

# Elasticsearch

5장 서비스 환경에 클러스터 구성

## 5.1 운영 환경을 위한 설정과 클러스터 구성

### 마스터 후보

- 노드의 역할에 master를 지정하면 해당 노드는 마스터 후보 노드가 된다. 마스터 후보 노드중에서 선거를 통해 마스터 노드가 선출된다.
- 마스터 노드는 클러스터를 관리하는 역할을 하고 인덱스 생성이나, 어떤 샤드를 어떤 노드에 할당할 것인지 등 중요한 작업을 수행한다.

### 데이터노드

- 실제 데이터를 들고 있는 노드다. CRUD, 검색, 집계와 같이 데이터와 관련된 작업을 수행한다.

### 인제스트노드

- 데이터가 색인되기 전에 전처리를 수행하는 인제스트 파이프라인을 수행하는 노드

### 조정 노드

- 클라이언트의 요청을 받아서 다른 노드에 요청을 분배하고 클라이언트에게 최종 응답을 돌려주는 노드.
- 마스터나 데이터 등 주요 역할을 수행하지 않고 조정 역할만 수행하는 노드는 조정 전용 노드라고 부른다.

### 원격 클러스터 클라이언트 노드

- 다른 엘라스틱서치 클러스터에 클라이언트로 붙을 수 있는 노드다. 노드 역할에 remote\_cluster\_client를 추가해 지정한다.
- 키바나 스택 모니터링 기능을 활용해서 모니터링 전용 클러스터를 구축한 뒤 메시지를 보내도록 하거나, 유료 기능인 클러스터 간 검색 기능등 활용한다.

### 데이터 티어 노드

- 데이터 노드를 용도 및 성능별로 hot-warm-cold-frozen 티어로 구분해 저장하는 데이터 티어 구조 채택 시 사용하는 역할이다.
- Data 역할 대신에 data\_content, data\_hot, data\_warm, data\_cold, data\_frozen 역할을 선택한다.

## 5.1 운영 환경을 위한 설정과 클러스터 구성

### config/elasticsearch.yml 설정

#### node.roles

- 노드 역할을 지정하는 설정이다. Master, data, ingest 등을 0개 이상 조합해 지정한다. [] 와 같이 비워두면 조정 전용 노드이다.
- 서비스 환경에서는 마스터 후보 역할과 데이터 역할을 분리해야 좋다.

#### discovery.send\_hosts

- 마스터 노드로 동작할 수 있는 노드 목록을 지정한다.

#### cluster.initial\_master\_nodes

- 클러스터를 처음 기동할 때 첫 마스터 선거를 수행할 후보 노드 목록을 지정한다. 이 값은 node.name에 기입한 값이어야 한다.

#### network.bind\_host 와 network.publish\_host

- network.bind\_host는 엘라스틱에 바인딩할 네트워크 주소
- network.publish\_host는 클러스터의 다른 노드에게 자신을 알릴 때 쓰는 주소
- 두 값은 network\_host에 지정한 값을 따른다.

#### http.port

- http 통신을 위해 사용하는 포트를 지정한다 (기본값은 9200-9300)

#### transport.port

- Transport 통신을 위해 사용하는 포트를 지정한다 (기본값은 9300-9400)

#### path.data

- 데이터 디렉토리로 여러 경로를 지정할 수 있다. 다만 7.13버전부터 지원 중단 선언하였다.

# 5.1 운영 환경을 위한 설정과 클러스터 구성

## 힙 설정

### 첫번째 대원칙

- 최소한 시스템 메모리의 절반 이하로 지정해야 한다. 시스템 메모리의 절반은 운영체제가 캐시로 쓰도록 나누는 것이 좋다.
- 루씬이 커널 시스템 캐시 많이 활용하기 때문이다.

### 두번째 대원칙

- 힙 크기를 32GB 이상 지정하지 않아야 한다는 것이다.
- JVM이 Compressed OOPS 라는 기능을 적용해 포인터를 32비트로 유지 할 수 있게 기능이 적용되는 힙 크기가 32GB 보다 살짝 아래 쪽이기 때문이다. Compressed OOPS를 사용하지 못하면 성능에 영향이 크다.

예를 들어 128GB 이상의 서버를 사용하는 경우 첫번째 원칙을 따라서 64GB는 OS가 사용해야 하고 두번째 원칙을 따라 힙을 약 32GB(32GB 보다 조금 낮게)

설정한다면 약 32GB의 메모리가 남는데 엘라스틱서치 프로세스를 2개 띄우거나, 나머지 32GB는 루씬이 사용하도록 뒤도 큰 낭비가 아니라고 한다.

keyword, long, date 등 text가 아닌 타입에 대한 정렬, 집계를 많이 해야 한다면 doc\_values를 사용하기때문에 **후자**가 유리하고  
Text 타입에 대한 정렬, 집계를 많이한다면 fieldData를 이용해야 하므로 **전자**가 유리하다.

## 스와핑

엘라스틱 서치는 스와핑을 사용하지 않도록 강력히 권고한다. 스와핑은 성능과 노드 안정성에 큰 영향을 미친다. 밀리세컨드 단위로 끝나는 할 GC를 분 단위 시간까지 걸리게 만들기도 한다.

## 스와핑 끄기

1. `sudo swapoff -a` -> 운영체제 재부팅 되면 사라짐
2. `sudo vim /etc/fstab`을 수정하여 swap부분 제거

## 5.1 운영 환경을 위한 설정과 클러스터 구성

### vm\_max\_map\_count

- 프로세스가 최대 몇 개 까지 메모리 맵 영역을 가질수 있는지 값 설정
- 루센은 mmap을 적극적으로 활용하기 때문에 값을 높일 필요가 있다. **262144**보다 작은 값이 지정되면 엘라스틱서치는 기동자체를 거부한다.
- Sysctl vm\_max\_map\_count 으로 확인하여 작다면 sudo vim /etc/sysctl.d/98-elasticsearch.conf에서 값 수정 (기본값이 **262144**)
- 운영체제의 일반적인 가이드는 시스템 메모리 / 128kb로 나눈 크기가 추천된다. 예를들어 시스템 메모리가 64gb 라면 524288로 계산
- Malloc이 mmap을 호출하도록 하는 기준인 m\_mmap\_threshol의 기본값이 128KB
- 실제로 대부분은 기본값인 **262144**를 사용한다고 한다.

### 파일 기술자

- 엘라스틱서치는 많은 파일 기술자를 필요로 한다. 이값을 최소 **65535** 이상으로 지정하도록 가이드한다.
- Sudo vim /ect/security/limits.conf에서 수정한다.

## 5.2 클러스터 구성 전략

### 마스터 후보 노드와 데이터 노드를 분리하는 이유

- 운영에서 발생할 수 있는 여러가지 문제를 보면 상대적으로 데이터 노드가 죽을 확률이 높다. 분리하지 않으면 마스터 역할이 제대로 수행되지 않아서 클러스터 안정성이 매우 크게 떨어지게 된다. 일부 데이터 노드에만 문제가 생기는 것으로 끝나는 상황이 클러스터 전체 장애로 커진다.

### 마스터 후보 노드와 투표 구성원(vote configuration)

- 마스터 노드는 클러스터를 관리하는 중요한 역할을 수행한다. 클러스터가 운여오디는 동안 항상 마스터 노드가 선출되어 있어야 하며 이런 마스터 노드를 선출하는 집단이 바로 투표 구성원이다.
- 마스터 후보 노드는 홀수대를 준비하는 것이 비용 대비 효용성이 좋다. 엘라스틱서치가 투표 구성원을 홀수로 유지하기 위해 투표 구성원에서 마스터 후보 노드를 하나 빼 두기 때문이다.

### 서버 자원이 많지 않은 경우

적어도 노드 3개가 확보되어야한다. 앞에서 성능이 낮은 장비에는 마스터노드, 높은 장비는 데이터 노드를 주고 최대한 분리하는 것이 좋다고 했지만 성능이 안좋으면 3대의 노드가 모두 마스터 후보 역할과 데이터 역할을 겸임하는 구성이 된다. 인덱스 설정에서 샤드의 복제본을 설정했다면 1대의 서버에 문제가 생기는 상황까지는 서비스 중단 없이 대응할 수 있다.