match_all": {}
      },
      "script": {
        "source": "1 / (1 + euclidean(params.vector, doc[params.field]))",
        "params": {
          "field": "my_vector",
          "vector": [1, 2]
        }
      }
    }
  }
}
```

CSS 当前支持的距离计算函数如下表所示：

函数签名	说明
euclidean(Float[], DocValues)	欧式距离函数。
cosine(Float[], DocValues)	余弦相似度函数。
innerproduct(Float[], DocValues)	内积函数。
hamming(String, DocValues)	汉明距离函数。只支持"dim_type"为"binary"的向量字段，输入的查询向量需要为 Base64 编码字符串格式。

7.1.5 向量检索的性能调优

写入性能优化

- 关闭副本，待数据导入完成后再开启副本，减少副本构建的开销。
- 调整“refresh_interval”为 120s 或者更大，避免频繁刷新索引生成大量小的 segments，同时减少 merge 带来的向量索引构建开销。
- 适当调大“native.vector.index_threads”的值（默认为 4），增加向量索引构建的线程数。

```
PUT _cluster/settings
{
  "persistent": {
    "native.vector.index_threads": 8
  }
}
```

查询性能优化

- 在批量导入场景下，数据写入完成后，执行 forcemerge 操作能有效提升查询效率。

```
POST index_name/_forcemerge?max_num_segments=1
```

- 如果向量索引所需堆外内存超过了熔断线，查询时索引的缓存管理器会控制索引的换进换出，导致查询变慢，此时可适当调大熔断线的配置。

```
PUT _cluster/settings
{
  "persistent": {
    "native.cache.circuit_breaker.cpu.limit": "75%"
  }
}
```

- 如果查询的 fetch 阶段开销较大，可通过配置“_source”减小 fdt 文件的大小，从而降低 fetch 开销。

```
PUT my index
{
  "settings": {
    "index": {
      "vector": "true"
    },
    "index.soft_deletes.enabled": false
  },
  "mappings": {
    "_source": {
      "excludes": ["my_vector"]
    },
    "properties": {
      "my_vector": {
        "type": "vector",
        "dimension": 128,
        "indexing": true,
        "algorithm": "GRAPH",
        "metric": "euclidean"
      }
    }
  }
}
```

7.1.6（可选）预构建与注册

在创建向量索引时，如果选择使用“IVF_GRAPH”和“IVF_GRAPH_PQ”的索引算法就需要对中心点向量进行预构建和注册。

背景信息

在向量索引加速算法中，IVF_GRAPH 和 IVF_GRAPH_PQ 适用于超大规模场景。这两种算法需要通过对子空间的切割缩小查询范围，子空间的划分通常采用聚类或者随机采样的方式。在预构建之前，需要通过聚类或者随机采样得到所有的中心点向量。

当完成生成中心点向量的工作之后，需要对中心点向量进行预构建和注册，以实现将中心点向量预构建 GRAPH 或者 GRAPH_PQ 索引，同时注册到 CSS 集群内，实现在多个节点间共享此索引文件。中心点索引在 shard 间复用能够有效减少训练的开销、中心点索引查询次数，提升写入以及查询的性能。

操作步骤

1. 选择启用向量检索的集群，单击操作列“Kibana”，登录 Kibana 界面。
2. 单击左侧导航栏的“Dev Tools”，进入操作界面。
3. 创建中心点索引表。
   - 创建的索引命名为 my_dict，注意该索引的 number_of_shards 数必须设置为 1，否则无法注册。
   - 当需要使用 IVF_GRAPH 索引时，中心点索引的 algorithm 设置为 GRAPH。
   - 当需要使用 IVF_GRAPH_PQ 索引时，中心点索引的 algorithm 设置为 GRAPH_PQ。

```
PUT my_dict
 {
   "settings": {
     "index": {
       "vector": true
     },
     "number_of_shards": 1,
     "number_of_replicas": 0
   },
   "mappings": {
     "properties": {
       "my_vector": {
         "type": "vector",
         "dimension": 2,
         "indexing": true,
         "algorithm": "GRAPH",
         "metric": "euclidean"
       }
     }
   }
 }
```

4. 写入中心点向量数据。

   参考导入向量数据将采样或者聚类得到的中心点向量写入上述创建的 my_dict 索引中。

5. 调用注册接口。

   将上述创建的 my_dict 索引注册具有全局唯一标识名称（dict_name）的 Dict 对象。

```
PUT _vector/register/my_dict
 {
   "dict_name": "my_dict"
 }
```

6. 创建 IVF_GRAPH 或 IVF_GRAPH_PQ 索引。

   在创建 IVF_GRAPH 或者 IVF_GRAPH_PQ 索引时，不再需要指定 dimension 以及 metric 信息，只需指定之前注册好的 dict 名称即可。

```
PUT my_index
 {
   "settings": {
     "index": {
       "vector": true
```

```
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "indexing": true,
        "algorithm": "IVF_GRAPH",
        "dict_name": "my_dict",
        "offload_ivf": false
      }
    }
  }
}
```

表7-9 Field mappings 参数

参数	说明
dict_name	指定依赖的中心点索引名称。该索引字段的向量维度和度量方式将与 dict 索引保持一致，不再需要额外指定。
offload_ivf	将底层索引实现的 IVF 倒排索引卸载到 ES 端实现，可以减少堆外内存的使用，以及减少写入/合并的性能开销，但是查询的性能也有一定的损失。采用默认值即可。 取值范围：true、false。 默认值：false。

7.1.7 管理向量索引缓存

CSS 的向量检索引擎使用 C++实现，使用的是堆外内存，该插件提供了接口对向量索引的缓存进行管理。

- **查看缓存统计信息**

```
GET /_vector/stats
```

在向量插件实现中，向量索引与 Lucene 其他类型索引一样，每一个 segment 构造并存储一份索引文件，在查询时，该索引文件会被加载到堆外内存中。插件使用缓存机制对这些堆外内存进行管理。上述 API 能够查询当前堆外内存使用量、缓存命中次数、加载次数等信息。

- **预加载向量索引**

```
PUT /_vector/warmup/{index_name}
```

使用上述接口能将指定 index_name 的向量索引预加载至堆外内存供查询使用。

- **清除缓存**

```
PUT /_vector/clear/cache
PUT /_vector/clear/cache/index_name
```

在使用向量索引时，缓存机制会限制堆外内存使用量。当总索引大小超出缓存大小限制时，将会发生索引项的换进换出，此时将会影响查询的性能。通过清除缓存 API 能够将不再使用的索引缓存清空，保证热数据索引的查询性能。

7.1.8 向量检索的客户端代码示例（Python）

Elasticsearch 提供了标准的 REST 接口，以及 Java、Python 等语言编写的客户端。

本节提供一份创建向量索引、导入向量数据和查询向量数据的 Python 代码示例，介绍如何使用客户端实现向量检索。

前提条件

客户端已经安装 python 依赖包。如果未安装可以执行如下命令安装：

```
# 根据集群实际版本填写，此处以 7.6 举例
pip install elasticsearch==7.6
```

代码示例

```
from elasticsearch import Elasticsearch
from elasticsearch import helpers

# 创建 Elasticsearch 客户端
def get_client(hosts: list, user: str = None, password: str = None):
    if user and password:
        return Elasticsearch(hosts, http_auth=(user, password), verify_certs=False,
ssl_show_warn=False)
    else:
        return Elasticsearch(hosts)

# 创建索引表
def create(client: Elasticsearch, index: str):
    # 索引 mapping 信息
    index_mapping = {
        "settings": {
            "index": {
                "vector": "true",  # 开启向量特性
                "number_of_shards": 1,  # 索引分片数，根据实际需求设置
                "number_of_replicas": 0,  # 索引副本数，根据实际需求设置
            }
        },
        "mappings": {
            "properties": {
                "my_vector": {
                    "type": "vector",
                    "dimension": 2,
                    "indexing": True,
                    "algorithm": "GRAPH",
                    "metric": "euclidean"
                }
                # 可根据需求添加其他字段
            }
        }
    }
    res = client.indices.create(index=index, body=index_mapping)
    print("create index result: ", res)

# 写入数据
```

```
def write(client: Elasticsearch, index: str, vecs: list, bulk_size=500):
    for i in range(0, len(vecs), bulk_size):
        actions = [
            {
                "_index": index,
                "my_vector": vec,
                # 可根据需求添加其他字段
            }
            for vec in vecs[i: i+bulk_size]
        ]
        success, errors = helpers.bulk(client, actions, request_timeout=3600)
        if errors:
            print("write bulk failed with errors: ", errors)  # 根据需求进行错误处理
        else:
            print("write bulk {} docs success".format(success))
    client.indices.refresh(index=index, request_timeout=3600)

# 查询向量索引
def search(client: Elasticsearch, index: str, query: list, size: int):
    # 查询语句，可根据需求选择合适的查询方式
    query body = {
        "size": size,
        "query": {
            "vector": {
                "my_vector": {
                    "vector": query,
                    "topk": size
                }
            }
        }
    }
    res = client.search(index=index, body=query_body)
    print("search index result: ", res)

# 删除索引
def delete(client: Elasticsearch, index: str):
    res = client.indices.delete(index=index)
    print("delete index result: ", res)

if __name__ == '__main__':
    # 对于非安全集群，使用：
    es_client = get_client(hosts=['http://x.x.x.x:9200'])

    # 对于开启了 https 的安全集群，使用：
    # es_client = get_client(hosts=['https://x.x.x.x:9200', 'https://x.x.x.x:9200'],
user='xxxxx', password='xxxxx')

    # 对于未开启 https 的安全集群，使用：
    # es_client = get_client(hosts=['http://x.x.x.x:9200', 'http://x.x.x.x:9200'],
user='xxxxx', password='xxxxx')

    # 测试索引名称
    index_name = "my_index"

    # 创建索引
```

```
create(es_client, index=index_name)

# 写入数据
data = [[1.0, 1.0], [2.0, 2.0], [3.0, 3.0]]
write(es_client, index=index_name, vecs=data)

# 查询索引
query_vector = [1.0, 1.0]
search(es_client, index=index_name, query=query_vector, size=3)

# 删除索引
delete(es_client, index=index_name)
```

7.2 存算分离

7.2.1 背景信息

云搜索服务支持将热数据存储在 SSD 来达到最佳的查询检索性能。将历史数据存储在 OBS 中降低数据的存储成本，该特性为存算分离特性。

使用场景

对于有海量数据写入和存储的场景，一般数据有明显的冷热区分，新写入的数据存储在 SSD 中，随着时间的推移，历史数据不再写入，查询 QPS 也降低，这时候可以调用云搜索服务提供的 API 将存储在 SSD 的热数据转储到 OBS，这个转储的过程称为冻结索引，也就是存算分离。

约束限制

- 目前仅 7.6.2 和 7.10.2 版本 Elasticsearch 集群和 1.3.6 版本 OpenSearch 集群支持存算分离。
- 由于存算分离的特性依赖 OBS，所以使用过程中要遵守 OBS 的“带宽”和“每秒请求数（QPS）”的使用限制。当超过限制时，集群中涉及到 OBS 查询的性能都会下降，例如恢复分片的速度变慢、查询数据时变慢等。

7.2.2 冻结索引

注意事项

- 在执行冻结操作前，需冻结的索引没有数据写入。在冻结操作执行前，会将索引配置为 read only，会导致写入数据出错。
- 在执行冻结操作后：
  - 索引变为只读。
  - 索引数据将会转储到 OBS，转移过程中，会占用网络带宽。
  - 转储后的索引，查询时延会增加。聚合时，由于查询复杂，数据读取多，时延变长会体现的更明显。

- 已冻结的索引不支持解冻，即不可回退为可写的索引。
- 冻结完成以后，会删除本地磁盘中的索引数据。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择需要冻结索引的集群，单击操作列“Kibana”，登录Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，进入操作页面。
4. 执行如下命令，将指定索引冻结到 OBS 中。

```
POST ${index_name}/_freeze_low_cost
```

表7-10 请求参数说明

参数名	说明
index_name	需要冻结的索引的名字。

返回结果如下：

```
{
   "freeze_uuid": "pdsRgUtSTymVDWR_HoTGFw"
}
```

表7-11 返回参数说明

参数名	说明
freeze_uuid	提交冻结请求后会启动一个异步任务，请求返回异步任务的 ID，使用该 ID 查询异步任务的进度。

说明

索引冻结请求下发后，会禁止索引的数据写入，冻结过程中，查询请求不受影响。在冻结完成后，会将索引先 close 再 open，在这段时间内，索引不可查询，集群可能短暂出现 red 状态，open 结束后恢复。

5. 执行如下命令获取冻结任务进度。

```
GET _freeze_low_cost_progress/${freeze_uuid}
```

表7-12 请求参数说明

参数名	说明
freeze_uuid	异步任务的 ID，该 ID 由 4 获取的。

返回结果如下：

```
{

 "stage" : "STARTED",
 "shards_stats" : {
  "INIT" : 0,
  "FAILURE" : 0,
  "DONE" : 0,
  "STARTED" : 3,
  "ABORTED" : 0
 },
 "indices" : {
  "data1" : [
   {
    "uuid" : "7OS-G1-tRke2jHZPlckexg",
    "index" : {
     "name" : "data1",
     "index_id" : "4b5PHXJITLaS6AurImfQ9A",
     "shard" : 2
    },
    "start_ms" : 1611972010852,
    "end_ms" : -1,
    "total time" : "10.5s",
    "total time in millis" : 10505,
    "stage" : "STARTED",
    "failure" : null,
    "size" : {
     "total bytes" : 3211446689,
     "finished_bytes" : 222491269,
     "percent" : "6.0%"
    },
    "file" : {
     "total_files" : 271,
     "finished_files" : 12,
     "percent" : "4.0%"
    },
    "rate_limit" : {
     "paused_times" : 1,
     "paused_nanos" : 946460970
    }
   },
   {
    "uuid" : "7OS-G1-tRke2jHZPlckexg",
    "index" : {
     "name" : "data1",
     "index_id" : "4b5PHXJITLaS6AurImfQ9A",
     "shard" : 0
    },
    "start_ms" : 1611972010998,
    "end_ms" : -1,
    "total_time" : "10.3s",
    "total_time_in_millis" : 10359,
    "stage" : "STARTED",
    "failure" : null,
    "size" : {
     "total_bytes" : 3221418186,
```

```
        "finished_bytes" : 272347118,
        "percent" : "8.0%"
      },
      "file" : {
        "total_files" : 372,
        "finished_files" : 16,
        "percent" : "4.0%"
      },
      "rate_limit" : {
        "paused_times" : 5,
        "paused_nanos" : 8269016764
      }
    },
    {
      "uuid" : "7OS-G1-tRke2jHZPlckexg",
      "index" : {
        "name" : "data1",
        "index_id" : "4b5PHXJITLaS6AurImfQ9A",
        "shard" : 1
      },
      "start ms" : 1611972011021,
      "end ms" : -1,
      "total time" : "10.3s",
      "total time in millis" : 10336,
      "stage" : "STARTED",
      "failure" : null,
      "size" : {
        "total_bytes" : 3220787498,
        "finished_bytes" : 305789614,
        "percent" : "9.0%"
      },
      "file" : {
        "total_files" : 323,
        "finished_files" : 14,
        "percent" : "4.0%"
      },
      "rate_limit" : {
        "paused_times" : 3,
        "paused_nanos" : 6057933087
      }
    }
  ]
 }
}
```


表7-13 返回参数说明

参数名	说明
stage	当前所处状态。取值包括： • INIT：刚启动或者正在初始化。 • FAILURE：失败。 • DONE：完成。

参数名	说明
	• STARTED：已启动。 • ABORTED：取消，预留字段。
shards_stats	处在各个状态的 shard 个数。
indices	每个索引的状态详情。

表7-14 indices 返回值说明

参数名	说明
uuid	freeze 的 uuid。
index	索引信息和 shard 信息。
start_ms	开始时间。
end_ms	结束时间，如果没有结束则显示为-1。
total_time	已花费时间。
total_time_in_millis	已花费时间毫秒数。
stage	当前 shard 所处的状态。
failure	失败原因，如果没有失败则显示为 null。
size.total_bytes	总共需要冻结的文件的字节数。
size.finished_bytes	已经完成冻结的字节数。
size.percent	已经完成冻结的字节数百分比。
file.total_bytes	总共需要冻结的文件个数。
file.finished_bytes	已经完成冻结的文件个数。
file.percent	已经完成冻结的文件个数百分比。
rate_limit.paused_times	达到限速导致冻结暂停的次数。
rate_limit.paused_nanos	达到限速导致冻结暂停的时间纳秒数。

冻结完成的索引会增加以下 settings，可参考表 7-15。

表7-15 冻结索引 settings

参数	说明
index.frozen_low_cost	标识该索引为冻结索引。取值为 true。

参数	说明
index.blocks.write	冻结后的索引禁止写入。取值为 true。
index.store.type	标识该索引的存储类型为 obs。取值为 obs。

6. 索引冻结后，会将数据进行缓存。执行如下命令获取当前缓存状态。关于缓存详见 7.2.3 配置缓存。

```
GET _frozen_stats
GET _frozen_stats/${node_id}
```

表7-16 请求参数说明

参数名	说明
node_id	获取单个节点的缓存状态，此参数为需要获取的节点 id。

返回结果如下：

```
{
  "_nodes" : {
    "total" : 3,
    "successful" : 3,
    "failed" : 0
  },
  "cluster_name" : "css-zzz1",
  "nodes" : {
    "7uwKO38RRoaON37YsXhCYw" : {
      "name" : "css-zzz1-ess-esn-2-1",
      "transport_address" : "10.0.0.247:9300",
      "host" : "10.0.0.247",
      "ip" : "10.0.0.247",
      "block_cache" : {
        "default" : {
          "type" : "memory",
          "block_cache_capacity" : 8192,
          "block_cache_blocksize" : 8192,
          "block_cache_size" : 12,
          "block_cache_hit" : 14,
          "block_cache_miss" : 0,
          "block_cache_eviction" : 0,
          "block_cache_store_fail" : 0
        }
      },
      "obs_stats" : {
        "list" : {
          "obs_list_count" : 17,
          "obs_list_ms" : 265,
          "obs_list_avg_ms" : 15
        },
        "get_meta" : {
```

```
      "obs_get_meta_count" : 79,
      "obs_get_meta_ms" : 183,
      "obs_get_meta_avg_ms" : 2
    },
    "get_obj" : {
      "obs_get_obj_count" : 12,
      "obs_get_obj_ms" : 123,
      "obs_get_obj_avg_ms" : 10
    },
    "put_obj" : {
      "obs_put_obj_count" : 12,
      "obs_put_obj_ms" : 2451,
      "obs_put_obj_avg_ms" : 204
    },
    "obs_op_total" : {
      "obs_op_total_ms" : 3022,
      "obs_op_total_count" : 120,
      "obs_op_avg_ms" : 25
    }
  },
  "reader_cache" : {
    "hit_count" : 0,
    "miss_count" : 1,
    "load_success_count" : 1,
    "load_exception_count" : 0,
    "total_load_time" : 291194714,
    "eviction_count" : 0
  }
},
"73EDpEqoQES749umJqxOzQ" : {
  "name" : "css-zzz1-ess-esn-3-1",
  "transport_address" : "10.0.0.201:9300",
  "host" : "10.0.0.201",
  "ip" : "10.0.0.201",
  "block_cache" : {
    "default" : {
      "type" : "memory",
      "block_cache_capacity" : 8192,
      "block_cache_blocksize" : 8192,
      "block_cache_size" : 12,
      "block_cache_hit" : 14,
      "block_cache_miss" : 0,
      "block_cache_eviction" : 0,
      "block_cache_store_fail" : 0
    }
  },
  "obs_stats" : {
    "list" : {
      "obs_list_count" : 17,
      "obs_list_ms" : 309,
      "obs_list_avg_ms" : 18
    },
    "get_meta" : {
      "obs_get_meta_count" : 79,
      "obs_get_meta_ms" : 216,
```

```
      "obs_get_meta_avg_ms" : 2
    },
    "get_obj" : {
      "obs_get_obj_count" : 12,
      "obs_get_obj_ms" : 140,
      "obs_get_obj_avg_ms" : 11
    },
    "put_obj" : {
      "obs_put_obj_count" : 12,
      "obs_put_obj_ms" : 1081,
      "obs_put_obj_avg_ms" : 90
    },
    "obs_op_total" : {
      "obs_op_total_ms" : 1746,
      "obs_op_total_count" : 120,
      "obs_op_avg_ms" : 14
    }
  },
  "reader_cache" : {
    "hit_count" : 0,
    "miss_count" : 1,
    "load_success_count" : 1,
    "load_exception_count" : 0,
    "total_load_time" : 367179751,
    "eviction_count" : 0
  }
},
"EF8WoLCUQbqJl1Pkqo9-OA" : {
  "name" : "css-zzz1-ess-esn-1-1",
  "transport_address" : "10.0.0.18:9300",
  "host" : "10.0.0.18",
  "ip" : "10.0.0.18",
  "block_cache" : {
    "default" : {
      "type" : "memory",
      "block_cache_capacity" : 8192,
      "block_cache_blocksize" : 8192,
      "block_cache_size" : 12,
      "block_cache_hit" : 14,
      "block_cache_miss" : 0,
      "block_cache_eviction" : 0,
      "block_cache_store_fail" : 0
    }
  },
  "obs_stats" : {
    "list" : {
      "obs_list_count" : 17,
      "obs_list_ms" : 220,
      "obs_list_avg_ms" : 12
    },
    "get_meta" : {
      "obs_get_meta_count" : 79,
      "obs_get_meta_ms" : 139,
      "obs_get_meta_avg_ms" : 1
    },
```

```
      "get_obj" : {
        "obs_get_obj_count" : 12,
        "obs_get_obj_ms" : 82,
        "obs_get_obj_avg_ms" : 6
      },
      "put_obj" : {
        "obs_put_obj_count" : 12,
        "obs_put_obj_ms" : 879,
        "obs_put_obj_avg_ms" : 73
      },
      "obs_op_total" : {
        "obs_op_total_ms" : 1320,
        "obs_op_total_count" : 120,
        "obs_op_avg_ms" : 11
      }
    },
    "reader_cache" : {
      "hit_count" : 0,
      "miss_count" : 1,
      "load_success_count" : 1,
      "load_exception_count" : 0,
      "total_load_time" : 235706838,
      "eviction_count" : 0
    }
  }
 }
}
```

7. 执行如下命令重置缓存状态。

```
POST _frozen_stats/reset
```

返回结果如下：

```
{
 "_nodes" : {
  "total" : 1,
  "successful" : 1,
  "failed" : 0
 },
 "cluster_name" : "Es-0325-007_01",
 "nodes" : {
  "mqTdk2YRSPyOSXfesREFSg" : {
   "result" : "ok"
  }
 }
}
```

说明

此命令用于性能问题的调试，如重置缓存状态后再次执行查询，可以清晰看到本次查询的缓存命令情况。在业务运行阶段不需要执行此命令。

8. 执行如下命令获取当前已经冻结的所有索引。

```
GET _cat/freeze_indices?stage=${STAGE}
```

表7-17 请求参数说明

参数名	说明
STAGE	支持如下赋值： • start：开始冻结但是还未完成的索引列表 • done： 已经完成冻结的索引列表 • unfreeze：没有冻结的索引列表 • 空或者其他值：正在冻结过程中或者已经冻结完成的所有索引列表

返回结果如下：

```
green open data2 0bNtxWDtRbOSkS4JYaUgMQ 3 0  5 0  7.9kb  7.9kb
green open data3 oYMLvw31QnyasqUNuyP6RA 3 0 51 0 23.5kb 23.5kb
```

说明

此命令的参数和返回值与开源 Elasticsearch 的_cat/indices 一致。

7.2.3 配置缓存

将数据转储到 OBS 后，为了尽可能的减少对 OBS 的访问请求，并提升 ES 的查询性能，系统将会缓存部分数据。第一次获取到数据时，会直接访问 OBS，之后将获取到的数据缓存在内存中，后续访问会先检查是否有缓存。数据缓存支持内存和文件。

ES 访问不同的文件访问的模式是不一样的，缓存系统支持多级缓存，分别使用不同的 block 大小来缓存不同的文件，如对 fdx，tip 文件，使用大量的小 block 缓存，对 fdt 文件，使用较少的大 block 缓存。

表7-18 针对缓存的所有配置

配置名	类型	说明
low_cost.obs.blockcache.names	Array	缓存系统支持多级缓存，分别用来缓存不同访问粒度的数据。此配置列出所有缓存的名字，即使不配置，系统也会默认有一个缓存，名字为 default。如果自定义配置，请确保有一个名字为 default 的缓存，其他名字任意。 默认值：default。
low_cost.obs.blockcache.<NAME>.type	ENUM	缓存的类型，支持 memory 和 file。 当使用 memory 类型的缓存时，会占用一定的内存大小。当使用 file 类型的缓存时，会使用磁盘作为缓存。建议使用超高 IO 型的磁盘提升缓存性能。 默认值：memory。
low_cost.obs.blockcache.<NAME>.blockshift	Integer	缓存每个 block 的大小，为字节左移数，即 2^x 字节。如配置为 16，表示 block 大小为 2^{16} 字

配置名	类型	说明
		节，等于 65536 字节，即 64K。 默认值：13（即 8K）。
low_cost.obs.blockcache.<NAME>.bank.count	Integer	缓存分区数。 默认值：1。
low_cost.obs.blockcache.<NAME>.number.blocks.perbank	Integer	每个缓存分区中包含的 block 数。 默认值：8192。
low_cost.obs.blockcache.<NAME>.exclude.file.types	Array	不缓存的文件后缀名。如果某些后缀既不包含在 exclude 列表，也不包含在 include 列表，则会使用 default 缓存。
low_cost.obs.blockcache.<NAME>.file.types	Array	缓存的文件后缀名。如果某些后缀既不包含在 exclude 列表，也不包含在 include 列表，则会使用 default 缓存。

以下为一个较为常见的缓存配置，该配置使用两级缓存，名字分别为 default 和 large。其中 default 缓存使用 64K 的 block 大小，并且一共有 30*4096 个 block，default 缓存用于缓存除 fdt 后缀的其他文件。large 缓存使用 2M 的 block 大小，一共有 5*1000 个 block，large 缓存用于缓存 fdx，dvd，tip 后缀的文件。

```
low_cost.obs.blockcache.names: ["default", "large"]
low_cost.obs.blockcache.default.type: file
low_cost.obs.blockcache.default.blockshift: 16
low_cost.obs.blockcache.default.number.blocks.perbank: 4096
low_cost.obs.blockcache.default.bank.count: 30
low_cost.obs.blockcache.default.exclude.file.types: ["fdt"]

low_cost.obs.blockcache.large.type: file
low_cost.obs.blockcache.large.blockshift: 21
low_cost.obs.blockcache.large.number.blocks.perbank: 1000
low_cost.obs.blockcache.large.bank.count: 5
low_cost.obs.blockcache.large.file.types: ["fdx", "dvd", "tip"]
```

表7-19 其他可配置参数

配置名	类型	说明
index.frozen.obs.max_bytes_per_sec	String	冻结过程中往 OBS 上传文件最大限速。动态配置，修改后立即生效。 默认值：150MB。
low_cost.obs.index.upload.threshold.use.multipart	String	冻结过程中文件大小超过此配置会使用 OBS 的分段上传。 默认值：1GB。
index.frozen.reader.cache.	Integer	此参数设置超时时间。

配置名	类型	说明
expire.duration.seconds		为了减少冻结后的索引占用的堆内存，在索引 shard 启动后，reader 会缓存一段时间，超时后关闭。 默认值：300s。
index.frozen.reader.cache.max.size	Integer	配置缓存最大值。 默认值：100。

7.2.4 查询冷数据性能提升

背景信息

在 Kibana 的 Discover 页面的首次查询时，由于此时无任何缓存，导致所有数据均需要从 OBS 上获取。当命中的文档数量较多时，需要耗费大量的时间从 OBS 上获取对应的时间字段以及文件元数据。如果将这一部分数据直接缓存在本地，即可大量提升查询性能，解决 Discover 页面首次查询慢的问题。

前提条件

此特性仅支持在 2023 年 02 月后创建的 7.6.2 和 7.10.2 版本集群以及 OpenSearch 集群。

查询冷数据本地缓存 API

您可以使用该 API 查询冷数据本地缓存的相关指标。

请求示例

```
GET /_frozen_stats/local_cache
GET /_frozen_stats/local_cache/{nodeId}
```

响应示例

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "elasticsearch",
  "nodes" : {
    "6by3lPy1R3m55Dcq3liK8Q" : {
      "name" : "node-1",
      "transport_address" : "127.0.0.1:9300",
      "host" : "127.0.0.1",
      "ip" : "127.0.0.1",
      "local_cache" : {
        "get_stats" : {
          "get_total_count" : 562,                    //从冷数据本地缓存查询数据的总次
```

```
数
          "get_hit_count" : 562,                                    //从冷数据本地缓存查询数据命中的
次数
          "get_miss_count" : 0,                                     //从冷数据本地缓存查询数据未命中
的次数
          "get_total_ns" : 43849200,                                //从冷数据本地缓存查询数据的总时
长
          "get_avg_ns" : 78023                                     //从冷数据本地缓存查询数据的平均
时长
        },
        "load_stats" : {
          "load_count" : 2,                                         //加载冷数据本地缓存的次数
          "load_total_ms" : 29,                                     //加载冷数据本地缓存的总时长
          "load_avg_ms" : 14,                                       //加载冷数据本地缓存的平均时长
          "load_fail_count" : 0,                                    //加载冷数据本地缓存的失败次数
          "load_overflow_count" : 0                                 //加载冷数据本地缓存时超过缓存
池的次数
        },
        "reload_stats" : {
          "reload_count" : 0,                                       //重新生成冷数据本地缓存的次数
          "reload_total_ms" : 0,                                    //重新生成冷数据本地缓存的总时长
          "reload_avg_ms" : 0,                                      //重新生成冷数据本地缓存的平均时
长
          "reload_fail_count" : 0                                   //重新生成冷数据本地缓存的失败
次数
        },
        "init_stats" : {
          "init_count" : 0,                                         //初始化冷数据本地缓存的次数
          "init_total_ms" : 0,                                      //初始化冷数据本地缓存的总时长
          "init_avg_ms" : 0,                                        //初始化冷数据本地缓存的平均时长
          "init_fail_count" : 0                                     //初始化冷数据本地缓存的失败次
数
        }
      }
    }
  }
}
```

配置参数

配置项	类型	单位	取值范围	scope	是否可动态修改	作用
low_cost.local_cache.max.capacity	Integer	无	10-5000，默认值500。	node	是	节点上，能够打开的冷数据缓存的最大数量。（每个 shard 对应一个缓存对象） 说明 • 堆内存使用率一直很高，可以尝试降低该值。

配置项	类型	单位	取值范围	scope	是否可动态修改	作用
						• 如果查询冷数据本地缓存的相关统计指标中load_overflow_count数值一直持续快速增加，建议调大该值。
index.low_cost.local_cache.threshold	Integer	%	0-100，默认值50。	index	是	启用冷数据本地缓存的阈值。 说明 • 如果date类型字段的占比小于此值。启用冷数据本地缓存date类型字段。否则不使用。 • 如果当前索引的date类型字段占据当前索引的绝大部分数据量，不建议使用此功能。
index.low_cost.local_cache.evict_time	字符串	天	1d-365d，默认值30d。	index	是	冷数据本地缓的淘汰时间。根据index.frozen_date（冻结成功的时间）判定。 说明 • 如果为旧集群冻结的索引，无index.frozen_date，则根据索引创建时间判定。 • 建议根据磁盘用量调整淘汰时长，节省磁盘空间。

修改参数

- 修改 low_cost.local_cache.max.capacity

```
PUT _cluster/settings
 {
   "persistent": {
     "low_cost.local_cache.max.capacity":1000
   }
 }
```

- 修改 index.low_cost.local_cache.threshold

```
PUT es_write_pref2-00000000021/_settings
 {
 "index.low_cost.local_cache.threshold":20
 }
```

- 修改 index.low_cost.local_cache.evict_time

```
PUT es_write_pref2-00000000021/_settings
 {
```

```
"index.low_cost.local_cache.evict_time":"7d"
}
```

7.2.5 监控 OBS 操作

为了更清晰的了解到存算分离的插件在 OBS 中的操作，新增了 OBS 实时速率的统计指标，并且将实时速率记录到系统索引中，帮助您实时了解 OBS 的使用情况。

前提条件

此特性仅支持在 2023 年 03 月后创建的 7.6.2 和 7.10.2 版本 Elasticsearch 集群以及 OpenSearch 集群。

功能介绍

- GET _frozen_stats/obs_rate 接口：查询 OBS 操作的实时速率。
- 新增系统索引.freeze_obs_rate-YYYY.mm.dd：存放 OBS 操作实时速率和 OBS 操作的数据，以便于了解整个 OBS 的操作趋势。
- 新增配置 low_cost.obs_rate_index.evict_time ：控制“.freeze_obs_rate-YYYY.mm.dd”索引的存放时间。

GET _frozen_stats/obs_rate 接口

- 计算方式：每 5 秒计算一次，查询前 5 秒内的平均 OBS 操作速率。
- 请求示例：

```
GET _frozen_stats/obs_rate
GET _frozen_stats/obs_rate/{nodeId}
```

“{nodeId}”为需要查询 OBS 操作速率的节点 ID。

- 响应示例：

```
{
  " _nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "elasticsearch",
  "nodes" : {
    "dflDvcSwTJ-fkiIlT2zE3A" : {
      "name" : "node-1",
      "transport_address" : "127.0.0.1:9300",
      "host" : "127.0.0.1",
      "ip" : "127.0.0.1",
      "update_time" : 1671777600482,                          // 当前统计值的更新时间。
      "obs_rate" : {
        "list_op_rate" : 0.0,                                 // obs list 操作的速率，单位：次/秒。
        "get_meta_op_rate" : 0.0,                             // obs get meta 操作的速率，单位：次/秒。
        "get_obj_op_rate" : 0.0,                              // obs get 操作的速率，单位：次/秒。
```

```
        "put_op_rate" : 0.0,                                    // obs put 操作的速率，单位：次/秒。
        "obs_total_op_rate" : 0.0,                               // obs 所有操作的速率，单位：次/秒。
        "obs_upload_rate" : "0.0 MB/s",                          // obs 上传数据的速率，MB/秒。
        "obs_download_rate" : "0.0 MB/s"                         // obs 下载数据的速率，MB/秒。
      }
    }
  }
}
```

系统索引

- 系统索引名称：“.freeze_obs_rate-YYYY.mm.dd” 。
- 示例：“.freeze_obs_rate-2023.01.23”

说明

该索引默认保留期限为 30 天。

配置项

配置项	类型	scope	是否可动态修改	说明
low_cost.obs_rate_index.evict_time	String	node	是	用于控制索引“.freeze_obs_rate-YYYY.mm.dd”的保留时间。 • 取值范围：1d~365d。 • 默认值：30d。 • 单位：天。

示例：修改“.freeze_obs_rate-YYYY.mm.dd”索引的保留时间。

```
PUT _cluster/settings
{
  "persistent": {
    "low_cost.obs_rate_index.evict_time":  "7d"
  }
}
```

7.3 流量控制 2.0

7.3.1 背景信息

特性介绍

云搜索服务支持流量控制特性，提供节点级别的流量控制功能，可提供单个节点基于黑白名单的访问限制、HTTPS 并发连接数限制、HTTP 最大连接数限制，基于节点内存的客户端写入流量反压控制，一键断流能力。同时也提供节点访问 IP 统计和 URL 的统计能力。每个功能配置独立的控制开关，默认关闭。所有参数配置为 null 可以恢复默认值。

开启客户端写入流量反压控制功能，会在节点堆内存过大时阻止大请求涌入节点占用内存，避免节点崩溃，减少节点不可用的风险。

- **HTTP/HTTPS 流控：**
  - HTTP/HTTPS 黑白名单设置 IP 和子网控制客户端 IP 访问，如果节点 IP 在黑名单中，则该客户端的连接将直接中断，节点不会处理任何请求。白名单规则优先于黑名单规则，如果客户端 IP 在黑白名单中都出现，客户端请求将不会被拒绝。
  - HTTP/HTTPS 并发连接数流控通过限制节点每秒中的 HTTP 连接总数来限制节点流量。
  - HTTP/HTTPS 新建连接数流控通过限制节点新建的连接数来限制节点流量。
- **内存流控：**

  基于节点堆内存使用量限制写入流量，将请求需要读取的内容反压在客户端，暂停请求的接收。同时尽力触发垃圾回收，按堆内存的空闲量继续读取请求。
- **请求采样统计：**

  可以记录客户端 IP 的访问和客户端的请求类型，用户可以基于统计值识别客户端 IP 的访问流量，分析当前客户端的写入和查询访问量。
- **一键断流：**

  可以切断节点的所有客户端连接，不包括 Kibana 访问和 CSS 后台运维、监控类接口。
- **流量控制：**

  提供单独的流量统计查看接口，记录当前客户端连接数以及客户端反压连接数，用户可以基于统计值评估流控配置阈值和衡量集群压力。
- **访问日志：**

  可以记录一段时间内节点接收的 HTTP/HTTPS 请求 URL 和 Body，用户可以基于访问日志信息分析当前的流量压力。

约束限制

- 目前仅 7.6.2 和 7.10.2 版本 Elasticsearch 集群支持流量控制特性。
- 2023 年 02 月后创建的 7.6.2 和 7.10.2 版本 ELasticsearch 集群仅支持流量控制 2.0 版本，之前创建的旧集群仅支持流量控制 1.0 版本，详情请见 7.4 流量控制 1.0。

7.3.2 HTTP/HTTPS 流控

通过在 Kibana 执行命令，可以开启或关闭集群的 HTTP/HTTPS 流控。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭 HTTP/HTTPS 流控。
   - **开启 HTTP/HTTPS 节点流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": true,
    "flowcontrol.http.allow": ["192.168.0.1/24", "192.168.2.1/24"],
    "flowcontrol.http.deny": "192.168.1.1/24",
    "flowcontrol.http.concurrent": 1000,
    "flowcontrol.http.newconnect": 1000,
    "flowcontrol.http.warmup_period": 0
  }
}
```

说明

当所有参数设置为 null 时，表示恢复配置默认值。

   - **关闭 HTTP/HTTPS 节点流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": false
  }
}
```

执行命令涉及的配置参数说明请见表 7-20。

表7-20 HTTP/HTTPS 流控的配置参数说明

配置名	类型	说明
flowcontrol.http.enabled	Boolean	HTTP/HTTPS 流控开关，默认关闭，开启会影响节点访问性能。 取值范围：true、false 默认值：false
flowcontrol.http.allow	List<String>	IP 地址访问白名单。 支持配置多个 IP 地址和掩码，或者 IP 地址列表形式，中间用英文逗号隔开。例如“xx.xx.xx.xx/24,xx.xx.xx.xx/24”或“xx.xx.xx.xx,xx.xx.xx.xx”形式。 默认值为空。

配置名	类型	说明
flowcontrol.http.deny	List<String>	IP 访问黑名单。 支持配置多个 IP 和掩码，或者 IP 列表形式，中间用英文逗号隔开。 默认值为空。
flowcontrol.http.concurrent	Integer	HTTP/HTTPS 请求的并发连接数阈值。 默认值：节点可用核数 * 400
flowcontrol.http.newconnect	Integer	HTTP/HTTPS 请求的每秒可以创建的新建连接数阈值。 默认值：节点可用核数 * 200
flowcontrol.http.warmup_period	Integer	HTTP/HTTPS 新建连接数达到最大速率的需要的时间，如果“flowcontrol.http.newconnect”配置为“100”且“flowcontrol.http.warmup_period”配置为“5000ms”，表示 5s 以后系统的新建连接数才可以达到每秒 100。 取值范围：0~10000 单位：ms 默认值：0

7.3.3 内存流控

背景信息

Elasticsearch 内部有熔断器机制，可以配置内存使用的阈值，当节点内存超过指定值时，会触发熔断，导致请求被中止或发生状态码为 429 的拒绝。而发生熔断拒绝的时间点是读取到整个客户端请求，该请求被节点完整读取，占用了一部分的堆内存。在请求被拒绝之前，为了防止请求完全进入节点，可以对客户端流量根据节点堆内存实时状态进行整流。

参数说明

在开启或关闭集群的内存流控时，执行命令涉及的配置参数如下：

表7-21 内存流控的配置参数说明

配置名	类型	说明
flowcontrol.memory.enabled	Boolean	内存流控开关，开启后持续监控内存使用情况。取值： • true

配置名	类型	说明
		• false（默认值）
flowcontrol.memory.heap_limit	String	限制节点全局堆内存的使用率。超过此值将进行流量反压。 取值范围：10%-100% 默认值：90%
flowcontrol.holding.in_flight_factor	Float	反压释放因子，原理类似于熔断器 network.breaker.inflight_requests.overhead 参数。内存达到限制值时，该值越大反压越强，写入流量将受限。 取值范围：≥0.5 默认值：1.0
flowcontrol.holding.max	TimeValue	每个请求最长的延迟时间，当延迟超过此值时可以设置断开该请求反压或断开请求链路。详见“flowcontrol.holding.max_strategy”配置。 取值范围：≥15s 默认值：60s
flowcontrol.holding.max_strategy	String	超过最大延迟时间后的运行策略。取值： • keep（默认值）：如果堆内存仍在高位，选择继续反压 - 何时执行请求仍由服务器根据实时内存自主决定。 • soft：如果堆内存仍在高位，也必须执行该请求。执行/拒绝权力交给 inFlight 熔断器。 • hard：如果堆内存仍在高位，丢弃该请求，同时断开该请求的客户端连接。
flowcontrol.memory.once_free_max	String	被暂停的请求队列一次性最大重新打开的内存，防止强压场景下短暂的低内存现象一次性冲挂集群。 取值范围：1%-50% 默认值：10%
flowcontrol.memory.nudges_gc	Boolean	写入压力过大时（1s 检查一次反压连接池，所有现有连接均被阻塞无法放开新的写入请求），是否尽力触发垃圾回收，保证写入稳定性。取值： • true（默认值） • false

说明

- “flowcontrol.memory.enabled”和“flowcontrol.memory.heap_limit”是最重要的2个参数配置。*enabled*是内存流控开关，*heap_limit*是指节点堆内存的阈值。
- “flowcontrol.memory.heap_limit”默认值90%是一个比较保守的阈值，即堆内存大于90%使用量时会停止读取客户端超过64KB的大请求，直至堆内存下降。如堆内存下降到85%，会开始允许最多一次读取5% × 堆内存最大值的客户端数据量。如果堆内存持续超过90%，则无法放开客户端连接的请求读取，此时会尝试触发GC算法进行垃圾回收，直到堆内存低于所设定的阈值。
- 日常使用时可以将“flowcontrol.memory.heap_limit”阈值设置为80%或以下，保证节点有一定的堆内存余量，供写入内存以外的行为使用，比如：Elasticsearch查询、Segment merge等。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录Kibana界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭内存流控。
   - 开启内存流控

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.memory.enabled": true,
         "flowcontrol.memory.heap_limit": "80%"
       }
     }
     ```

   - 关闭集群内存流控

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.memory.enabled": false
       }
     }
     ```

7.3.4 请求采样统计

背景信息

开启请求统计可以记录客户端IP的访问和客户端的请求类型，用户可以基于统计值识别客户端IP的访问流量，分析当前客户端的写入和查询访问量。

表7-22 请求统计的配置参数说明

配置名	类型	说明
flowcontrol.log.access.enabled	Boolean	是否统计最近访问ES集群的客户端IP地址及其请求数量统计。取值：

配置名	类型	说明
		• true • false（默认值）
flowcontrol.log.access.count	Integer	统计最近 IP 地址访问集群的客户端来源个数。 取值范围：0-100 默认值：10

说明

- IP 统计分为请求类型统计和日志记录开关。
- “flowcontrol.log.access.enabled”控制是否对客户端请求进行访问统计，会统计 bulk 写入和 search、msearch 查询请求的个数。

7.3.5 一键断流

一键断流可以切断节点上除运维接口外的所有连接，用于应对突发流量场景下的集群异常，达到快速恢复集群的目的。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭一键断流。
   - 开启一键断流

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.break.enabled": true
       }
     }
     ```

   - 关闭一键断流

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.break.enabled": false
       }
     }
     ```

7.3.6 访问统计、查看流量控制信息

流量控制提供单独的接口查看节点的流量控制情况。

操作步骤

1. 登录云搜索服务管理控制台。

2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令查询流量控制情况。
   - 查看所有节点的流量控制情况

     ```
     GET /_nodes/stats/filter/v2
     ```

   - 查看所有节点的流量控制详细情况

     ```
     GET /_nodes/stats/filter/v2?detail
     ```

   - 查看某个具体节点的流量控制情况

     ```
     GET /_nodes/{nodeId}/stats/filter/v2
     ```

     *{nodeId}*为需要查看流量控制的节点 ID。

     响应示例：

     ```
     {
       "_nodes" : {
         "total" : 1,
         "successful" : 1,
         "failed" : 0
       },
       "cluster_name" : "css-xxxx",
       "nodes" : {
         "d3qnVIpPTtSoadkV0LQEkA" : {
           "name" : "css-xxxx-ess-esn-1-1",
           "host" : "192.168.x.x",
           "timestamp" : 1672236425112,
           "flow_control" : {
             "http" : {
               "current_connect" : 52,
               "rejected_concurrent" : 0,
               "rejected_rate" : 0,
               "rejected_black" : 0,
               "rejected_breaker" : 0
             },
             "access_items" : [
               {
                 "remote_address" : "10.0.0.x",
                 "search_count" : 0,
                 "bulk_count" : 0,
                 "other_count" : 4
               }
             ],
             "holding_requests" : 0
           }
         }
       }
     }
     ```

表7-23 响应参数说明

参数名	说明
current_connect	节点实际的 HTTP 连接数据信息， 没有开启流控这个配置也会记录，等同于 GET /_nodes/stats/http 接口 current_open

参数名	说明
	值， 可以看到节点当前的客户端连接数。
rejected_concurrent	HTTP 流控开启生效，关闭后不清零，开启流控期间拒绝的并发连接数。
rejected_rate	HTTP 流控开启生效，关闭后不清零，开启流控期间拒绝的新建连接数。
rejected_black	HTTP 流控开启生效，关闭后不清零，配置黑名单拒绝的请求数。
rejected_breaker	一键断流开启，拒绝的新建连接数。
remote_address	IP 地址统计，基于配置值统计节点访问的 IP 地址和请求数量。
search_count	该客户端以_search、_msearch 访问的次数。
bulk_count	该客户端以_bulk 访问的次数。
other_count	该客户端其他请求的访问次数。

7.3.7 临时访问统计日志

背景信息

流量控制提供 2 种方式查看访问日志。

- 一种是提供单独的 API 开启和查看访问日志，API 参数配置记录访问日志时间和大小，访问日志内容通过 Rest 接口返回。
- 一种是通过日志打印的方式记录访问日志，开启后用户的访问日志会以文件的方式打印到后端日志中，用户通过查看日志文件查看访问日志。本章节将会介绍此种方式临时访问日志记录。

临时访问日志记录在开启或关闭访问日志时，执行命令涉及的配置参数如下：

表7-24 访问日志的配置参数说明

配置名	类型	说明
duration_limit	String	访问日志记录时间。 取值范围：10~120 单位：s 默认值：30
capacity_limit	String	访问日志记录大小。统计开启访问日志后记录的请求大小，当统计的大小大于该配置值，访问日志记录终止。

配置名	类型	说明
		取值范围：1~5 单位：MB 默认值：1

说明

“duration_limit”和“capacity_limit”只要有一个参数达到阈值，访问日志记录就会停止。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启访问日志。
   - 开启集群所有节点的访问日志

     ```
     PUT /_access_log?duration_limit=30s&capacity_limit=1mb
     ```
   - 开启集群中某一节点的访问日志

     ```
     PUT /_access_log/{nodeId}?duration_limit=30s&capacity_limit=1mb
     ```

     {nodeId}为需要开启访问日志的节点 ID。
4. 执行命令查看访问日志。
   - 查看集群所有节点的访问日志 API

     ```
     GET /_access_log
     ```
   - 查看集群中某一节点的访问日志 API

     ```
     GET /_access_log/{nodeId}
     ```

     *{nodeId}*为需要开启访问日志的节点 ID。

     响应示例：

     ```
     {
       "_nodes" : {
         "total" : 1,
         "successful" : 1,
         "failed" : 0
       },
       "cluster_name" : "css-flowcontroller",
       "nodes" : {
         "8x-ZHu-wTemBQwpcGivFKg" : {
           "name" : "css-flowcontroller-ess-esn-1-1",
           "host" : "10.0.0.98",
           "count" : 2,
           "access" : [
             {
               "time" : "2021-02-23 02:09:50",
               "remote_address" : "/10.0.0.98:28191",
               "url" : "/_access/security/log?pretty",
               "method" : "GET",
               "content" : ""
     ```

```
      },
      {
        "time" : "2021-02-23 02:09:52",
        "remote_address" : "/10.0.0.98:28193",
        "url" : "/_access/security/log?pretty",
        "method" : "GET",
        "content" : ""
      }
    ]
  }
 }
}
```

表7-25 响应参数说明

参数名	说明
name	节点名称。
host	节点对应的 IP 地址。
count	统计周期内，访问节点的请求数量。
access	统计周期内，访问节点的请求详情。参数说明请参见表 7-26。

表7-26 access

参数名	说明
time	记录请求时间。
remote_address	请求对应的源 IP 地址和端口。
url	请求的原始 URL
method	对应请求 Path 的方法。
content	对应请求的内容。

5. 执行命令删除访问日志。
   - 删除集群所有节点的访问日志 API

     ```
     DELETE /_access_log
     ```

   - 删除集群中某一节点的访问日志 API

     ```
     DELETE /_access_log/{nodeId}
     ```

     *{nodeId}*为需要开启访问日志的节点 ID。

7.3.8 开启访问日志记录到文件

流量控制功能支持集群访问日志记录，并写入后台日志文件中，可通过日志备份功能备份到 OBS 中进行查看，如下是开启访问日志记录到文件的命令：

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.log.file.enabled": true
  }
}
```

表7-27 参数说明

参数	类型	说明
flowcontrol.log.file.enabled	Boolean	是否将每个请求的日志详情记录到后台日志文件。取值： • true • false（默认值）

说明

- 开启访问日志记录到文件后，客户端访问集群节点，会记录为{集群名_access_log.log}文件，通过日志备份功能可以详细查看访问日志。
- 定位问题完成后，建议关闭此开关。

7.4 流量控制 1.0

7.4.1 背景信息

特性介绍

云搜索服务支持流量控制特性，提供节点级别的流量控制功能，可提供单个节点基于黑白名单的访问限制、HTTP 并发连接数限制、HTTP 最大连接数限制、基于请求 Path 的堆内存最大使用量流控能力、基于 CPU 最大占用率流控能力，一键断流能力，同时也提供节点访问 IP 统计和 URL 的采样统计能力。每个功能配置独立的控制开关，默认关闭。所有参数配置为 null 可以恢复默认值。

开启流控功能会使请求在入口处直接阻塞，可以缓解节点高并发场景下的集群压力，降低 P99 时延，减少节点不可用的风险。

- **HTTP/HTTPS 流控**
  - HTTP/HTTPS 黑白名单设置 IP 和子网控制客户端 IP 访问，如果节点 IP 在黑名单中，则该客户端的连接将直接中断，节点不会处理任何请求。白名单规则优先于黑名单规则，如果客户端 IP 在黑白名单中都出现，客户端请求将不会被拒绝。

  - HTTP/HTTPS 并发连接数流控通过限制节点每秒中的 HTTP 连接总数来限制节点流量。
  - HTTP/HTTPS 新建连接数流控通过限制节点新建的连接数来限制节点流量。
- **内存流控**

  基于节点堆内存使用量限制请求 Path，支持设置内存流控白名单、全局内存使用阈值和基于单个 Path 设置堆内存使用阈值。全局内存流控阈值优先于单个 Path 内存阈值，白名单配置 Path 不参与内存流控。
- **Path 全局免流控白名单**

  可以根据客户需要，设置 Path 白名单全局免流控，当用户需要自定义插件时，可适当配置。
- **请求采样统计**

  可以记录客户端 IP 的访问数量和采样用户的请求 Path，用户可以基于统计值识别客户端 IP 的访问流量和分析哪些请求 Path 访问量大。
- **流量控制**

  提供单独的流量统计查看接口，记录触发流控的数量，用户可以基于统计值评估流控配置阈值和衡量集群压力。
- **访问日志**

  可以记录一段时间内节点接收的 HTTP/HTTPS 请求 URL 和 Body，用户可以基于访问日志信息分析当前的流量压力。
- **CPU 流控**

  基于节点配置的最大 CPU 占用率来限制节点访问流量。
- **一键断流**

  可以切断节点的所有访问流量，不包括 kibana 访问和 elasticsearch monitor 类接口。

约束限制

- 目前仅 7.6.2 和 7.10.2 版本支持流量控制特性。
- 开启流量控制功能会消耗部分节点性能。
- 开启流控会直接拒绝超过阈值的用户请求。
- 内存流控和 CPU 流控都是基于请求 Path 的流控，Path 长度和个数不应该配置过多，否则影响集群性能。

7.4.2 HTTP/HTTPS 流控

背景信息

通过在 Kibana 执行命令，可以开启或关闭集群的 HTTP/HTTPS 流控。执行命令涉及的配置参数如下：

表7-28 HTTP/HTTPS 流控的配置参数说明

配置名	类型	说明
flowcontrol.http.enabled	Boolean	HTTP/HTTPS 流控开关，默认关闭，开启会影响节点访问性能。 取值范围：true、false 默认值：false
flowcontrol.http.allow	List<String>	IP 地址访问白名单。 支持配置多个 IP 地址和掩码，或者 IP 地址列表形式，中间用英文逗号隔开。例如“xx.xx.xx.xx/24,xx.xx.xx.xx/24”或“xx.xx.xx.xx,xx.xx.xx.xx”形式。 默认值为空。
flowcontrol.http.deny	List<String>	IP 访问黑名单。 支持配置多个 IP 和掩码，或者 IP 列表形式，中间用英文逗号隔开。 默认值为空。
flowcontrol.http.concurrent	Integer	HTTP/HTTPS 请求的并发连接数阈值。 默认值：节点可用核数 * 400
flowcontrol.http.newconnect	Integer	HTTP/HTTPS 请求的每秒可以创建的新建连接数阈值。 默认值：节点可用核数 * 200
flowcontrol.http.warmup_period	Integer	HTTP/HTTPS 新建连接数达到最大速率的需要的时间，如果“flowcontrol.http.newconnect”配置为“100”且“flowcontrol.http.warmup_period”配置为“5000ms”，表示 5s 以后系统的新建连接数才可以达到每秒 100。 取值范围：0~10000 单位：ms 默认值：0

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭 HTTP/HTTPS 流控。
   - **开启 HTTP/HTTPS 节点流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": true,
    "flowcontrol.http.allow": ["192.168.0.1/24", "192.168.2.1/24"],
    "flowcontrol.http.deny": "192.168.1.1/24",
    "flowcontrol.http.concurrent": 1000,
    "flowcontrol.http.newconnect": 1000,
    "flowcontrol.http.warmup_period": 0
  }
}
```

说明

当所有参数设置为 null 时，表示恢复配置默认值。

- **关闭 HTTP/HTTPS 节点流控**

```
PUT / cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": false
  }
}
```

7.4.3 内存流控

背景信息

Elasticsearch 内部有熔断器机制，可以配置内存使用的阈值，当节点内存超过指定值，触发熔断，请求操作终止。但是 Elasticsearch 在调用 API 时没有判断当前的堆内存使用量，如果在请求处理过程中计算，即使熔断也会造成堆内存的消耗，频繁熔断会导致节点不可用，同时熔断器不支持单个请求的熔断阈值配置。但是，当在 Rest 请求入口处设置堆内存使用限制时，可以阻断 API 的请求，达到保护节点的目的。而内存流控可以配置节点全局流控和基于单个请求 Path 的精细化内存控制，其中单个请求 Path 的流控通过配置请求 Path 和堆内存阈值，在请求处理前判断配置的堆内存阈值，超过阈值中断当前的请求 Path。

说明

- 开启内存流控会消耗部分请求性能。
- 开启内存流控会导致 Kibana 的部分 search 请求失败。
- ES 551 版本开启内存流控会导致_mget 请求被拦截，Kibana 访问异常，可以把_mget 请求加入请求白名单规避。

在开启或关闭集群的内存流控时，执行命令涉及的配置参数如下：

表7-29 内存流控的配置参数说明

配置名	类型	说明
flowcontrol.memory.enabled	Boolean	内存流控开关，默认关闭，开启内存流控对节点性能有些许影响。

配置名	类型	说明
		取值范围：true、false 默认值：false
flowcontrol.memory.allow_path	List<String>	内存流控白名单 Path。 配置的路径不参与内存流控，可以支持通配符配置。集群控制的查询类接口默认放通，不参与内存流控，避免内存达到阈值，不能查询集群信息。 例如： • "flowcontrol.memory.allow_path": "/index/_search", • "flowcontrol.memory.allow_path": "/index*/_search", • "flowcontrol.memory.allow_path": ["/index/_search", "/index1/_bulk"], 支持最大配置 10 个 Path，每个 Path 最大长度限制小于 32。 默认值为空。
flowcontrol.memory.heap_limit	String	限制节点全局堆内存的最大使用率。不能低于堆内存的 10%。 取值范围：10%-100% 默认值：90%
flowcontrol.memory.*.filter_path	String	配置需要进行内存流控的访问 Path，控制单个请求 Path 的堆内存使用阈值。 默认值 "**", 表示匹配所有的路径。如果只配置了单路径“flowcontrol.memory.heap_limit”，没有配置“flowcontrol.memory.*.filter_path”，表示除去白名单外的所有 path 都影响。白名单规则优先于单路径规则，如果一个路径同时配置了

配置名	类型	说明
		“flowcontrol.memory.allow_path”和“flowcontrol.memory.*.filter_path”，此请求路径会被允许。 例如同时配置了“flowcontrol.memory.allow_path”和“flowcontrol.memory.*.filter_path”，其中flowcontrol.memory.*.filter_path="abc/_search"且flowcontrol.memory.allow_path="abc/_search"，此种情况 abc/_search 将不被流控。 最大长度：32
flowcontrol.memory.*.heap_limit	String	配置请求 Path 的堆内存阈值，堆内存超过阈值触发流控。 取值范围：0%-100% 默认值：90%

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭内存流控。
   - **开启内存流控**

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.memory.enabled": true,
         "flowcontrol.memory.allow_path": "/index/_search",
         "flowcontrol.memory.heap_limit": "85%"
       }
     }
     ```

   - **开启单个请求 Path 的内存流控**

     基于单个索引和请求 Path 设置堆内存使用阈值，可以基于此规则做优先级调度。

     ```
     PUT /_cluster/settings
     {
       "persistent": {
     ```

```
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory": {
      "flowcontrol_search": {
        "filter_path": "index1/_search",
        "heap_limit": "50%"
      },
      "flowcontrol_bulk": {
        "filter_path": "index*/_bulk",
        "heap_limit": "50%"
      }
    }
  }
}
```

- **删除单个请求 Path 的内存流控配置**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory": {
      "flowcontrol_search": {
        "filter_path": null,
        "heap_limit": null
      }
    }
  }
}
```

- **关闭集群内存流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": false
  }
}
```

7.4.4 Path 全局免流控白名单

背景信息

在添加集群的 Path 全局免流控白名单时，执行命令涉及的配置参数如下：

表7-30 Path 全局免流控白名单的配置参数说明

配置名	类型	说明
flowcontrol.path.white_list	List<String>	Path 全局免流控白名单，配置的路径不参与内存、CPU 流控和一键断流，IP 流控除外。 支持最大配置 10 个 Path，每个 Path 最大长度限制小于 32。 默认值为空。

配置名	类型	说明
		说明 一般不建议配置此值，仅在自定义插件时根据业务需求配置。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令添加 Path 全局免流控白名单。

```
PUT _cluster/settings
{
  "persistent": {
    "flowcontrol.path.white_list": "xxxx"
  }
}
```

7.4.5 请求采样统计

背景信息

开启请求采样统计可以记录访问节点的 IP 地址和数量，同时可以采样请求的 Path，记录请求 URL 和 Body，用于获取访问量大的客户端 IP 地址和请求 Path。

在开启或关闭集群的请求采样统计时，执行命令涉及的配置参数如下：

表7-31 请求采样统计的配置参数说明

配置名	类型	说明
flowcontrol.statics.enabled	Boolean	请求采样统计开关。开启请求采样统计对节点性能会有影响。 取值范围：true、false 默认值：false
flowcontrol.statics.threshold	Integer	统计最近时间访问的请求数量。配置为 100，表示会统计出最近访问最多的 100 个 IP 地址和基于采样统计的访问最多的 100 个 URL。 最小值：10 最大值：1000 默认值：100
flowcontrol.statics.sample_frequency	Integer	Path 采样频率。配置为 100，表示每 100 个请求采样统计一次。 最小值：50

配置名	类型	说明
		默认值：100

说明

- IP 统计和 URL 采样统计基于访问时间缓存策略，节点会记录最近访问的 IP 和请求 URL，如果缓存空间达到设置的阈值（flowcontrol.statics.threshold 配置值），访问时间距离现在最久的记录将被清除掉。
- URL 采样统计当前基于 URL hash 值确认访问 Path 的一致性。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭请求采样统计。
   - **开启采样统计**

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.statics.enabled": true,
         "flowcontrol.statics.threshold": 100,
         "flowcontrol.statics.sample_frequency": 50
       }
     }
     ```

   - **关闭采样统计**

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.statics.enabled": false
       }
     }
     ```

7.4.6 流控控制

流量控制提供单独的接口查看节点的流量控制情况。

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令查询流量控制情况。
   - 查看所有节点的流量控制情况

     ```
     GET /_nodes/stats/filter
     ```

   - 查看某个具体节点的流量控制情况

     ```
     GET /_nodes/{nodeId}/stats/filter
     ```

     {nodeId}为需要查看流量控制的节点 ID。

   响应示例：

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-flowcontroller",
  "nodes" : {
    "ElBRNCMbTj6L1C-Wke-Dnw" : {
      "name" : "css-flowcontroller-ess-esn-1-1",
      "host" : "10.0.0.133",
      "timestamp" : 1613979513747,
      "flow_control" : {
        "transport" : {
          "concurrent_req" : 0,
          "rejected_concurrent" : 0,
          "rejected_new" : 0,
          "rejected_deny" : 0
        },
        "http" : {
          "concurrent req" : 0,
          "rejected concurrent" : 0,
          "rejected new" : 0,
          "rejected deny" : 0
        },
        "memory" : {
          "memory allow" : 41,
          "memory_rejected" : 0
        },
        "cpu": {
          "rejected_cpu" : 0
        }
        "ip_address" : [
          {
            "ip" : "/10.0.0.198",
            "count" : 453
          },
          {
            "ip" : "/198.19.49.1",
            "count" : 42
          }
        ],
        "url_sample" : [
          {
            "url" : "/*/_search?pretty=true",
            "method" : "GET",
            "remote_address" : "/10.0.0.198:16763",
            "count" : 1
          }
        ]
      }
  }
}
```

返回值以 Node 级别分开，http 记录并发和新建连接数据统计，memory 记录内存流控统计，ip_address 记录最近最多访问的客户端 IP，url_sample 记录采样的最近最多请求 URL。cpu 记录 CPU 流控统计。

表7-32 响应参数说明

参数名	说明
concurrent_req	节点实际的 TCP 连接数据信息， 没有开启流控这个配置也会记录，参考 GET /_nodes/stats/http 接口 current_open 值，但是会比这个值小，这里忽略了白名单 IP 和内部节点 IP。
rejected_concurrent	HTTP 流控开启生效，关闭后不清零，开启流控期间拒绝的并发连接数。
rejected_new	HTTP 流控开启生效，关闭后不清零，开启流控期间拒绝的新建连接数。
rejected_deny	HTTP 流控开启生效，关闭后不清零，配置黑名单拒绝的请求数。
memory_allow	内存流控开启生效，关闭后不清零，内存流控允许的请求数，触发内存流控后允许的请求数量，allow_path 白名单中通过请求不会被记录，如果 allow_path 配置为“**”，所有请求都不会被记录。
memory_rejected	内存流控开启生效，关闭后不清零，内存流控拒绝的请求数，触发内存流控后拒绝的请求数量，allow_path 白名单中通过请求不会被记录，如果 allow_path 配置为“**”，所有请求都不会被记录。
rejected_cpu	CPU 流控开启生效，关闭后不清零，超过 CPU 流控阈值拒绝的请求数。
ip_address	IP 地址统计，基于配置值统计节点访问的 IP 地址和请求数量。参数说明请参见表 7-33。
url_sample	请求 Path 采样统计，基于配置时间和采样间隔统计相同请求 URL 数量。参数说明请参见表 7-34。

表7-33 ip_address

参数名	说明
ip	访问节点的源 IP 地址。
method	对应 IP 地址的访问次数统计。

表7-34 url_sample

参数名	说明
url	请求的采样统计，记录访问节点的请求 URL。
method	对应请求 Path 的方法。
remote_address	请求对应的源 IP 地址和端口。
count	对应请求 Path 的采样统计次数。

7.4.7 访问日志

背景信息

流量控制提供两种方式查看访问日志。

- 一种是提供单独的 API 开启和查看访问日志，API 参数配置记录访问日志时间和大小，访问日志内容通过 Rest 接口返回。
- 一种是通过日志打印的方式记录访问日志，开启后用户的访问日志会以文件的方式打印到后端日志中，用户通过查看日志文件查看访问日志。

开启访问日志会影响集群性能。

在开启或关闭访问日志时，执行命令涉及的配置参数如下：

表7-35 访问日志的配置参数说明

配置名	类型	说明
duration_limit	String	访问日志记录时间。 取值范围：10~120 单位：s 默认值：30
capacity_limit	String	访问日志记录大小。统计开启访问日志后记录的请求大小，当统计的大小大于该配置值，访问日志记录终止。 取值范围：1~5 单位：MB 默认值：1

说明

duration_limit 和 capacity_limit 只要有一个参数达到阈值，访问日志记录就会停止。

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启访问日志。
   - **开启集群所有节点的访问日志**

     ```
     PUT /_access_log?duration_limit=30s&capacity_limit=1mb
     ```
   - **开启集群中某一节点的访问日志**

     ```
     PUT /_access_log/{nodeId}?duration_limit=30s&capacity_limit=1mb
     ```

     {nodeId}为需要开启访问日志的节点 ID。
4. 执行命令查看访问日志 API。
   - 查看集群所有节点的访问日志 API

     ```
     GET /_access_log
     ```
   - 查看集群中某一节点的访问日志 API

     ```
     GET /_access_log/{nodeId}
     ```

     {nodeId}为需要开启访问日志的节点 ID。

   响应示例：

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-flowcontroller",
  "nodes" : {
    "8x-ZHu-wTemBQwpcGivFKg" : {
      "name" : "css-flowcontroller-ess-esn-1-1",
      "host" : "10.0.0.98",
      "count" : 2,
      "access" : [
        {
          "time" : "2021-02-23 02:09:50",
          "remote_address" : "/10.0.0.98:28191",
          "url" : "/_access/security/log?pretty",
          "method" : "GET",
          "content" : ""
        },
        {
          "time" : "2021-02-23 02:09:52",
          "remote_address" : "/10.0.0.98:28193",
          "url" : "/_access/security/log?pretty",
          "method" : "GET",
          "content" : ""
        }
      ]
    }
  }
}
```

表7-36 响应参数说明

参数名	说明
name	节点名称。
host	节点对应的 IP 地址。
count	统计周期内，访问节点的请求数量。
access	统计周期内，访问节点的请求详情。参数说明请参见表 7-37。

表7-37 access

参数名	说明
time	记录请求时间。
remote_address	请求对应的源 IP 地址和端口。
url	请求的原始 URL
method	对应请求 Path 的方法。
content	对应请求的内容。

5. 执行命令开启或关闭访问日志文件记录。

   访问日志功能提供配置开关可以记录用户的所有访问日志。日志默认会记录到后台的 acces_log.log 日志文件中。日志文件单个文件最大支持 250M，最多保存 5 个文件。访问日志文件可以通过日志备份到 OBS 中查看。

   - **开启访问日志文件记录**

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.accesslog.enabled": true
        }
     }
     ```

   - **关闭访问日志文件记录**

     ```
     PUT /_cluster/settings
     {
       "persistent": {
         "flowcontrol.accesslog.enabled": false
        }
     }
     ```

7.4.8 CPU 流控

背景信息

CPU 流控可以基于当前节点的 CPU 占用率实现流量控制。

CPU 流控通过配置节点的最大 CPU 占用率来避免流量冲击下节点掉线风险，可以基于流量阈值预估 CPU 占用率最大值。当节点 CPU 超过配置阈值后，CPU 流控会丢弃节点请求，达到保护集群的目的，节点内流量和 elasticsearch monitor 类接口不会被流控。

在开启或关闭 CPU 流控时，执行命令涉及的配置参数如下：

表7-38 CPU 流控的配置参数说明

配置名	类型	说明
flowcontrol.cpu.enabled	Boolean	CPU 流控开关，开启会影响节点访问性能。 • 取值范围：true、false • 默认值：false
flowcontrol.cpu.percent_limit	Integer	节点最大 CPU 占用率配置。 • 取值范围：0~100 • 默认值：90
flowcontrol.cpu.allow_path	List	CPU 流控白名单，CPU 流控基于请求 path 做流控，配置 allow_path 白名单，请求将不参与 CPU 流控。 • 默认值为空。 • 单 path 最大支持 32 个字符，最多支持配置 10 个请求路径，支持通配符配置。 例如，配置“auto_*/_search”将不限制所有的 auto_前缀索引的 search 请求。
flowcontrol.cpu.*.filter_path	String	配置需要进行 CPU 流控的访问 Path，控制单个请求 Path 的 CPU 使用阈值。 • 最大长度 32。 例如： `"flowcontrol.cpu.search.filter_path": "/index/_search", "flowcontrol.cpu.search.limit": 60,` • 默认值 "**", 表示匹配所有的路径。如果只配置了单路径 limit，没有配置 filter_path，表示除去白名单外的所有 path 都影响。白名单规则优先于单路径规则，如果一个路径同时配置了 allow_path 和 filter_path, 此请求路径会被允许。

配置名	类型	说明
		例如，同时配置了 filter_path 和 allow_path，其中 filter_path="abc/_search", allow_path="abc/_search"，此种情况 abc/_search 将不被流控。
flowcontrol.cpu.*.limit	Integer	配置请求 Path 的 CPU 阈值，CPU 超过阈值触发流控。 • 取值范围：0~100 • 默认值：90

操作步骤

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭 CPU 流控。
   - **开启 CPU 流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.cpu.enabled": true,
    "flowcontrol.cpu.percent_limit": 80,
    "flowcontrol.cpu.allow path": ["index/_search"]
  }
}
```

   - **关闭 CPU 流控**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.cpu.enabled": false
  }
}
```

7.4.9 一键断流

一键断流可以切断节点上除运维接口外的所有流量，用于应对突发流量场景下的集群异常，达到快速恢复集群的目的。

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面选择目标集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行命令开启或关闭一键断流。
   - **开启一键断流**

```
PUT /_cluster/settings
{
  "persistent": {
```

```
    "flowcontrol.break.enabled": true
  }
}
```

- **关闭一键断流**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.break.enabled": false
  }
}
```

7.5 大查询隔离

7.5.1 背景信息

大查询隔离特性针对查询请求进行独立管理，将高内存、长耗时的查询请求进行隔离，保证节点内存安全。在节点堆内存过高使用时，触发中断控制程序，根据选择的中断策略将其中一条大查询请求进行中断，取消其正在运行的查询任务。

大查询隔离同时设置了全局查询超时特性，用户可实时配置所有查询请求的超时时间，中断超时查询请求。

说明

目前仅 7.6.2 和 7.10.2 版本支持大查询隔离特性。

7.5.2 操作步骤

大查询隔离特性和全局超时特性默认关闭，用户可根据需要实时配置，配置后立即生效。以下是详细的配置方法：

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择待配置大查询隔离的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 在 Kibana 的左侧导航中选择“Dev Tools”，执行如下命令开启大查询隔离和全局超时的特性开关。

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.enabled": true,
    "search.isolator.time.enabled": true
  }
}
```

以上两个开关是独立功能，分别具有以下独立的参数配置：

表7-39 大查询隔离和全局超时的参数配置

特性开关	配置参数	参数说明
search.isolator.enabled	search.isolator.memory.task.limit search.isolator.time.management	单个分片查询任务被定义为大查询任务的阈值。
	search.isolator.memory.pool.limit search.isolator.memory.heap.limit search.isolator.count.limit	触发隔离池内查询任务中断的阈值。 说明 参数“search.isolator.memory.heap.limit”定义了节点实际堆内存的使用限制，包括写入和查询等操作，超过限制时将选取隔离池内的大查询任务进行中断。
	search.isolator.strategy search.isolator.strategy.ratio	中断隔离池中某一条查询任务的选取策略。
search.isolator.time.enabled	search.isolator.time.limit	全局查询任务超时设置。

4. 大查询隔离和全局超时的分别具有独立的参数配置，可以根据实际场景执行不同的命令进行配置。
   - 单个分片查询任务被定义为大查询任务的阈值。

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.memory.task.limit": "50MB",
    "search.isolator.time.management": "10s"
  }
}
```

表7-40 参数说明

参数名	数据类型	说明
search.isolator.memory.task.limit	String	查询任务用于聚合等操作向节点申请的大内存，申请内存超过此阈值将进行隔离观察。 • 取值范围：0b~节点最大堆内存 • 默认值：50MB 说明 可以通过如下命令查询集群堆内存使用情况和最大值。 **GET _cat/nodes?&h=id,ip,port,r,ramPercent,ramCurrent,heapMax,heapCurrent**

参数名	数据类型	说明
search.isolator.time.management	String	查询任务创建至今的时长（即开始使用集群资源进行查询），超过阈值将被隔离观察。 • 取值范围：≥ 0ms • 默认值：10s

- 触发隔离池内查询任务中断的阈值。

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.memory.pool.limit": "50%",
    "search.isolator.memory.heap.limit": "90%",
    "search.isolator.count.limit": 1000
  }
}
```

表7-41 参数说明

参数名	数据类型	说明
search.isolator.memory.pool.limit	String	当前节点最大堆内存百分比，当隔离池所有大查询任务申请的内存超过此阈值将触发中断控制程序，取消执行隔离池其中一条大查询任务。 • 取值范围：0.0~100.0% • 默认值：50%
search.isolator.memory.heap.limit	String	当前节点堆内存的实际使用阈值，当节点堆内存使用超过阈值百分比时触发中断控制程序，取消执行隔离池其中一条大查询任务。 • 取值范围：0.0~100.0% • 默认值：90%
search.isolator.count.limit	Integer	当前节点隔离池的大查询任务数阈值，被观察的查询任务数超过此阈值将触发中断控制程序，不再接受新的大查询。如果继续触发大查询请求，则直接取消该请求。 • 取值范围：10~50000 • 默认值：1000

说明

根据业务设置“search.isolator.memory.pool.limit”，“search.isolator.count.limit”参数时，可结合“search.isolator.memory.task.limit”，“search.isolator.time.management”两个参数控制查询任务进入到隔离池的数量。

- 中断隔离池中某一条查询任务的选取策略。

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.strategy": "fair",
    "search.isolator.strategy.ratio": "0.5%"
  }
}
```

参数名	数据类型	说明
search.isolator.strategy	String	触发中断控制程序时大查询选取的策略。根据策略选取一条查询进行中断。 说明 大查询隔离池每秒检查一次，直至堆内存下降到安全范围。 取值范围：fair、mem-first、time-first • mem-first 策略是指触发中断时，选取隔离池中堆内存使用最大的一条查询任务进行中断。 • time-first 策略是指触发中断时，选取隔离池中已运行时间最长的一条查询任务进行中断。 • fair 策略是综合考虑内存和时间两种因素，如果分片查询的堆内存申请大小相差不超过“最大堆内存乘 search.isolator.strategy.ratio”的大小，则认为时间较长的查询更应该中断。否则认为堆内存使用较大的查询更应该中断。 默认值：fair
search.isolator.strategy.ratio	String	fair 策略的阈值，仅当“search.isolator.strategy”值为“fair”生效。综合考虑大查询的运行时间及内存，当大查询任务内存相差不超过此阈值时，考虑选取运行时间长的大查询进行中断。当查询内存相差超过此阈值时，选取大内存查询任务进行中断。 • 取值范围：0.0-100.0% • 默认值：1%

- 全局查询任务超时设置。

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.time.limit": "120s"
```

```
  }
}
```

参数名	数据类型	说明
search.isolator.time.limit	String	当全局查询超时功能开启时，所有已创建的查询任务超过此时长将被取消执行。 • 取值范围：≥ 0ms • 默认值：120s

7.6 索引监控

7.6.1 背景信息

索引监控提供了丰富的监控指标，用于监控集群索引的运行状况和变化趋势，衡量业务使用情况，同时可以针对可能存在的风险及时处理，保障集群的稳定运行。

索引监控会采集索引的 stats 信息保存到集群的监控索引中(monitoring-eye-css-[yyyy-mm-dd])，索引默认保存一周。

目前仅 7.6.2 和 7.10.2 版本 ELasticsearch 集群支持索引监控能力。

7.6.2 启用索引监控

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择需要启用索引监控的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行如下命令打开索引监控开关：

```
PUT  cluster/settings
{
  "persistent": {
    "css.monitoring.index.enabled": "true"
  }
}
```

4. （可选）如果需要监控单个索引，可以在 Kibana 的“Dev Tools”执行如下命令：

```
PUT _cluster/settings
{
  "persistent": {
    "css.monitoring.index.enabled": "true",
    "css.monitoring.index.interval": "30s",
    "css.monitoring.index.indices": ["index_name"],
    "css.monitoring.history.duration": "3d"
  }
}
```

表7-42 配置参数说明

参数名	数据类型	说明
css.monitoring.index.enabled	Boolean	索引监控控制开关，设置为 true 将打开集群索引监控功能。 默认值：false
css.monitoring.index.interval	Time	索引监控的监控数据采集时间间隔。 最小值：1s 默认值：10s
css.monitoring.index.indices	String	索引监控的索引名称，默认监控所有索引，可以配置监控单个索引，也可以配置通配符监控某一类索引。 例如： • “"css.monitoring.index.indices": ["index_name"]”表示只监控“index_name”索引。 • “"css.monitoring.index.indices": ["log_*"]”表示监控以“log_”开头的索引。 • “"css.monitoring.index.indices": ["index1", "index2"]”表示监控“index1”,"index2"两个索引。 默认值：*（表示监控所有索引）
css.monitoring.history.duration	Time	监控数据存储的索引保留时间，默认保存一周。 最小值：1d 默认值：7d

须知

索引监控不会监控 monitoring-eye-css-* 开始的索引，避免使用的索引名称匹配到监控索引。

7.6.3 查看索引读写流量

索引监控提供了查询接口，方便查询一段时间内的索引读写流量。

前提条件

已创建好集群，且已 7.6.2 启用索引监控。

操作步骤

1. 登录云搜索服务管理控制台。

2. 在“集群管理”页面，选择已创建的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dev Tools”，执行如下命令查看索引读写流量：
   - 查看所有索引读写流量

     ```
     GET  /_cat/monitoring
     ```

   - 查看某一索引的读写流量

     ```
     GET  /_cat/monitoring/{indexName}
     ```

     {indexName}为需要查看读写流量的索引名称。
   - 查看索引不同时间段的读写流量

     ```
     GET _cat/monitoring?begin=1650099461000
     GET _cat/monitoring?begin=2022-04-16T08:57:41
     GET _cat/monitoring?begin=2022-04-16T08:57:41&end=2022-04-17T08:57:41
     ```

表7-43 请求参数说明

参数名	是否必选	说明
begin	否	查看监控的起始时间，UTC 时间，默认是当前时间的前 5 分钟。 支持时间格式：strict_date_optional_time|epoch_millis 默认值：当前时间减去 5 分钟。
end	否	查看监控的结束时间，UTC 时间，默认是当前时间。 支持时间格式： strict_date_optional_time|epoch_millis 默认值：当前时间。

说明

不支持查看系统索引，以“.”开头的索引是系统索引。

返回信息示例：

```
index   begin            end              status pri rep init unassign
docs.count docs.deleted store.size pri.store.size delete.rate indexing.rate
search.rate
test 2022-03-25T09:46:53.765Z 2022-03-25T09:51:43.767Z yellow  1   1  0    1
9        0     5.9kb       5.9kb        0/s          0/s         0/s
```

表7-44 返回信息的参数说明

参数名	说明
index	索引名称。
begin	查看监控数据的起始时间。
end	查看监控数据的结束时间。
status	查询监控时间间隔内的索引状态。

参数名	说明
pri	查询监控时间间隔内的索引的 shard 数量。
rep	查询监控时间间隔内的索引副本数量。
init	查询监控时间间隔内的索引的初始化数量。
unassign	查询监控时间间隔内的索引的未分配数量。
docs.count	查询监控时间间隔内的文档数量。
docs.deleted	查询监控时间间隔内的文档删除数量。
store.size	查询监控时间间隔内存储的索引大小。
pri.store.size	查询监控时间间隔内的索引主分片的大小。
delete.rate	监控时间间隔内的索引每秒删除数量。
indexing.rate	监控时间间隔内的索引每秒写入数量。
search.rate	监控时间间隔内的索引每秒查询数量。

7.6.4 查看索引监控

方便查看索引的监控信息，CSS 索引监控预置了 kibana 的 **Dashboard** 和 **Visualizations** 图表。

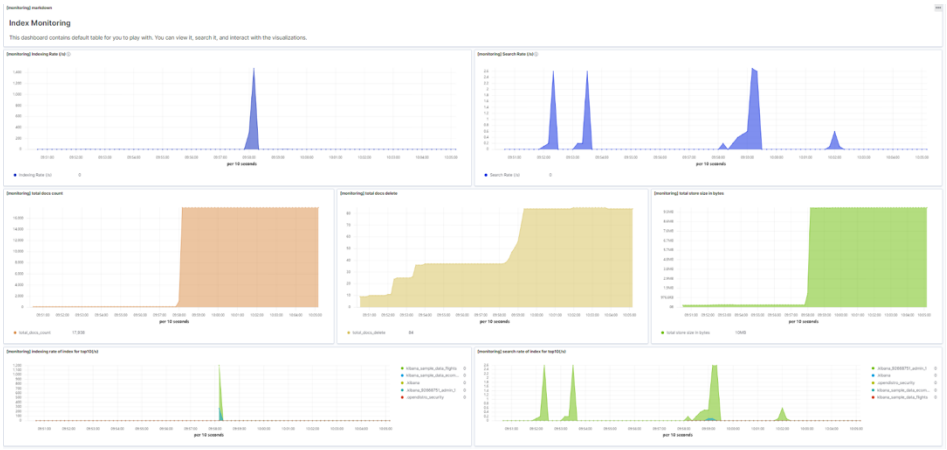
前提条件

已创建好集群，且已 7.6.2 启用索引监控。

查看 Dashboard 图表

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择已创建的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 单击左侧导航栏的“Dashboard”，打开 dashboard 界面。
4. 单击 “[Monitoring] Index monitoring Dashboard”查看预置 dashboard。

图7-1 预置 dashboard 图表



预置 dashboard 展示集群的每秒读写数量和每秒读写数量前 10 的索引情况。

表7-45 预置图表说明

图表名称	说明
[monitoring] markdown	Markdown 图表，简单说明 dashboard 表达的内容。
[monitoring] Indexing Rate (/s)	集群每秒写入文档数。
[monitoring] Search Rate (/s)	集群每秒查询次数。
[monitoring] indexing rate of index for top10	每秒写入文档数最多的 Top10 索引。
[monitoring] search rate of index for top10	每秒查询次数最多的 Top10 索引。
[monitoring] total docs count	集群总文档数量变化。
[monitoring] total docs delete	集群总的删除文档数量变化。
[monitoring] total store size in bytes	集群总文档所占用的存储空间变化。
[monitoring] indices store_size for top10	占用存储空间最多的 Top10 索引。
[monitoring] indices docs_count for top10	文档数量最多的 Top10 索引。
[monitoring] indexing time in millis of index for top10(ms)	单位时间内文档写入时延最大的 Top10 索引（ms）。
[monitoring] search query time in millis of index for top10(ms)	单位时间内索引查询时间最大的 Top10 索引（ms）。
[monitoring] segment count of	索引 Segment 数量最多的 Top10 索引。

图表名称	说明
index for top10	
[monitoring] segment memory in bytes of index for top10	索引 Segment 占用堆内存最大的 Top10 索引。

须知

索引监控不允许删除 monitoring-eye-css-* 的 index pattern，否则会导致监控图表异常。

自定义 Visualizations 图表

索引监控定期把 index/stats 信息存储到 monitoring-eys-css 索引中，通过使用 kibana 图表功能可以绘制自定义的图表。

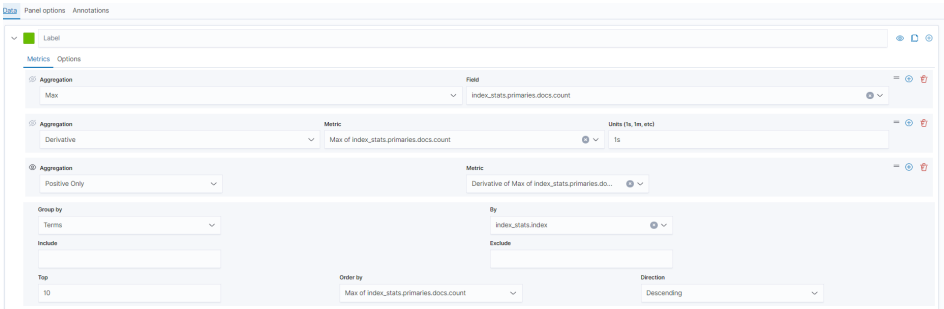
下面以文档数据为例介绍通过图表查看文档数量的变化趋势的操作步骤：

1. 登录云搜索服务管理控制台。
2. 选择已创建的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 选择左侧的“Visualize”进入图形列表。
4. 单击“Create visualization”，选择“TSVB” ，进入 TSVB 绘图页面。
5. 设置图表参数，查看 Visualizations。

   如图所示，在“Data”页签，配置参数。

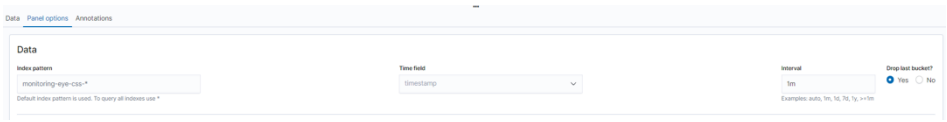
   - “Field”选择“index_stats.primaries.docs.count”表示主分片的文档数量。
   - “Aggreation”选择“Derivative”表示聚合桶之间的差异，“Units”设置为“1s”表示使用每秒代替速率。
   - “Aggreation”选择“Positive Only”避免数字重置后出现负数。
   - 当需要区分不同索引的情况时，则将“Group by”设置为“Terms”，“By”设置为“index_stats.index”，最终得到的聚合结果将按照索引名称区分。

图7-2 TSVB 绘图页面



- 当需要查看不同时间段的数据时，则需要将时间聚合间隔设置好，否则将导致数据显示不完整。在“Panel options”页签，将“Interval”设置为“1m”或“30m”，即可调整“timestamp”的时间间隔。

图7-3 设置时间间隔



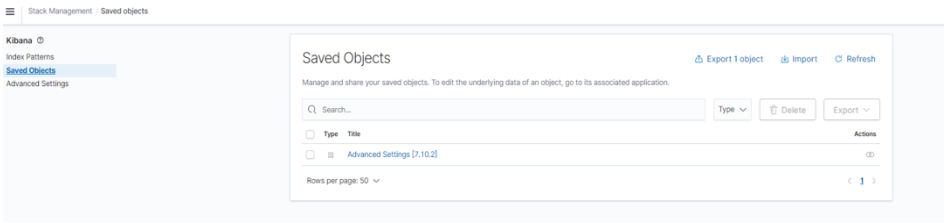
导入索引监控图表

kibana 提供图表的导入和导出功能，如果出现索引监控图表不显示，可以在 kibana 重新导入图表加载监控视图。

下面介绍如何在 kibana 中导入图表：

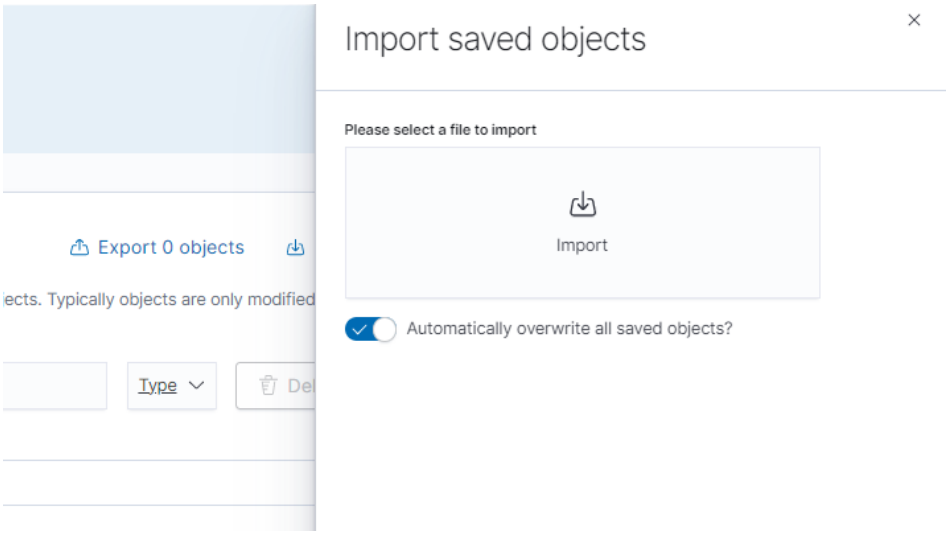
1. 参考 7.6.5 kibana-monitor 创建“monitoring-kibana.ndjson”文件。
2. 登录 kibana，选择“Management > Stack Management > Saved objects ”。

图7-4 选择 Saved objects



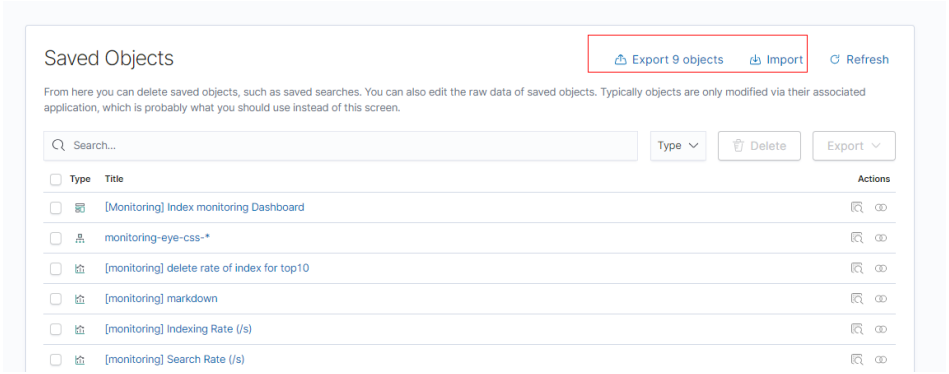
3. 单击“Import ”，选择上传 1 中创建的“monitoring-kibana.ndjson”文件。

图7-5 上传文件



4. 上传完成选择 done，索引监控图表即导入成功。

图7-6 索引监控图表导入成功



7.6.5 kibana-monitor

kibana-monitor 配置文件内容如下。

建议保存为“monitoring-kibana.ndjson”文件。

```
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] segment memory in bytes of index for
top10","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] segment
memory in bytes of index for
top10\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-11e7-
af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-469d-
11e7-af02-
69e470af7417\",\"color\":\"#68BC00\",\"split mode\":\"terms\",\"split color mode\":
\"kibana\",\"metrics\":[{\"id\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"max\",\"field\":\"index stats.total.segments.memory in by
tes\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"bytes\",\"c
```

```
hart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"stacked\":\"n
one\",\"label\":\"segments memory in bytes
\",\"type\":\"timeseries\",\"terms_field\":\"index_stats.index\",\"terms_order_by\"
:\"61ca57f2-469d-11e7-af02-
69e470af7417\"}],\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-
css-
*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"ax
is_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"tooltip_mode\":\"show_all
\",\"default_index_pattern\":\"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false}}"},"id":"3ae5d820
-6628-11ed-8cd7-
973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIwNiwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] segment count of index for
top10","uiStateJSON":"{}","version":1,"visState":"{\"aggs\":[],\"params\":{\"axis_f
ormatter\":\"number\",\"axis_position\":\"left\",\"axis_scale\":\"normal\",\"defaul
t_index_pattern\":\"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"filter\":{\"language\":\"kuery\",\"query\
":\"\"},\"id\":\"61ca57f0-469d-11e7-af02-
69e470af7417\",\"index pattern\":\"monitoring-eye-css-
*\",\"interval\":\"\",\"isModelInvalid\":false,\"series\":[{\"axis position\":\"rig
ht\",\"chart type\":\"line\",\"color\":\"rgba(231,102,76,1)\",\"fill\":0.5,\"format
ter\":\"number\",\"id\":\"61ca57f1-469d-11e7-af02-
69e470af7417\",\"label\":\"segment count of index for
top10\",\"line width\":1,\"metrics\":[{\"field\":\"index stats.total.segments.count
\",\"id\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"max\"}],\"point_size\":1,\"separate_axis\":0,\"split_colo
r_mode\":\"kibana\",\"split_mode\":\"terms\",\"stacked\":\"none\",\"terms_field\":\
"index_stats.index\",\"terms_order_by\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"timeseries\"}],\"show_grid\":1,\"show_legend\":1,\"time_f
ield\":\"timestamp\",\"tooltip_mode\":\"show_all\",\"type\":\"timeseries\"},\"title
\":\"[monitoring] segment count of index for
top10\",\"type\":\"metrics\"}"},"id":"45d571c0-6626-11ed-8cd7-
973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIwNywyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring]
markdown","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring]
markdown\",\"type\":\"markdown\",\"params\":{\"fontSize\":12,\"openLinksInNewTab\":
false,\"markdown\":\"### Index Monitoring \\nThis dashboard contains default table
for you to play with. You can view it, search it, and interact with the
visualizations.\"},\"aggs\":[]}"},"id":"b2811c70-a5f1-11ec-9a68-
ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIwOCwyXQ=="}
{"attributes":{"description":"number of document being indexing for primary and
replica
shards","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"title":"[monitoring]
Indexing Rate
(/s)","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] Indexing
Rate (/s)\",\"type\":\"metrics\",\"params\":{\"id\":\"61ca57f0-469d-11e7-af02-
69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-469d-11e7-
af02-
69e470af7417\",\"color\":\"rgba(0,32,188,1)\",\"split_mode\":\"everything\",\"metri
cs\":[{\"id\":\"61ca57f2-469d-11e7-af02-
```

```
69e470af7417\",\"type\":\"max\",\"field\":\"indices_stats._all.total.indexing.index
_total\"},{\"unit\":\"1s\",\"id\":\"fed72db0-a5f8-11ec-aa10-
992297d21a2e\",\"type\":\"derivative\",\"field\":\"61ca57f2-469d-11e7-af02-
69e470af7417\"},{\"unit\":\"\",\"id\":\"14b66420-a5f9-11ec-aa10-
992297d21a2e\",\"type\":\"positive_only\",\"field\":\"fed72db0-a5f8-11ec-aa10-
992297d21a2e\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"nu
mber\",\"chart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"sta
cked\":\"none\",\"label\":\"Indexing Rate
(/s)\",\"type\":\"timeseries\",\"split_color_mode\":\"rainbow\",\"hidden\":false}],
\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-css-
*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"ax
is_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"default_index_pattern\":\
"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"legend_position\
":\"bottom\"},\"aggs\":[]}"},"id":"de4f8ab0-a5f8-11ec-9a68-
ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIwOSwyXQ=="}
{"attributes":{"description":"number of search request being executed in primary
and replica
shards","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"title":"[monitoring]
Search Rate
(/s)","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] Search
Rate (/s)\",\"type\":\"metrics\",\"params\":{\"id\":\"61ca57f0-469d-11e7-af02-
69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-469d-11e7-
af02-
69e470af7417\",\"color\":\"rgba(0,33,224,1)\",\"split mode\":\"everything\",\"metri
cs\":[{\"id\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"max\",\"field\":\"indices_stats._all.total.search.query_t
otal\"},{\"unit\":\"1s\",\"id\":\"b1093ac0-a5f7-11ec-aa10-
992297d21a2e\",\"type\":\"derivative\",\"field\":\"61ca57f2-469d-11e7-af02-
69e470af7417\"},{\"unit\":\"\",\"id\":\"c17db930-a5f7-11ec-aa10-
992297d21a2e\",\"type\":\"positive_only\",\"field\":\"b1093ac0-a5f7-11ec-aa10-
992297d21a2e\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"nu
mber\",\"chart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"sta
cked\":\"none\",\"split_color_mode\":\"rainbow\",\"label\":\"Search Rate
(/s)\",\"type\":\"timeseries\",\"filter\":{\"query\":\"\",\"language\":\"kuery\"}}]
,\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-css-
*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"ax
is_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"default_index_pattern\":\
"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"legend_position\
":\"bottom\"},\"aggs\":[]}"},"id":"811df7a0-a5f8-11ec-9a68-
ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIxMCwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] total docs
count","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] total
docs count\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-
11e7-af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-
469d-11e7-af02-
69e470af7417\",\"color\":\"rgba(218,139,69,1)\",\"split_mode\":\"everything\",\"spl
it_color_mode\":\"kibana\",\"metrics\":[{\"unit\":\"\",\"id\":\"61ca57f2-469d-11e7-
af02-
69e470af7417\",\"type\":\"max\",\"field\":\"indices_stats._all.total.docs.count\"}]
,\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"number\",\"chart_t
```

ype\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"stacked\":\"none\",\"label\":\"total_docs_count\",\"type\":\"timeseries\"}],\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-css-*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"axis_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"tooltip_mode\":\"show_all\",\"default_index_pattern\":\"monitoring-eye-css-*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"legend_position\":\"bottom\"}}"},"id":"eea89780-664b-11ed-8cd7-973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-01T12:41:01.165Z","version":"WzIxMSwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"title":"[monitoring] total docs delete","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] total docs delete\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-11e7-af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-469d-11e7-af02-69e470af7417\",\"color\":\"rgba(214,191,87,1)\",\"split_mode\":\"everything\",\"split_color_mode\":\"kibana\",\"metrics\":[{\"id\":\"61ca57f2-469d-11e7-af02-69e470af7417\",\"type\":\"max\",\"field\":\"indices_stats._all.total.docs.deleted\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"number\",\"chart type\":\"line\",\"line width\":1,\"point size\":1,\"fill\":0.5,\"stacked\":\"none\",\"label\":\"totol docs delete\",\"type\":\"timeseries\",\"hidden\":false}],\"time_field\":\"timestamp\",\"index pattern\":\"monitoring-eye-css-*\",\"interval\":\"\",\"axis position\":\"left\",\"axis formatter\":\"number\",\"axis scale\":\"normal\",\"show legend\":1,\"show grid\":1,\"tooltip mode\":\"show all\",\"default index pattern\":\"monitoring-eye-css-*\",\"default timefield\":\"timestamp\",\"isModelInvalid\":false,\"drop last bucket\":1,\"legend_position\":\"bottom\"}}"},"id":"cfbb4e20-664c-11ed-8cd7-973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-01T12:41:01.165Z","version":"WzIxMiwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"title":"[monitoring] total store size in bytes","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] total store size in bytes\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-11e7-af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-469d-11e7-af02-69e470af7417\",\"color\":\"#68BC00\",\"split_mode\":\"everything\",\"split_color_mode\":\"kibana\",\"metrics\":[{\"id\":\"61ca57f2-469d-11e7-af02-69e470af7417\",\"type\":\"max\",\"field\":\"indices_stats._all.total.store.size_in_bytes\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"bytes\",\"chart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"stacked\":\"none\",\"label\":\"total store size in bytes\",\"type\":\"timeseries\"}],\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-css-*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"axis_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"tooltip_mode\":\"show_all\",\"default_index_pattern\":\"monitoring-eye-css-*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"legend_position\":\"bottom\",\"background_color_rules\":[{\"id\":\"7712e550-664f-11ed-8b5d-8db37e5b4cc4\"}],\"bar_color_rules\":[{\"id\":\"77680a30-664f-11ed-8b5d-8db37e5b4cc4\"}]}}"},"id":"c7f72ae0-664e-11ed-8cd7-973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-01T12:41:01.165Z","version":"WzIxMywyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"

```
title":"[monitoring] indexing rate of index for
top10(/s)","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring]
indexing rate of index for
top10(/s)\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-
11e7-af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-
469d-11e7-af02-
69e470af7417\",\"color\":\"#68BC00\",\"split_mode\":\"terms\",\"metrics\":[{\"id\":
\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"max\",\"field\":\"index_stats.total.indexing.index_total\
"},{\"unit\":\"1s\",\"id\":\"541ed8f0-a5ee-11ec-aa10-
992297d21a2e\",\"type\":\"derivative\",\"field\":\"61ca57f2-469d-11e7-af02-
69e470af7417\"},{\"unit\":\"\",\"id\":\"67ec1f50-a5ee-11ec-aa10-
992297d21a2e\",\"type\":\"positive_only\",\"field\":\"541ed8f0-a5ee-11ec-aa10-
992297d21a2e\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"nu
mber\",\"chart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"sta
cked\":\"none\",\"label\":\"indexing_rate\",\"type\":\"timeseries\",\"split_filters
\":[{\"color\":\"#68BC00\",\"id\":\"81004200-a5ee-11ec-aa10-
992297d21a2e\",\"filter\":{\"query\":\"\",\"language\":\"kuery\"}}],\"filter\":{\"q
uery\":\"\",\"language\":\"kuery\"},\"terms_field\":\"index_stats.index\",\"terms_o
rder_by\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"terms_size\":\"10\",\"terms_direction\":\"desc\",\"split_color_mod
e\":\"rainbow\"}],\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-
css-
*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"ax
is_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"default_index_pattern\":\
"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"tooltip_mode\":\
"show_all\"}}"},"id":"943b3e00-a5ef-11ec-9a68-
ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIxNCwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] search rate of index for
top10(/s)","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring]
search rate of index for
top10(/s)\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-
11e7-af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"61ca57f1-
469d-11e7-af02-
69e470af7417\",\"color\":\"rgba(99,157,12,1)\",\"split_mode\":\"terms\",\"metrics\"
:[{\"id\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"max\",\"field\":\"index_stats.total.search.query_total\"}
,{\"unit\":\"1s\",\"id\":\"fdfdfad0-a5ef-11ec-aa10-
992297d21a2e\",\"type\":\"derivative\",\"field\":\"61ca57f2-469d-11e7-af02-
69e470af7417\"},{\"unit\":\"\",\"id\":\"0aaa26a0-a5f0-11ec-aa10-
992297d21a2e\",\"type\":\"positive_only\",\"field\":\"fdfdfad0-a5ef-11ec-aa10-
992297d21a2e\"}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"nu
mber\",\"chart_type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"sta
cked\":\"none\",\"label\":\"search
rate\",\"type\":\"timeseries\",\"terms_field\":\"index_stats.index\",\"terms_order_
by\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"split_color_mode\":\"rainbow\"}],\"time_field\":\"timestamp\",\"in
dex_pattern\":\"monitoring-eye-css-
*\",\"interval\":\"\",\"axis_position\":\"left\",\"axis_formatter\":\"number\",\"ax
is_scale\":\"normal\",\"show_legend\":1,\"show_grid\":1,\"default_index_pattern\":\
"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"isModelInvalid\":false,\"tooltip_mode\":\
```

```
"show_all\"}}"},"id":"ab503550-a5ef-11ec-9a68-
ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIxNSwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] indices store_size for
top10","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring] indices
store_size for
top10\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"id\":\"61ca57f0-469d-11e7-
af02-69e470af7417\",\"type\":\"timeseries\",\"series\":[{\"id\":\"38474c50-a5f5-
11ec-aa10-
992297d21a2e\",\"color\":\"#68BC00\",\"split_mode\":\"terms\",\"metrics\":[{\"id\":
\"38474c51-a5f5-11ec-aa10-
992297d21a2e\",\"type\":\"max\",\"field\":\"index_stats.total.store.size_in_bytes\"
}],\"separate_axis\":0,\"axis_position\":\"right\",\"formatter\":\"bytes\",\"chart_
type\":\"line\",\"line_width\":1,\"point_size\":1,\"fill\":0.5,\"stacked\":\"none\"
,\"label\":\"store_size for
index\",\"type\":\"timeseries\",\"terms_field\":\"index_stats.index\",\"terms_order
_by\":\"38474c51-a5f5-11ec-aa10-
992297d21a2e\",\"filter\":{\"query\":\"\",\"language\":\"kuery\"},\"split_color_mod
e\":\"rainbow\"}],\"time_field\":\"timestamp\",\"index_pattern\":\"monitoring-eye-
css-
*\",\"interval\":\"\",\"axis position\":\"left\",\"axis formatter\":\"number\",\"ax
is scale\":\"normal\",\"show legend\":1,\"show grid\":1,\"default index pattern\":\
"monitoring-eye-css-
*\",\"default timefield\":\"timestamp\",\"isModelInvalid\":false,\"filter\":{\"quer
y\":\"\",\"language\":\"kuery\"},\"bar color rules\":[{\"id\":\"7d9d3cb0-a5f5-11ec-
aa10-992297d21a2e\"}],\"tooltip mode\":\"show all\"}}"},"id":"c78119a0-a5f5-11ec-
9a68-ada9d754c566","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIxNiwyXQ=="}
{"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"{}"},"
title":"[monitoring] search query time in millis of index for
top10(ms)","uiStateJSON":"{}","version":1,"visState":"{\"title\":\"[monitoring]
search query time in millis of index for
top10(ms)\",\"type\":\"metrics\",\"aggs\":[],\"params\":{\"axis_formatter\":\"numbe
r\",\"axis_max\":\"\",\"axis_min\":\"\",\"axis_position\":\"left\",\"axis_scale\":\
"normal\",\"default_index_pattern\":\"monitoring-eye-css-
*\",\"default_timefield\":\"timestamp\",\"id\":\"61ca57f0-469d-11e7-af02-
69e470af7417\",\"index_pattern\":\"monitoring-eye-css-
*\",\"interval\":\"\",\"isModelInvalid\":false,\"series\":[{\"axis_position\":\"rig
ht\",\"chart_type\":\"line\",\"color\":\"#68BC00\",\"fill\":0.5,\"formatter\":\"num
ber\",\"id\":\"61ca57f1-469d-11e7-af02-
69e470af7417\",\"label\":\"index_query_time_in_millis\",\"line_width\":1,\"metrics\
":[{\"field\":\"index_stats.total.search.query_time_in_millis\",\"id\":\"61ca57f2-
469d-11e7-af02-69e470af7417\",\"type\":\"max\"},{\"unit\":\"1s\",\"id\":\"42c92b10-
6645-11ed-925a-6de90846447d\",\"type\":\"derivative\",\"field\":\"61ca57f2-469d-
11e7-af02-
69e470af7417\"}],\"point_size\":1,\"separate_axis\":0,\"split_color_mode\":\"kibana
\",\"split_mode\":\"terms\",\"stacked\":\"none\",\"terms_field\":\"index_stats.inde
x\",\"terms_order_by\":\"61ca57f2-469d-11e7-af02-
69e470af7417\",\"type\":\"timeseries\"}],\"show_grid\":1,\"show_legend\":1,\"time_f
ield\":\"timestamp\",\"tooltip_mode\":\"show_all\",\"type\":\"timeseries\",\"backgr
ound_color\":null,\"filter\":{\"query\":\"\",\"language\":\"kuery\"},\"legend_posit
ion\":\"right\"}}"},"id":"c8109100-6627-11ed-8cd7-
973626cf6f70","references":[],"type":"visualization","updated_at":"2022-12-
01T12:41:01.165Z","version":"WzIxNywyXQ=="}
```

{"attributes":{"description":"","hits":0,"kibanaSavedObjectMeta":{"searchSourceJSON":"{\"query\":{\"language\":\"kuery\",\"query\":\"\"},\"filter\":[]}"},"optionsJSON":"{\"hidePanelTitles\":false,\"useMargins\":true}","panelsJSON":"[{\"gridData\":{\"x\":0,\"y\":0,\"w\":48,\"h\":5,\"i\":\"971ed6c6-81b9-491b-9f08-e3ae9c382abd\"},\"panelIndex\":\"971ed6c6-81b9-491b-9f08-e3ae9c382abd\",\"embeddableConfig\":{},\"panelRefName\":\"panel_0\"},{\"gridData\":{\"x\":0,\"y\":5,\"w\":24,\"h\":15,\"i\":\"5a6982e7-0c6c-4733-8a2d-e4c57cdf7397\"},\"panelIndex\":\"5a6982e7-0c6c-4733-8a2d-e4c57cdf7397\",\"embeddableConfig\":{},\"panelRefName\":\"panel_1\"},{\"gridData\":{\"x\":24,\"y\":5,\"w\":24,\"h\":15,\"i\":\"662476f4-739c-4a05-858c-2ee8230cf410\"},\"panelIndex\":\"662476f4-739c-4a05-858c-2ee8230cf410\",\"embeddableConfig\":{},\"panelRefName\":\"panel_2\"},{\"gridData\":{\"x\":0,\"y\":20,\"w\":16,\"h\":15,\"i\":\"d89c38e2-33f3-4592-b503-20460a6a7a57\"},\"panelIndex\":\"d89c38e2-33f3-4592-b503-20460a6a7a57\",\"embeddableConfig\":{},\"panelRefName\":\"panel_3\"},{\"gridData\":{\"x\":16,\"y\":20,\"w\":16,\"h\":15,\"i\":\"1f693b49-79fa-4807-94e8-0c12f51e54f8\"},\"panelIndex\":\"1f693b49-79fa-4807-94e8-0c12f51e54f8\",\"embeddableConfig\":{},\"panelRefName\":\"panel_4\"},{\"gridData\":{\"x\":32,\"y\":20,\"w\":16,\"h\":15,\"i\":\"616b143d-74e9-4dac-98ba-5849536f0fba\"},\"panelIndex\":\"616b143d-74e9-4dac-98ba-5849536f0fba\",\"embeddableConfig\":{},\"panelRefName\":\"panel_5\"},{\"gridData\":{\"x\":0,\"y\":35,\"w\":24,\"h\":11,\"i\":\"cfa82f27-1b8d-49ba-a7b9-d8809d3b258c\"},\"panelIndex\":\"cfa82f27-1b8d-49ba-a7b9-d8809d3b258c\",\"embeddableConfig\":{},\"panelRefName\":\"panel_6\"},{\"gridData\":{\"x\":24,\"y\":35,\"w\":24,\"h\":11,\"i\":\"135d13eb-aab6-43ca-9029-7d26e91d90e3\"},\"panelIndex\":\"135d13eb-aab6-43ca-9029-7d26e91d90e3\",\"embeddableConfig\":{},\"panelRefName\":\"panel_7\"},{\"gridData\":{\"x\":0,\"y\":46,\"w\":24,\"h\":11,\"i\":\"28a77de1-9110-49e8-b273-724f880b1653\"},\"panelIndex\":\"28a77de1-9110-49e8-b273-724f880b1653\",\"embeddableConfig\":{},\"panelRefName\":\"panel_8\"},{\"gridData\":{\"x\":24,\"y\":46,\"w\":24,\"h\":11,\"i\":\"80ece867-cf23-4935-bfbc-430afa51bcca\"},\"panelIndex\":\"80ece867-cf23-4935-bfbc-430afa51bcca\",\"embeddableConfig\":{},\"panelRefName\":\"panel_9\"},{\"gridData\":{\"x\":0,\"y\":57,\"w\":24,\"h\":11,\"i\":\"2ba970aa-c9c4-491b-bdd3-c1b1ee9bc8d3\"},\"panelIndex\":\"2ba970aa-c9c4-491b-bdd3-c1b1ee9bc8d3\",\"embeddableConfig\":{},\"panelRefName\":\"panel_10\"},{\"gridData\":{\"x\":24,\"y\":57,\"w\":24,\"h\":11,\"i\":\"f2e1b6ab-ddf7-492e-aaca-9460f11aa4aa\"},\"panelIndex\":\"f2e1b6ab-ddf7-492e-aaca-9460f11aa4aa\",\"embeddableConfig\":{},\"panelRefName\":\"panel_11\"},{\"gridData\":{\"x\":0,\"y\":68,\"w\":24,\"h\":11,\"i\":\"dd14182d-d8b9-47f2-bf36-6cba3b09586c\"},\"panelIndex\":\"dd14182d-d8b9-47f2-bf36-6cba3b09586c\",\"embeddableConfig\":{},\"panelRefName\":\"panel_12\"},{\"gridData\":{\"x\":24,\"y\":68,\"w\":24,\"h\":11,\"i\":\"a47f9333-52b7-49b7-8cac-f470cf405131\"},\"panelIndex\":\"a47f9333-52b7-49b7-8cac-f470cf405131\",\"embeddableConfig\":{},\"panelRefName\":\"panel_13\"}]","timeRestore":false,"title":"[Monitoring] Index monitoring Dashboard","version":1},"id":"524eb000-a5f2-11ec-9a68-ada9d754c566","references":[{"id":"b2811c70-a5f1-11ec-9a68-ada9d754c566","name":"panel_0","type":"visualization"},{"id":"de4f8ab0-a5f8-11ec-9a68-ada9d754c566","name":"panel_1","type":"visualization"},{"id":"811df7a0-a5f8-11ec-9a68-ada9d754c566","name":"panel_2","type":"visualization"},{"id":"eea89780-664b-11ed-8cd7-973626cf6f70","name":"panel_3","type":"visualization"},{"id":"cfbb4e20-664c-11ed-8cd7-973626cf6f70","name":"panel_4","type":"visualization"},{"id":"c7f72ae0-664e-11ed-8cd7-973626cf6f70","name":"panel_5","type":"visualization"},{"id":"943b3e00-

```
a5ef-11ec-9a68-
ada9d754c566","name":"panel_6","type":"visualization"},{"id":"ab503550-a5ef-11ec-
9a68-ada9d754c566","name":"panel_7","type":"visualization"},{"id":"c78119a0-a5f5-
11ec-9a68-ada9d754c566","name":"panel_8","type":"visualization"},{"id":"225f6020-
a5f1-11ec-9a68-
ada9d754c566","name":"panel_9","type":"visualization"},{"id":"17d49220-662a-11ed-
8cd7-973626cf6f70","name":"panel_10","type":"visualization"},{"id":"c8109100-6627-
11ed-8cd7-973626cf6f70","name":"panel_11","type":"visualization"},{"id":"45d571c0-
6626-11ed-8cd7-
973626cf6f70","name":"panel_12","type":"visualization"},{"id":"3ae5d820-6628-11ed-
8cd7-
973626cf6f70","name":"panel_13","type":"visualization"}],"type":"dashboard","update
d_at":"2022-12-01T12:41:01.165Z","version":"WzIxOCwyXQ=="}
{"exportedCount":16,"missingRefCount":0,"missingReferences":[]}
```

7.7 集群监控增强

7.7.1 P99 时延监控

背景信息

Elasticsearch 社区针对 search 请求的监控都是平均时延，无法有效反应集群的实际 search 情况，此特性新增对集群 search 请求的 P99 时延监控。

约束限制

P99 时延监控目前仅支持 7.6.2 和 7.10.2 版本集群。

获取监控信息

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择需要启用索引监控的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 在左侧导航栏，选择“Dev Tools”，执行以下命令获取当前集群的 P99 时延：

```
GET /search/stats/percentile
```

返回样例如下：

```
{
  "overall" : {
    "1.0" : 2.0,
    "5.0" : 2.0,
    "25.0" : 6.5,
    "50.0" : 19.5,
    "75.0" : 111.0,
    "95.0" : 169.0,
    "99.0" : 169.0,
    "max" : 169.0,
    "min" : 2.0
  },
  "last_one_day" : {
```

```
    "1.0" : 2.0,
    "5.0" : 2.0,
    "25.0" : 6.5,
    "50.0" : 19.5,
    "75.0" : 111.0,
    "95.0" : 169.0,
    "99.0" : 169.0,
    "max" : 169.0,
    "min" : 2.0
  },
  "latest" : {
    "1.0" : 26.0,
    "5.0" : 26.0,
    "25.0" : 26.0,
    "50.0" : 26.0,
    "75.0" : 26.0,
    "95.0" : 26.0,
    "99.0" : 26.0,
    "max" : 26.0,
    "min" : 26.0
  }
}
```

说明

- 其中“overall”表示集群从启动到当前时间的统计值，“last_one_day”表示最近一天的统计值，“latest”表示从上次重置到当前时间的统计值。
- P99 时延的计算是近似值，不提供精确值，越靠近两端的统计值越准确，即 99%的时延比 50%的时延更准确。
- 如果重启集群，P99 时延数据将被清空，P99 时延数据将从集群重启成功后重新计算。

其他操作

- 自定义百分百数值。

  您可以自行指定百分百数值：

  ```
  GET /search/stats/percentile
  {
    "percents": [1, 50, 90]
  }
  ```

- 重置 latest 统计值。

  您可以执行以下命令重置 latest 统计值：

  ```
  POST /search/stats/reset
  ```

  返回样例：

  ```
  {
    "nodes" : {
      "css-c9c8-ess-esn-1-1" : "ok"
    }
  }
  ```

7.7.2 HTTP 状态码监控

背景信息

外部通过 HTTP 访问 Elasticsearch 都会返回 response 和相应的状态码，开源 Elasticsearch 服务端没有对状态码进行统计，无法准确知道调用 ES 接口的实际状态。用户无法通过监控知道整个集群的请求情况。HTTP 状态码监控提供监控集群的 HTTP 状态码的能力。

前提条件

HTTP 状态码监控目前仅 7.6.2 和 7.10.2 版本集群支持。

获取状态码

1. 登录云搜索服务管理控制台。
2. 在“集群管理”页面，选择需要启用索引监控的集群，单击操作列“Kibana”，登录 Kibana 界面。
3. 在左侧导航栏，选择“Dev Tools”。
4. 在 Dev Tools 的 Console 界面中执行根据集群版本执行对应的命令。
   - 7.6.2 版本集群，请执行以下命令获取状态码统计：

     ```
     GET /_nodes/http_stats
     ```

     返回样例:

     ```
     {
        "_nodes" : {
          "total" : 1,
          "successful" : 1,
          "failed" : 0   },
         "cluster_name" : "css-8362",
         "nodes" : {
          "F9IFdQPARaOJI7oL7HOXtQ" : {
             "http_code" : {
                "200" : 114,
                "201" : 5,
                "429" : 0,
                "400" : 7,
                "404" : 0,
                "405" : 0
              }
            }
          }
      }
     ```

   - 7.10.2 版本集群，请执行以下命令获取状态码统计：

     ```
     GET _nodes/stats/http
     ```

     返回样例:

     ```
     {
     // ...
       "cluster_name" : "css-2985",
       "nodes" : {
     ```

```
// ...
   "omvR9_W-TsGApraMApREjA" : {

// ...
     "http" : {
       "current_open" : 4,
       "total_opened" : 37,
       "http_code" : {
         "200" : 25,
         "201" : 7,
         "429" : 0,
         "400" : 3,
         "404" : 0,
         "405" : 0
       }
     }
   }
  }
}
```

8 监控

8.1 集群支持的监控指标

功能说明

本节定义了云搜索服务上报云监控服务的监控指标的命名空间，监控指标列表和维度定义。用户可以通过云监控服务提供的管理控制台或 API 接口来检索云搜索服务产生的监控指标和告警信息。

命名空间

SYS.ES

监控指标

- 监控的**指标 ID**、**指标名称**、**指标含义**以及**取值范围**参见表 8-1。
- 监控的**测量对象**：CSS 集群。本章节介绍 Elasticsearch 集群的监控指标。
- 监控的**监控周期**（原始指标）：1 分钟

说明

累计值：从节点启动时开始叠加数值，当节点重启后清零重新累计。

表8-1 云搜索服务支持的监控指标

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
status	集群健康状态	该指标用于统计测量监控对象的状态。	0,1,2,3 • 0：集群是100%可用的。 • 1：数据是完整的，部分副本缺失。高可	CSS集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
			用性在某种程度上弱化，存在风险，请及时关注集群情况。 • 2：数据缺失，集群使用时将出现异常。 • 3：没有获取到集群状态。		
disk_util	磁盘使用率	该指标用于统计测量对象的磁盘使用率。 单位：百分比	0-100%	CSS集群	1 分钟
max_jvm_heap_usage	最大 JVM 堆使用率	CSS 集群中各个节点的 JVM 堆使用率的最大值。 单位：百分比。	0-100%	CSS集群	1 分钟
max_jvm_young_gc_time	最大 JVM Young GC 耗时	CSS 集群中各个节点的 JVM Young GC 耗时累计值的最大值。 单位：ms	≥ 0 ms	CSS集群	1 分钟
max_jvm_young_gc_count	最大 JVM Young GC 次数	CSS 集群中各个节点的 JVM Young GC 次数累计值的最大值。	≥ 0	CSS集群	1 分钟
max_jvm_old_gc_time	最大 JVM Old GC 耗时	CSS 集群中各个节点的 JVM Old GC 耗时累计值	≥ 0 ms	CSS集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		的最大值。 单位：ms			
max_jvm_old_gc_count	最大 JVM Old GC 次数	CSS 集群中各个节点的 JVM Old GC 次数累计值的最大值。	≥ 0	CSS 集群	1 分钟
total_fs_size	文件系统总大小	CSS 集群的文件系统总大小。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
free_fs_size	文件系统可用大小	CSS 集群的文件系统可用大小。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
max_cpu_usage	最大 CPU 利用率	CSS 集群中各个节点的 CPU 利用率的最大值。 单位：百分比	0-100%	CSS 集群	1 分钟
max_cpu_time_of_jvm_process	最大 JVM 进程使用的 CPU 时间	CSS 集群中各个节点 JVM 进程使用 CPU 的时间累计值的最大值。 单位：ms	≥ 0 ms	CSS 集群	1 分钟
max_virtual_memory_size_of_jvm_process	最大 JVM 进程使用的虚拟内存大小	CSS 集群中各个节点 JVM 进程可使用的虚拟内存大小的最大值。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
max_current_opened_http_count	最大当前打开的 HTTP 连接数	CSS 集群中各个节点打开且尚未关闭的 HTTP	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		连接数的最大值。			
max_total_opened_http_count	最大全部打开的 HTTP 连接数	CSS 集群中各个节点打开过的 HTTP 连接数累计值的最大值。	≥ 0	CSS 集群	1 分钟
indices_count	索引数量	CSS 集群的索引数量。	≥ 0	CSS 集群	1 分钟
total_shards_count	分片数量	CSS 集群的分片数量。	≥ 0	CSS 集群	1 分钟
primary_shards_count	主分片数量	CSS 集群的主分片数量。	≥ 0	CSS 集群	1 分钟
docs_count	文档数量	CSS 集群的文档数量。	≥ 0	CSS 集群	1 分钟
docs_deleted_count	被删除的文档数量	CSS 集群的被删除的文档数量。	≥ 0	CSS 集群	1 分钟
nodes_count	节点数量	CSS 集群的节点数量。	≥ 0	CSS 集群	1 分钟
data_nodes_count	数据节点数量	CSS 集群的数据节点数量。	≥ 0	CSS 集群	1 分钟
coordinating_nodes_count	协调节点数量	CSS 集群的协调节点数量。	≥ 0	CSS 集群	1 分钟
master_nodes_count	Master 节点数量	CSS 集群的 Master 节点数量。	≥ 0	CSS 集群	1 分钟
ingest_nodes_count	Client 节点数量	CSS 集群的 Client 节点数量。	≥ 0	CSS 集群	1 分钟
max_load_average	最大节点 Load 值	CSS 集群中各个节点在操作系统中 1	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		分钟平均排队任务数的最大值。			
avg_cpu_usage	平均 CPU 使用率	CSS 集群中各节点 CPU 利用率的平均值。 单位：百分比	0-100%	CSS 集群	1 分钟
avg_load_average	平均节点 Load 值	CSS 集群中各节点在操作系统中 1 分钟平均排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_jvm_heap_usage	平均 JVM 堆使用率	CSS 集群中各节点 JVM 堆内存使用率的平均值。 单位：百分比	0-100%	CSS 集群	1 分钟
max_open_file_descriptors	已打开的最大文件描述符数	CSS 集群中各个节点已打开的文件描述符数的最大值。	≥ 0	CSS 集群	1 分钟
avg_open_file_descriptors	已打开的平均文件描述符数	CSS 集群中各节点已打开的文件描述符数的平均值。	≥ 0	CSS 集群	1 分钟
sum_max_file_descriptors	最大允许的文件描述符数	CSS 集群中各节点最大允许的文件描述符数之和。	≥ 0	CSS 集群	1 分钟
sum_open_file_descriptors	已打开的文件描述符数	CSS 集群中各节点已打开的文件描	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		述符数之和。			
sum_thread_pool_write_queue	Write 队列中总排队任务数	写入线程池中的排队任务数。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_search_queue	Search 队列中总排队任务数	CSS 集群中各节点在搜索线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_force_merge_queue	ForceMerge 队列中总排队任务数	CSS 集群中各节点在强制合并线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_write_rejected	Write 队列中总的已拒绝任务数	CSS 集群中各节点在写入线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_search_rejected	Search 队列中总的已拒绝任务数	CSS 集群中各节点在搜索线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_force_merge_rejected	Forcemerge 队列中总的已拒绝任务数	CSS 集群中各节点在强制合并线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_search_queue	Search 队列中最大排队任务数	CSS 集群中各个节点在搜索线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
max_thread_pool_force_	ForceMerge 队列中最	CSS 集群中各个节点在	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
merge_queue	大排队任务数	强制合并线程池中的排队任务数的最大值。			
sum_thread_pool_write_threads	Write 线程池总大小	CSS 集群中各节点写入线程池的大小之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_search_threads	Search 线程池总大小	CSS 集群中各节点搜索线程池的大小之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_force_merge_threads	ForceMerge 线程池总大小	CSS 集群中各节点强制合并线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_write_queue	Write 队列中平均排队任务数	CSS 集群中各节点在写入线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_search_queue	Search 队列中平均排队任务数	CSS 集群中各节点在搜索线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_force_merge_queue	ForceMerge 队列中平均排队任务数	CSS 集群中各节点在强制合并线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_search_threads	Search 线程池平均大小	CSS 集群中各节点搜索线程池的大小的平均值。	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
avg_thread_pool_write_threads	Write 线程池平均大小	CSS 集群中各节点写入线程池的大小的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_force_merge_threads	ForceMerge 线程池平均大小	CSS 集群中各节点强制合并线程池的大小的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_write_rejected	Write 队列中平均已拒绝任务数	CSS 集群中各节点写入线程池中的已拒绝任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_flush_queue	Flush 队列中总排队任务数	CSS 集群中各个节点在 Flush 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_flush_rejected	Flush 队列中总的已拒绝任务数	CSS 集群中各节点在 Flush 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_flush_queue	Flush 队列中最大排队任务数	CSS 集群中各个节点在 Flush 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_flush_threads	Flush 线程池总大小	CSS 集群中各节点 Flush 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_flush_queue	Flush 队列中平均排队任务数	CSS 集群中各节点在 Flush 线程池	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		中的排队任务数的平均值。			
avg_thread_pool_flush_threads	Flush 线程池平均大小	CSS 集群中各节点在 Flush 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_generic_queue	Generic 队列中总排队任务数	CSS 集群中各个节点在 Generic 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_generic_rejecte	Generic 队列中总的已拒绝任务数	CSS 集群中各节点在 Generic 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_generic_queue	Generic 队列中最大排队任务数	CSS 集群中各个节点在 Generic 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_generic_threads	Generic 线程池总大小	CSS 集群中各节点 Generic 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_generic_queue	Generic 队列中平均排队任务数	CSS 集群中各节点在 Generic 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_generic_threads	Generic 线程池平均	CSS 集群中各节点在 Generic 线程	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
	大小	池中的排队任务数的平均值。			
sum_thread_pool_management_queue	Management 队列中总排队任务数	CSS 集群中各个节点在 Management 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_management_rejected	Management 队列中总的已拒绝任务数	CSS 集群中各节点在 Management 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_management_queue	Management 队列中最大排队任务数	CSS 集群中各个节点在 Management 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_management_threads	Management 线程池总大小	CSS 集群中各节点 Management 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_management_queue	Management 队列中平均排队任务数	CSS 集群中各节点在 Management 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_management_threads	Management 线程池平均大小	CSS 集群中各节点在 Management 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_refresh_queue	Refresh 队列中总排队任务数	CSS 集群中各个节点在 Refresh 线程	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		池中的排队任务数之和。			
sum_thread_pool_refresh_rejected	Refresh 队列中总的已拒绝任务数	CSS 集群中各节点在 Refresh 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_refresh_queue	Refresh 队列中最大排队任务数	CSS 集群中各个节点在 Refresh 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_refresh_threads	Refresh 线程池总大小	CSS 集群中各节点 Refresh 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_refresh_queue	Refresh 队列中平均排队任务数	CSS 集群中各节点在 Refresh 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_refresh_threads	Refresh 线程池平均大小	CSS 集群中各节点在 Refresh 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_searcher_queue	OBS Searcher 队列中总排队任务数	CSS 集群中各个节点在 OBS Searcher 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_searcher_reject	OBS Searcher 队列中总的	CSS 集群中各节点在 OBS Searcher	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
ed	已拒绝任务数	线程池中的已拒绝任务数之和。			
max_thread_pool_obs_searcher_queue	OBS Searcher 队列中最大排队任务数	CSS 集群中各个节点在 OBS Searcher 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_searcher_threads	OBS Searcher 线程池总大小	CSS 集群中各节点 OBS Searcher 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_searcher_queue	OBS Searcher 队列中平均排队任务数	CSS 集群中各节点在 OBS Searcher 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_searcher_threads	OBS Searcher 线程池平均大小	CSS 集群中各节点在 OBS Searcher 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_queue	OBS 队列中总排队任务数	CSS 集群中各个节点在 OBS 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_rejected	OBS 队列中总的已拒绝任务数	CSS 集群中各节点在 OBS 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_obs_queue	OBS 队列中最大排队任务数	CSS 集群中各个节点在 OBS 线程池中的排队任	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		务数的最大值。			
sum_thread_pool_obs_threads	OBS 线程池总大小	CSS 集群中各节点 OBS 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_queue	OBS 队列中平均排队任务数	CSS 集群中各节点在 OBS 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_threads	OBS 线程池平均大小	CSS 集群中各节点在 OBS 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_upload_queue	OBS Upload 队列中总排队任务数	CSS 集群中各个节点在 OBS Upload 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_upload_rejected	OBS Upload 队列中总的已拒绝任务数	CSS 集群中各节点在 OBS Upload 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_obs_upload_queue	OBS Upload 队列中最大排队任务数	CSS 集群中各个节点在 OBS Upload 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_upload_threads	OBS Upload 线程池总大小	CSS 集群中各节点 OBS Upload 线程池的大小之和。	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
avg_thread_pool_obs_upload_queue	OBS Upload 队列中平均排队任务数	CSS 集群中各节点在 OBS Upload 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_upload_threads	OBS Upload 线程池平均大小	CSS 集群中各节点在 OBS Upload 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_download_queue	OBS Download 队列中总排队任务数	CSS 集群中各个节点在 OBS Download 线程池中的排队任务数之和。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_download_rejected	OBS Download 队列中总的已拒绝任务数	CSS 集群中各节点在 OBS Download 线程池中的已拒绝任务数之和。	≥ 0	CSS 集群	1 分钟
max_thread_pool_obs_download_queue	OBS Download 队列中最大排队任务数	CSS 集群中各个节点在 OBS Download 线程池中的排队任务数的最大值。	≥ 0	CSS 集群	1 分钟
sum_thread_pool_obs_download_threads	OBS Download 线程池总大小	CSS 集群中各节点 OBS Download 线程池的大小之和。	≥ 0	CSS 集群	1 分钟
avg_thread_pool_obs_download_queue	OBS Download 队列中平均排队任	CSS 集群中各节点在 OBS Download 线	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
	务数	程池中的排队任务数的平均值。			
avg_thread_pool_obs_download_threads	OBS Download 线程池平均大小	CSS 集群中各节点在 OBS Download 线程池中的排队任务数的平均值。	≥ 0	CSS 集群	1 分钟
min_free_fs_size	最小可用存储空间	CSS 集群中各个节点可用存储空间的最小值。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
avg_jvm_old_gc_count	JVM 老年代平均 GC 次数	CSS 集群中各个节点“老年代”垃圾回收的运行次数的累计值的平均值。	≥ 0	CSS 集群	1 分钟
avg_jvm_old_gc_time	JVM 老年代平均 GC 时间	CSS 集群中各个节点执行“老年代”垃圾回收所花费的时间累计值的平均值。 单位：ms	≥ 0 ms	CSS 集群	1 分钟
avg_jvm_young_gc_count	JVM 年轻代平均 GC 次数	CSS 集群中各个节点“年轻代”垃圾回收的运行次数的累计值的平均值。	≥ 0	CSS 集群	1 分钟
avg_jvm_young_gc_time	JVM 年轻代平均 GC 时间	CSS 集群中各个节点执行“年轻代”垃圾回	≥ 0 ms	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		收所花费的时间累计值的平均值。 单位：ms			
avg_max_file_descriptors	最大允许的文件描述符数-平均值	CSS 集群中各节点最大允许的文件描述符数的平均值。	≥ 0	CSS 集群	1 分钟
avg_mem_free_in_bytes	平均可用内存空间	CSS 集群中各节点未使用的内存容量的平均值。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
avg_mem_free_percent	平均可用内存比例	CSS 集群中各节点未使用的内存比例的平均值。 单位：百分比	0-100%	CSS 集群	1 分钟
avg_mem_used_in_bytes	平均已用内存空间	CSS 集群中各节点已使用的内存容量的平均值。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
avg_mem_used_percent	平均已用内存比例	CSS 集群中各节点已使用的内存比例的平均值。 单位：百分比	0-100%	CSS 集群	1 分钟
max_mem_free_in_bytes	最大可用内存空间	CSS 集群中各个节点未使用的内存容量的最大	≥ 0 bytes	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		值。 单位：byte			
max_mem_free_percent	最大可用内存比例	CSS 集群中各个节点未使用的内存比例的最大值。 单位：百分比	0-100%	CSS 集群	1 分钟
max_mem_used_in_bytes	最大已用内存空间	CSS 集群中各个节点已使用的内存容量的最大值。 单位：byte	≥ 0 bytes	CSS 集群	1 分钟
max_mem_used_percent	最大已用内存比例	CSS 集群中各个节点已使用的内存比例的最大值。 单位：百分比	0-100%	CSS 集群	1 分钟
sum_jvm_old_gc_count	JVM 老年代总 GC 次数	CSS 集群中各个节点“老年代”垃圾回收的运行次数的累计值之和。	≥ 0	CSS 集群	1 分钟
sum_jvm_old_gc_time	JVM 老年代总 GC 时间	CSS 集群中各个节点执行“老年代”垃圾回收所花费的时间累计值之和。 单位：ms	≥ 0ms	CSS 集群	1 分钟
sum_jvm_young_gc_count	JVM 年轻代总 GC 次数	CSS 集群中各个节点“年轻代”垃圾回收的运	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
		行次数的累计值之和。			
sum_jvm_young_gc_time	JVM 年轻代总 GC 时间	CSS 集群中各个节点执行“年轻代”垃圾回收所花费的时间累计值之和。 单位：ms	≥ 0 ms	CSS 集群	1 分钟
sum_current_opened_http_count	当前已打开 HTTP 连接数	CSS 集群中各个节点打开且尚未关闭的 HTTP 连接数之和。	≥ 0	CSS 集群	1 分钟
sum_total_opened_http_count	历史已打开 HTTP 连接数	CSS 集群中各个节点打开过的 HTTP 连接数累计值之和。	≥ 0	CSS 集群	1 分钟
IndexingLatency	平均索引延迟	分片完成索引操作所需的平均时间。 单位：ms	≥ 0 ms	CSS 集群	1 分钟
IndexingRate	平均索引速率	入库 TPS，集群每秒平均索引操作数。	≥ 0	CSS 集群	1 分钟
SearchLatency	平均查询延迟	分片完成搜索操作所需的平均时间。 单位：ms。	≥ 0 ms	CSS 集群	1 分钟
SearchRate	平均查询速率	查询 QPS，集群每秒平均查询操作数。	≥ 0	CSS 集群	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
task_max_running_time	最大 Task 运行时长	该指标为集群运行的所有 task 中，运行时长最长的 task 的耗时	≥ 0 ms	CSS 集群	1 分钟
number_of_pending_tasks	Pending Task 排队任务数	CSS 集群中待 Master 处理的 task 的排队任务数。	≥ 0	CSS 集群	1 分钟

维度

表8-2 维度说明

Key	Value
cluster_id	CSS 集群

8.2 Logstash 集群支持的监控指标

功能说明

本节定义了云搜索服务上报云监控服务的监控指标的命名空间，监控指标列表和维度定义。用户可以通过云监控服务提供管理控制台或 API 接口来检索云搜索服务产生的监控指标和告警信息。

命名空间

SYS.ES

CSS.CUSTOM

监控指标

- 监控的**指标 ID**、**指标名称**、**指标含义**以及**取值范围**参见表 8-3。
- 监控的**测量对象**：CSS 集群。本章节介绍 Logstash 集群的监控指标。Elasticsearch 集群所支持的监控指标请见 8.1 集群支持的监控指标。

- 监控的**监控周期**（原始指标）：1 分钟

说明

累计值：从节点启动时开始叠加数值，当节点重启后清零重新累计。

表8-3 云搜索服务支持的监控指标

指标 ID	指标名称	指标含义	取值范围
max_jvm_heap_usage	最大 JVM 堆使用率	CSS 集群中各个节点的 JVM 堆使用率的最大值。 单位：百分比	0-100%
max_jvm_young_gc_time	最大 JVM Young GC 耗时	CSS 集群中各个节点的 JVM Young GC 耗时累计值的最大值。 单位：ms	≥ 0 ms
max_jvm_young_gc_count	最大 JVM Young GC 次数	CSS 集群中各个节点的 JVM Young GC 次数累计值的最大值。	≥ 0
max_jvm_old_gc_time	最大 JVM Old GC 耗时	CSS 集群中各个节点的 JVM Old GC 耗时累计值的最大值。 单位：ms	≥ 0 ms
max_jvm_old_gc_count	最大 JVM Old GC 次数	CSS 集群中各个节点的 JVM Old GC 次数累计值的最大值。	≥ 0
max_cpu_usage	最大 CPU 利用率	CSS 集群中各个节点的 CPU 利用率的最大值。 单位：百分比	0-100%
max_load_average	最大节点 Load 值	CSS 集群中各个节点在操作系统中 1 分钟平均排队任务数的最大值。	≥ 0
avg_cpu_usage	平均 CPU 使用率	CSS 集群中各节点 CPU 利用率的平均值。 单位：百分比	0-100%
avg_load_average	平均节点 Load 值	CSS 集群中各节点在操作系统中 1 分钟平均排队任务数的平均	≥ 0

指标 ID	指标名称	指标含义	取值范围
		值。	
avg_jvm_heap_usage	平均 JVM 堆使用率	CSS 集群中各节点 JVM 堆内存使用率的平均值。 单位：百分比	0-100%
avg_jvm_old_gc_count	JVM 老年代平均 GC 次数	CSS 集群中各个节点“老年代”垃圾回收的运行次数的累计值的平均值。	≥ 0
avg_jvm_old_gc_time	JVM 老年代平均 GC 时间	CSS 集群中各个节点执行“老年代”垃圾回收所花费的时间累计值的平均值。 单位：ms	≥ 0 ms
avg_jvm_young_gc_count	JVM 年轻代平均 GC 次数	CSS 集群中各个节点“年轻代”垃圾回收的运行次数的累计值的平均值。	≥ 0
avg_jvm_young_gc_time	JVM 年轻代平均 GC 时间	CSS 集群中各个节点执行“年轻代”垃圾回收所花费的时间累计值的平均值。 单位：ms	≥ 0 ms
sum_events_in	集群下所有节点经过 input 插件的数据总数	该指标用于统计所有节点经过 input 插件的数据总数。	≥ 0
sum_events_filtered	集群下所有节点经过 filtere 插件的数据总数	该指标用于统计所有节点经过过滤处理的数据总数。	≥ 0
sum_events_out	集群下所有节点经过 out 插件的数据总数	该指标用于统计所有节点经过 out 插件的数据总数。	≥ 0
events_in	当前节点经过 input 插件的数据数	该指标用于统计当前节点经过 input 插件的数据数。	≥ 0
events_filtered	当前节点经过 filtered 插件的数据数	该指标用于统计当前节点过滤处理的数据数。	≥ 0
events_out	当前节点经过	该指标用于统计当前	≥ 0

指标 ID	指标名称	指标含义	取值范围
	out 插件的数据数	节点经过 out 插件的数据数。	
logstash_pipeline_events_in	当前管道监控周期内经过 input 插件的数据数。	该指标用于统计当前管道监控周期内经过 input 插件的数据数。	≥ 0
logstash_pipeline_events_filtered	当前管道监控周期内经过 filtered 插件的数据数。	该指标用于统计当前管道监控周期内过滤处理的数据数。	≥ 0
logstash_pipeline_events_out	当前管道监控周期内经过 out 插件的数据数。	该指标用于统计当前管道监控周期内经过 out 插件的数据数。	≥ 0

维度

表8-4 维度说明

Key	Value
cluster_id	CSS 集群

8.3 节点支持的监控指标

功能说明

本节定义了云搜索服务上报云监控服务的节点监控指标的命名空间，监控指标列表和维度定义。用户可以通过云监控服务提供的管理控制台或 API 接口来检索云搜索服务产生的监控指标和告警信息。

命名空间

SYS.ES

监控指标

- 监控的**指标 ID**、**指标名称**、**指标含义**以及**取值范围**参见表 8-5。
- 监控的**测量对象：CSS 集群 - 云服务节点**
- 监控的**监控周期**（原始指标）：1 分钟

说明

累计值：从节点启动时开始叠加数值，当节点重启后清零重新累计。

表8-5 云搜索服务节点支持的监控指标

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
jvm_heap_usage	JVM 堆使用率	节点 JVM 堆内存使用率。 单位：百分比	0-100%	CSS 集群 - 云服务节点	1 分钟
cpu_usage	CPU 利用率	CPU 利用率。 单位：百分比	0-100%	CSS 集群 - 云服务节点	1 分钟
load_average	节点 Load 值	操作系统中 1 分钟平均排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
open_file_descriptors	已打开的文件描述符数	节点已打开的文件描述符数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
max_file_descriptors	最大允许的文件描述符数	最大允许的文件描述符数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_write_queue	Write 队列中总排队任务数	写入线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_search_queue	Search 队列中总排队任务数	搜索线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_force_merge_queue	ForceMerge 队列中总排队任务数	强制合并线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_write_rejected	Write 队列中总的已拒绝	写入线程池中的已拒绝任务	≥ 0	CSS 集群 - 云服务	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
	任务数	数。		节点	
thread_pool_search_rejected	Search 队列中总的已拒绝任务数	搜索线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_force_merge_rejected	ForceMerge 队列中总的已拒绝任务数	强制合并线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_write_threads	Write 线程池总大小	写入线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_search_threads	Search 线程池总大小	搜索线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_force_merge_threads	ForceMerge 线程池总大小	强制合并线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_flush_queue	Flush 队列中总排队任务数	Flush 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_flush_rejected	Flush 队列中总的已拒绝任务数	Flush 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_flush_threads	Flush 线程池总大小	Flush 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_generic_queue	Generic 队列中总排队任务数	Generic 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_generic_rejected	Generic 队列中总的已拒	Generic 线程池中的已拒绝	≥ 0	CSS 集群 -	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
	绝任务数	任务数。		云服务节点	
thread_pool_generic_threads	Generic 线程池总大小	Generic 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_management_queue	Management 队列中总排队任务数	Management 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_management_rejected	Management 队列中总的已拒绝任务数	Management 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_management_threads	Management 线程池总大小	Management 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_refresh_queue	Refresh 队列中总排队任务数	Refresh 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_refresh_rejected	Refresh 队列中总的已拒绝任务数	Refresh 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_refresh_threads	Refresh 线程池总大小	Refresh 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_searcher_queue	OBS Searcher 队列中总排队任务数	OBS Searcher 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_searcher_rejected	OBS Searcher 队列中总的已拒绝任务数	OBS Searcher 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
thread_pool_obs_searcher_threads	OBS Searcher 线程池总大小	OBS Searcher 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_queue	OBS 队列中总排队任务数	OBS 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_rejected	OBS 队列中总的已拒绝任务数	OBS 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_threads	OBS 线程池总大小	OBS 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_upload_queue	OBS Upload 队列中总排队任务数	OBS Upload 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_upload_rejected	OBS Upload 队列中总的已拒绝任务数	OBS Upload 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_upload_threads	OBS Upload 线程池总大小	OBS Upload 线程池的大小。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_download_queue	OBS Download 队列中总排队任务数	OBS Download 线程池中的排队任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_download_rejected	OBS Download 队列中总的已拒绝任务数	OBS Download 线程池中的已拒绝任务数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
thread_pool_obs_download_threads	OBS Download 线程池总大小	OBS Download 线程池的大小。	≥ 0	CSS 集群 - 云服务	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
				节点	
free_fs_size	文件系统可用大小	CSS 集群的文件系统可用大小。 单位：byte	≥ 0 bytes	CSS 集群 - 云服务节点	1 分钟
total_fs_size	文件系统总大小	CSS 集群的文件系统总大小。 单位：byte	≥ 0 bytes	CSS 集群 - 云服务节点	1 分钟
jvm_old_gc_count	JVM 老年代总 GC 次数	“老年代”垃圾回收的运行次数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
jvm_old_gc_time	JVM 老年代总 GC 时间	执行“老年代”垃圾回收所花费的时间。 单位：ms	≥ 0 ms	CSS 集群 - 云服务节点	1 分钟
jvm_young_gc_count	JVM 年轻代总 GC 次数	“年轻代”垃圾回收的运行次数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
jvm_young_gc_time	JVM 年轻代 GC 时间	执行“年轻代”垃圾回收所花费的时间。 单位：ms	≥ 0 ms	CSS 集群 - 云服务节点	1 分钟
mem_free_in_bytes	可用内存空间	节点未使用的内存容量。 单位：byte	≥ 0 bytes	CSS 集群 - 云服务节点	1 分钟
mem_free_percent	可用内存比例	节点未使用的内存比例。	≥ 0	CSS 集群 - 云服务节点	1 分钟
mem_used_in_bytes	已用内存空间	节点已使用的内存容量。 单位：byte	≥ 0 bytes	CSS 集群 - 云服务	1 分钟

指标 ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
				节点	
current_opened_http_count	当前已打开 HTTP 连接数	节点当前打开的 HTTP 连接数。	≥ 0	CSS 集群 - 云服务节点	1 分钟
total_opened_http_count	全部打开的 HTTP 连接数	节点全部打开的 HTTP 连接数。	≥ 0	CSS 集群 - 云服务节点	1 分钟

维度

表8-6 维度说明

Key	Value
cluster_id	CSS 集群

8.4 配置集群监控

云搜索服务支持通过云监控服务 CES 对已创建成功的集群进行日常监控。配置集群监控后，就可以在 CES 管理控制台直观查看集群的监控指标数据。

配置集群监控的操作流程：

1. 配置告警规则：根据实际业务需要对监控指标设置自定义告警规则，当监控指标超过设置的阈值时，会以邮箱、HTTP、HTTPS 等方式通知您。
2. 配置监控对象：为集群或集群中某个节点配置监控指标。
3. 查看监控指标：您可以选择不同的监控时间周期，查看监控指标数据变化情况。

前提条件

- 集群处于“可用”或“处理中”状态。
- 集群正常运行时长大于 10 分钟。

推荐配置的监控指标

- 监控集群的 **cpu、jvm 使用情况**，推荐重点配置如下监控指标：平均 JVM 堆使用率、最大 JVM 堆使用率、平均 CPU 使用率、最大 CPU 利用率。

- 监控集群的**写入、查询延迟和吞吐量情况**，推荐重点配置如下监控指标：平均索引延迟、平均索引速率、平均查询延迟、平均查询速率。
- 监控集群的**写入、查询的排队队列和拒绝情况**，推荐重点配置如下监控指标：Write 队列中总排队任务数、Search 队列中总排队任务数、Write 队列中总的已拒绝任务数、Search 队列中总的已拒绝任务数。

配置告警规则

1. 登录云监控服务 CES 管理控制台。
2. 左侧导航栏选择“告警 > 告警规则”，进入告警规则列表页面。
3. 在“资源类型”列，筛选“云搜索服务”，查看是否有满足要求的告警规则。

   如果没有，请参考云监控服务 CES 的“创建告警规则和通知”章节，新建 CSS 服务的告警规则。其中，“资源类型”和“维度”参数的填写说明请参见表 8-7，其他参数可以根据 CES 服务的参数说明自定义。

表8-7 告警内容的配置说明

参数	参数解释	配置说明
资源类型	配置告警规则监控的服务名称。	选择**云搜索服务**。
维度	用于指定告警规则对应指标的维度名称。	CSS 支持 2 个维度，根据实际需要选择维度。 • **CSS 集群**：以集群维度指定告警规则。 • **CSS 集群 - 云服务节点**：以集群中的某个节点维度指定告警规则。

配置监控对象

1. 参考云监控服务 CES 的“创建监控面板”章节，创建一个监控面板。如果已有监控面板，可以跳过该步骤。
2. 参考云监控服务 CES 的“添加监控视图”章节，添加 CSS 监控视图。

   其中，“资源类型”和“维度”参数的填写说明请参见表 8-8，其他参数可以根据 CES 服务的参数说明自定义。

表8-8 监控视图的配置说明

参数	参数解释	配置说明
资源类型	添加监控视图的服务名称。	选择**云搜索服务**。
维度	指定监控的维度名称。	CSS 支持 2 个维度，根据实际需要选择维度。 • **CSS 集群**：以集群维度监控。 • **CSS 集群 - 云服务节点**：以集群中的某个节点维度监控。

查看监控指标

1. 登录云搜索服务管理控制台。
2. 选择目标集群，单击操作列“监控信息”查看监控指标。
3. 选择待查看的时间段页签。
4. 查看监控指标数据。

9 审计

9.1 支持云审计的关键操作

公有云平台提供了云审计服务。通过云审计服务，您可以记录与云搜索服务相关的操作事件，便于日后的查询、审计和回溯。

前提条件

已开通云审计服务。

支持审计的关键操作列表

表9-1 支持审计的关键操作列表

操作名称	资源类型	事件名称
创建集群	cluster	createCluster
删除集群	cluster	deleteCluster
扩容集群	cluster	roleExtendCluster
重启集群	cluster	rebootCluster
配置自定义词库	cluster	loadLexicon
删除自定义词库	cluster	deleteLexicon
设置集群快照的基础配置	cluster	updateSnapshotPolicy
设置自动创建快照策略	cluster	updateAutoSnapshotPolicy
集群升级	cluster	upgradeCluster
升级重试	cluster	retryAction
手动创建快照	snapshot	createSnapshot
恢复快照	snapshot	restoreSnapshot

操作名称	资源类型	事件名称
删除快照	snapshot	deleteSnapshot

9.2 在 CTS 事件列表查看云审计事件

操作场景

用户进入云审计服务创建管理类追踪器后，系统开始记录云服务资源的操作。在创建数据类追踪器后，系统开始记录用户对 OBS 桶中数据的操作。云审计服务管理控制台会保存最近 7 天的操作记录。

本节介绍如何在云审计服务管理控制台查看或导出最近 7 天的操作记录：

- 在新版事件列表查看审计事件
- 在旧版事件列表查看审计事件

使用限制

- 用户通过云审计控制台只能查询最近 7 天的操作记录。如果需要查询超过 7 天的操作记录，您必须配置转储到对象存储服务(OBS)或云日志服务(LTS)，才可在 OBS 桶或 LTS 日志组里面查看历史事件信息。否则，您将无法追溯 7 天以前的操作记录。
- 云上操作后，1 分钟内可以通过云审计控制台查询管理类事件操作记录，5 分钟后才可通过云审计控制台查询数据类事件操作记录。

在新版事件列表查看审计事件

1. 登录管理控制台。
2. 单击左上角 ☰，选择“管理与部署 > 云审计服务 CTS”，进入云审计服务页面。
3. 单击左侧导航树的“事件列表”，进入事件列表信息页面。
4. 事件列表支持通过高级搜索来查询对应的操作事件，您可以在筛选器组合一个或多个筛选条件：
   - 事件名称：输入事件的名称。
   - 事件 ID：输入事件 ID。
   - 资源名称：输入资源的名称，当该事件所涉及的云资源无资源名称或对应的 API 接口操作不涉及资源名称参数时，该字段为空。
   - 资源 ID：输入资源 ID，当该资源类型无资源 ID 或资源创建失败时，该字段为空。
   - 云服务：在下拉框中选择对应的云服务名称。
   - 资源类型：在下拉框中选择对应的资源类型。
   - 操作用户：在下拉框中选择一个或多个具体的操作用户。

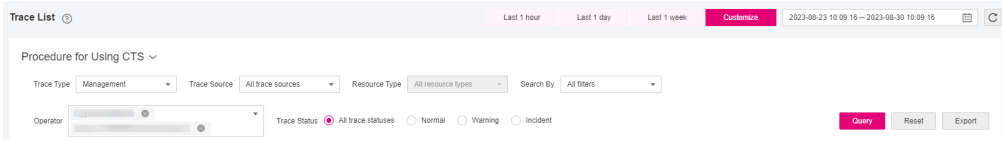
- 事件级别：可选项为“normal”、“warning”、“incident”，只可选择其中一项。
  - normal：表示操作成功。
  - warning：表示操作失败。
  - incident：表示比操作失败更严重的情况，例如引起其他故障等。
- 时间范围：可选择查询最近 1 小时、最近 1 天、最近 1 周的操作事件，也可以自定义最近 7 天内任意时间段的操作事件。

5. 在事件列表页面，您还可以导出操作记录文件、刷新列表、设置列表展示信息等。
   - 在搜索框中输入任意关键字，按下 Enter 键，可以在事件列表搜索符合条件的数据。
   - 单击“导出”按钮，云审计服务会将查询结果以.xlsx 格式的表格文件导出，该.xlsx 文件包含了本次查询结果的所有事件，且最多导出 5000 条信息。
   - 单击按钮，可以获取到事件操作记录的最新信息。
   - 单击按钮，可以自定义事件列表的展示信息。启用表格内容折行开关，可让表格内容自动折行，禁用此功能将会截断文本，默认停用此开关。
6. 关于事件结构的关键字段详解，请参见“云审计服务事件参考 > 事件结构”章节和“云审计服务事件参考 > 事件样例”章节。
7. （可选）在新版事件列表页面，单击右上方的“返回旧版”按钮，可切换至旧版事件列表页面。

在旧版事件列表查看审计事件

1. 登录管理控制台。
2. 单击左上角，选择“管理与部署 > 云审计服务 CTS”，进入云审计服务页面。
3. 单击左侧导航树的“事件列表”，进入事件列表信息页面。
4. 用户每次登录云审计控制台时，控制台默认显示新版事件列表，单击页面右上方的“返回旧版”按钮，切换至旧版事件列表页面。
5. 事件列表支持通过筛选来查询对应的操作事件，如图 9-1 所示。当前事件列表支持四个维度的组合查询，详细信息如下：

图9-1 筛选框

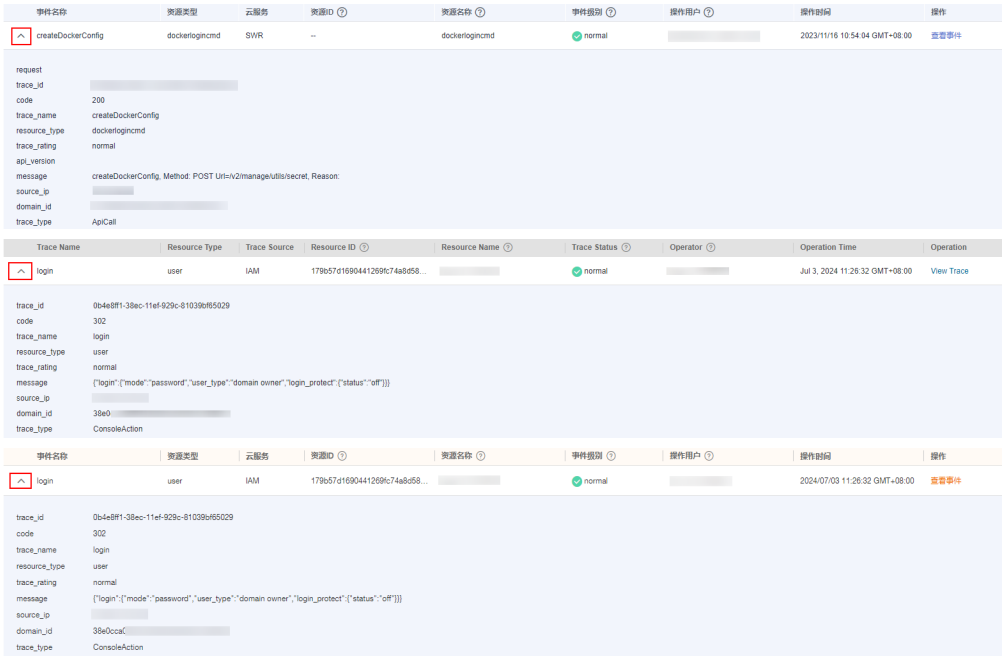


- 事件类型、事件来源、资源类型和筛选类型，在下拉框中选择查询条件。
  - 筛选类型按资源 ID 筛选时，还需手动输入某个具体的资源 ID。

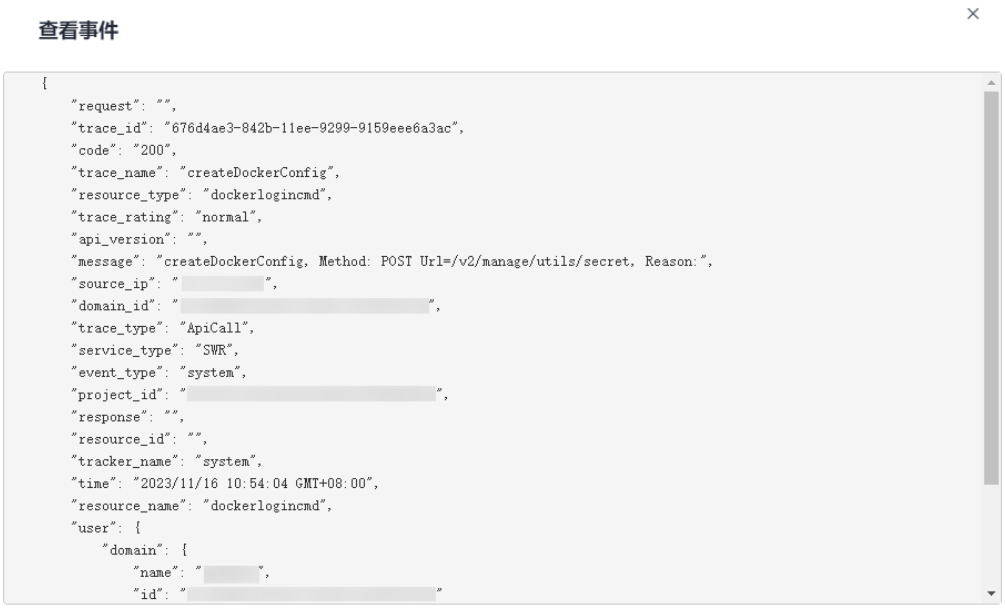
- 筛选类型按事件名称筛选时，还需选择某个具体的事件名称。
- 筛选类型按资源名称筛选时，还需选择或手动输入某个具体的资源名称。

- 操作用户：在下拉框中选择某一具体的操作用户，此操作用户指用户级别，而非租户级别。
- 事件级别：可选项为“所有事件级别”、“Normal”、“Warning”、“Incident”，只可选择其中一项。
- 时间范围：可选择查询最近 1 小时、最近 1 天、最近 1 周的操作事件，也可以自定义最近 7 天内任意时间段的操作事件。
- 单击“导出”按钮，云审计服务会将查询结果以 CSV 格式的表格文件导出，该 CSV 文件包含了本次查询结果的所有事件，且最多导出 5000 条信息。

6. 选择完查询条件后，单击“查询”。
7. 在事件列表页面，您还可以导出操作记录文件和刷新列表。
   - 单击“导出”按钮，云审计服务会将查询结果以 CSV 格式的表格文件导出，该 CSV 文件包含了本次查询结果的所有事件，且最多导出 5000 条信息。
   - 单击按钮，可以获取到事件操作记录的最新信息。
8. 在需要查看的事件左侧，单击 展开该记录的详细信息。



9. 在需要查看的记录右侧，单击“查看事件”，会弹出一个窗口显示该操作事件结构的详细信息。



10. 关于事件结构的关键字段详解，请参见《云审计服务用户指南》中的“云审计服务事件参考 > 事件结构”章节和“云审计服务事件参考 > 事件样例”章节。
11. （可选）在旧版事件列表页面，单击右上方的“体验新版”按钮，可切换至新版事件列表页面。

10 常见问题

10.1 产品咨询

10.1.1 云搜索服务如何保证数据和业务运行安全

云搜索服务主要从以下几个方面保障数据和业务运行安全：

- 网络隔离

  整个网络划分为 2 个平面，即业务平面和管理平面。两个平面采用物理隔离的方式进行部署，保证业务、管理各自网络的安全性。
  - 业务平面：主要是集群的网络平面，支持为用户提供业务通道，对外提供数据定义、索引、搜索能力。
  - 管理平面：主要是管理控制台，用于管理云搜索服务。
- 主机安全

  云搜索服务提供如下安全措施：
  - 通过 VPC 安全组来确保 VPC 内主机的安全。
  - 通过网络访问控制列表（ACL），可以允许或拒绝进入和退出各个子网的网络流量。
  - 内部安全基础设施（包括网络防火墙、入侵检测和防护系统）可以监视通过 IPsec VPN 连接进入或退出 VPC 的所有网络流量。
- 数据安全

  在云搜索服务中，通过多副本、集群跨 az 部署、索引数据第三方（OBS）备份功能保证用户的数据安全。

10.1.2 云搜索服务有哪些存储选项

云搜索服务采用 EVS 和本地磁盘存储用户的索引。在集群创建过程中，用户可指定 EVS 的类型及规格（即卷大小）。

- 支持 EVS 类型有普通 I/O、高 I/O、超高 I/O。
- 针对不同的 ECS，其对应的 EVS 卷大小限制根据创建集群选择的节点规格而定。

10.1.3 云搜索服务存储容量的上限是多少

创建集群过程中，最少可创建 1 个节点，最多可创建 200 个节点，其中每个节点（对应一个 ECS）可挂载一定数量的 EVS。

可参考不同 ECS 挂载 EVS 卷大小的不同，计算出云搜索服务存储容量的总大小，EVS 卷大小根据创建集群选择的节点规格而定。

10.1.4 有哪些工具可以使用云搜索服务

管理云搜索服务，或使用其搜索引擎的 API，提供了如下三种方式。可以基于已构建好的请求消息发起请求。

- curl

  curl 是一个命令行工具，用来执行各种 URL 操作和信息传输。curl 充当的是 HTTP 客户端，可以发送 HTTP 请求给服务端，并接收响应消息。curl 适用于接口调试。关于 curl 详细信息请参见 https://curl.haxx.se/。

- 编码

  通过编码调用接口，组装请求消息，并发送处理请求消息。

- REST 客户端

  Mozilla Firefox、Google Chrome 都为 REST 提供了图形化的浏览器插件，发送处理请求消息。

  - 针对 Firefox，请参见 Firefox REST Client。
  - 针对 Chrome，请参见 Postman。

10.1.5 申请的集群节点磁盘空间会有哪些开销

占用集群节点磁盘空间的日志及文件如下所示：

- 日志文件：Elasticsearch 日志
- 数据文件：Elasticsearch 索引文件
- 其他文件：集群配置文件
- 操作系统：默认余留 5%的存储空间

10.1.6 云搜索服务使用的数据压缩算法是什么？

云搜索服务支持的数据压缩算法有两种：一种是 Elasticsearch 默认的 **LZ4 算法**，另一种是 **best_compression 算法**。

- **LZ4 算法**

  LZ4 算法是 Elasticsearch 的默认压缩算法，该算法对数据的解压/压缩效率很快，但压缩率较低一些。

  压缩算法的实现流程：压缩过程以至少 4 个 bytes 为扫描窗口查找匹配，每次移动 1byte 进行扫描，遇到重复的就进行压缩。该算法适用于读取量大、写入量小的场景。

- **best_compression 算法**

除了默认的 LZ4 算法，云搜索服务还支持自定义 best_compression 算法。该算法适用于写入量大、索引存储成本高的场景，例如日志场景、时序分析场景等，可以大大降低索引的存储成本。

执行如下命令，可以将默认压缩算法（LZ4 算法）切换为 best_compression 算法：

```
PUT index-1
{
   "settings": {
      "index": {
         "codec": "best compression"
      }
   }
}
```

两者比较，LZ4 算法在解压/压缩速率方面更快一些，而 best_compression 算法在压缩率和解压率方面则更优秀一些。

10.2 集群管理

10.2.1 区域和可用区

10.2.1.1 什么是区域和可用区

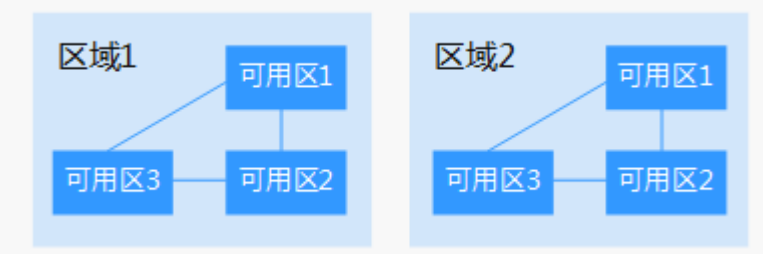
什么是区域、可用区？

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个 Region 内共享弹性计算、块存储、对象存储、VPC 网络、弹性公网 IP、镜像等公共服务。Region 分为通用 Region 和专属 Region，通用 Region 指面向公共租户提供通用云服务的 Region；专属 Region 指只承载同一类业务或只面向特定租户提供业务服务的专用 Region。
- 可用区（AZ，Availability Zone）：一个 AZ 是一个或多个物理数据中心的集合，有独立的风火水电，AZ 内逻辑上再将计算、网络、存储等资源划分成多个集群。一个 Region 中的多个 AZ 间通过高速光纤相连，以满足用户跨 AZ 构建高可用性系统的需求。

图 10-1 阐明了区域和可用区之间的关系。

图10-1 区域和可用区



如何选择区域？

建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。

如何选择可用区？

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

区域和终端节点

当您通过 API 使用资源时，您必须指定其区域终端节点。更多信息，请参考《云搜索服务 API 参考》的“终端节点”章节。

10.2.1.2 如何查看集群所分布的可用区？

在集群的“基本信息”页面，可以获取集群所分布的可用区信息。

1. 登录云搜索服务管理控制台。
2. 选择“集群管理 > Elasticsearch”，进入集群列表页面。
3. 单击集群名称进入集群“基本信息”页面，集群配置信息中的“可用区”即集群所分布的可用区。

图10-2 集群配置信息

10.2.2 集群版本

10.2.2.1 Filebeat 版本与集群版本的关系

- 非安全模式集群：不限制。
- 安全模式集群：需使用跟集群版本配套的 filebeat oss 版本，请参考 https://www.elastic.co/cn/downloads/past-releases#filebeat-oss 进行下载。

10.2.3 安全模式集群

10.2.3.1 如何获取 CSS 服务的安全证书？

CSS 服务只有启用 HTTPS 访问的安全集群才能下载安全证书（CloudSearchService.cer）。安全证书不支持在公网环境下使用。

1. 登录云搜索服务控制台。
2. 选择“集群管理”进入集群列表。
3. 单击对应集群的名称，进入集群基本信息页面。
4. 在“基本信息”页面，单击“HTTPS 访问”后面的“下载证书”。

10.2.3.2 如何转换 CER 安全证书的格式？

启用了 HTTPS 访问的安全集群可以下载 CSS 服务安全证书（CloudSearchService.cer）。而大多数软件支持“.pem”或“.jks”格式的证书，因此要对安全证书进行格式转换。

- 将安全证书从“.cer”格式转换为“.pem”格式。

```
openssl x509 -inform pem -in CloudSearchService.cer -out newname.pem
```

- 将安全证书从“.cer”格式转换为“.jks”格式。

```
keytool -import -alias newname -keystore ./truststore.jks -
file ./CloudSearchService.cer
```

其中，*newname* 是由用户自定义的证书名称。

执行命令后，会提示设置证书密码，并确认密码。请保存该密码，后续接入集群会使用。

10.2.3.3 CSS 集群支持修改安全组吗？

集群创建成功后，支持修改集群的安全组。

须知

- 进行安全组切换前，请确保业务接入时需要的 9200 端口已经放开，错误的安全组配置可能会导致业务无法访问，请谨慎操作。
- 建议在业务空闲时操作。
- 2023 年 2 月之前创建的集群无法进行安全组修改，建议至新集群后，进行安全组修改。

1. 登录云搜索服务控制台。
2. 选择“集群管理”进入集群列表。
3. 单击对应集群的名称，进入集群基本信息页面。
4. 在“安全组”右侧，单击“更改安全组”。

图10-3 更改安全组

5. 在“更改安全组”弹窗中，选择需要替换的安全组，单击“确定”启动更改任务。

10.2.4 参数配置

10.2.4.1 Elasticsearch 集群如何设置 search.max_buckets 参数？

问题描述

CSS 默认聚合操作中允许的最大 bucket 数量为 10000，如果执行聚合操作时需要返回超过 10000 个 bucket，可以通过修改 search.max_buckets 的值来允许更多的 bucket 返回。但是需要注意，增加 search.max_buckets 的值也会增加集群的负载和内存使用，因此需要谨慎使用。

解决方案

在 Kibana 的“Dev Tools”页面执行如下命令：

```
PUT _cluster/settings
{
   "persistent": {
      "search.max_buckets": 20000
   }
}
```

10.2.4.2 如何修改 Elasticsearch 集群的 TLS 算法？

目前 CSS 在 7.6.2 及以上版本支持修改 TLS 算法。

1. 登录云搜索服务控制台。
2. 选择“集群管理”进入集群列表。
3. 选择需要修改的集群，单击集群名称，进入集群基本信息页面。
4. 选择“参数配置”，单击“编辑”，展开“自定义”，单击“添加”。

   在自定义参数中添加“参数名”称为 **opendistro_security.ssl.http.enabled_ciphers**，“参数值”为 **['TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256', 'TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384']**

   说明

   如果“参数值”为多个的算法协议，需要一个中括号包围。如果“参数值”为单个的算法协议，需要单引号引起来。

5. 修改完成后，单击上方的“提交”弹出“提交配置”窗口，确认参数无误后勾选“参数修改后需要手动重启才能生效”，单击“确定”。

   当下方的参数修改列表显示“作业状态”为“成功”时，表示修改保存成功。
6. 返回集群列表，单击集群操作列的“更多 > 重启”重启集群，使修改的配置生效。

10.2.4.3 如何开启 Elasticsearch 集群的安全审计日志？

Elasticsearch 集群的安全审计日志功能默认是关闭的。目前 CSS 在 7.6.2 及以上版本的 Elasticsearch 集群支持开启安全审计日志功能。

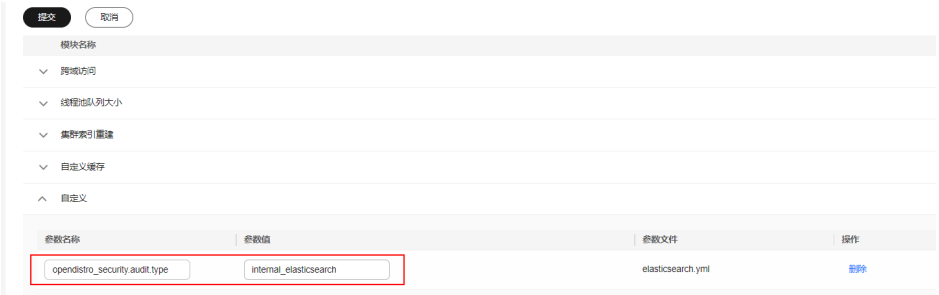
说明

集群须为安全集群。

1. 登录云搜索服务控制台。
2. 选择“集群管理 >Elasticsearch”进入集群列表。
3. 选择需要开启的集群，单击集群名称，进入集群基本信息页面。
4. 选择“参数配置”，单击“编辑”，展开“自定义”，单击“添加”。

   在自定义参数中添加“参数名称”为 **opendistro_security.audit.type**，“参数值”为 **internal_elasticsearch**。

   图10-4 自定义参数



5. 修改完成后，单击上方的“提交”弹出“提交配置”窗口，确认参数无误后勾选“参数修改后需要手动重启才能生效”，单击“确定”。

当下方的参数修改列表显示“作业状态”为“成功”时，表示修改保存成功。

6. 返回集群列表，单击集群操作列的“更多 > 重启”重启集群，使修改的配置生效。
7. 重启成功后，单击集群操作列的“Kibana”，进入 Kibana 页面，输入用户名及密码后，进入“Dev Tools”页面。
8. 在“Console”中输入 **GET _cat/indices?v** 命令，启动命令后查看结果，有 .*audit*相关的索引表示开启成功。

10.3 开源搜索引擎咨询

10.3.1 如何批量设置索引副本数为 0？

1. 登录集群 Kibana 界面，在 Kibana 的左侧导航中选择“Dev Tools”。
2. 执行命令 **PUT /*/_settings{"number_of_replicas":0}** 。

说明

*可能会匹配安全索引，不建议执行。建议执行批量操作需要的对应索引。如：**PUT /test*/_settings{"number_of_replicas":0}**。

10.3.2 为什么新创建的索引分片全部被分配到一个 node 节点上？

原因分析

新建索引分片被集中分配于一个 node 节点上可能有以下原因：

- 之前索引的分配导致某个节点上的 shards 数量过少，新建索引 shards 分配被 balance.shard 参数主导，为了平衡所有索引的全部分片，将 shards 集中分配在数量过少的节点上。
- 节点扩容，当新节点加入时新节点上的 shards 数量为 0，此时集群会自动进行 rebalance，但是 rebalance 需要时间，此时新建索引很容易会被 balande.shard 参数主导，平衡所有索引的分片，即都分配在新节点上看起来更加平衡。

涉及集群平衡性 shard 分配主要有两个配置参数：

cluster.routing.allocation.balance.index（默认值 0.45f）

cluster.routing.allocation.balance.shard（默认值 0.55f）

说明

- “balance.index”：值越大，shard 分配越倾向于使得每个索引的所有分片在节点上均匀分布，如 a 索引共有 6 个 shards，数据节点有 3 个，该配置值倾向于让 a 索引 2、2、2 平衡分配。
- “balance.shard”：值越大，shard 分配越倾向于使得所有分片（所有索引的）在节点上平衡，如索引 a 有 2 个 shards，索引 b 有 4 个 shards，该配置倾向于所有 6 个分片进行 2、2、2 平衡分配。
- balance.index 和 balance.shard 共同负责 shards 分配。

解决方案

当新建的索引分片被全部分配在一个 node 节点上时，有以下 2 种解决办法：

1. 扩容集群需要新建索引时，按照如下所示设置对应参数。

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

   即单个索引在每个节点上最多分配 2 个 shards。其中，具体每个节点最多分配多少个 shards，请根据集群数据节点个数、索引分片（主、副）的数量自行决定。

2. 如果是 shards 集中分配在数量过少的节点上导致索引 shards 分配到同一个节点上，可以使用 POST _cluster/reroute 的 move 命令迁移分片到其他节点，rebalance 模块会自动分配其他更合适的分片与其交换节点。根据具体业务使用场景可以适当调节 balance.index，balance.shard 配置。

10.3.3 Elasticsearch 7.x 集群如何在 index 下创建 type？

在 Elasticsearch 7.x 版本中，去掉了 type 概念，在 7.x 及以后的版本中，index 都不再支持创建 type。

如果需要强制使用，可以在命令中添加“include_type_name=true”强制使用 type 类型。

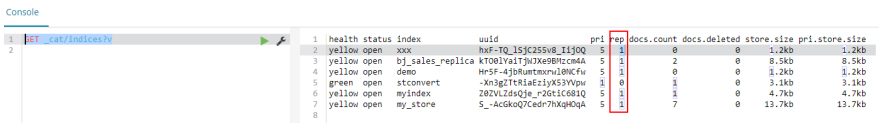
```
PUT _template/urldialinfo_template?include_type_name=true
```

执行命令后，界面会有提示：

```
"#! Deprecation: [types removal] Specifying include_type_name in put index template requests is deprecated. The parameter will be removed in the next major version. "
```

10.3.4 如何配置 CSS 集群双副本？

1. 在 Kibana 里执行 **GET _cat/indices?v** 命令确认集群副本的数目。如果 rep 参数列为 1，说明是双副本。

2. 如果不是，可以如下执行命令设置副本数。

   PUT /index/_settings

   {

   "number_of_replicas" : 1　　//表示需要设置的副本数

   }

说明

index 为需要修改的索引名称，需根据实际情况进行修改。

10.3.5 json 里设置了 1 个分片，是否可以通过修改配置，达到 4 分片，2 副本的效果

索引一旦创建成功，主 shards 数量不可变。

修改副本数，可通过在 Kibana 中执行以下命令：

```
PUT /indexname/_settings
{
"number_of_replicas" : 1      //表示需要设置的副本数
}
```

说明

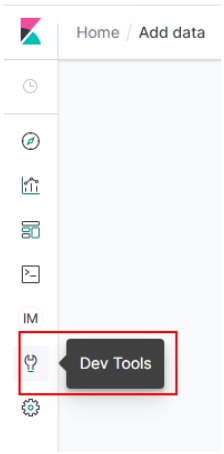
index 为需要修改的索引名称，需根据实际情况进行修改。

10.3.6 Elasticsearch 集群分片过多会有哪些影响

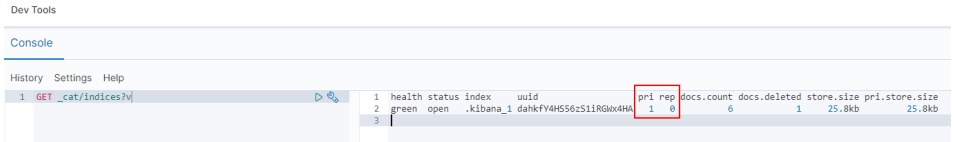
1. 集群创建分片的速度随着集群分片数量增多而逐渐减低。
2. 触发 Elasticsearch 自动创建 index 时，创建速度变慢会导致大量写入请求堆积在内存中，严重时可导致集群崩溃。
3. 分片过多时，如果不能及时掌控业务的变化，可能经常遇到单分片记录超限、写入拒绝等问题。

10.3.7 在 CSS 的控制台界面怎么查看集群的分片数以及副本数？

1. 登录云搜索控制台。
2. 在集群管理页面，选择需要查看的集群操作列的 Kibana。
3. 登录 Kibana 界面，选择 **Dev Tools**。



4. 在 **Dev Tools** 的 Console 界面中执行 **GET _cat/indices?v** 命令，查询集群分片数和副本数。如图，pri 列表示该索引分片数，rep 列表示副本数。索引一旦创建，pri 无法修改的，rep 可以动态修改。



10.3.8 Elasticsearch 集群在 Kibana 如何查询索引数据

在 Kibana 可以通过 API 查询索引数据，命令如下：

```
GET indexname/_search
```

返回数据如下图所示：

图10-5 返回数据

表10-1 参数说明

参数	描述
“took”	耗时几毫秒。
“time_out”	是否超时。
“_shard”	数据被拆到了 5 个分片上，搜索时使用了 5 个分片，5 个分片都成功地返回了数据，失败了 0 个，跳过了 0 个。
“hits.total”	查询结果的数量，3 个 document。
“max_score”	就是 document 对于一个 search 的相关度的匹配分数，越相关，就越匹配，分数也越高。
“hits.hits”	包含了匹配搜索的 document 的详细数据。

10.3.9 CSS 是否支持停止集群

CSS 不支持停止集群功能。如果是迁移集群的场景，用户需要将旧集群的功能先暂停，确认集群迁移成功后，再将旧集群删除。可以采用如下方式处理：

- 如果使用的集群版本支持，可以使用切断节点上除运维接口外的所有流量。
- 如果使用的集群版本不支持流量控制功能，可以关闭所有业务索引的读写。例如所有业务索引以 log 开头，在 Kibana 的“Dev Tools”页面执行以下命令：

```
PUT log*/_settings
{
  "index.blocks.read": true,
  "index.blocks.write": true,
  "index.blocks.metadata": true
}
```

10.3.10 Elasticsearch 集群中某个客户端节点的 node.roles 为 i 表示该节点是 ingest 节点吗？

问题描述

集群某个客户端节点的“node.roles”为“i”表示该节点是 ingest 节点吗？

- 如果客户端节点是 ingest 节点，那么集群中是否存在 Coordinating only node，所有节点都是 Coordinating node 分摊来客户端请求吗？
- 如果没有 ingest 业务时，那么客户端节点是不是就处于空闲状态？

解决方案

集群节点的“node.roles”为“i”时，表示集群的客户端节点上启用了 ingest 节点模式。

- Elasticsearch 的“coordinating only node”在 CSS 服务中称为“client node”，如果集群中没有设置 client node，则所有节点都是 client node 共同分摊客户端请求。
- ingest 节点相当于一套 ELK，用于数据转换，当没有 ingest 业务时，客户端节点也不会闲置。

10.3.11 Elasticsearch 集群设置默认分页返回最大条数

解决方案

- 方法 1：

  打开 Kibana，在 devtools 界面执行如下命令：

```
PUT _all/_settings?preserve_existing=true
{
"index.max_result_window" : "10000000"
}
```

- 方法 2：

  后台执行如下命令进行设置：

```
curl -XPUT 'http://localhost:9200/_all/_setting?preserve_existing=true'-d
{
"index.max_result_window":"1000000"
}
```

⚠ 注意

该配置会相应的消耗内存与 CPU，请谨慎设置。

10.3.12 如何更新 Elasticsearch 集群生命周期策略？

Elasticsearch 生命周期实现使用的是 Open Distro 的 ISM。此处简单介绍不涉及 ISM template 的策略更新步骤，如果要配置有关 ISM template 的策略可以参考 Open Distro 文档。

1. 当创建一个 policy 时，系统会往**.opendistro-ism-config** 索引中写入一条数据，这条数据的“_id”就是 policy 的名字，内容是 policy 的定义。

   图10-6 写入一条数据

2. 将 policy 和索引绑定以后，系统会再往**.opendistro-ism-config** 索引中写入一条数据。这条数据的初始状态如下图所示。

图10-7 数据初始状态


```
{
  "_index" : ".opendistro-ism-config",
  "_type" : "_doc",
  "_id" : "FABkSF5GSTCmR0QkW41HVw",
  "_score" : 1.0,
  "_source" : {
    "managed_index" : {
      "name" : "data1",
      "enabled" : true,
      "index" : "data1",
      "index_uuid" : "FABkSF5GSTCmR0QkW41HVw",
      "schedule" : {
        "interval" : {
          "start_time" : 1641432652693,
          "period" : 1,
          "unit" : "Minutes"
        }
      },
      "last_updated_time" : 1641432652694,
      "enabled_time" : 1641432652694,
      "policy_id" : "policy1",
      "policy_seq_no" : null,
      "policy_primary_term" : null,
      "policy" : null,
      "change_policy" : null
    }
  }
}
]
```


3. 执行 explain 命令，此时返回的内容只有一条 policy 的 id。

```
GET _opendistro/_ism/explain/data2
{
  "data2" : {
    "index.opendistro.index_state_management.policy_id" : "policy1"
  }
}
```

之后 Open Distro 会执行一个初始化的流程，将 policy 的内容填到这条数据中，初始化以后的数据如下图所示。

图10-8 初始化后数据

初始化结束后，policy 中的 min_index_age 都会被复制过来。

说明

如果此时去更新 policy 的内容，已经完成初始化流程的索引是完全不感知的，因为他已经将旧的 policy 的内容复制了一份，更新 policy 的时候不会去更新复制的那部分内容。

4. 修改完 policy 以后，执行 change_policy API 完成策略更新，如下所示。

```
POST _opendistro/_ism/change_policy/data1
{
  "policy_id": "policy1"
}
```

参考信息

关于如何创建、使用生命周期策略，请参见。

10.3.13 如何设置云搜索服务的慢查询日志的阈值？

云搜索服务的慢查询日志设置和 elasticsearch 保持一致，通过 _settings 接口设置。例如，您可以在 Kibana 中执行如下样例，设置索引级别。

```
PUT /my_index/_settings
{
    "index.search.slowlog.threshold.query.warn": "10s",
    "index.search.slowlog.threshold.fetch.debug": "500ms",
```

```
    "index.indexing.slowlog.threshold.index.info": "5s"
}
```

- 查询慢于 10 秒输出一个 WARN 日志。
- 获取慢于 500 毫秒输出一个 DEBUG 日志。
- 索引慢于 5 秒输出一个 INFO 日志。

详细可参考官网：https://www.elastic.co/guide/cn/elasticsearch/guide/current/logging.html

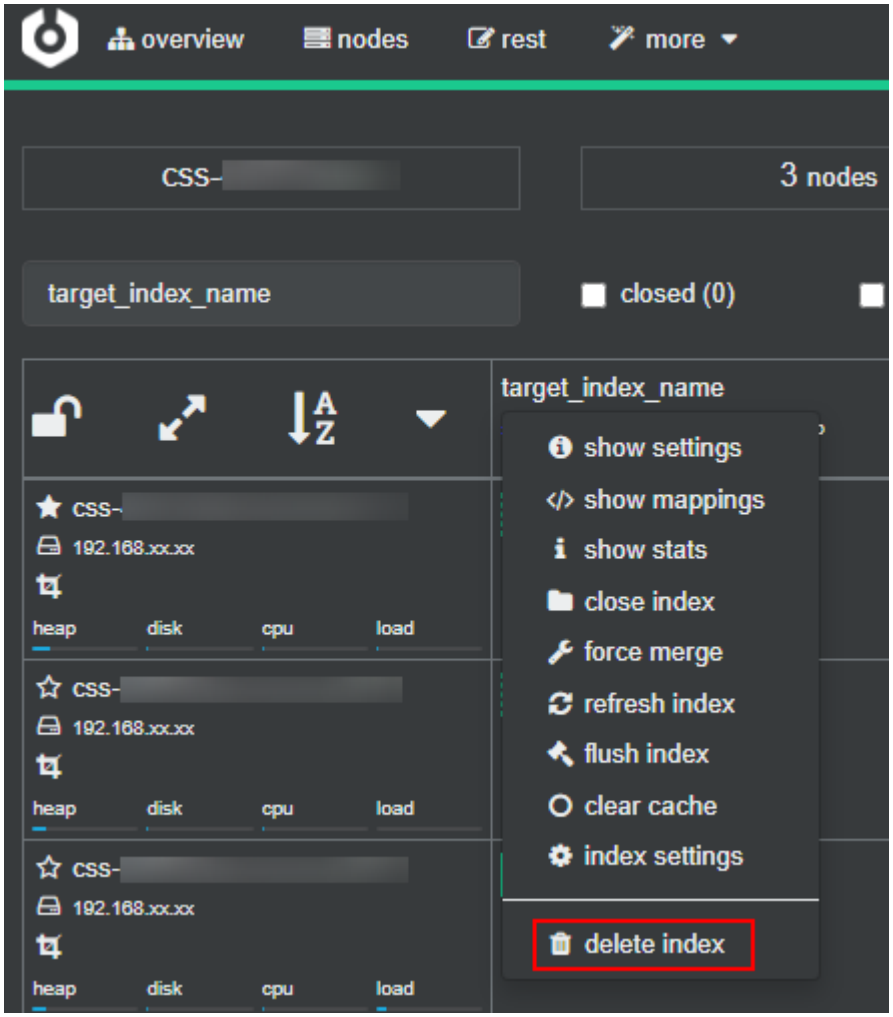
10.3.14 如何清理索引数据？

- 自动化定期清理

  可以写定时任务调用清理索引的请求，定期执行。CSS 支持 Opendistro Index State Management。详见：https://opendistro.github.io/for-elasticsearch-docs/docs/im/ism/。

- 手动清理
  - 登录 Kibana，在 Dev tools 里执行 **DELETE /索引名**命令。
  - 登录 Cerebro, 过滤检索出目标索引名，单击索引名称，选择**“delete index”**并在弹框中选择 **“Confirm”。**

图10-9 Cerebro 删除索引



- 登录弹性云服务器，删除单条索引数据命令。

  curl -XDELETE http://IP:9200/索引名

  删除某一天 logstash 的所有数据命令，例如删除 19 号所有数据。

  非安全模式集群：**curl -XDELETE 'http://IP:9200/logstash-2017.06.19*'**

  安全模式集群：**curl -XDELETE -u username:password 'https://IP:9200/logstash-2017.06.19' -k**

说明

- username：管理员账户名默认为 admin。
- password：创建集群时设置的密码。
- IP：任意一个集群节点的 IP 地址。

10.3.15 CSS 集群如何清理缓存？

- **清理 fielddata**

进行聚合和排序时，会使用 **fielddata** 数据结构，会占用较大内存。

a. 在 Kibana 执行如下命令，查看索引的 **fielddata** 占用情况。

```
DELETE /_search/scroll
{
"scroll_id" :
"DXF1ZXJ5QW5kRmV0Y2gBAAAAAAAAAD4WYm9laVYtZndUQlNsdDcwakFMNjU1QQ=="
}
```

b. 当 **fielddata** 占用内存过高时，可以执行如下命令清理 **fielddata**。

```
POST /test/_cache/clear?fielddata=true
```

“test”为 fielddata 占用内存较高的索引名称。

- **清理 segment**

  每个 segment 的 FST 结构都会被加载到内存中，并且这些内存是不会被垃圾回收的。因此如果索引的 segment 数量过大，会导致内存使用率较高，建议定期进行清理。

  a. 在 Kibana 执行如下命令，查看各节点的 segment 数量和占用内存大小。

  ```
  GET /_cat/nodes?v&h=segments.count,segments.memory&s=segments.memory:desc
  ```

  b. 如果 segment 占用内存过高时，可以通过删除部分不用的索引、关闭索引或定期合并不再更新的索引等方式释放内存。

- **清理 cache**

  在 Kibana 执行如下命令清理 cache。

  ```
  POST _cache/clear
  ```

10.3.16 使用 delete_by_query 命令删除数据后，为什么磁盘使用率反而增加？

使用 **delete_by_query** 命令删除数据并不是真正意义上的物理删除，它仅仅是对数据增加了删除标记。当再次搜索时，会搜索全部数据后再过滤掉带有删除标记的数据。

因此，该索引所占的空间并不会因为执行磁盘删除命令后马上释放掉，只有等到下一次段合并时才真正的被物理删除，这个时候磁盘空间才会释放。

相反，在查询带有删除数据时需要占用磁盘空间，这时执行磁盘删除命令不但没有被释放磁盘空间，反而磁盘使用率上升了。

10.4 集群插件使用

10.4.1 云搜索服务是否支持 SearchGuard 插件的安装？

不支持。

云搜索服务提供了安全模式集群与 SearchGuard 插件功能一样。

10.4.2 Elasticsearch 集群原生的 script dotProduct 无法执行

问题原因

使用向量功能时，由于原生 Elasticsearch 向量功能是在 x-pack 插件中，目前云搜索服务没有集成 x-pack 组件。因此，Elasticsearch 集群原生的 script dotProduct 无法执行。

解决方案

建议您使用云搜索服务自研的向量检索功能。云搜索服务基于自研的向量搜索引擎，结合 Elasticsearch 的插件机制，高效集成了向量检索能力，能够同时满足高性能、高精度、低成本、多模态等多种高维向量检索场景。更多信息请参考。

说明

向量检索功能支持 7.6.2 和 7.10.2 版本的集群。

10.5 集群访问/集群连接

10.5.1 自行搭建的 Kibana 和 Cerebro 可以访问 CSS 集群吗？

支持自建 Kibana 和 Cerebro 接入 CSS 集群。

- 自建 Kibana 对接 CSS 集群：请参见 10.9.2 自建 Kibana 如何对接云搜索服务的 Elasticsearch？。
- 自建 Cerebro 对接 CSS 集群：直接启动自建 Cerebro，启动后填写集群的内网 IP 即可。
  - 安全模式的集群填写：https://***内网 IP***:9200
  - 非安全模式的集群填写：http://***内网 IP***:9200

10.5.2 9200 和 9300 端口是否都开放？

都开放。9200 端口为外部访问 Elasticsearch 集群端口，9300 为节点之间通讯端口。

访问 9300 端口有以下几种方式：

- 如果是同 VPC 同子网内可直接访问。
- 如果是同 VPC 下跨子网访问，需要单独申请路由配置。
- 如果是不同的 VPC 不同的子网访问，需要先通过对等连接，打通两个 VPC 网络，然后单独申请路由配置，联通两个子网。

10.5.3 如何使用 NAT 网关实现云搜索服务公网访问

开通公网访问云搜索服务操作视图：

1.获取云搜索服务信息

2.配置 NAT 网关

3.修改云搜索服务安全组规则

4.通过公网访问云搜索服务

⚠ 注意

如果非安全模式集群使用此功能，则会把集群数据直接暴露到公网，请禁用此功能。

获取云搜索服务信息

步骤 1 登录云搜索服务管理控制台。

步骤 2 在“集群管理”页面，单击集群名称，进入集群基本信息页面。

步骤 3 在“配置信息”模块获取内网访问地址、VPC 和子网信息，如图 10-10 所示。

图10-10 获取信息

----结束

配置 NAT 网关

步骤 1 创建 NAT 网关。

1. 登录控制台，在“服务列表”搜索“NAT 网关”，进入网络控制台页面。
2. 单击“购买 NAT 网关”，配置 NAT 网关的相关信息。详细请参考《NAT 网关用户指南》**购买 NAT 网关**。

说明

“虚拟私有云”和“子网”配置为获取云搜索服务信息获取的信息。

3. 配置完成后，单击“立即购买”。

步骤 2 添加 DNAT 规则。

1. NAT 网关购买成功后，在 NAT 控制台，单击购买成功的 NAT 网关“名称”，进入 NAT 网关详情页面。
2. 选择“DNAT 规则”页签，单击“添加 DNAT 规则”。详细请参考《NAT 网关用户指南》**添加 DNAT 规则**。

说明

- 弹性公网 IP：可以根据自己业务在弹性公网 IP 页面创建。
- 公网端口：可以自定义。
- 私网 IP：云搜索服务的内网访问 IP，即获取云搜索服务信息获取的“内网访问地址”。
- 私网端口：9200
- 如果创建的集群包含多个“内网访问地址”，则需要添加多个 DNAT 规则。

3. 添加完成后，单击“确定”。

----结束

修改云搜索服务安全组规则

步骤 1 登录云搜索服务管理控制台，选择对应的集群，单击集群名称，进入集群“基本信息”页面。

步骤 2 在“基本信息”页面，单击“安全组”跳转到安全组基本信息页面。

步骤 3 在安全组“基本信息”页面，选择“入方向规则”页签。

步骤 4 单击“添加规则”，添加 9200 端口入方向规则。

步骤 5 配置完成后，单击“确定”。

----结束

通过公网访问云搜索服务

在浏览器中输入 https://IP:port 或者 http://IP:port，访问云搜索服务。

- IP:port：弹性公网 IP:端口号，即创建 DNAT 规则设置的弹性公网 IP 和公网端口。

- 如果集群开启了**安全模式**，请输入 https://IP:port，并且输入安全模式的用户名和密码。
- 如果集群未开启**安全模式**，请输入 http://IP:port。

10.5.4 新建集群是否可以使用老集群 IP 地址？

原集群的 IP 地址无法更换为新集群的 IP 地址。

如果集群 IP 发生变化，可能原因如下：

- 集群内网 IP 发生变化。

  确认集群是否进行扩/缩容操作。扩容操作会增加集群内网 IP。缩容操作会减少集群 IP，如减少的节点有业务运行，会发生故障。
- 集群公网 IP 发生变化。

  确认集群是否开启安全模式。6.5.4 版本之前的 Elasticsearch 集群在创建时支持开启安全模式、公网访问，6.5.4 及之后版本的 Elasticsearch 集群支持开启“安全模式”，开启后集群会增加一个公网 IP。对于已经绑定的公网 IP，可以在集群的中通过单击“公网访问”参数右侧的“解绑”，解绑公网 IP。
- 用户侧本地 IP 发生变化。

  对已经配置了公网访问的集群，可以在集群的中通过单击“访问控制”右侧的“设置”，设置访问控制开关和访问白名单。只有包含在访问白名单里的 IP 才能访问集群。

10.5.5 CSS 集群是否支持采用 x-pack-sql-jdbc 进行客户端连接并查询？

不支持。

目前云搜索服务没有集成 x-pack 组件。

10.5.6 ECS 无法连接到集群

遇到该问题，请按照如下操作步骤排查解决。

1. 先确认 ECS 实例和集群是否在同一个 VPC。
   - 如果在，执行步骤 2。
   - 如果不在，需要重新创建 ECS 实例，使之和集群在同一个 VPC 下。
2. 查看集群的安全组的出方向和入方向是否已允许 9200 端口（TCP 协议），或者允许的端口范围已包含 9200 端口（TCP 协议）。
   - 如果是，执行步骤 3。
   - 如果不是，请前往 VPC 页面，设置“安全组”的出方向和入方向已允许 9200 端口或允许的端口范围已包含 9200 端口。
3. 查看 ECS 实例是否添加安全组。
   - 如果有，检查安全组的配置规则是否满足要求，在集群“基本信息”页面，可以查看“安全组信息”。然后执行步骤 4。

- 如果没有，从 ECS 的实例详情页面，进入 VPC 页面，选择“安全组”，添加安全组。

4. 在 ECS 实例上，测试是否可以正常连接到集群。

   ssh *< 节点的内网访问地址和端口号>*

   说明

   当集群包含多个节点时，需要逐个节点测试是否可以正常连接到该集群中的每个节点。

   - 如果可以通信，说明网络是正常的。
   - 如果端口不通，请联系技术支持协助排查。

10.6 集群迁移

10.6.1 Elasticsearch 是否支持不同 VPC 之间的数据迁移？

Elasticsearch 不支持直接迁移不同 VPC 之间的数据，但是可以通过以下 2 种方式进行迁移。

方法一：

可以使用备份与恢复功能迁移集群数据。

方法二：

1. 打通 VPC 网络，建立对等连接。
2. 打通网络后，使用 Logstash 进行数据迁移。

10.6.2 如何跨 Region 迁移 CSS 集群？

CSS 集群不支持直接迁移，但可以通过 OBS 桶备份和恢复的方式进行数据迁移实现集群迁移。

- 如果 OBS 桶在同一个区域，请参考进行集群迁移。
- 如果 OBS 桶跨区域，请先参考进行跨区域复制 OBS 桶，再参考进行集群迁移。

说明

- 在跨区域复制之前，要保证目标集群设置的快照文件夹为空，否则无法将快照信息刷新到目标集群的快照列表中。
- 每次迁移都需要将文件夹置空。

10.7 集群备份与恢复

10.7.1 如何查询快照信息？

前提条件

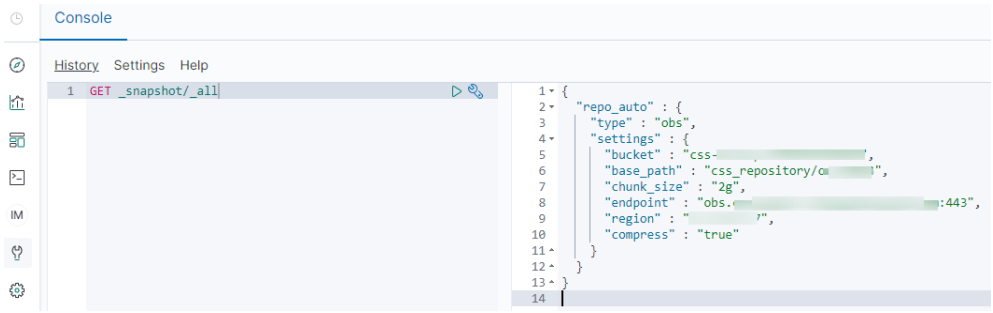
集群开启了快照，并且设置了快照信息。

快照查询

1. 在云搜索服务的“集群管理”页面上，单击集群“操作”列的“Kibana”访问集群。
2. 在 Kibana 的左侧导航中选择“Dev Tools”，单击“Get to work”，进入 Console 界面。

   Console 左侧区域为输入框，右侧为结果输出区域， 为执行命令按钮。
3. 执行命令 **GET _snapshot/_all**，查询所有仓库信息，如下图所示。

   图10-11 查询所有仓库信息



   - bucket：OBS 桶名。
   - base_path：路径名称。前缀默认固定，后面是集群名称。
   - endpoint：OBS 域名。
   - region：所在 region。
4. 查询指定快照信息。
   a. 执行 **GET _snapshot/repo_auto/_all** 命令，查询当前仓库下面所有的快照列表。

图10-12 快照信息

- snapshot：快照名称。
- state：快照状态。
- start_time、start_time_in_millis、end_time、end_time_in_millis：快照时间。
- shards：shards 个数。total 表示总共的个数。failed 表示失败的个数。successful 表示成功的个数。

b. 执行 **GET _snapshot/repo_auto/$snapshot-xxx**，查询指定快照信息。
   - **$snapshot-xxx** 需根据实际情况替换为具体的快照名称。
   - **repo_auto** 后面跟快照名称，也可以跟通配符。

5. （可选）删除指定快照信息。

   如果要删除指定的快照，执行 **DELETE _snapshot/repo_auto/$snapshot-xxx**。

   $snapshot-xxx 需根据实际情况替换为具体的快照名称。

10.7.2 集群被删除后是否还能恢复？

如果被删除的集群启用过快照功能，且 OBS 桶中创建的快照并未被删除，则可以通过 OBS 桶中存储的快照信息恢复集群。否则，被删除的集群无法被恢复，因此请谨慎操作删除任务。

通过 OBS 桶中存储的快照信息恢复被删除集群的操作步骤：

1. 登录云搜索服务管理控制台。
2. 单击右上角的“创建集群”新建一个集群，创建集群过程中集群快照开关关闭，待集群创建完成后开启集群快照。

须知

新集群和被删集群要在同一个 region 下，集群的版本要等于或高于被删集群，新集群的节点数至少要大于被删集群节点数的一半，否则集群可能恢复失败。

3. 当新建集群的“集群状态”会变为“可用”时，单击集群名称进入“基本信息”页面。
4. 在左侧导航栏选择“集群快照”，进入“集群快照”管理页面。启用集群快照功能。其中，“OBS 桶”和“备份路径”填写被删除集群存放快照信息的 OBS 桶和路径。

   保存配置后，在快照管理列表中，**等待几分钟后**，可以看到被删除集群的快照信息。**如果没有显示，可再次编辑快照基础配置，修改备份路径为其他再改回正确备份路径后保存，再次重试。**

   说明

   如果要在其他已创建好的集群上恢复被删集群的数据，也需要将已创建好的集群快照的“OBS 桶”和“备份路径”参数配置为被删除集群存放快照信息的 OBS 桶和路径。

5. 单击快照“操作”列的“恢复”，弹出“恢复”页面。
6. 在“恢复”页面配置集群的恢复参数。

   “索引”：指定需要进行恢复的索引名称，默认为空。如保持默认值，即不指定索引名称，则表示恢复所有的索引数据。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?特殊字符。支持使用“*”匹配多个索引，比如 index*，表示恢复快照中名称前缀是 index 的所有索引。

   “索引名称匹配模式”：在恢复时，可以根据文本框中定义的过滤条件去恢复符合条件的索引，过滤条件请使用正则表达式。默认值“index_(.+)”表示所有的索引。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?,特殊字符。

   “索引名称替换模式”：索引重命名的规则。默认值“restored_index_$1”表示在所有恢复的索引名称前面加上“restored_”。0～1024 个字符，不能包含空格和大写字母，且不能包含"\<|>/?,特殊字符。

   说明

   “索引名称匹配模式”和“索引名称替换模式”需要同时设置才会生效。

   “集群”：选择需要进行恢复的集群名称，可选择当前集群或者其他集群。只能选择处于“可用”状态的集群，如果快照所属的集群处于“不可用”状态，那么也无法将快照恢复到本集群。恢复到其他集群时，目标集群中的版本不低于本集群。如果已选择其他集群，且该集群中存在同名的索引，则恢复完成后，该同名的索引中的数据将会被覆盖，请谨慎操作。

   “覆盖目标集群同名桶 shard 结构索引”：默认不覆盖，通过快照恢复数据，是以快照文件覆盖的形式进行数据恢复。覆盖目标集群同名的索引后，可能导致目标集群同名索引数据丢失，请谨慎勾选操作。

图10-13 恢复快照

7. 单击“确定”开始恢复。恢复成功，快照列表中“任务状态”将变更为“恢复成功”，索引数据将根据快照信息重新生成。

图10-14 恢复成功

10.8 集群监控与运维

10.8.1 用户平时需要关注云搜索服务的哪些监控指标

用户需要关注的监控指标为磁盘使用率和集群健康状态。用户可以登录到云监控服务，根据实际应用场景配置告警提示，当收到告警，可采取相应措施消除告警。

配置示例：

- 如果在某段时间内（如 5min），磁盘使用率出现多次（如 5 次）不低于某特定值（如 85%）的情况，则发出相应告警。
- 如果在某段时间内（如 5min），集群健康状态出现多次（如 5 次）大于 0 的情况，则发出相应告警。

采取措施：

- 收到与磁盘使用率有关的告警时，可以调查磁盘空间消耗，查看是否可以从集群节点中删除数据或是将数据存档到其他系统以释放空间，或者扩容磁盘。

- 收到与集群健康状态有关的告警时，可以查看集群的分片分配是否正常以及 Shard 是否已丢失，在 Cerebro 上查看进程是否发生重启。

10.8.2 Elasticsearch 集群平均已用内存比例达到 98%

问题现象

查看集群监控发现，Elasticsearch 集群“平均已用内存比例”一直处于 98%，用户担心内存比例过高是否对集群有影响。

问题原因

在 Elasticsearch 集群中，Elasticsearch 会占用 50%内存，另外 50%内存会被 Lucene 用于缓存文件，因此节点内存占用会一直很高，平均已用内存比例达到 98%是正常现象，请您放心使用。

解决方案

您可以关注“最大 JVM 堆使用率”和“平均 JVM 堆使用率”这两个指标来监控集群内存使用情况。

10.8.3 如何查看集群总磁盘使用率？

在集群的“基本信息”页面，可以查看集群总磁盘使用率。

1. 登录云搜索服务管理控制台。
2. 选择“集群管理 > Elasticsearch”，进入集群列表页面。
3. 单击集群名称进入集群“基本信息”页面，集群配置信息中“集群存储使用量”与“集群存储容量”的比值，即集群总磁盘使用率。

图10-15 集群基本信息

10.8.4 单节点的使用率过高是否会影响集群的业务？

问题现象

查看集群监控发现，Elasticsearch 集群“磁盘使用率”达到 80%以上，用户担心单节点使用率过高会对集群业务产生影响。

业务影响

- 单节点使用率超过 85%：会导致新的分片无法分配。
- 单节点使用率超过 90%：Elasticsearch 会尝试将对应节点中的分片迁移到其他磁盘使用率比较低的数据节点中。
- 单节点使用率超过 95%：系统会对 Elasticsearch 集群中对应节点里每个索引强制设置 read_only_allow_delete 属性，此时该节点上的所有索引将无法写入数据，只能读取和删除对应索引。

单节点使用率过高，可通过操作动态调整集群节点的数量和容量。新扩容的节点不会立即分配索引，可打开 cerebro 看下节点索引分配情况，做进一步观察，也可以修改 indices.recovery.max_bytes_per_sec 和 cluster.routing.allocation.cluster_concurrent_rebalance 两个参数值增加索引分配速度。

10.9 Kibana 使用

10.9.1 如何修改登录 Kibana 和 Cerebro 的管理员密码

当您想要更换登录 Kibana 和 Cerebro 的管理员密码，或者忘记管理员密码时，可以对密码进行重置。

1. 在集群管理列表，选择需要重置密码的集群，单击集群名称，进入集群基本信息页面。
2. 在“配置信息”区域，单击“重置密码”后的“重置”，设置并确认新的管理员密码。

说明

- 可输入的字符串长度为 8~32 个字符。
- 密码至少包含大写字母、小写字母、数字和特殊字符四类中的三类。其中支持的特殊字符有：~!@#$%^&*()-_=+\|[{}];:,<.>/?
- 不能与管理员账户名或倒序的管理员账户名相同。
- 建议定期修改密码。

10.9.2 自建 Kibana 如何对接云搜索服务的 Elasticsearch？

自建 Kibana 对接云搜索服务的 ES，需满足如下条件：

- 本地环境需要支持外网访问。
- 通过同 vpc 下 ECS 服务搭建 Kibana，本地公网访问 Kibana 即可。
- 只支持 OSS 版本的 Kibana 镜像连接到云搜索服务的 Elasticsearch。

Kibana 配置文件参考：

- 安全模式：

```
elasticsearch.username: "***"
elasticsearch.password: "***"
elasticsearch.ssl.verificationMode: none
server.ssl.enabled: false
server.rewriteBasePath: false
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: https://10.0.0.xxx:9200
elasticsearch.requestHeadersWhitelist: ["securitytenant","Authorization"]
opendistro_security.multitenancy.enabled: true
opendistro_security.multitenancy.tenants.enable_global: true
opendistro_security.multitenancy.tenants.enable_private: true
opendistro_security.multitenancy.tenants.preferred: ["Private", "Global"]
opendistro_security.multitenancy.enable_filter: false
```

说明

- 安全模式需要安装插件 opendistro_security_kibana，详细请参考 https://github.com/opendistro-for-elasticsearch/security-kibana-plugin/tags?after=v1.3.0.0。
- 安装的插件版本需要和集群版本保持一致，可通过 **GET _cat/plugins** 获取到集群安全插件的版本号。

- 非安全模式：

```
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: http://10.0.0.xxx:9200
```

10.9.3 Kibana 是否支持导出数据功能？

Kibana 导出数据需要依赖 SQL Workbench 插件，目前云搜索服务只有 Elasticsearch 7.6.2 及以上的版本支持。

在 Kibana 的 SQL Workbench 里，输入 Elasticsearch SQL 语句可以查询数据，也可以“Download”导出数据，支持自定义导出 1~200 条数据，缺省导出 200 条数据。

图10-16 SQL Workbench