

# elasticsearch集群高可用与防止脑裂探讨

elasticsearch(简称ES)是一个分布式搜索引擎, 也是一个文档型的数据库。

多台ES可以组建集群来实现集群的高可用, 但建议使用3台或3台以上ES节点来实现集群, 主要是因为集群高可用与脑裂的原因。下面我们来探讨一下:

## 三台ES集群的搭建方法



1, 三台都确认jdk已经安装, 这里使用centos上默认的openjdk

```
# java -version
openjdk version "1.8.0_161"
OpenJDK Runtime Environment (build 1.8.0_161-b14)
OpenJDK 64-Bit Server VM (build 25.161-b14, mixed mode)
```

2, 三台都下载并安装ES

```
# wget
https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-6.5.2.rpm
# rpm -ivh elasticsearch-6.5.2.rpm
```

3, 三台分别配置ES

```
# grep -v '#' /etc/elasticsearch/elasticsearch.yml
cluster.name: elk-cluster
node.name: 10.1.1.x
path.data: /var/lib/elasticsearch
path.logs: /var/log/elasticsearch
network.host: 0.0.0.0
http.port: 9200
discovery.zen.ping.unicast.hosts: ["10.1.1.11", "10.1.1.12", "10.1.1.13"]
discovery.zen.minimum_master_nodes: 2
```

三台ES配置就这一个参数不一样，使用不同的IP来指定不同节点

#### 4, 三台都启动服务并验证端口

```
# systemctl start elasticsearch
# systemctl enable elasticsearch

# netstat -ntlp | grep java
tcp6      0  0  :::9200      :::*          LISTEN
5329/java
tcp6      0  0  :::9300      :::*          LISTEN
5329/java
```

#### 5, 验证集群OK

```
# curl http://10.1.1.x:9200/_cluster/health?pretty
10.1.1.x可替换成任意ES其中1台
{
  "cluster_name" : "elk-cluster",
  "status" : "green",
  "timed_out" : false,
  "number_of_nodes" : 3,
  "number_of_data_nodes" : 3,
  "active_primary_shards" : 0,
  "active_shards" : 0,
  "relocating_shards" : 0,
  "initializing_shards" : 0,
  "unassigned_shards" : 0,
  "delayed_unassigned_shards" : 0,
  "number_of_pending_tasks" : 0,
  "number_of_in_flight_fetch" : 0,
  "task_max_waiting_in_queue_millis" : 0,
  "active_shards_percent_as_number" : 100.0
}
```

集群正常运行

集群中有3个节点

```
}
```

# 高可用与防止脑裂讨论

配置文件中有一个重点参数: `discovery.zen.minimum_master_nodes`, 你可以理解为最小允许成为1个集群的节点数(默认情况下所有节点都可侯选为master, 所以在此不讨论master, 只讨论节点即可)。

此参数在3节点集群中 `discovery.zen.minimum_master_nodes` 值的说明:

值	说明
1	如果1个节点关机,则剩下2个节点继续正常运行集群
1	如果2个节点关机,则剩下1个节点继续正常运行集群
1	如果1个节点网络失联(服务还在), 则它单独为1个集群。剩下2个节点也为1个集群( <b>脑裂产生</b> )
2	如果1个节点关机,则剩下2个节点继续正常运行集群( <b>防止单点故障,高可用</b> )
2	如果2个节点关机, 则剩下1个节点无法运行集群
2	如果1个节点网络失联(服务还在), 它无法为集群。剩下2个节点正常运行集群。( <b>防止了脑裂</b> )
3	如果1个节点关机, 集群全完蛋

**综上所述:** N个节点集群, 建议配置使用 $N/2+1$ ( $N/2$ 取整, 不四舍五入). 如5节点集群, 值建议配置为3.

## 验证方法

按照上表所述一个一个验证:

- 模拟1个节点关机,就在那台机器上使用 `init 0` 模拟, 然后使用 `curl http://10.1.1.x:9200/_cluster/health?pretty` 验证集群状态与集群节点数

说明: 10.1.1.x代表剩下节点的IP

- 模拟节点失联, 就在那台机器上使用 `ifconfig 网卡名 down` 模拟, 但因为有一个未配置的默认参数 `discovery.zen.ping_timeout` 的值为30的原因, 会造成使用curl测试会要等待半分钟以上时间

## 结论

---

ES集群经过合理的配置可以实现高可用与防止脑裂。主要配置参数为:

- `discovery.zen.minimum_master_nodes` 理解为最小允许成为1个集群的节点数
- `discovery.zen.ping_timeout` 理解为当1个节点与集群失联的超时时间