**ES数据迁移文档**

V1.0.0

吴志彬 @ Enlink 数据中心

**目录**

[1. 概述 3](#_Toc28424)

[1.1. 文档内容 3](#_Toc24222)

[1.2. 数据备份、恢复与索引的迁移 3](#_Toc21759)

[2. 数据备份 4](#_Toc30946)

[2.1. 文件挂载 4](#_Toc13522)

[3. 在ES一个节点创建备份与恢复 5](#_Toc5157)

[3.1. 创建仓库 5](#_Toc24308)

[3.2. 创建快照 5](#_Toc7585)

[3.3. 查看备份信息 7](#_Toc14334)

[状态码说明 7](#_Toc2728)

[3.4. 恢复 7](#_Toc32403)

[4. 使用reindex备份索引 8](#_Toc11145)

[5. 使用elastucsearch-dunp工具进行备份 9](#_Toc25907)

[拷贝映射 9](#_Toc9860)

[拷贝数据 9](#_Toc18390)

# 概述

## 文档内容

本文档介绍ES的数据备份、恢复与索引的迁移

## 数据备份、恢复与索引的迁移

在数据量越来越大的时候，硬盘或SSD总有满的一天或者ES需要升级而保留之前的数据情况下或者不小心删除了某个索引而造成的数据丢失，我们需要对数据做保留与恢复

数据备份：将数据已其他格式保存在其他位置，而不会因为ES 节点的升级或误操作而丢失

数据恢复：在其他节点中使用已保存的备份进行恢复的操作

索引的迁移就是数据备份与恢复的一整个操作，这里我们使用的是一个工具。方便直接的将一个节点上的索引移动到另一个节点

# 数据备份

## 文件挂载

每个节点都要安装nfs 服务

|  |
| --- |
| yum install -y nfs-utils |

启动服务

|  |
| --- |
| service nfs start |
| service rpcbind start |

将做备份的节点做“文件存放服务器”

将同步备份的节点做“挂载服务器”

在对应的路径下创建文件夹（本文档是将 /usr/local/mount/backups 文件夹做共享文件夹）

修改“文件存放服务器”中的 /etc/exports 文件，在其中添加命令 /usr/local/mount/backups \*(rw,)

在“挂载服务器”上使用命令进行挂载

|  |
| --- |
| mount -t nfs <”文件存放服务器的IP”>:/usr/local/mount/backups /usr/local/mount/backups |

# 在ES一个节点创建备份与恢复

## 创建仓库

新建文件夹做仓库使用，这是使用的共享文件夹

修改共享文件夹的权限

|  |
| --- |
| chown -R es:es /usr/local/mount/backups |

在ES配置文件(elasticsearch.yml)中添加仓库路径

|  |
| --- |
| path.repo: ["/usr/local/mount/backups"] |

使用API创建仓库

*表 3-1 创建仓库*

|  |
| --- |
| PUT /\_snapshot/<仓库名称>  {  "type": "fs",  "settings": {  "location": "/usr/local/mount/backups"  }  } |

## 创建快照

|  |
| --- |
| PUT /\_snapshot/<仓库名称>/<快照名称> |
| PUT /\_snapshot/backups/snapshot\_1 |

返回的结果

*表 3-2 创建快照返回的结果*

|  |
| --- |
| {  "snapshot": {  "snapshot": "snapshot\_1", <快照名称>  "uuid": "F\_CXAnKOSkSJiOgoSUUHgw",  "version\_id": 5000099,  "version": "5.0.0",  "indices": [ <备份的索引>  ".monitoring-data-2",  ".kibana",  "car2",  ".monitoring-es-2-2019.01.10"  ],  "state": "SUCCESS", <[状态码](#_状态码说明)>  "start\_time": "2019-01-10T03:13:17.162Z",  "start\_time\_in\_millis": 1547089997162,  "end\_time": "2019-01-10T03:13:17.278Z",  "end\_time\_in\_millis": 1547089997278,  "duration\_in\_millis": 116,  "failures": [],  "shards": {  "total": 8,  "failed": 0,  "successful": 8  }  }  } |

指定索引名称进行备份(推荐以这种方式进行备份，ES或者Kibana会自动创建索引，以至于恢复的时候造成冲突)

*表 3-3 指定索引名称备份*

|  |
| --- |
| PUT /\_snapshot/backups/snapshot\_5?wait\_for\_completion=true  {  "indices": "car1,car2,car-\*", ①  "ignore\_unavailable": false, ②  "include\_global\_state": false ③  } |

①：指定的索引，也可以用\*表示所有索引

②ignore\_unavaliable：true。忽略不存在的索引。如果该项未设置并且缺少索引，则快照失败

③include\_global\_state: false。可以防止将集群全局状态存储为快照的一部分。默认情况下，如果参与快照的一个或多个索引没有可用的所有主分片，则快照失败

## 查看备份信息

显示的即进行备份的返回结果

查看所有快照

|  |
| --- |
| GET \_snapshot/backups/\_all |

查看正在进行的快照

|  |
| --- |
| GET \_snapshot/backups/\_current |

### 状态码说明

1.IN\_PROGRESS：快照正在进行

2.SUCCESS：快照已完成，所有分片都已成功存储

3.FAILED：快照以错误结束，无法存储任何数据

4. PARTIAL：存储了全局集群状态，但至少有一个分片的数据未成功存储。

5.INCOMPATIBLE：快照使用旧版本创建，因此与集群版本不兼容

## 恢复

|  |
| --- |
| POST \_snapshot/<仓库名>/<快照名>/\_restore |

在恢复期间更改索引设置

|  |
| --- |
| POST \_snapshot/<仓库名>/<快照名>/\_restore  {  "indices":"index\_1",  "index\_setting":{  "index.number\_of\_replicas":0  },  "ignore\_index\_setting":[  "index.refresh\_interval"  ]  } |

# 使用reindex备份索引

在需要进行备份的节点的elasticsearch.yml中添加白名单

|  |
| --- |
| reindex.remote.whitelist: ["192.168.226.134:9200" ,"IP:PORT"] |

进行备份

|  |
| --- |
| POST \_reindex  {  "source": {  "remote": {  "host": "http://192.168.226.134:9200"  },  "index": "car-1"  },  "dest": {  "index": "car-1"  }  } |

# 使用elastucsearch-dunp工具进行备份

安装elasticsearch-dump

|  |
| --- |
| yum install -y epel-release |
| yum install -y nodejs |
| yum install -y npm |
| npm install elasticdump |
| cd node\_modeles/elasticdump/bin 之后便可执行操作 |

#### 拷贝映射

|  |
| --- |
| ./elasticdump --input=http://192.168.226.133:9200/rally-metrics-2019-01 --output=http://192.168.226.134:9200/rally-metrics-2019-01 --type=mapping |

|  |
| --- |
| ./elasticdump --input=http://<IP>:<PORT>/<index> --output=http://<IP>:<PORT>/<index> --type=mapping |

#### 拷贝数据

|  |
| --- |
| ./elasticdump --input=http://192.168.226.133:9200/rally-metrics-2019-01 --output=http://192.168.226.134:9200/rally-metrics-2019-01 --type=data |

|  |
| --- |
| ./elasticdump --input=http://<IP>:<PORT>/<index> --output=http://<IP>:<PORT>/<index> --type=data |