

云搜索服务

常见问题

文档版本 23
发布日期 2024-03-04



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

1 产品咨询	1
1.1 云搜索服务如何保证数据和业务运行安全	1
1.2 云搜索服务有哪些存储选项	1
1.3 云搜索服务存储容量的上限是多少	2
1.4 有哪些工具可以使用云搜索服务	2
1.5 申请的集群节点磁盘空间会有哪些开销	2
1.6 云搜索服务使用的数据压缩算法是什么？	2
2 计费相关	4
2.1 如何退订云搜索服务集群？	4
2.2 云搜索服务的包周期资源如何续费？	4
2.3 购买了折扣套餐包，如何使用？	5
2.4 折扣套餐包是如何扣费的？	5
2.5 折扣套餐包到期了怎么续费？	5
2.6 套餐包到期时间不同，如何将按需计费转为包年包月计费模式？	6
3 集群管理	7
3.1 区域和可用区	7
3.1.1 什么是区域和可用区	7
3.1.2 如何查看集群所分布的可用区？	8
3.2 集群版本	9
3.2.1 Filebeat 版本与集群版本的关系	9
3.3 安全模式集群	9
3.3.1 如何获取 CSS 服务的安全证书？	10
3.3.2 如何转换 CER 安全证书的格式？	10
3.3.3 CSS 集群支持修改安全组吗？	11
3.4 参数配置	11
3.4.1 Elasticsearch 集群如何设置 search.max_buckets 参数？	12
3.4.2 如何修改 Elasticsearch 集群的 TLS 算法？	12
3.4.3 如何开启 Elasticsearch 集群的安全审计日志？	12
4 开源搜索引擎咨询	14
4.1 如何批量设置索引副本数为 0？	14
4.2 为什么新创建的索引分片全部被分配到一个 node 节点上？	14
4.3 Elasticsearch 7.x 集群如何在 index 下创建 type？	15

4.4 如何配置 CSS 集群双副本？	15
4.5 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.....	16
4.6 Elasticsearch 集群分片过多会有哪些影响.....	16
4.7 在 CSS 的控制台界面怎么查看集群的分片数以及副本数？	16
4.8 Elasticsearch 集群在 Kibana 如何查询索引数据.....	17
4.9 CSS 是否支持停止集群.....	18
4.10 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？	18
4.11 Elasticsearch 集群设置默认分页返回最大条数.....	19
4.12 如何更新 Elasticsearch 集群生命周期策略？	19
4.13 如何设置云搜索服务的慢查询日志的阈值？	21
4.14 如何清理索引数据？	22
4.15 CSS 集群如何清理缓存？	23
4.16 使用 delete_by_query 命令删除数据后，为什么磁盘使用率反而增加？	24
5 集群插件使用.....	25
5.1 云搜索服务是否支持 SearchGuard 插件的安装？	25
5.2 Elasticsearch 集群原生的 script dotProduct 无法执行.....	25
6 集群访问/集群连接.....	26
6.1 自行搭建的 Kibana 和 Cerebro 可以访问 CSS 集群吗？	26
6.2 9200 和 9300 端口是否都开放？	26
6.3 如何使用 NAT 网关实现云搜索服务公网访问.....	26
6.4 新建集群是否可以使用老集群 IP 地址？	28
6.5 CSS 集群是否支持采用 x-pack-sql-jdbc 进行客户端连接并查询？	29
6.6 ECS 无法连接到集群.....	29
7 集群迁移.....	30
7.1 Elasticsearch 是否支持不同 VPC 之间的数据迁移？	30
7.2 如何跨 Region 迁移 CSS 集群？	30
8 集群备份与恢复.....	31
8.1 如何查询快照信息？	31
8.2 集群被删除后是否还能恢复？	32
9 集群监控与运维.....	35
9.1 用户平时需要关注云搜索服务的哪些监控指标.....	35
9.2 Elasticsearch 集群平均已用内存比例达到 98%.....	35
9.3 如何查看集群总磁盘使用率？	36
9.4 单节点的使用率过高是否会影响集群的业务？	36
10 Kibana 使用.....	38
10.1 如何修改登录 Kibana 和 Cerebro 的管理员密码.....	38
10.2 自建 Kibana 如何对接云搜索服务的 Elasticsearch？	39
10.3 Kibana 是否支持导出数据功能？	40

1 产品咨询

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。

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。

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。

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](#)。

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

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。

2 计费相关

2.1 如何退订云搜索服务集群？

包年包月集群退订

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，选择需要退订的集群。
3. 在操作列单击“更多 > 退订/释放”。
4. 在“退订集群”对话框中，输入需要退订/释放的集群名称，单击“确定”。
进入退订资源页面，可以在该页面核对资源信息以及退费金额。
5. 填写退订原因，勾选相关协议后，单击“退订”。
在弹出确认退订提示信息后，再次单击“退订”。

按需计费集群退订

1. 登录云搜索服务管理控制台。
2. 在集群管理页面，选择需要退订的集群。
3. 在操作列单击“更多 > 删除”。
4. 在“删除集群”对话框中，输入需要删除的集群名称，单击“确定”。

折扣套餐包退订

已购买的折扣套餐包不支持退订。

2.2 云搜索服务的包周期资源如何续费？

云搜索服务支持按年按月续费资源，具体续费操作如下：

已有集群时如何续费

如果是已经建好的包周期集群，需要续订包周期，可以执行如下步骤：

1. 在云搜索服务的集群管理列表页，找到需要续订的计费模式为“包年/包月”的集群。

2. 单击操作列的“更多 > 续费”，确定后进入续费页面。
3. 选择续费时长，并支付费用，完成续费。

新建集群时如何自动续费

如果是新建集群时，想要实现自动续费包周期，可以执行如下步骤：

在创建集群页面，设置“订购周期”后，勾选“自动续费”。待集群到期后将会自动续订下一个周期。

图 2-1 设置自动续费



更多详细的包周期续费管理请参见[续费管理](#)。

2.3 购买了折扣套餐包，如何使用？

说明

折扣套餐包功能已于2022年11月下线，已购买的套餐包不影响使用。

折扣套餐包购买成功后系统不会自动创建集群，您需要前往云搜索服务管理控制台创建与折扣套餐包中区域、节点规格、节点存储类型或带宽类型相同的集群才能使用该折扣套餐。创建集群的具体操作步骤请参见[创建Elasticsearch类型集群（安全模式）](#)或[创建Elasticsearch类型集群（非安全模式）](#)。

2.4 折扣套餐包是如何扣费的？

说明

折扣套餐包功能已于2022年11月下线，已购买的套餐包不影响使用。

折扣套餐包的计费方式是用户预先购买一定的云搜索服务节点实例、存储类型或者带宽类型、使用小时数配额。在使用过程中，系统优先扣减折扣套餐所购买的配额，配额内使用量不再收费，超出配额的使用量以按需计费方式收费。

所购买的集群的区域、节点规格、存储类型或者带宽类型，必须和套餐包购买的匹配，否则无法使用套餐包。

2.5 折扣套餐包到期了怎么续费？

说明

折扣套餐包功能已于2022年11月下线，已购买的套餐包不影响使用。

- 如果之前购买折扣套餐包时未选择自动续费，折扣套餐包到期后，将无法续费，集群会转换成按需计费。
- 如果之前购买折扣套餐包时选择了自动续费，折扣套餐包到期后，会自动续费套餐包。

云搜索服务支持按需计费和包年包月2种计费模式，您可以根据业务需求，将按需计费转换为包年包月计费模式。更多信息，请参见[按需集群转包年包月计费模式](#)。

2.6 套餐包到期时间不同，如何将按需计费转为包年包月计费模式？

套餐包到期时间相同

套餐包到期之后，您可以直接将按需计费转换为包年包月计费模式，转换后立即生效，具体操作请参见[按需集群转包年包月计费模式](#)。

套餐包到期时间不相同

到期时间不同的套餐包，建议您在套餐包到期后及时续费，续费时间可根据其他套餐包到期时间来决定，原则上让所有资源尽量同时到期，可以按照以下方式进行续费。

1. 在控制台上方菜单栏，选择“费用与成本 > 我的套餐”。
2. 在“资源包”页面，选择需要续费的套餐资源包，单击操作列的“续费”，进入续费页面。
3. 选择续费时长，并支付费用，完成续费。

套餐包全部到期之后，可以直接将按需计费转换为包年包月计费模式。

3 集群管理

3.1 区域和可用区

3.1.1 什么是区域和可用区

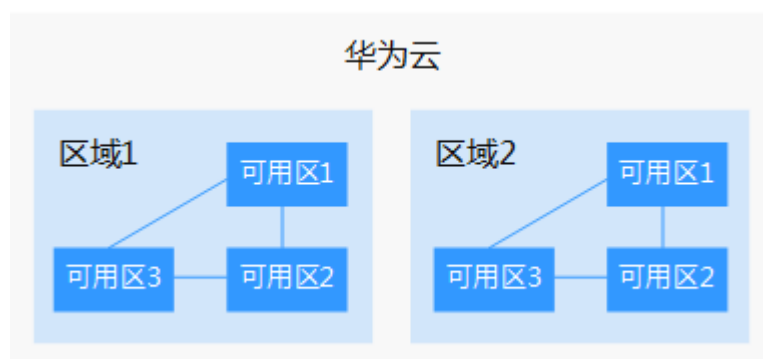
什么是区域、可用区？

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3-1阐明了区域和可用区之间的关系。

图 3-1 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。更多信息请参见[华为云全球站点](#)。

如何选择区域？

选择区域时，您需要考虑以下几个因素：

- 地理位置

一般情况下，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。

- 在除中国大陆以外的亚太地区有业务的用户，可以选择“中国-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
- 在非洲地区有业务的用户，可以选择“非洲-约翰内斯堡”区域。
- 在欧洲地区有业务的用户，可以选择“欧洲-巴黎”区域。
- 在拉丁美洲地区有业务的用户，可以选择“拉美-圣地亚哥”区域。

说明

“拉美-圣地亚哥”区域位于智利。

- 资源的价格

不同区域的资源价格可能有差异，请参见[华为云服务价格详情](#)。

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[地区和终端节点](#)。

3.1.2 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 3-2 集群配置信息

配置信息	
区域	
可用区	
虚拟私有云	vpc
子网	subnet-
安全组	dws 更改安全组
安全模式	启用
重置密码	重置
企业项目	
公网访问	-- 绑定
HTTPS访问	开启 下载证书
内网访问IPv4地址	192.

3.2 集群版本

3.2.1 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。

3.3 安全模式集群

3.3.1 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”。

图 3-3 下载证书

配置信息	
区域	
可用区	
虚拟私有云	vpc-
子网	subnet
安全组	dws 更改安全组
安全模式	启用
重置密码	重置
企业项目	default
公网访问	-- 绑定
HTTPS访问	开启 下载证书
内网访问IPv4地址	192

3.3.2 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 der -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。

3.3.3 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[使用备份与恢复迁移集群数据](#)至新集群后，进行安全组修改。

- 登录云搜索服务控制台。
- 选择“集群管理”进入集群列表。
- 单击对应集群的名称，进入集群基本信息页面。
- 在“安全组”右侧，单击“更改安全组”。

图 3-4 更改安全组



- 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。

3.4 参数配置

3.4.1 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10,000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
  }
}
```

3.4.2 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。

3.4.3 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。

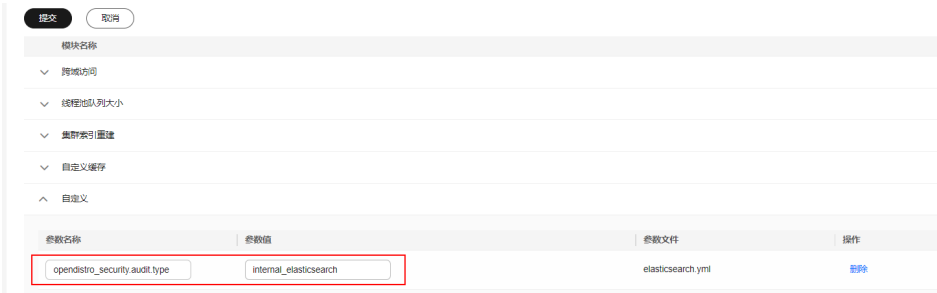
📖 说明

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 3-5 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。

4 开源搜索引擎咨询

4.1 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}。

4.2 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balance.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：

- 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。
- 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。

4.3 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。

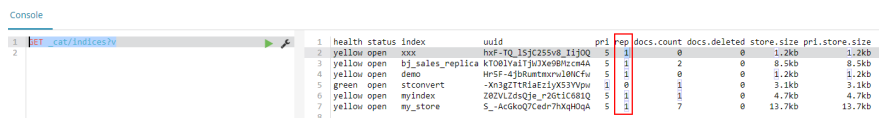
```
PUT _template/urlinfo_template?include_type_name=true
```

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."
```

4.4 如何配置 CSS 集群双副本?

- 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。



	health	status	index	uuid	pri	rep	docs.count	docs.deleted	store.size	pri.store.size
1	yellow	open	xxx	hxf-TQ_15jC255v0_I1j0Q	5	1	0	0	1.2kb	1.2kb
2	yellow	open	bj_sales_replica	k70b1val17juXa90tcm4A	5	1	2	0	8.5kb	8.5kb
3	yellow	open	demo	he5r-dj0kumtenu0l0McFu	5	1	0	0	1.2kb	1.2kb
4	green	open	stconvert	-Xn3gZTrRlaEziyX53VpW	1	0	1	0	3.1kb	3.1kb
5	yellow	open	myindex	Z8ZVLzdsQje_r26t1C68lQ	5	1	1	0	4.7kb	4.7kb
6	yellow	open	my_store	S_-4ccknQ7Cedr7hAq6Dqk	5	1	7	0	13.7kb	13.7kb

- 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。

4.5 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

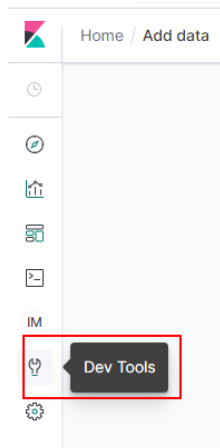
index为需要修改的索引名称，需根据实际情况进行修改。

4.6 Elasticsearch 集群分片过多会有哪些影响

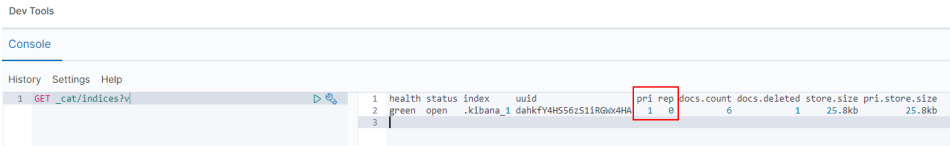
1. 集群创建分片的速度随着集群分片数量增多而逐渐减低。
2. 触发Elasticsearch自动创建index时，创建速度变慢会导致大量写入请求堆积在内存中，严重时可导致集群崩溃。
3. 分片过多时，如果不能及时掌控业务的变化，可能经常遇到单分片记录超限、写入拒绝等问题。

4.7 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。



4.8 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 4-1 返回数据



表 4-1 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。

参数	描述
“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。

4.9 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**流量控制功能**，可以使用**一键断流**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
}
```

4.10 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。

4.11 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。

4.12 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 4-2 写入一条数据

```
{
  "_index": ".opendistro-ism-config",
  "_type": "_doc",
  "id": "policy1",
  "_score": 1.0,
  "source": {
    "policy": {
      "policy_id": "policy1",
      "description": "A simple default policy that changes the replica count between hot and cold states.",
      "last_updated_time": 1641432150329,
      "schema_version": 1,
      "error_notification": null,
      "default_state": "hot",
      "states": [
        {
          "name": "hot",
          "actions": [ ],
          "transitions": [
            {
              "state_name": "delete",
              "conditions": {
                "min_index_age": "2d"
              }
            }
          ]
        },
        {
          "name": "delete",
          "actions": [
            {
              "delete": { }
            }
          ],
          "transitions": [ ]
        }
      ]
    }
  }
}
```

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 4-3 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 4-4 初始化后数据

```

1  "index": ".opendistro-ism-config",
2  "type": "doc",
3  "id": "FABKSF5GSTCmRQKw41HWw",
4  "score": 1.0,
5  "source": {
6    "managed_index": {
7      "name": "data1",
8      "enabled": true,
9      "index": "data1",
10     "index_uuid": "FABKSF5GSTCmRQKw41HWw",
11     "schedule": {
12       "interval": {
13         "start_time": 1641432652693,
14         "period": 1,
15         "unit": "Minutes"
16       }
17     },
18     "last_updated_time": 1641432652694,
19     "enabled_time": 1641432652694,
20     "policy_id": "policy1",
21     "policy_seq_no": 3,
22     "policy_acyclic_term": 1,
23     "policy": {
24       "policy_id": "policy1",
25       "description": "A simple default policy that changes the replica count between hot and cold states.",
26       "last_updated_time": 1641432150329,
27       "schema_version": 1,
28       "error_notification": null,
29       "default_state": "hot",
30       "states": [
31         {
32           "name": "hot",
33           "actions": [ ],
34           "transitions": [
35             {
36               "state_name": "delete",
37               "conditions": {
38                 "min_index_age": "2d"
39               }
40             }
41           ]
42         },
43         {
44           "name": "delete",
45           "actions": [
46             {
47               "delete": { }
48             }
49           ],
50           "transitions": [ ]
51         }
52       ]
53     },
54     "change_policy": null
55   }
56 }

```

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[管理索引生命周期](#)。

#### 4.13 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"
}
```

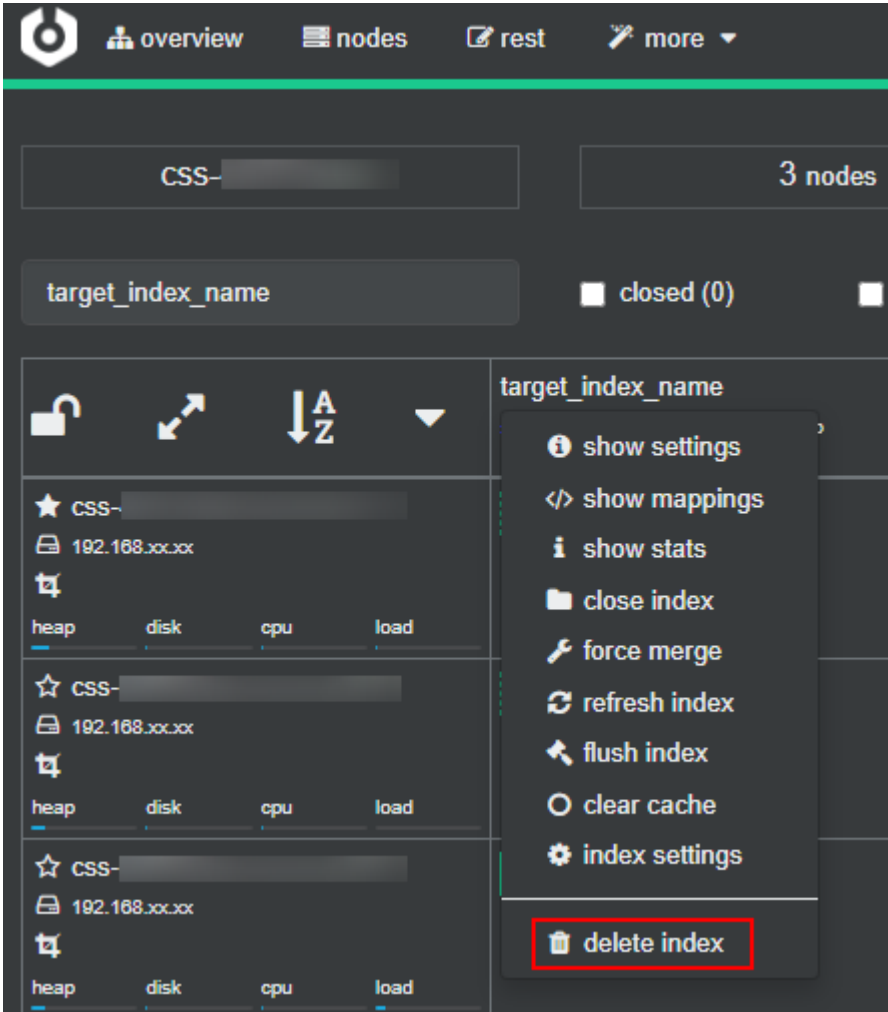
- 查询慢于10秒输出一个WARN日志。
- 获取慢于500毫秒输出一个DEBUG日志。
- 索引慢于5秒输出一个INFO日志。

详细可参考官网：<https://www.elastic.co/guide/cn/elasticsearch/guide/current/logging.html>

## 4.14 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 4-5 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。

## 4.15 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" :
 "DXF1ZXJ5QW5kRmV0Y2gBAAAAAAAAAAD4WYm9laVYtZndUQlNsdDcwakFMNjU1QQ=="
}
```

- 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
```

## 4.16 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。

# 5 集群插件使用

## 5.1 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。安全模式详情参考[安全模式简介](#)。

## 5.2 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。更多信息请参见[向量检索](#)。

#### 说明

向量检索功能支持7.6.2和7.10.2版本的集群。

# 6 集群访问/集群连接

## 6.1 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[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

## 6.2 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。

## 6.3 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6-1所示。

图 6-1 获取信息

配置信息

区域

可用区

虚拟私有云

vpc

子网

subnet

安全组

dw

更改安全组

安全模式

启用

重置密码

重置

企业项目

default

HTTPS访问

关闭

内网访问IPv4地址

192.

----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

**步骤1** 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。

----结束

## 6.4 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[基本信息](#)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[基本信息](#)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。

## 6.5 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。

## 6.6 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。

# 7 集群迁移

## 7.1 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。具体操作请参见[使用备份与恢复迁移集群数据](#)。

### 方法二：

1. 打通VPC网络，建立对等链接。具体操作请参见[对等连接简介](#)。
2. 打通网络后，使用Logstash进行数据迁移。具体操作请参见[使用Logstash迁移集群数据](#)。

## 7.2 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[备份与恢复索引](#)进行集群迁移。
- 如果OBS桶跨区域，请先参考[配置跨区域复制](#)进行跨区域复制OBS桶，再参考[备份与恢复索引](#)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。

# 8 集群备份与恢复


## 8.1 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。

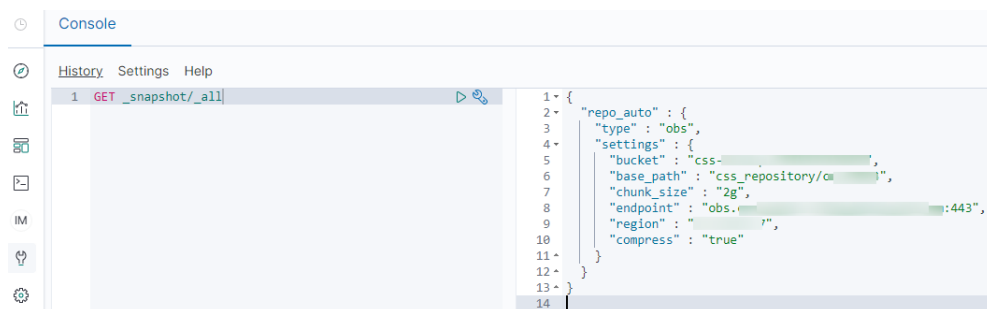
### 快照查询

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 8-1 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 8-2 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。

## 8.2 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。

图 8-3 选择恢复快照



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。

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 8-4 恢复成功

创建快照

快照名称 ▼ | 请输入快照名称 🔍 🔍

| 名称ID <span>🔍</span>                        | 快照状态 <span>🔍</span> | 任务状态 <span>🔍</span> | 快照类型 <span>🔍</span> | 快照创建时间 <span>🔍</span>         | 操作                                    |
|--------------------------------------------|---------------------|---------------------|---------------------|-------------------------------|---------------------------------------|
| <span>▼</span> snapshot-bee8 <span></span> | <span>●</span> 可用   | <span>●</span> 恢复成功 | Manual              | 2024/03/15 17:04:49 GMT+08:00 | <a href="#">恢复</a> <a href="#">删除</a> |

# 9 集群监控与运维

## 9.1 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。告警配置方法，请参见[配置告警规则](#)。

### 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。

## 9.2 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。

9.3 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 9-1 集群基本信息

| 基本信息         |                               |
|--------------|-------------------------------|
| 集群名称         | CSS- <div></div>              |
| ID           | <div></div>                   |
| 集群版本         | 7.10.2                        |
| 集群状态         | <div></div> 可用                |
| 任务状态         | --                            |
| 创建时间         | 2024/03/23 11:33:19 GMT+08:00 |
| 集群存储容量 (GB)  | 200                           |
| 集群存储使用量 (GB) | 10                            |

9.4 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[扩容](#)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 Kibana 使用

## 10.1 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。

图 10-1 重置密码

## 配置信息

|       |                            |
|-------|----------------------------|
| 区域    |                            |
| 可用区   |                            |
| 虚拟私有云 | vpc-                       |
| 子网    | subnet-                    |
| 安全组   | dws- <a href="#">更改安全组</a> |
| 安全模式  | 启用                         |
| 重置密码  | <a href="#">重置</a>         |

## 10.2 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

## 10.3 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。

Elasticsearch SQL语句请参见[查询Elasticsearch SQL](#)

图 10-2 SQL Workbench

