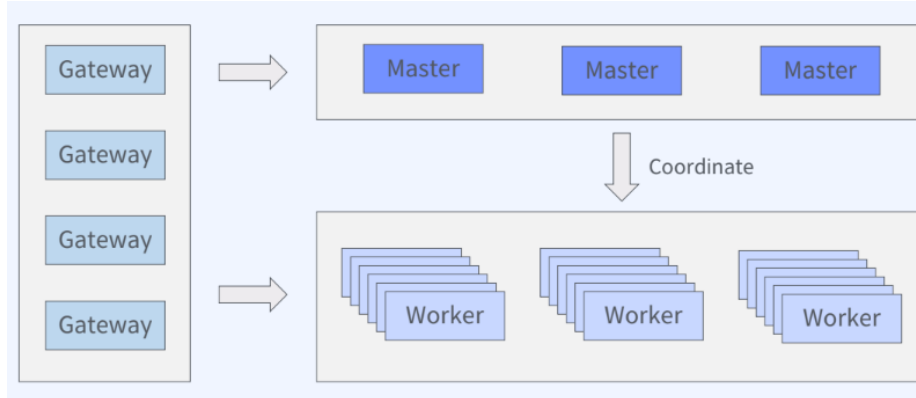


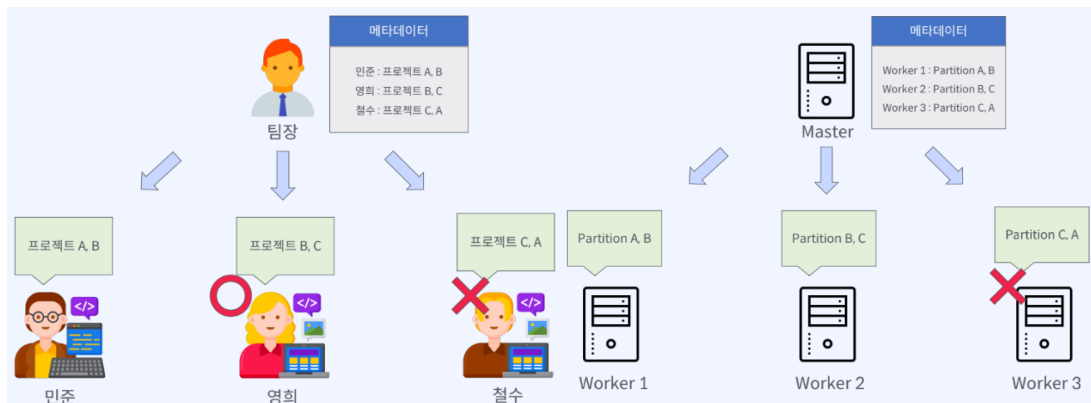
Hadoop Cluster 구축, 고려사항

Cluster란?

- 여러대의 컴퓨터들이 연결되어 하나의 시스템처럼 동작하는 컴퓨터들의 집합



worker : 작업을 수행하는 노드. 데이터를 처리하거나 조회, 검색서비스를 제공한다.



특정 노드에 문제가 발생 되더라도 다른 서버가 처리할 수 있도록 지원한다.

Cluster 규모 결정

1. 스토리지 용량으로 결정하기

- 저장될 데이터 크기 예측
- 복제 전략 결정
- 저장 기간 고려
- 필요한 노드 수 결정

2. 데이터 수집 속도로 결정하기

- 수집속도 예측
- 처리속도 예측