

# Apache Kafka 구성 및 관리

## 2 Kafka in Real Environment

## Broker 설치 구성 전용 서버 권장

2.

Kafka in Real  
Environment

Broker는 분리된 각각의 전용 서버에 분리하여 설치 구성하는 것을 권장

- N개의 Broker가 있으면, Replication Factor(RF)는 최대 N까지 사용하여 Topic 생성 가능
- **Mission Critical Topic**에는 **Replication Factor**는 보통 **3** 을 많이 사용
  - ➔ 3 개의 Broker로 구성하고 하나의 Broker가 장애 상태시, RF 3인 Topic 생성 불가능
  - ➔ 따라서, **Broker는 4 개 이상 하나의 Cluster로 구성하는 것을 권장**
- **데이터 유실 방지**를 위해서, **min.insync.replicas**는 **2** 를 많이 사용
  - ➔ 3 개의 Broker로 구성하고 1 개의 Broker가 장애 상태시, Topic에 Write는 가능
  - ➔ 1 개의 Broker가 추가로 장애시, 데이터 유실의 가능성이 높아짐
  - ➔ 따라서, **Broker는 4 개 이상 하나의 Cluster로 구성하는 것을 권장**

## Broker CPU

처리량에 따라서 Thread 관련 파라미터 튜닝 필수

Broker는 CPU를 많이 사용하지는 않으나, 처리량에 따라서 Thread 파라미터 튜닝이 필요하며, Thread 증가에 따라 CPU 사용량이 증가함

- 설정을 잘 못 하거나, 메모리가 충분하지 않거나, Bug로 인해 CPU 사용률이 높을 수도 있음
- Reference Architecture에는 Dual 12-core Sockets를 권장
- Broker 파라미터
  - `num.io.threads` (기본값 : 8) : Disk 개수보다 크게 설정
  - `num.network.threads` (기본값 : 3) : TLS를 사용할 경우 두 배 이상으로 설정
  - `num.recovery.threads.per.data.dir` (기본값 : 1) : Broker 시작시 빠른 기동을 위해서, core수까지만 설정
  - `num.replica.fetchers` (기본값 : 1) : Source Broker에서 메시지를 복제하는데 사용되는 Thread 수. 빠르게 복제하기 위해 값을 증가(Latency를 만족하는). Broker의 CPU 사용률과 네트워크 사용률이 올라감
  - `num.cleaner.threads` (기본값 : 1) : Disk 개수 혹은 core 개수까지만 설정

## Open File Descriptors Broker는 File을 많이 사용

2.

Kafka in Real  
Environment

Broker는 많은 수의 Partition을 지원하므로 상대적으로 소규모 배포에서도 Open File Handle 수가 기본값을 쉽게 초과할 수 있음

- 최소  
`ulimit -n 100000`

## Broker Memory

### Broker의 JVM Heap + OS Page Cache

#### Broker는 JVM Heap을 많이 사용하지 않음

- Broker의 Heap 메모리는 운영환경의 경우 대부분 6 GB 까지 할당
- 매우 큰 Cluster 혹은 매우 많은 Partition이 필요한 경우 12 GB 이상 사용
- Broker의 OS만을 위해서는 보통 1 GB 정도 할당
- Broker 는 OS Page Cache를 많이 사용
  - OS Page Cache를 통해서 Zero Copy<sup>1)</sup> 전송을 수행
  - 많을수록 성능에 유리
- 운영환경용 Broker 메모리는 최소 32 GB 이상 권장하며, 처리량에 따라서 64 GB 이상 사용 권장

1) Chapter 02. Apache Kafka 심화 개념 및 이해 - 01. Producer Acks, Batch, Page Cache, Flush 참고

## Broker Java Heap Memory Heap 메모리 설정 방법

2.

Kafka in Real  
Environment

**kafka-server-start** 스크립트에 Java Heap 설정하는 옵션이 있음

- **\$KAFKA\_HEAP\_OPTS**  
-Xms6g -Xmx6g
- **\$KAFKA\_JVM\_PERFORMANCE\_OPTS**  
-server -XX:MetaspaceSize=96m -XX:MinMetaspaceFreeRatio=50  
-XX:MaxMetaspaceFreeRatio=80 **-XX:+UseG1GC** -XX:MaxGCPauseMillis=20  
-XX:InitiatingHeapOccupancyPercent=35 -XX:G1HeapRegionSize=16M

## Broker Network

전체 처리 메시지(복제 포함)가 네트워크를 통해 전달

- 처리량이 작은 Application일 경우, Broker의 NW은 1 Gigabit(Gb) 으로 충분
- 처리량이 큰 Application일 경우, Broker의 NW은 10 Gigabit(Gb) 이상 필요
- Producer에서 압축 옵션을 사용하면 네트워크를 보다 효율적으로 사용 가능
- Internal 과 External 트래픽 간 분리 가능

KIP-103: Separation of Internal and External traffic

<https://cwiki.apache.org/confluence/display/KAFKA/KIP-103%3A+Separation+of+Internal+and+External+traffic>

## Broker Disk

### Kafka 성능에 큰 영향

## 2.

Kafka in Real Environment

### Kafka Broker의 data log 용 Disk는 OS 용 Disk와 분리 권장

- Broker의 data log용으로 여러 개의 **Local Disk** 사용을 권장(**RAID10 권장, JBOD 사용 가능**)
- **SSD** Disk를 권장
- **XFS 파일시스템**을 사용해야 함
- mount시에 **noatime** 옵션 사용 - Linux가 각 파일에 마지막으로 액세스한 시간을 기록하는 파일 시스템 메타데이터를 유지 관리하는 방식을 off
- Broker의 파라미터 중 **log.dirs** 에 콤마(,)로 구분한 디렉토리들로 정의
- 하나의 Partition은 하나의 volume에서 생성됨
- Partition들은 log.dirs 의 디렉토리에 round-robin 방식으로 분배
- **NAS 사용 불가**



## Virtual Memory

### Kafka 성능에 큰 영향

## 2.

Kafka in Real Environment

- Memory swapping 최소화  
`vm.swappiness=1` (기본값 : 60)
- Blocking Flush (synchronous) 빈도 감소  
`vm.dirty_ratio=80` (기본값 : 20)
- Non-Blocking background flushes (asynchronous) 빈도 증가  
`vm.dirty_background_ratio=5` (기본값 : 10)
- 이 파라미터들은 `/etc/sysctl.conf` 에 설정하고 `sysctl -p` 명령어로 Load함

## Zookeeper 권장 사양 홀수 개로 구성

2.

Kafka in Real  
Environment

### HW 권장 사양

- 최소 2 Core – 권장 4 Core
- Memory  
8 GB
- Transaction log (dataLogDir)  
512 GB SSD
- Database snapshots (dataDir)  
2 TB SSD RAID 10

## Zookeeper Disk Auto Purge 옵션

# 2.

Kafka in Real  
Environment

### ZooKeeper 의 server.properties 내의 purge snapshots 파라미터 설정

- **autopurge.snapRetainCount** : 보존할 SnapShot 개수(권장 3)
- **autopurge.purgeInterval** : Purge Interval(권장 24)
- 위의 권장 옵션은 24 시간마다 3 개를 제외한 모든 스냅샷을 제거하는 설정
- 미션 크리티컬 시스템의 경우에는 3-5개 Zookeeper 노드에 추가 스토리지를 사용하여 보존할 SnapShot 수를 조정하는 경우도 있음

## Summary

### Broker, Zookeeper 설정

2.

Kafka in Real  
Environment

- Broker는 분리된 각각의 전용 서버에 분리하여 설치 구성하는 것을 권장
- 처리량에 따라서 Thread 관련 파라미터 튜닝 필수
- Broker는 File을 많이 사용
- Broker Heap Memory
- Broker Disk 옵션