Yandex

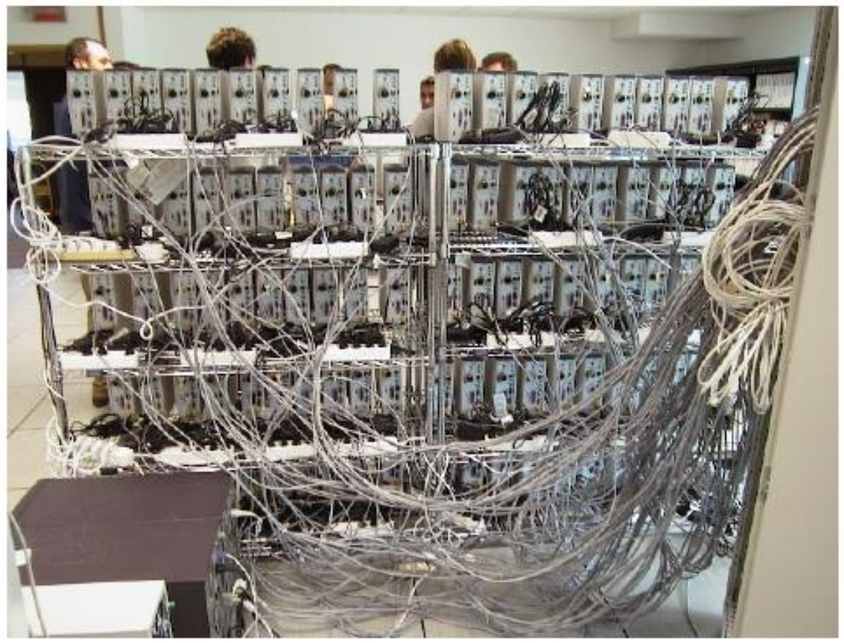
Distributed Systems

Unreliable Components

분산 시스템의 구성

Architecture

Fail-Recovery + Fair-Loss Link + Asynchronous



hadoop

Fail-Stop + Perfect Link + Synchronous

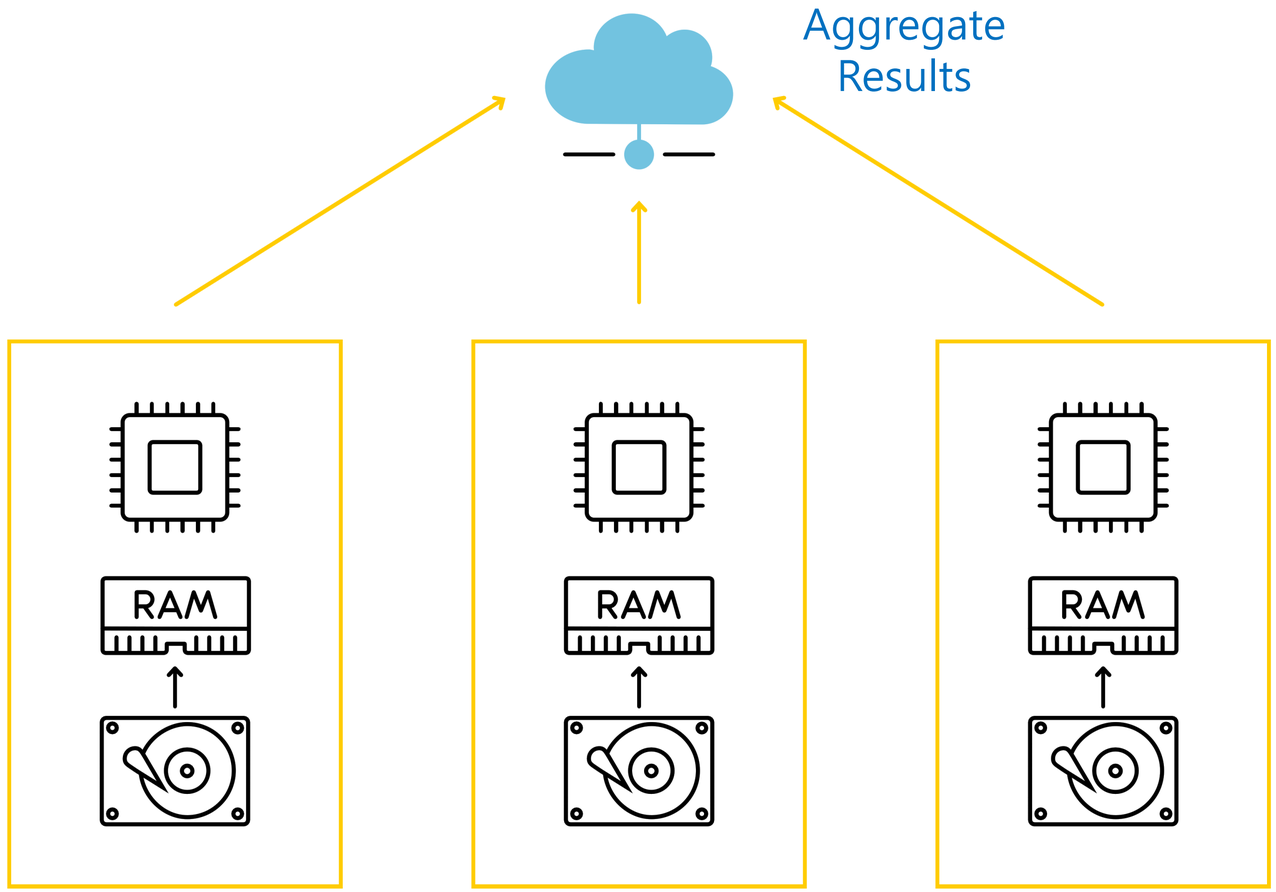
OpenMP

MPI

슈퍼 컴퓨터

Byzantine-Failure + Byzantine Link + Asynchronous

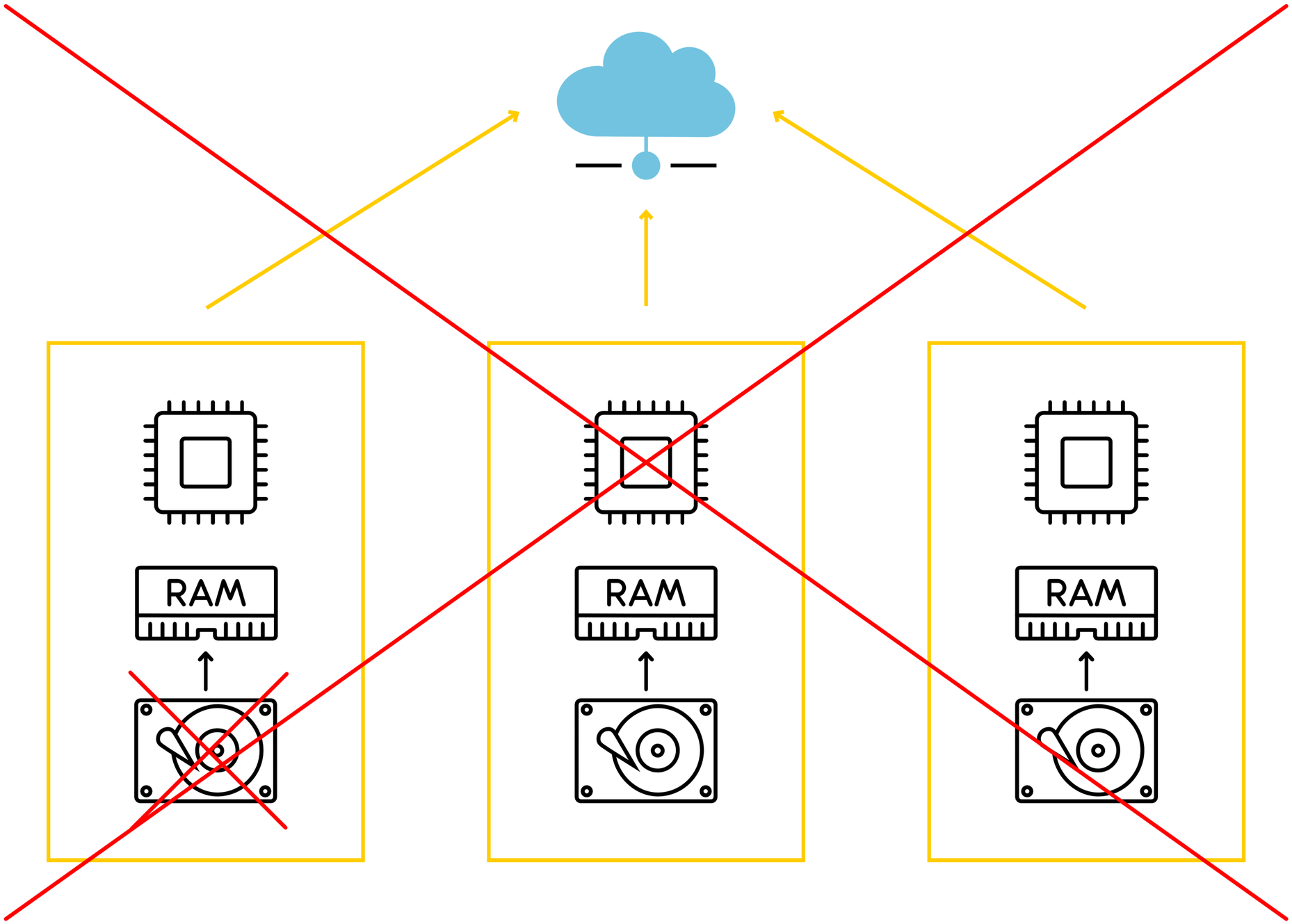
Grid Computing



Node Failures:

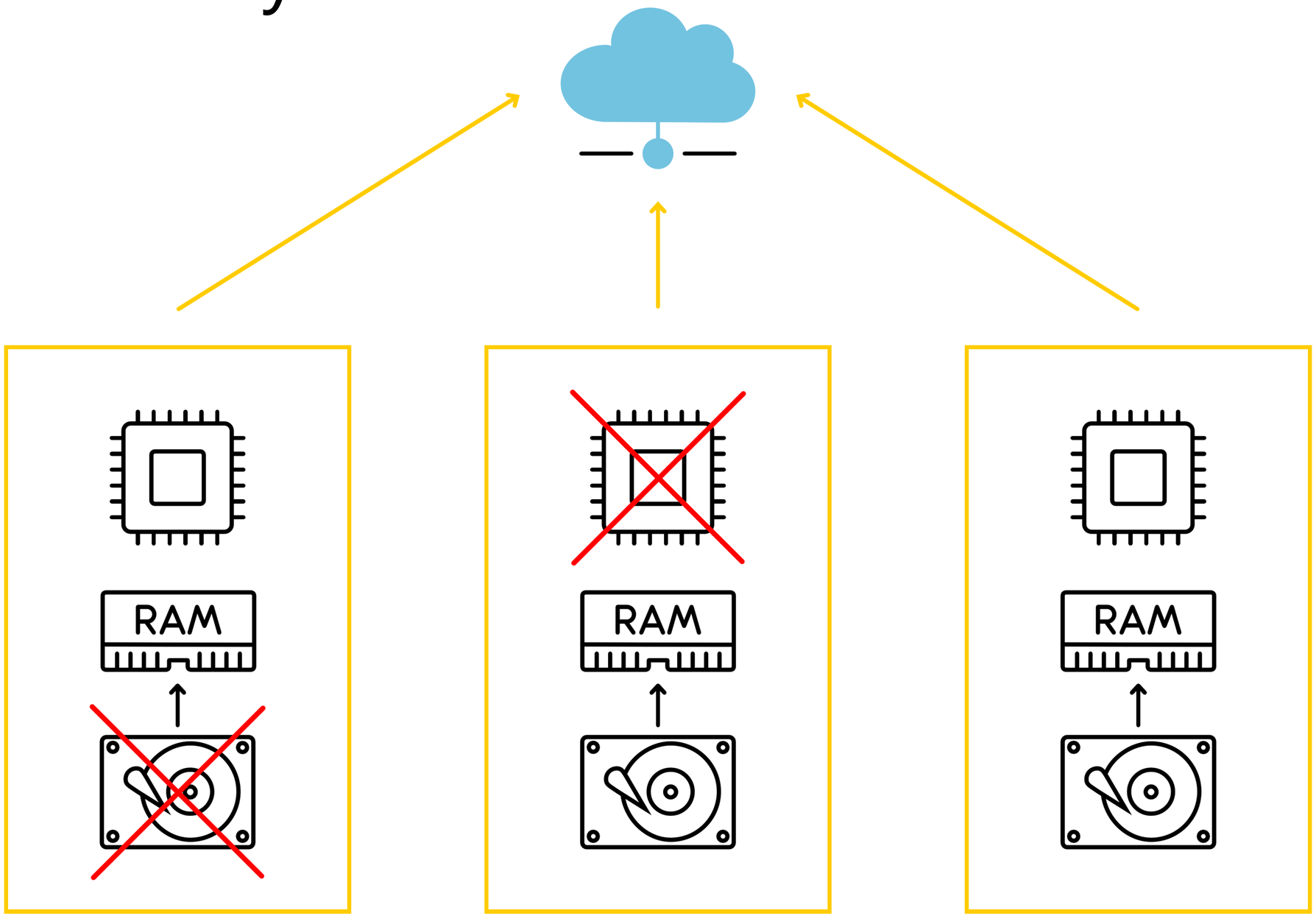
- Fail-Stop
- Fail-Recovery
- Byzantine

Fail-Stop



Node(machine)에 문제가 생길 경우, 관리자가 물리적 + 시스템적 조치를 취해야 한다.

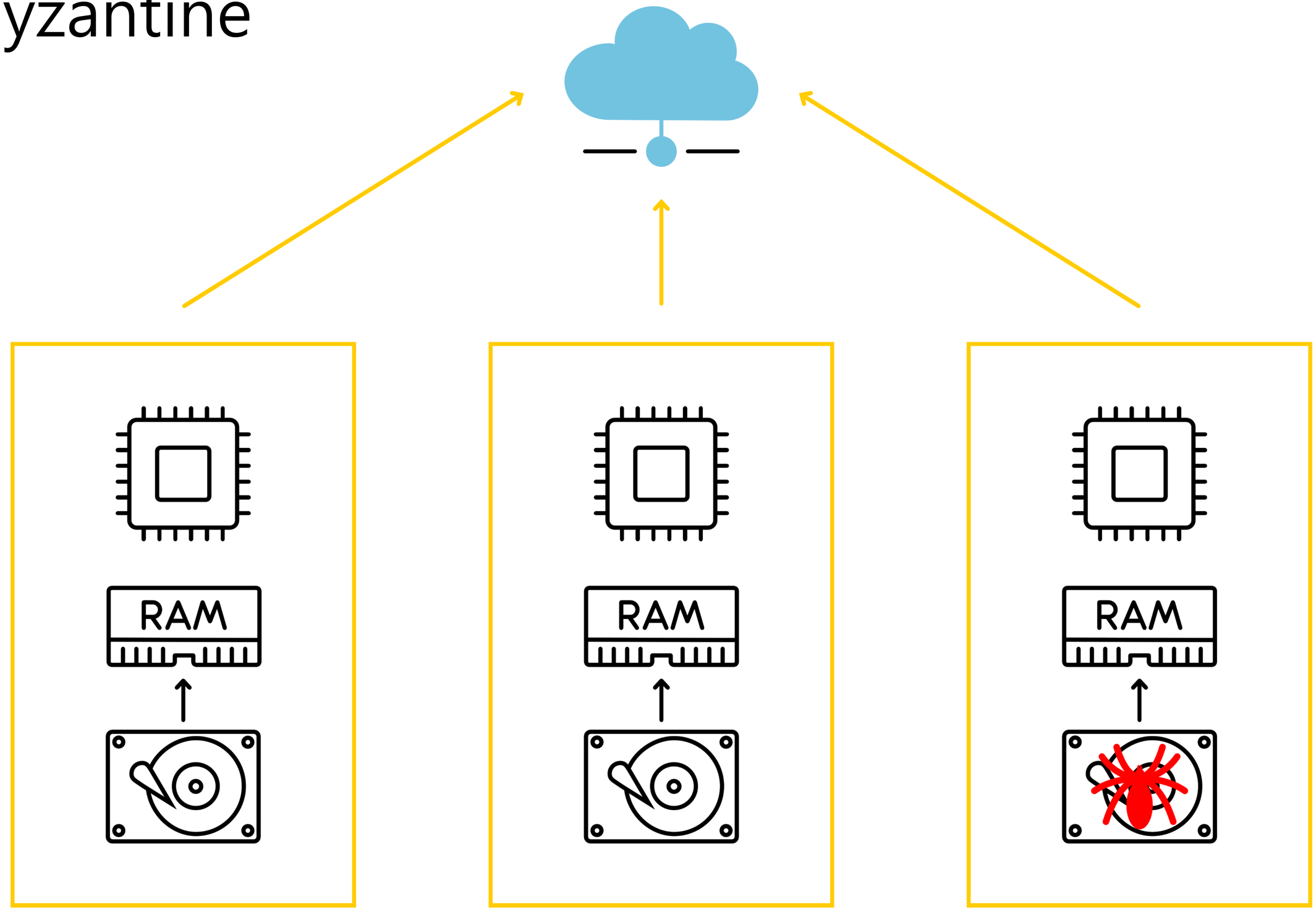
Fail-Recovery



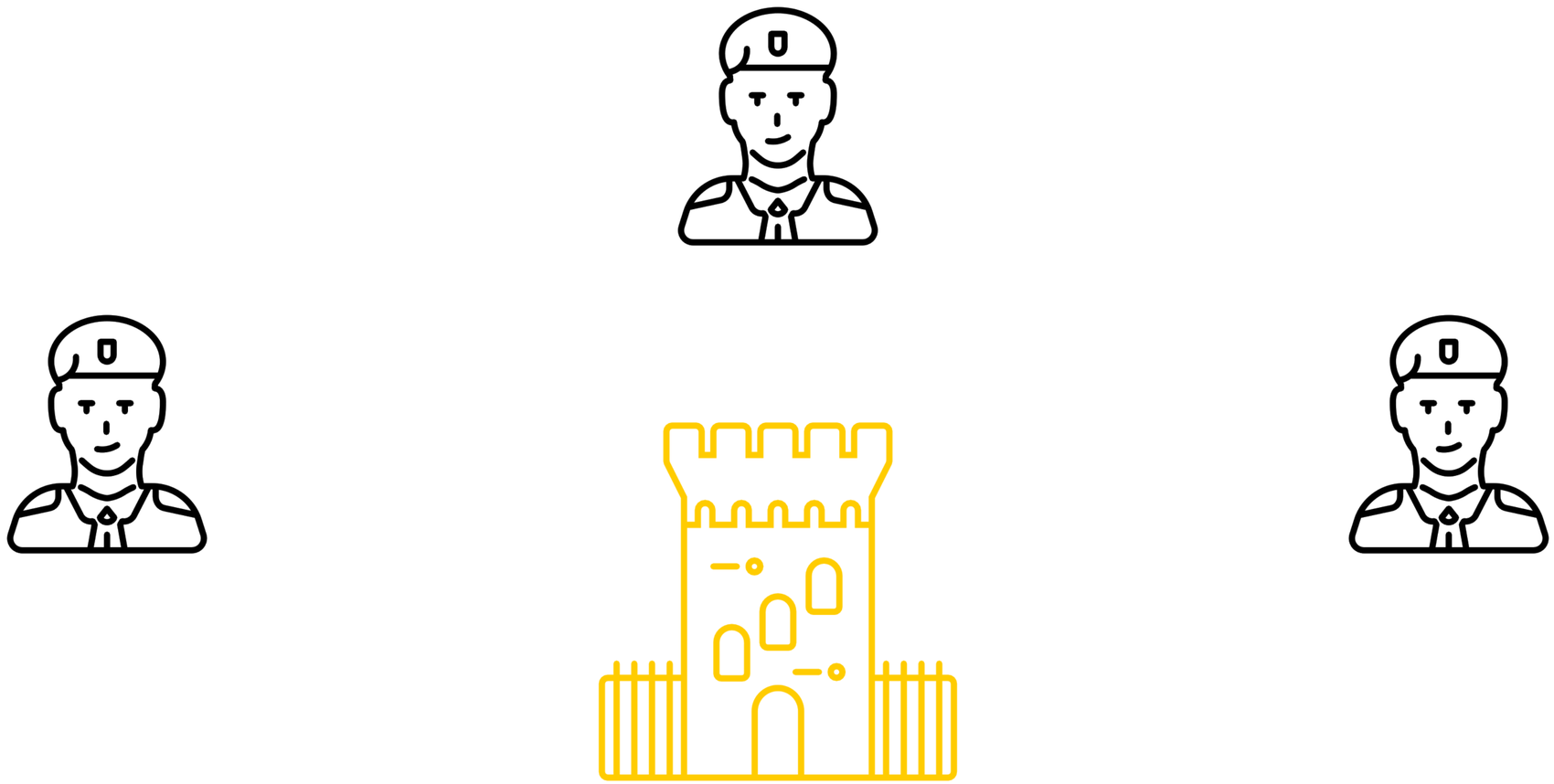
During Computation, nodes can arbitrarily crash and return back
Node가 서비스 중에 이탈하거나 다시 돌아올 수 있다.

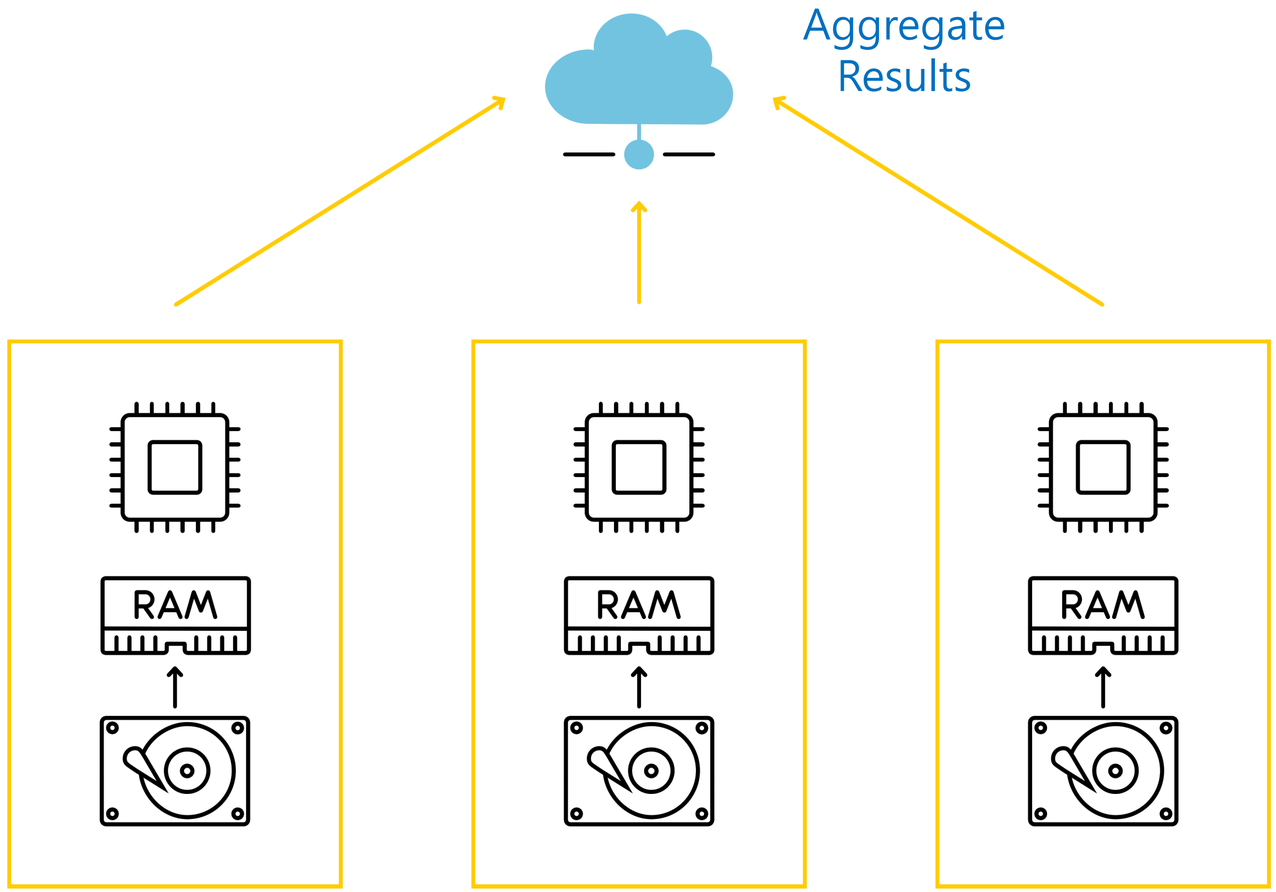
e.g.) machine의 일부분이 고장나면 시스템이 스스로 해당 machine을 제외하여 서비스를 수행하고 물리적 수리를 끝내면 관리자 관여없이 시스템이 알아서 서비스에 연결한다.

Byzantine



The Byzantine Generals Problem



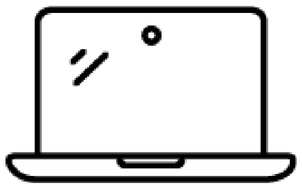


Link Failures:

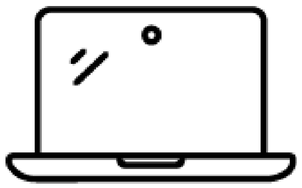
- Perfect Link
- Fair-loss Link
- Byzantine

Fair-Loss Link

Packet Loss가 발생할 수 있다. / 재전송을 통해 해결



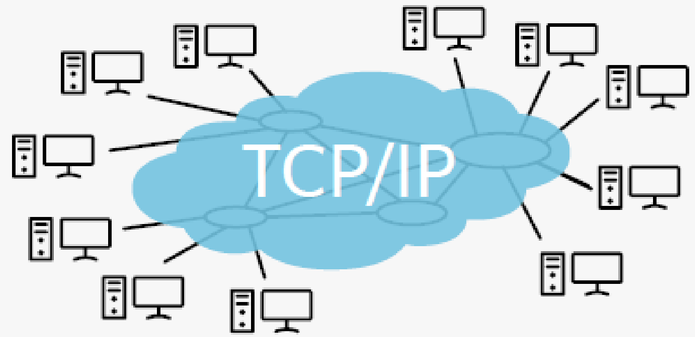
$send\ (a_1, a_2, ..., a_k)$

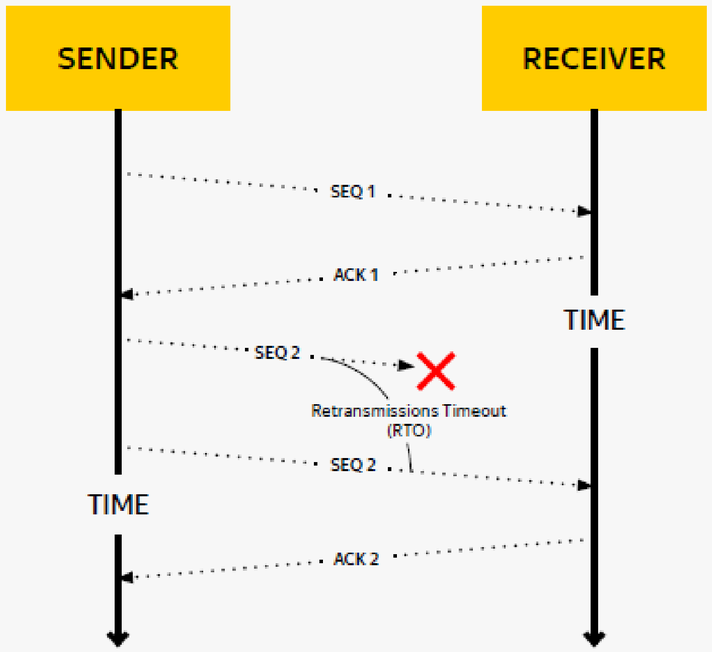


$receive\ (b_1, b_2, ..., b_l)$

$l \leq k$

$\{b_i\} \in \{a_j\}$





[A] Synchronous Systems

> Every message between nodes is delivered within limited time;
> Clock drift is limited;
> Each instruction execution is also limited.

1 clock skew

machine마다 시간이 다를 수 있다.

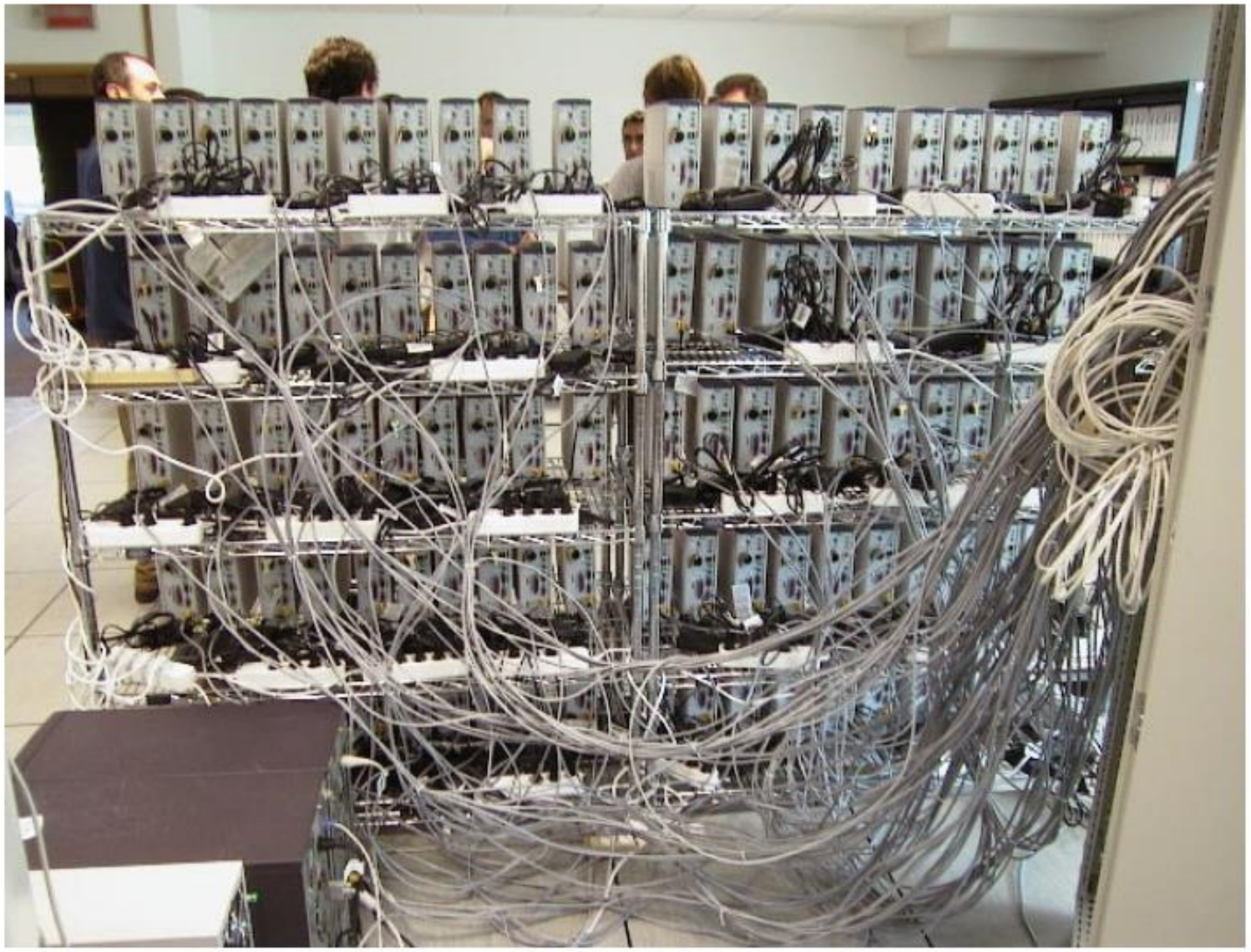
2 clock drift

각 machine의 clock rate가 조금씩 달라져
원래 시간과 맞지 않는 현상

Logical clock을 통해 해결

Lamport logical clock이 대표적

Summary



You can provide a thorough description of a distributed system according to its robustness to node failures, link failures and clock synchronization model used in the system.

BigDATAteam