

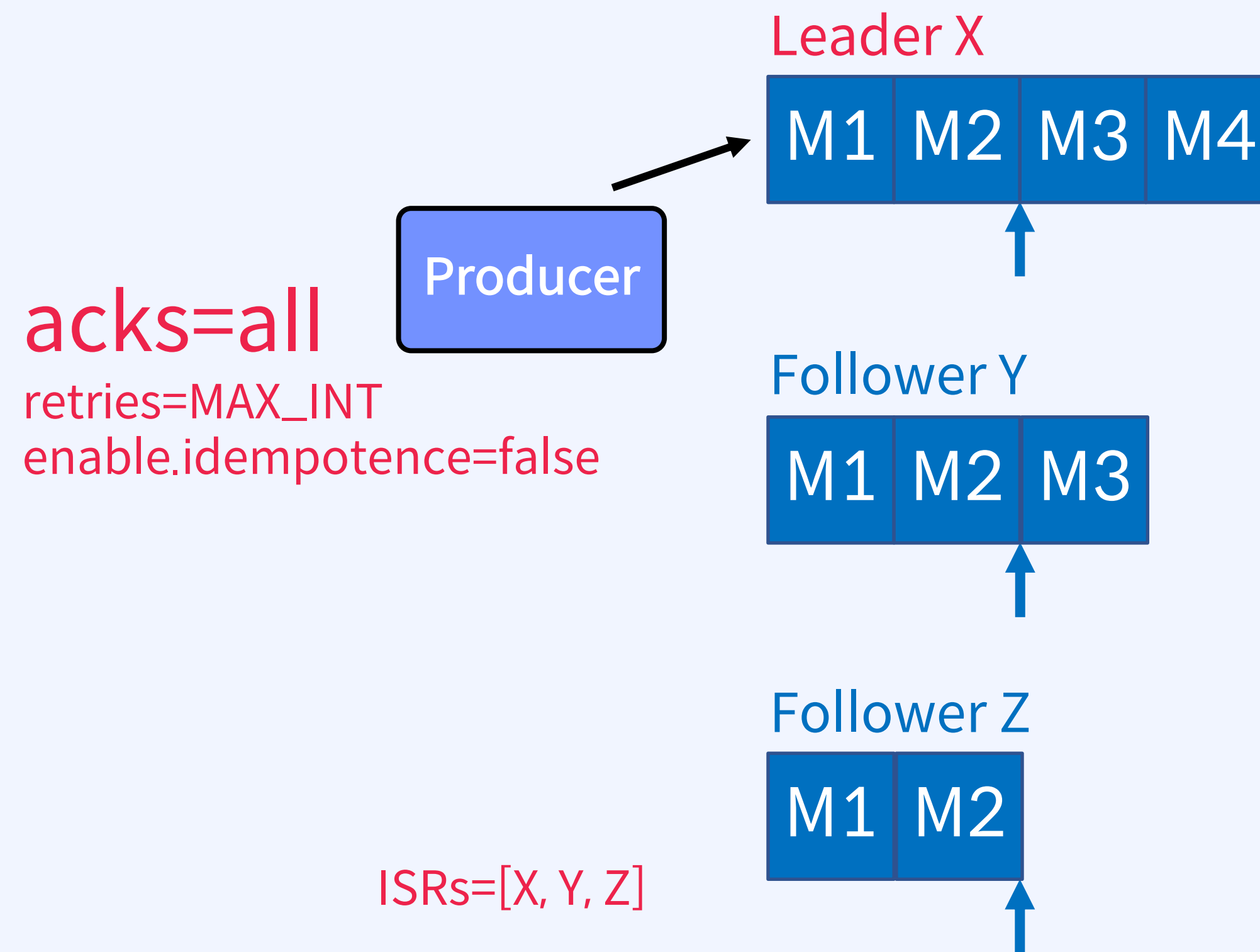
Apache Kafka 심화 개념 및 이해

3 Replica Recovery

Replica Recovery acks=all 의 중요성

3. Replica Recovery

3개의 Replica로 구성된 하나의 Partition
Producer가 4개의 메시지(M1, M2, M3, M4)를 보냈음



메시지 M1, M2가 ISR 리스트
전체에 복제되었고 Commit되었음

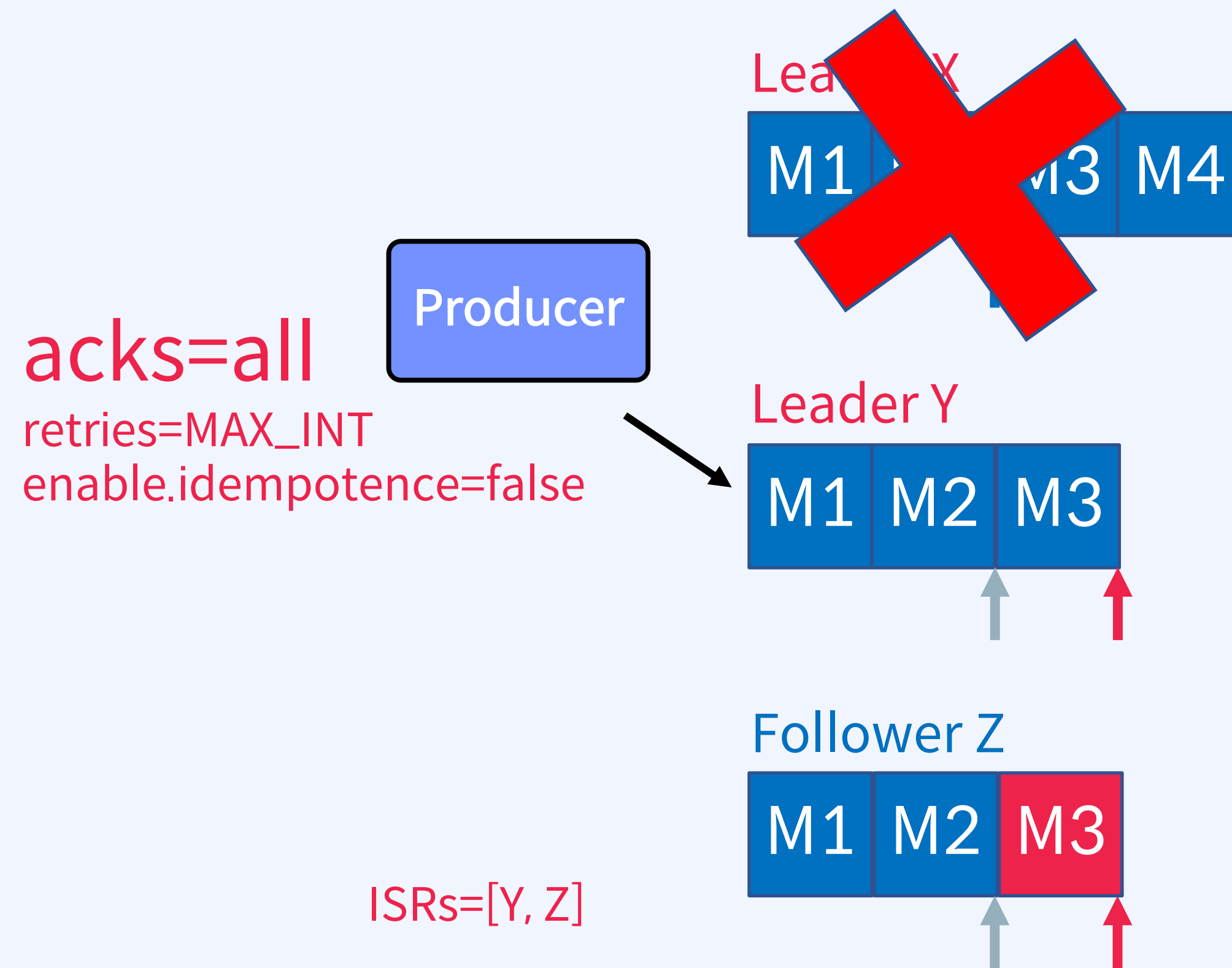
Y는 M3를 복제했지만 Commit은
못한 상태

Z는 M2까지만 복제했음

Replica Recovery acks=all 의 중요성

3. Replica Recovery

Broker X가 장애가 나면, 새로운 Leader가 선출됨
Controller가 Y를 Leader로 선출했다고 가정



M3, M4를 Commit하기 전에
Broker X 장애 발생

Y가 Leader로 선출

Leader Epoch가 0에서 1로 증가

Z는 M3를 fetch함

Y는 High Water Mark를 진행

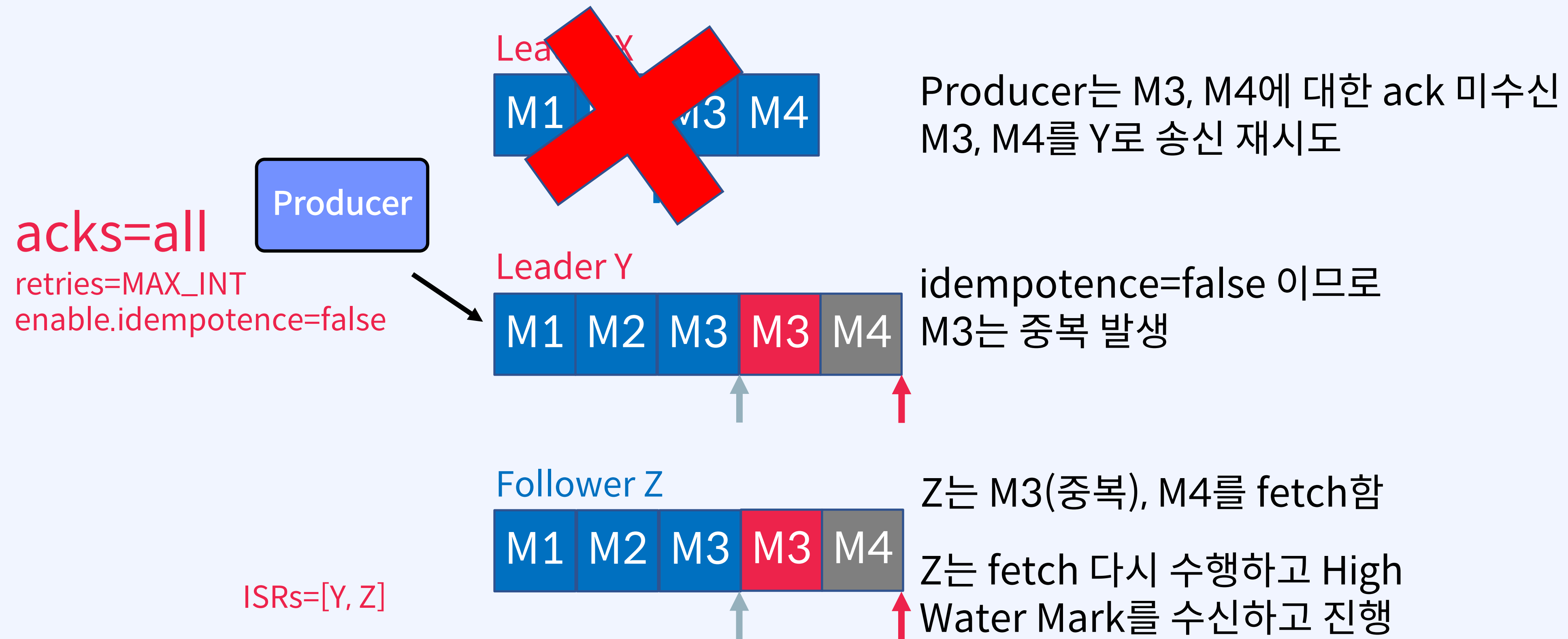
Z는 fetch 다시 수행하고 High
Water Mark를 수신하고 진행

M3는 Commit
된 적이 없는데 왜
Y는 M3를 가지고
있을 수 있는가?

Replica Recovery acks=all 의 중요성

3. Replica Recovery

Broker X는 M3, M4에 대한 ack를 Producer에게 보내지 못했음
Producer는 재시도에 의해 M3, M4를 다시 보냄



Replica Recovery

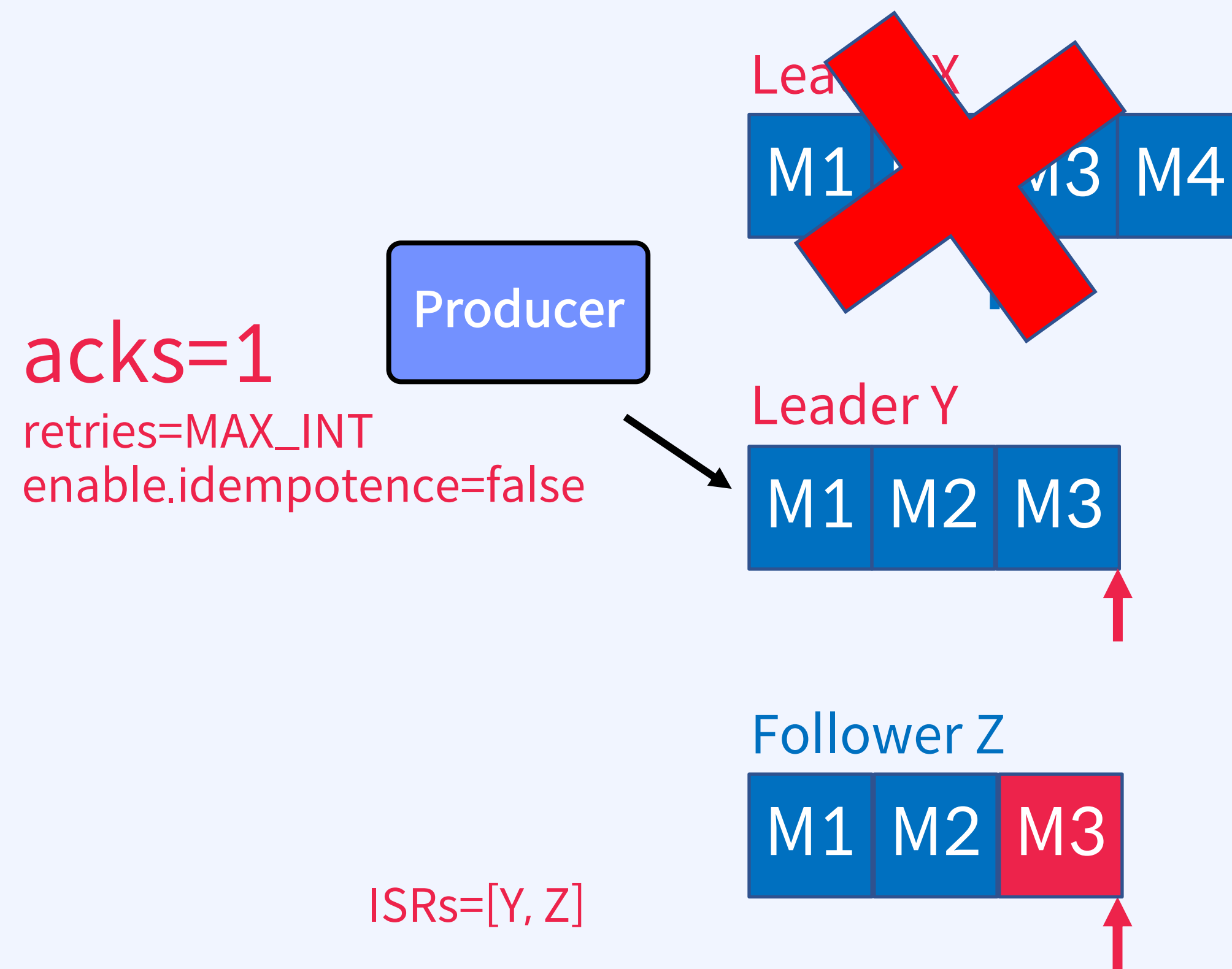
만약 acks=1 이었다면?

3.

Replica Recovery

만약, acks=1 이었다면?

Y, Z가 복제를 못했던 M4는 어떻게 될까?



Leader X가 장애나기 전에
Producer는 M4에 대한
ack를 수신

Y가 Leader로 선출

Leader Epoch가 0에서 1로 증가

Z는 M3를 fetch함

Y는 High Water Mark를 진행

Z는 fetch 다시 수행하고 High
Water Mark를 수신하고 진행

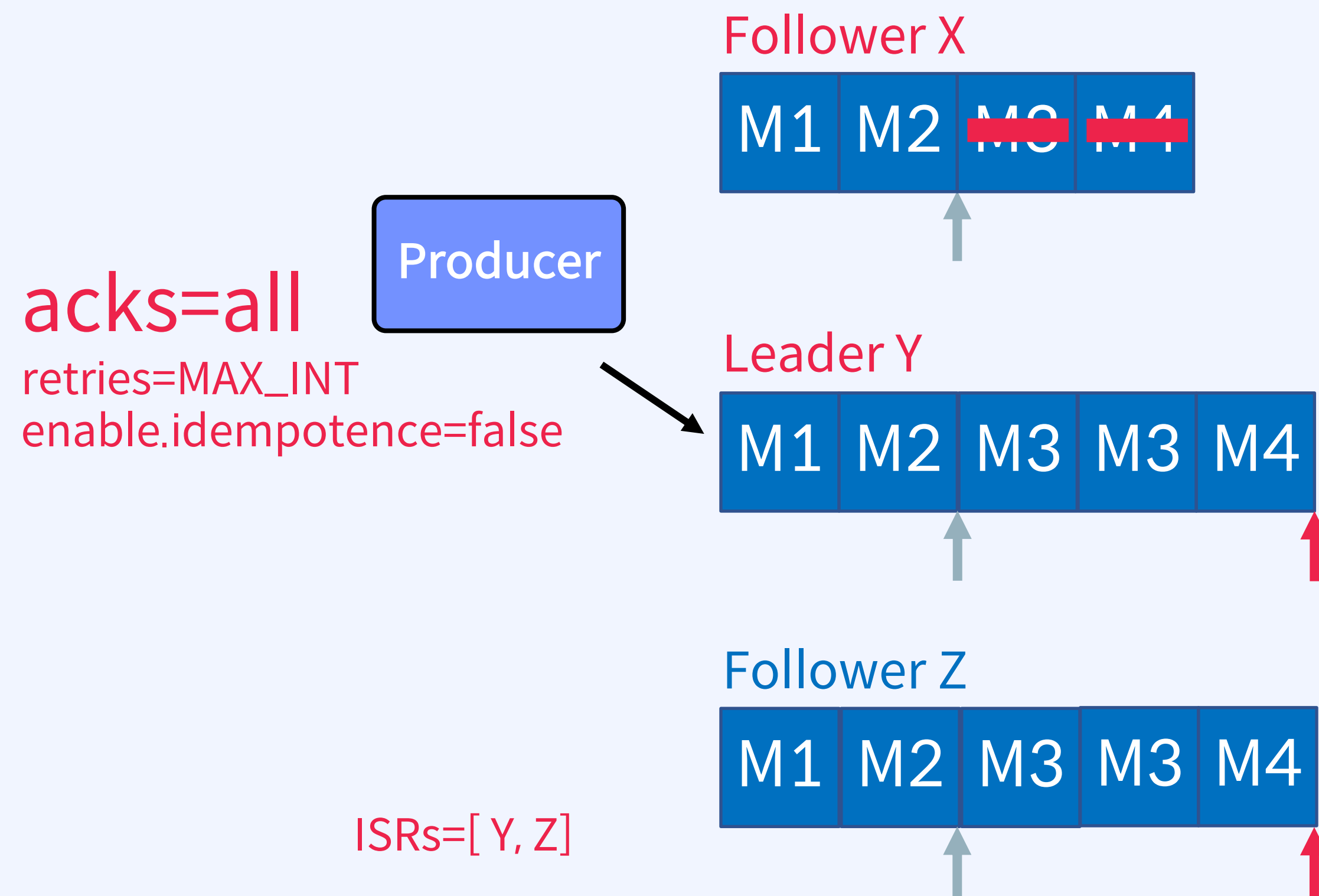
Producer가 M4
송신을 재시도하지
않게 되므로, M4는
영원히 잃어버리게
되는 결과 초래

Replica Recovery

장애가 발생했던 X가 복구되면?

장애가 발생했던 X가 복구되면?

X는 Follower가 되어서 Y로부터 복제함



X가 복구되면 Zookeeper에 연결

X는 Controller로부터 metadata를 받음

X는 Leader Y로부터 Leader Epoch를 fetch

X는 Leadership이 변경된 시점부터 Truncate

Replica Recovery

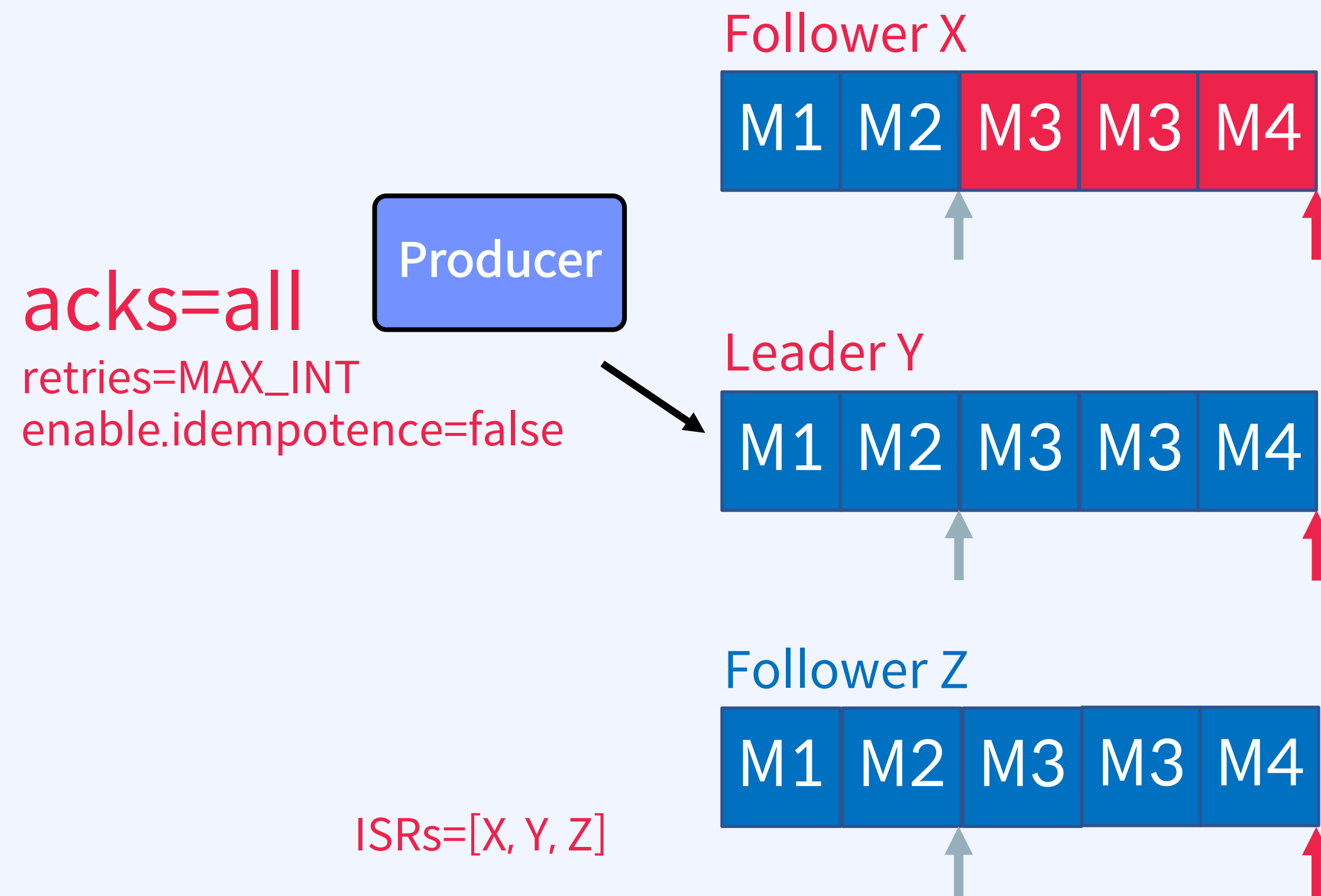
장애가 발생했던 X가 복구되면?

3.

Replica Recovery

장애가 발생했던 X가 복구되면?

X는 Follower가 되어서 Y로부터 복제함



X는 Leader Y로부터 복제

복제가 한번 일어나면, ISR 리스트에 복귀

Availability 와 Durability 가용성과 내구성 중 선택?

3. Replica Recovery

Topic 파라미터 - **unclean.leader.election.enable**

ISR 리스트에 없는 Replica를 Leader로 선출할 것인지에 대한 옵션 (default : false)

ISR 리스트에 Replica가 하나도 없으면 Leader 선출을 안 함 - 서비스 중단

ISR 리스트에 없는 Replica를 Leader로 선출함 - 데이터 유실

Topic 파라미터 - **min.insync.replicas**

최소 요구되는 ISR의 개수에 대한 옵션 (default : 1)

ISR 이 min.insync.replicas 보다 적은 경우,
Producer는 NotEnoughReplicas 예외를 수신

Producer에서 **acks=all**과 함께 사용할 때 더 강력한 보장 + **min.insync.replicas=2**

n개의 Replica가 있고, min.insync.replicas=2 인 경우 n-2개의 장애를 허용할 수 있음

Availability 와 Durability 가용성과 내구성 중 선택?

3. Replica Recovery

데이터 유실이 없게 하려면?

Topic : **replication.factor** 는 2 보다 커야 함(**최소 3이상**)

Producer : **acks** 는 **all** 이어야 함

Topic : **min.insync.replicas** 는 1 보다 커야 함(**최소 2 이상**)

데이터 유실이 다소 있더라도 가용성을 높게 하려면?

Topic : **unclean.leader.election.enable** 를 **true** 로 설정

Summary

가용성과 내구성 관련 파라미터

3.

Replica Recovery

- replication.factor
- acks
- min.insync.replicas
- unclean.leader.election.enable