

Database Systems

Lecture 20 – Ch. 21. Parallel DBMS

24

Beomseok Nam (남범석)
bnam@skku.edu

Parallel vs. Distributed

■ Parallel DBMSs:

synchronized clock 갖고 있다
가기기자

→ Nodes are physically close to each other.

→ Nodes connected with high-speed LAN. (100g, 200g)

→ Nodes access Distributed Shared Memory (DSM)

복용 차이

synchronized clock 갖고 있지 않음

■ Distributed DBMSs:

단일 어제지사

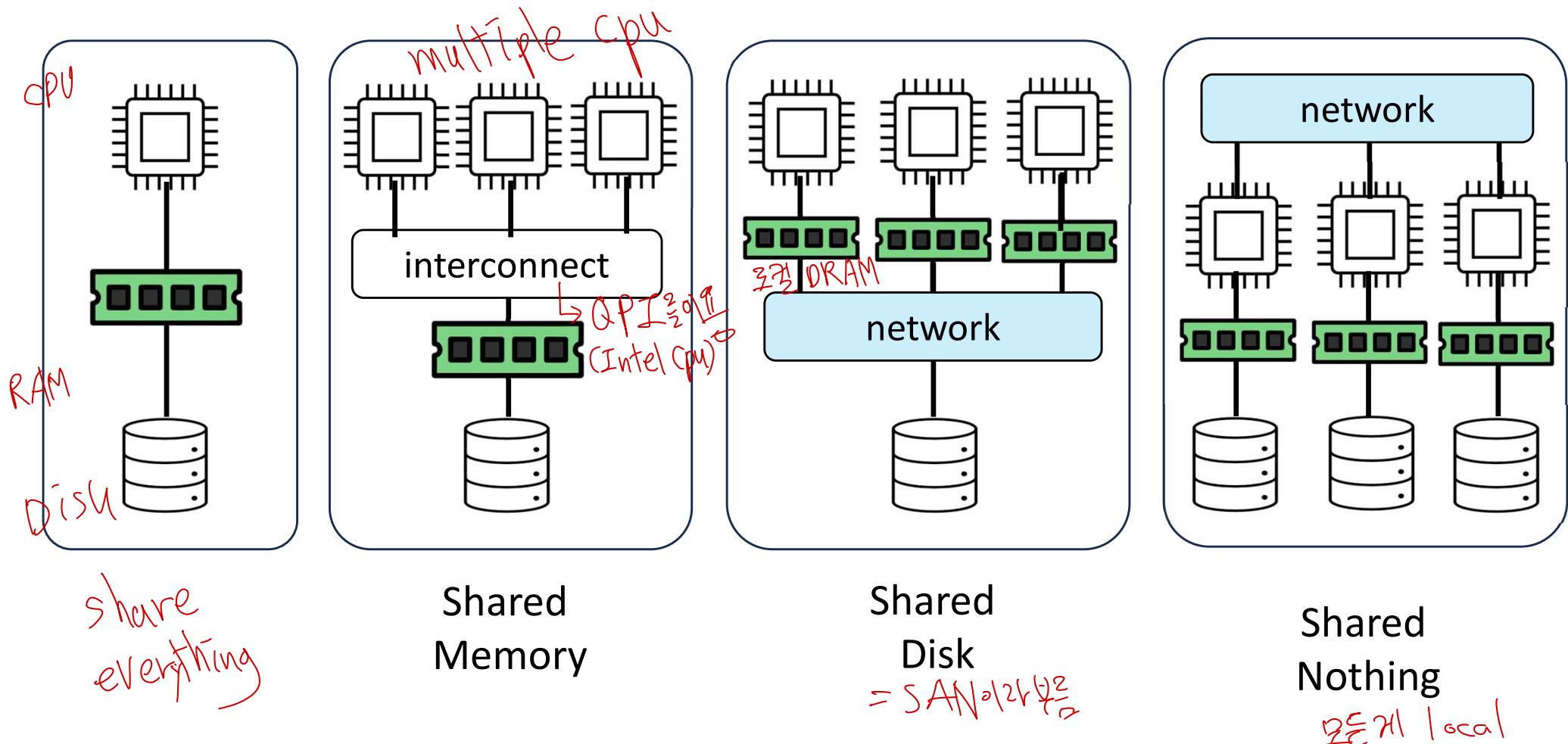
→ Nodes can be far from each other.

→ Nodes connected using public network.

1g

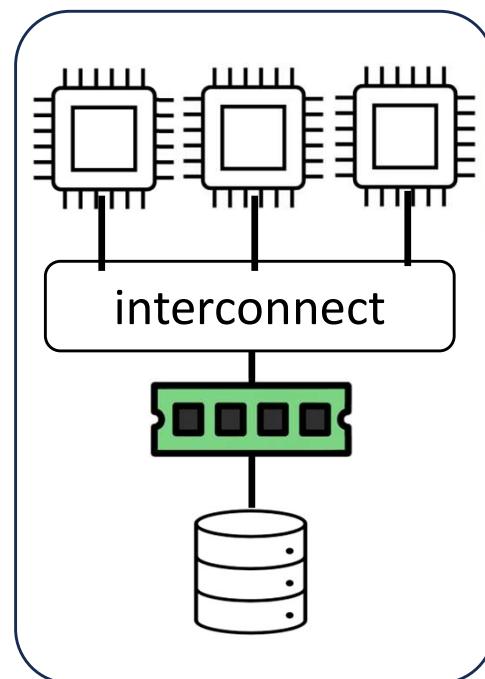
→ Nodes do not share memory

System Architecture



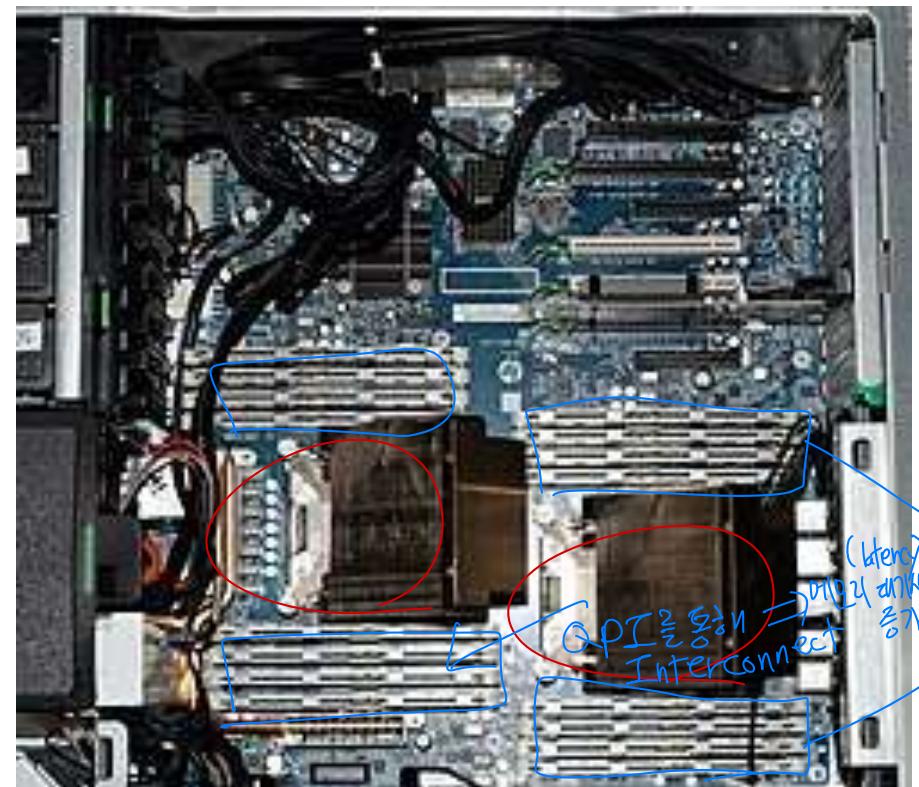
Shared Memory

- Memory shared by multiple processors on the same machine
- NUMA (Non-Uniform Memory Access) architecture



Shared
Memory

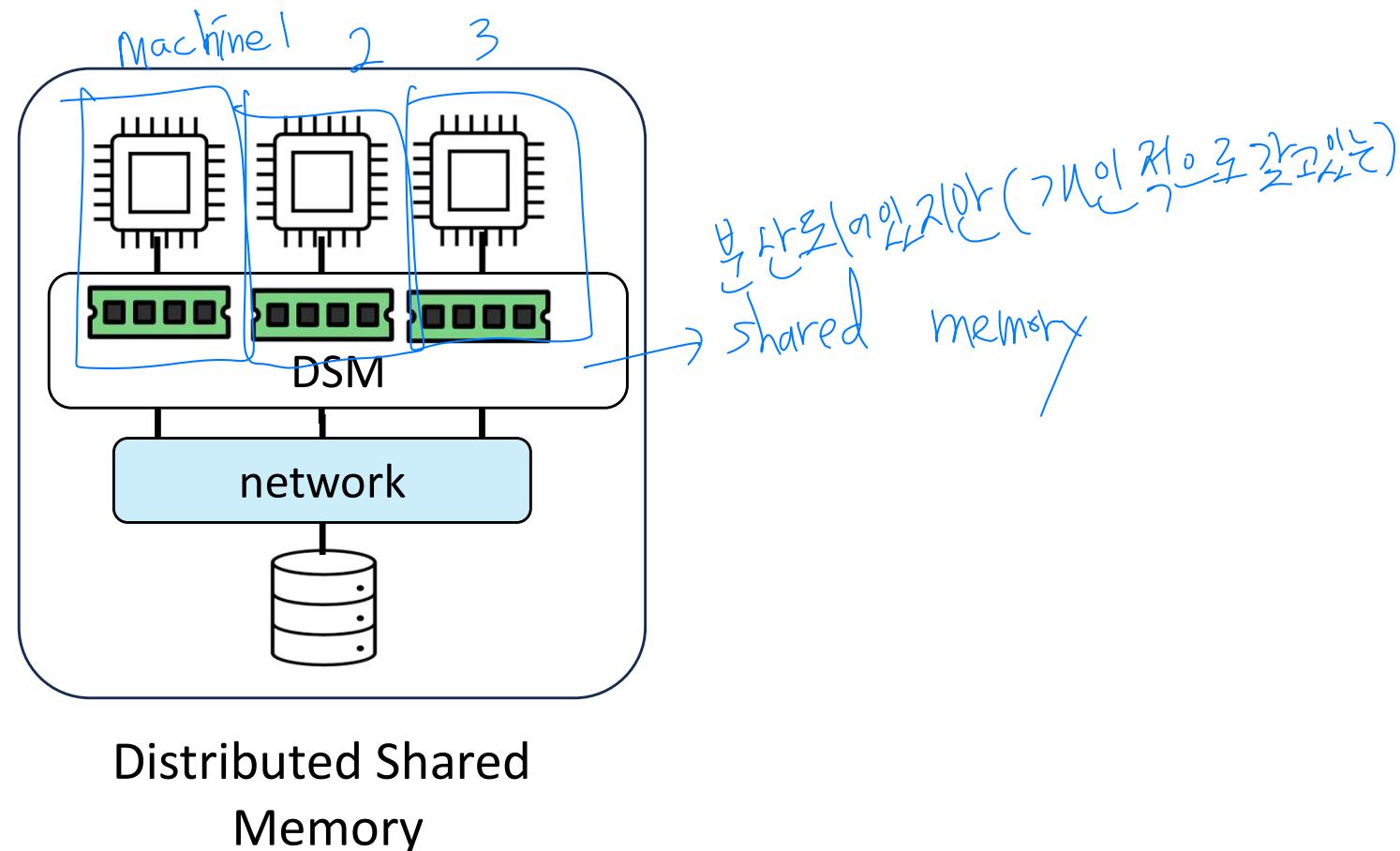
local memory $\xrightarrow{\text{latency}}$
remote memory $\xrightarrow{\text{latency}}$
QPI $\xrightarrow{\text{latency}}$ $\xrightarrow{\text{latency}}$ $\xrightarrow{\text{latency}}$
 $\xrightarrow{\text{latency}}$ $\xrightarrow{\text{latency}}$ $\xrightarrow{\text{latency}}$



motherboard of a workstation
with two CPU sockets

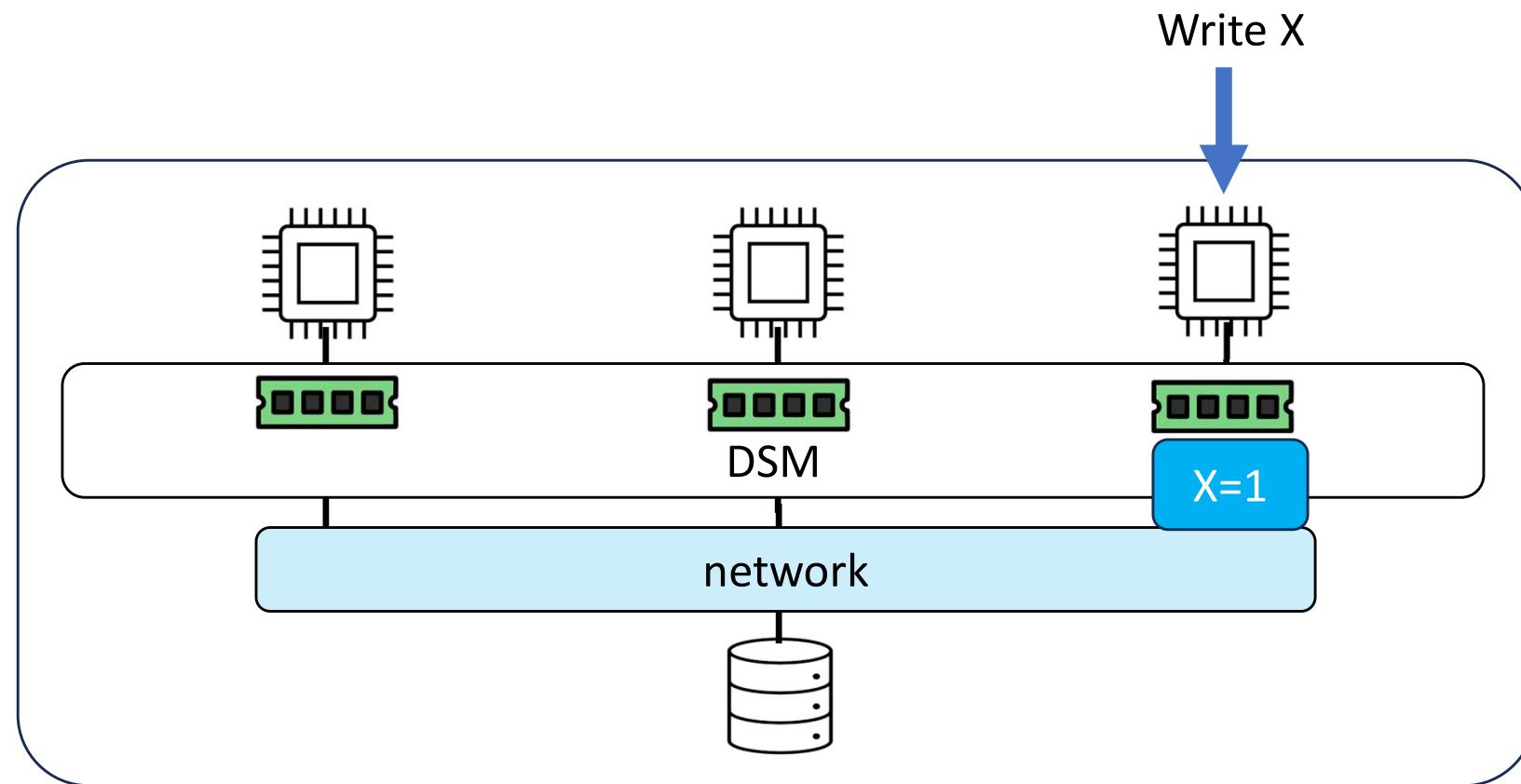
Distributed Shared Memory

- A software abstraction that allows processes on different machines (nodes) to share memory
 - TreadMarks, Ivy, Munin, ORACLE cache fusion



Distributed Shared Memory

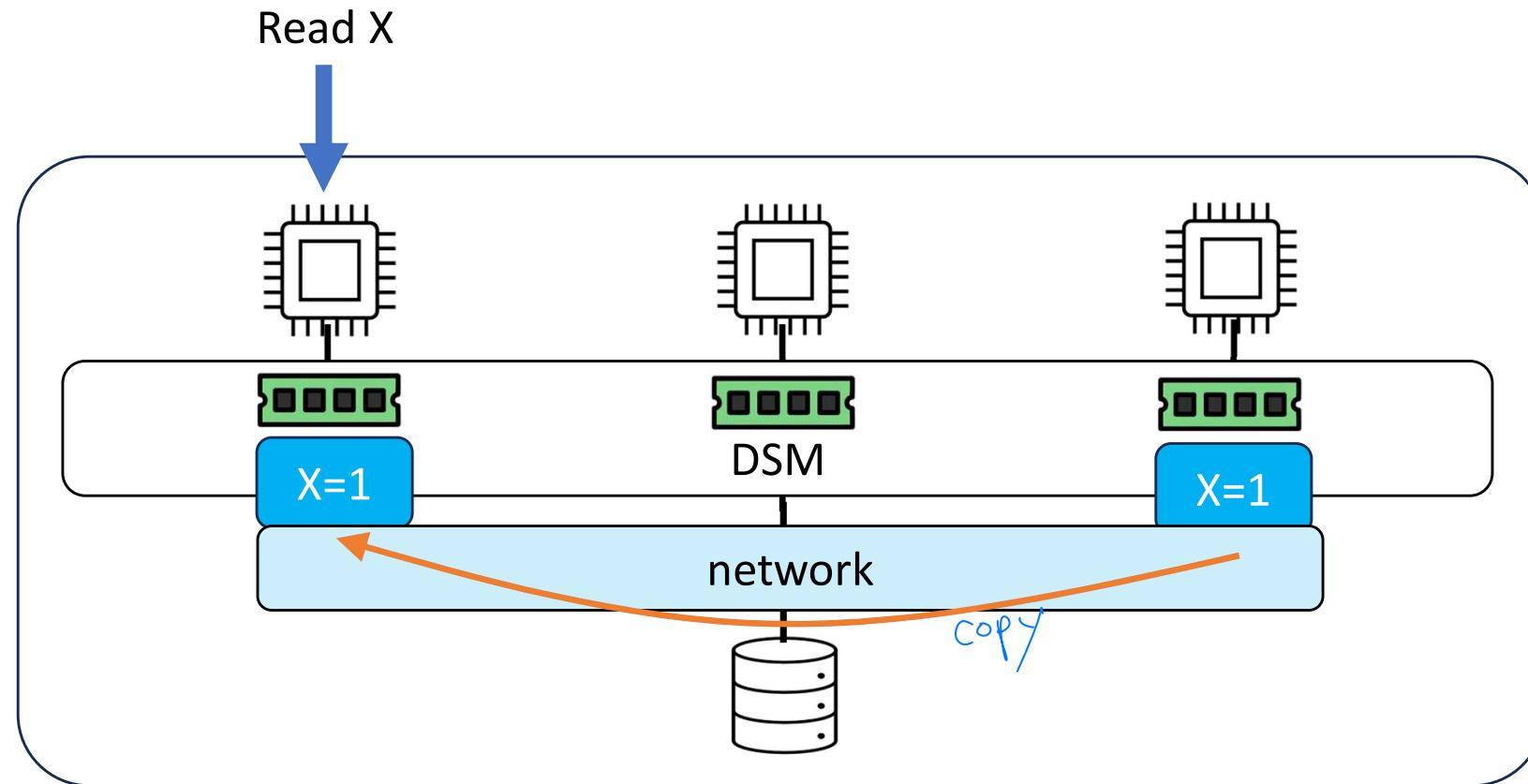
- Cache Coherence



Distributed Shared
Memory

Distributed Shared Memory

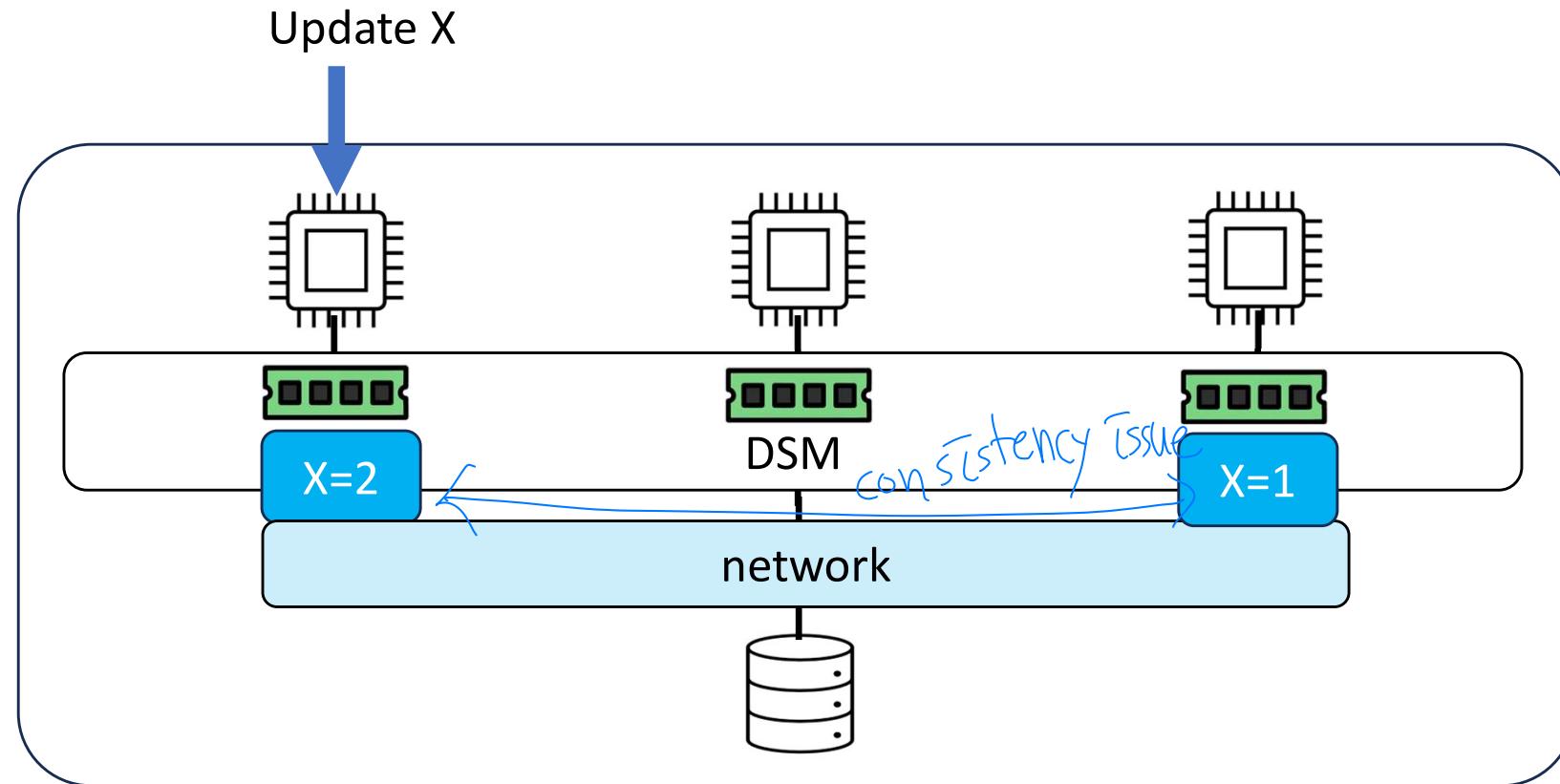
- Cache Coherence



Distributed Shared
Memory

Distributed Shared Memory

- Cache Coherence

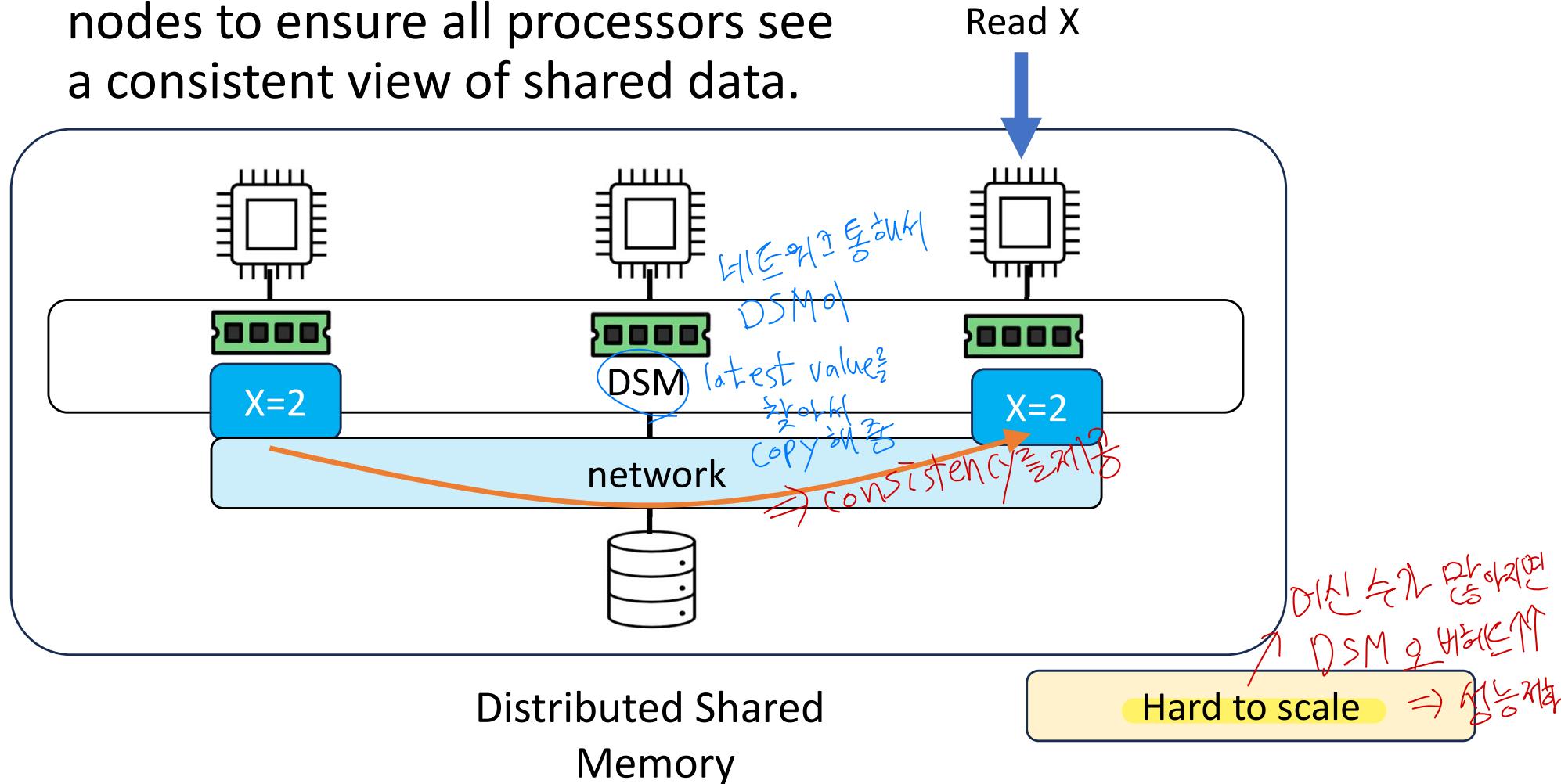


Distributed Shared
Memory

Distributed Shared Memory

■ Cache Coherence Issue

- DSM tracks memory updates across nodes to ensure all processors see a consistent view of shared data.

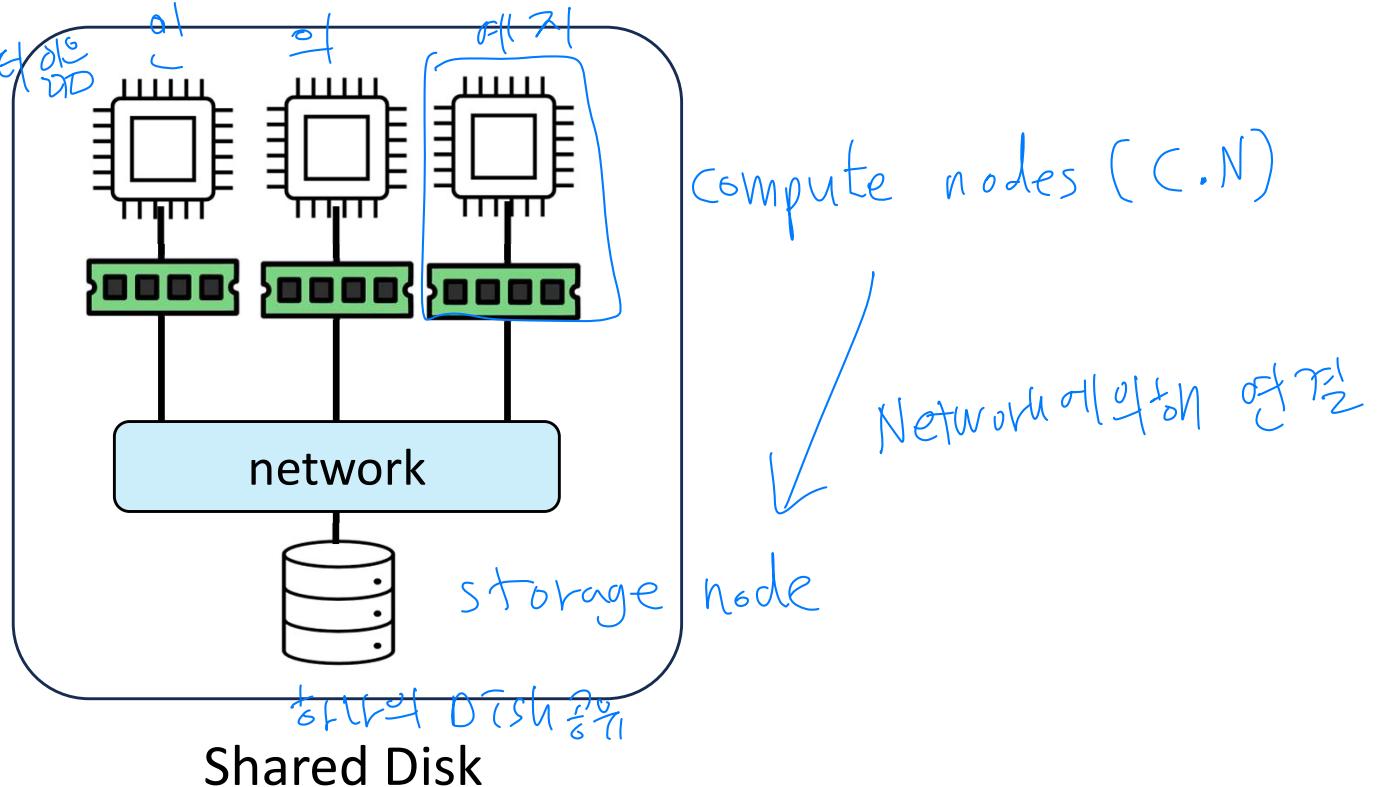


Shared Disk (Disaggregated storage system)

- All CPUs can access a single logical disk directly via an interconnect, but each has its own private memory.

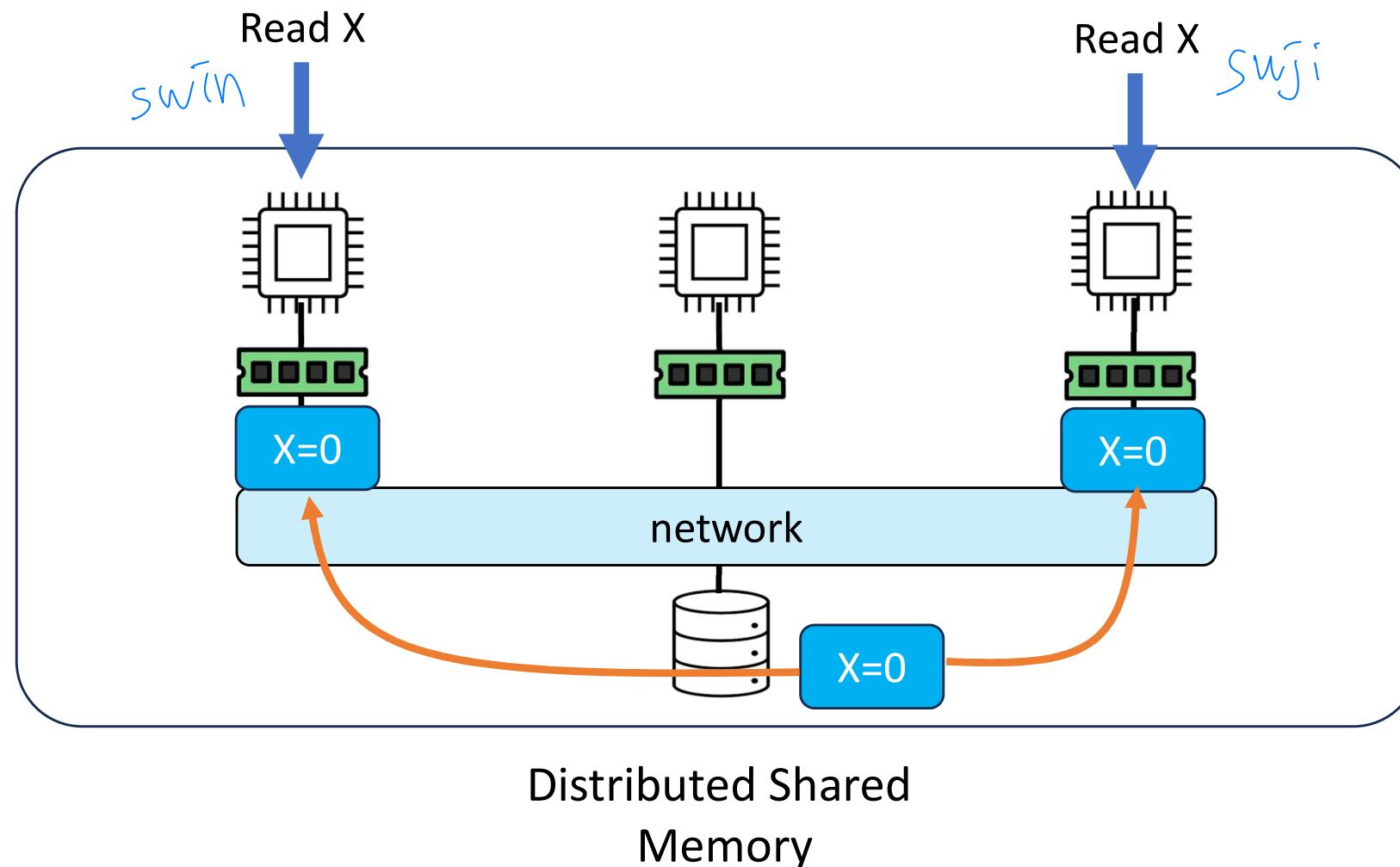
shared Disk 고 사용
→ C.N이 하나의 저장소를 공유하도록 도와주고 관리함
Cluster file systems (e.g., NFS, OCFS2, GFS2, etc)
(Network File system)

(여기) 인의 디자인
어느 시스템 간 속도로 같은 디자인



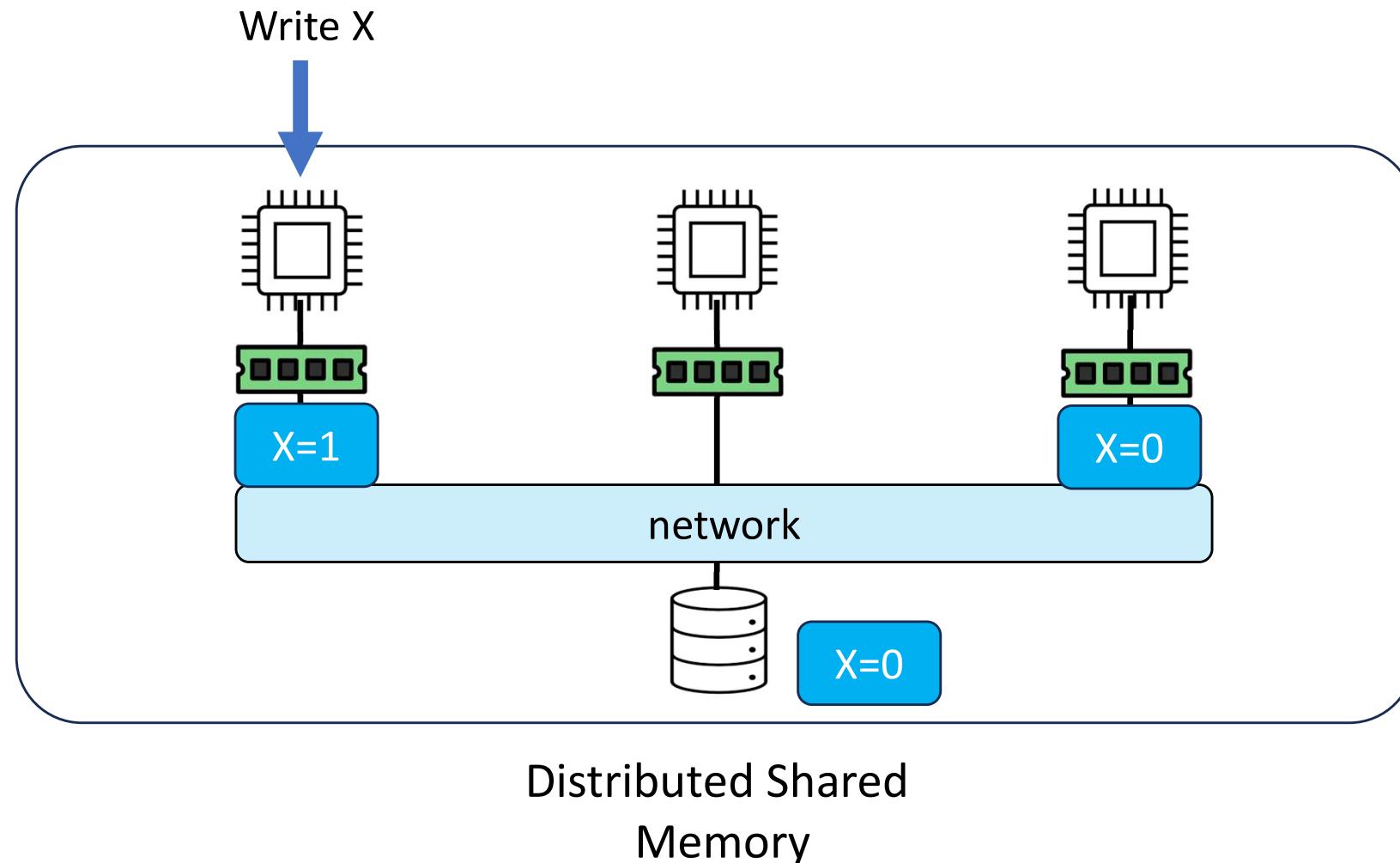
Shared Disk

- Same Cache Coherence issue exists in page cache



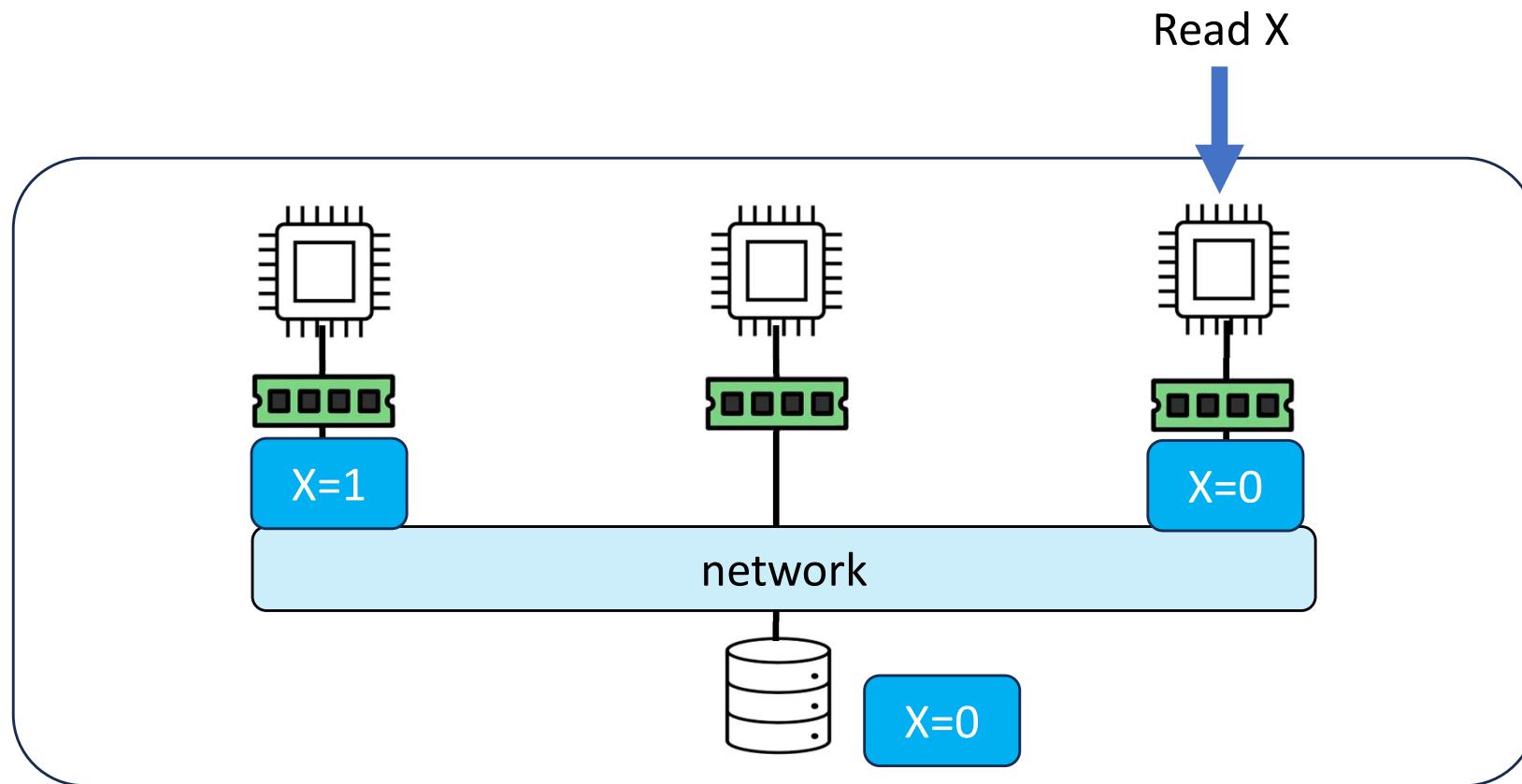
Shared Disk

- Same Cache Coherence issue exists in page cache



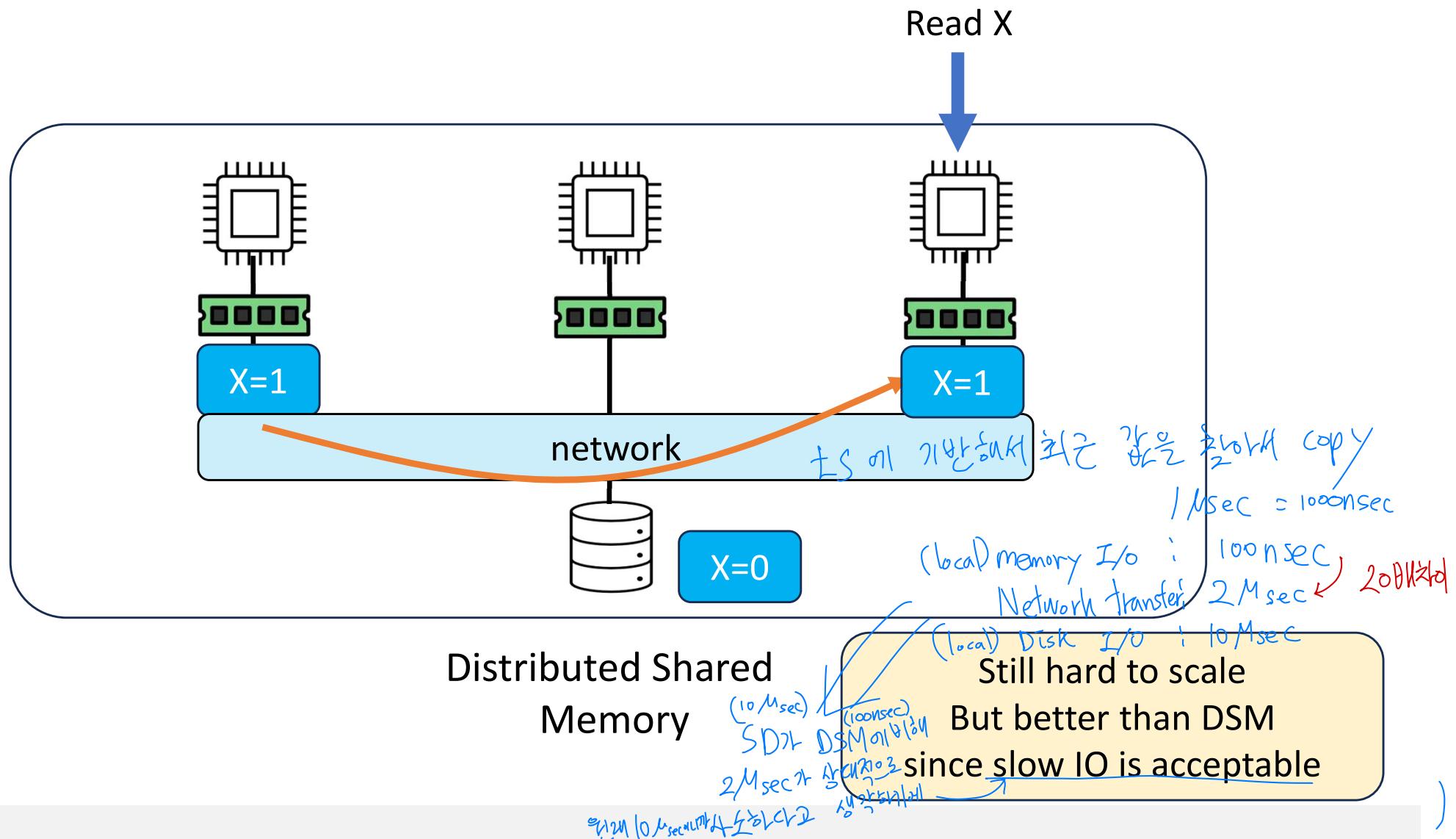
Shared Disk

- Same Cache Coherence issue exists in page cache



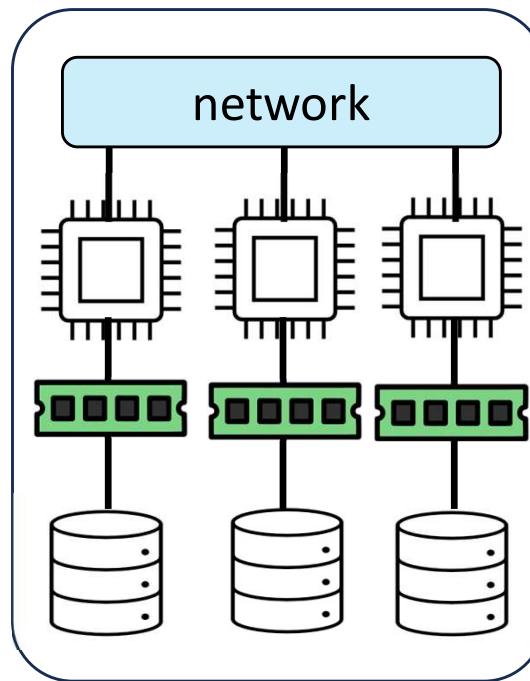
Distributed Shared
Memory

Shared Disk



Shared Nothing (distributed system)

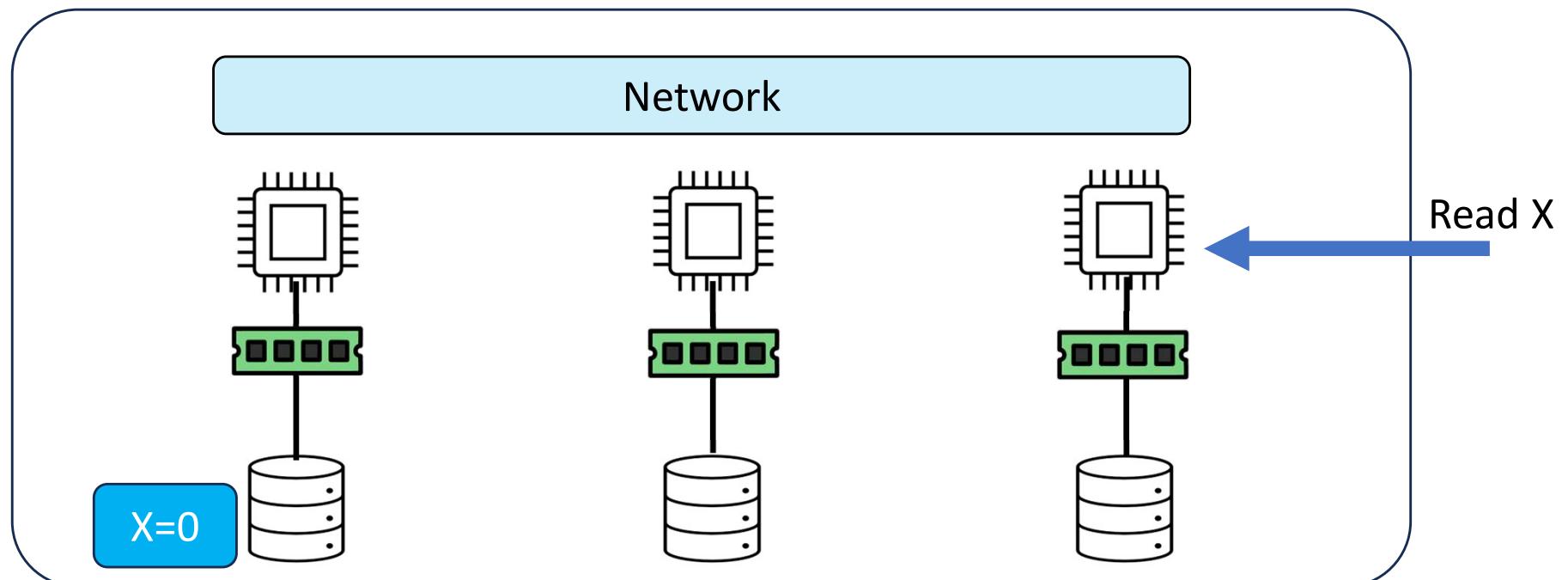
- Each node has its own CPU, memory, and local disk.
- Nodes only communicate with each other via network.



Shared
Nothing

Shared Nothing

- Google's Choice
- MapReduce: Move Computation, not Data



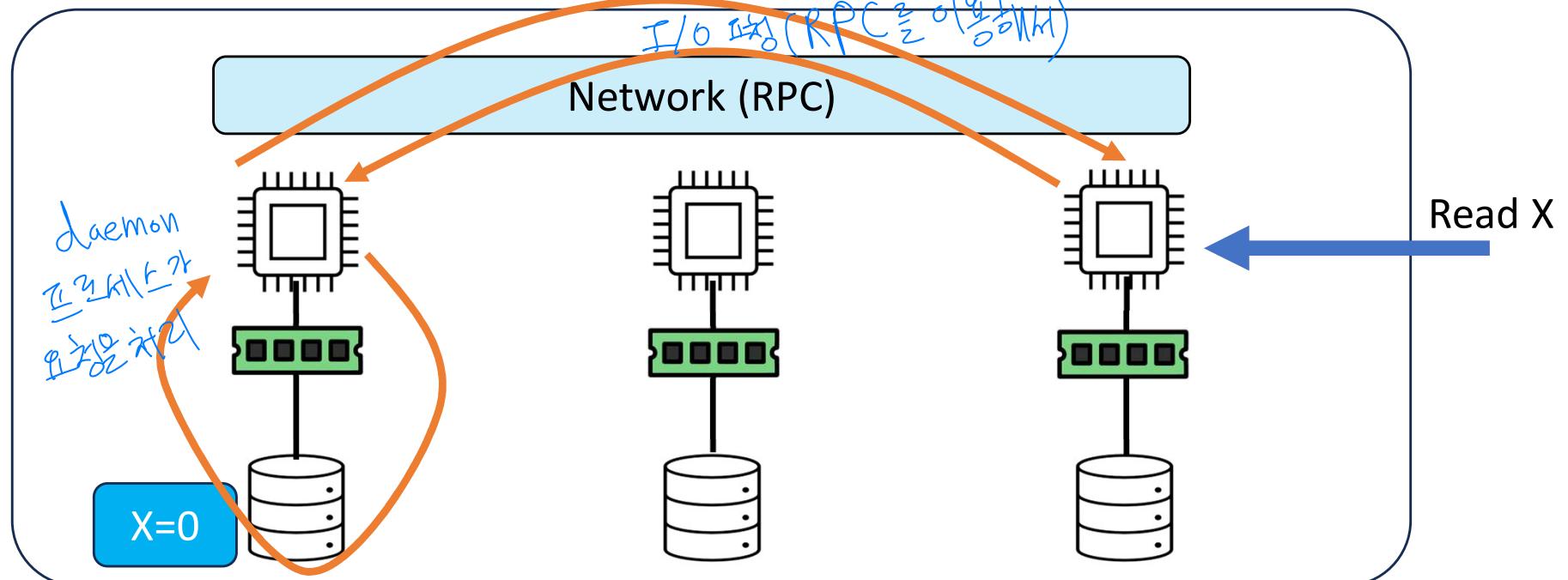
Shared
Nothing

Shared Nothing

- Google's Choice

SAN: busy (크기가 큰 흔적 데이터를 Network를 통해 전송하기가)
디버깅을
이용하는
제한된 고리를 네트워크 통해 전송

- MapReduce: Move Computation, not Data

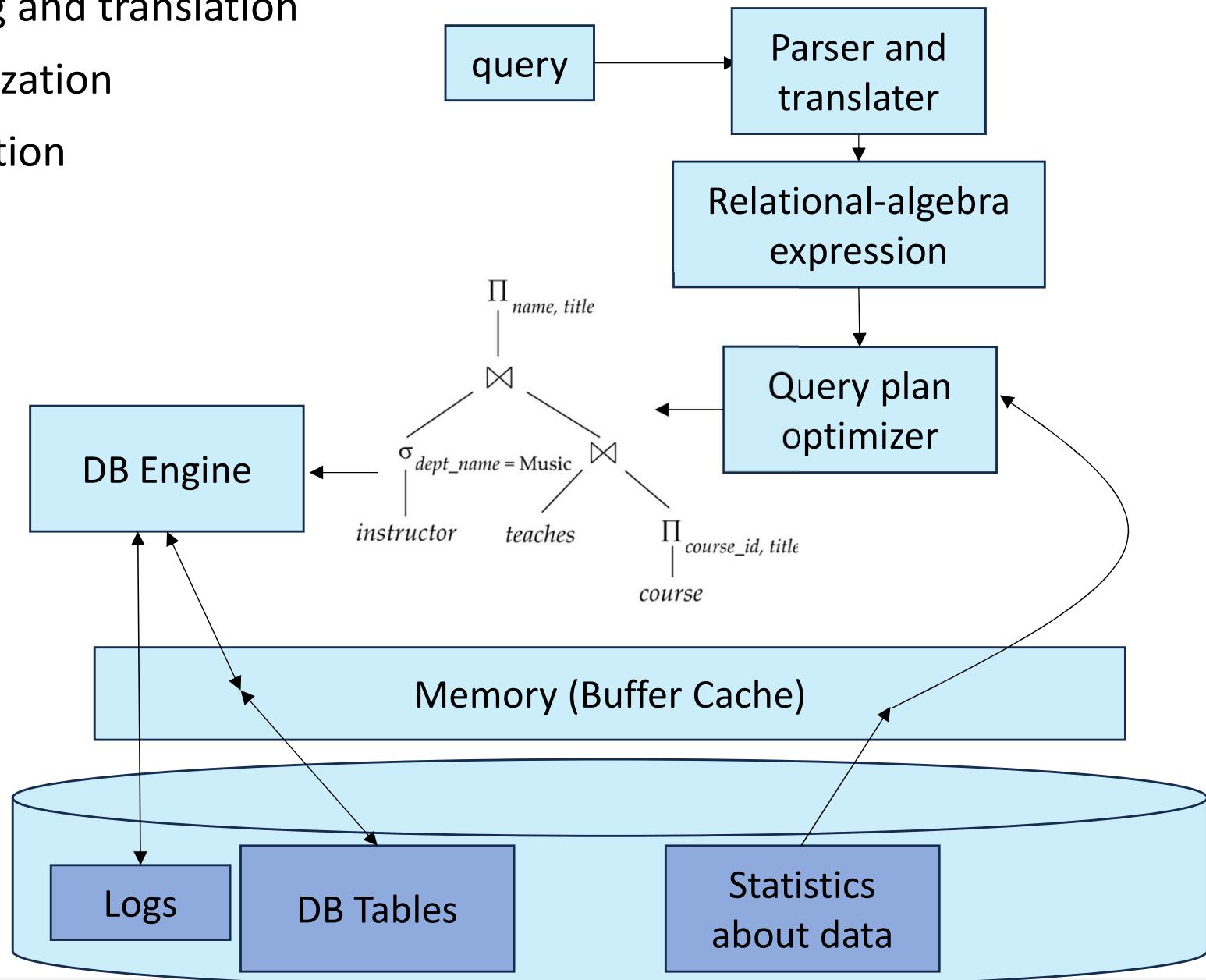


Shared
Nothing

Parallel Query Processing

Basic Steps in DB Query Processing

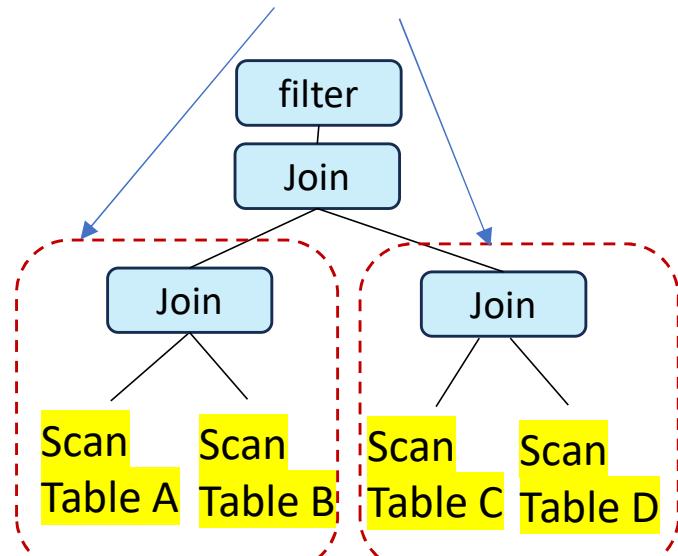
1. Parsing and translation
2. Optimization
3. Evaluation



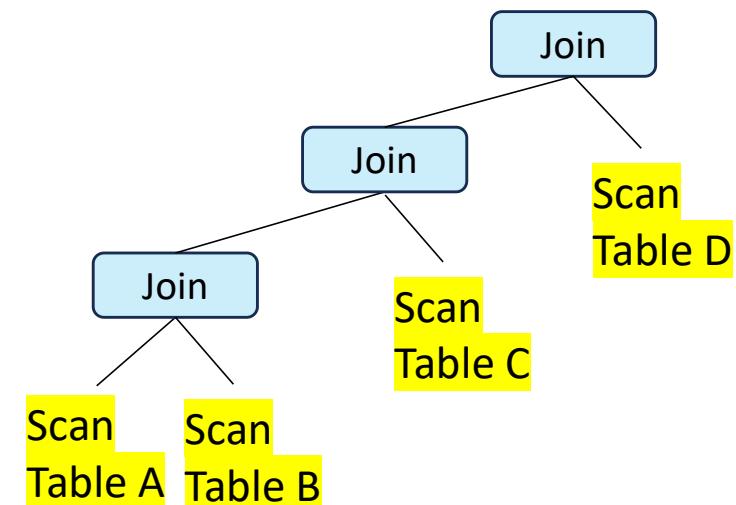
Query Planning

- Query execution plan is represented as a tree structure
 - Intermediate outputs generated by join are often very large
 - **Left Deep Plan** performs well in a single node
 - **Bush Plan** is known to perform better in distributed database systems

Can be executed concurrently



(a) Bush Plan

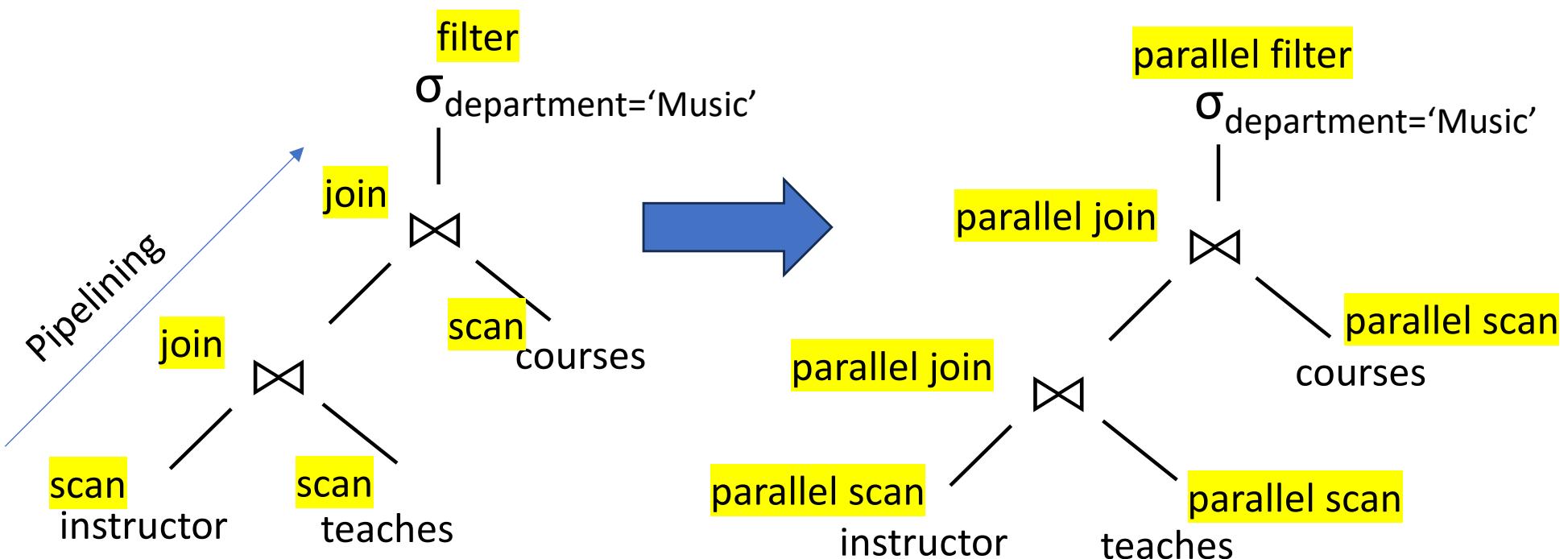


(b) Left Deep Plan

Query Planning in Parallel DBMS

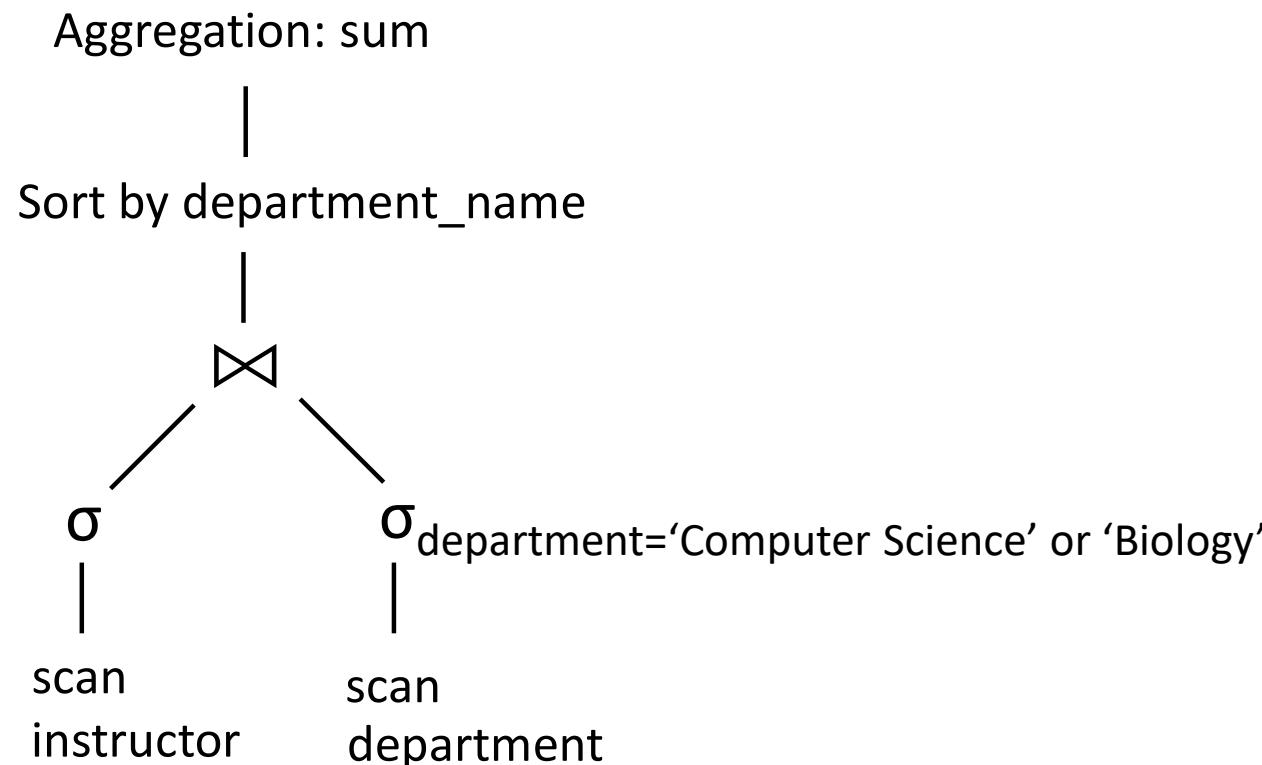
■ Parallel DBMS

- Multiple servers share the same storage and DB tables
- Query operators are executed in parallel



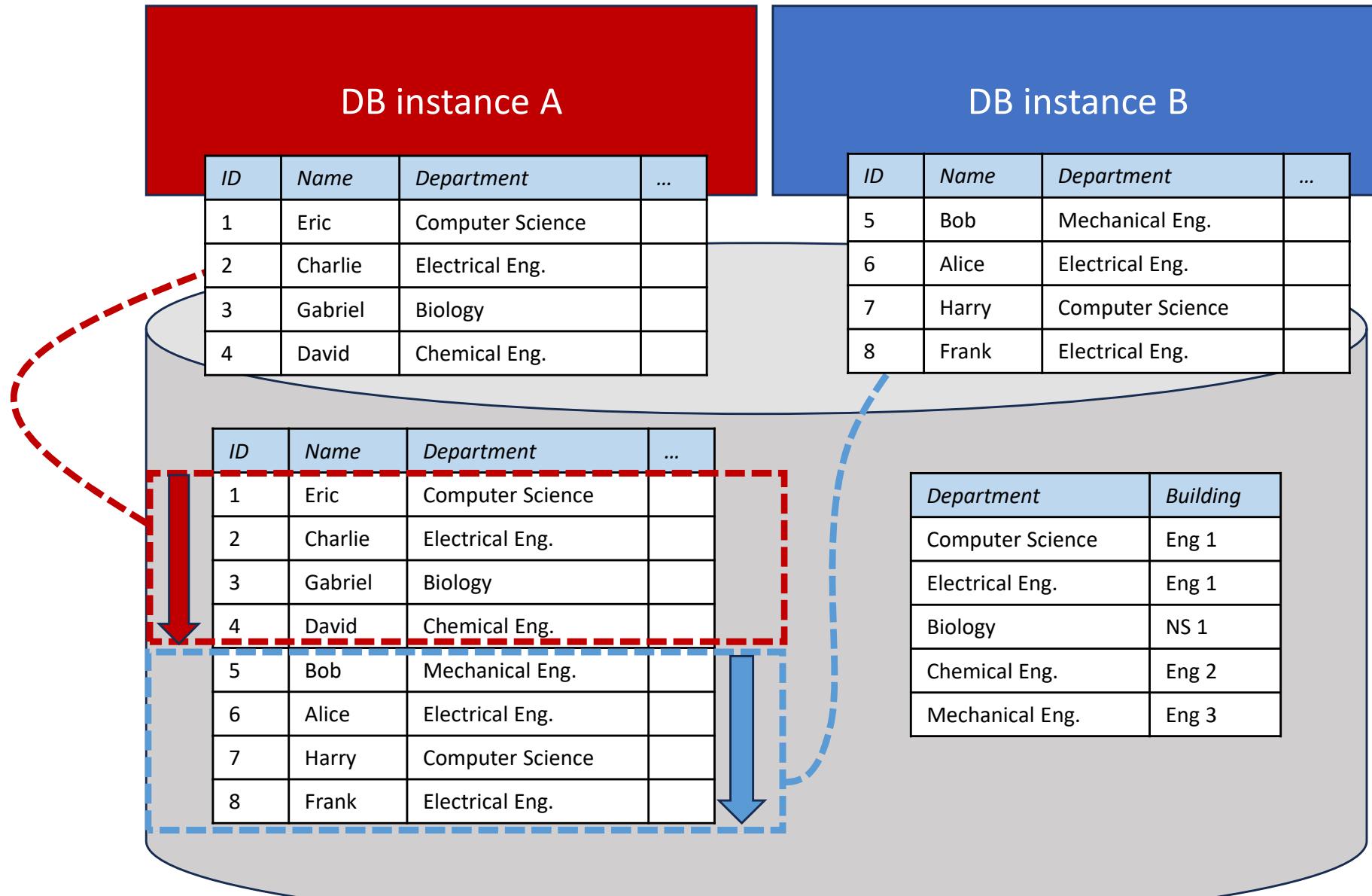
Parallel Query Processing

- ```
SELECT department_name, SUM(salary)
 FROM instructor JOIN department
 WHERE department_name IN ('Computer Science', 'Biology')
 GROUP BY department_name;
```



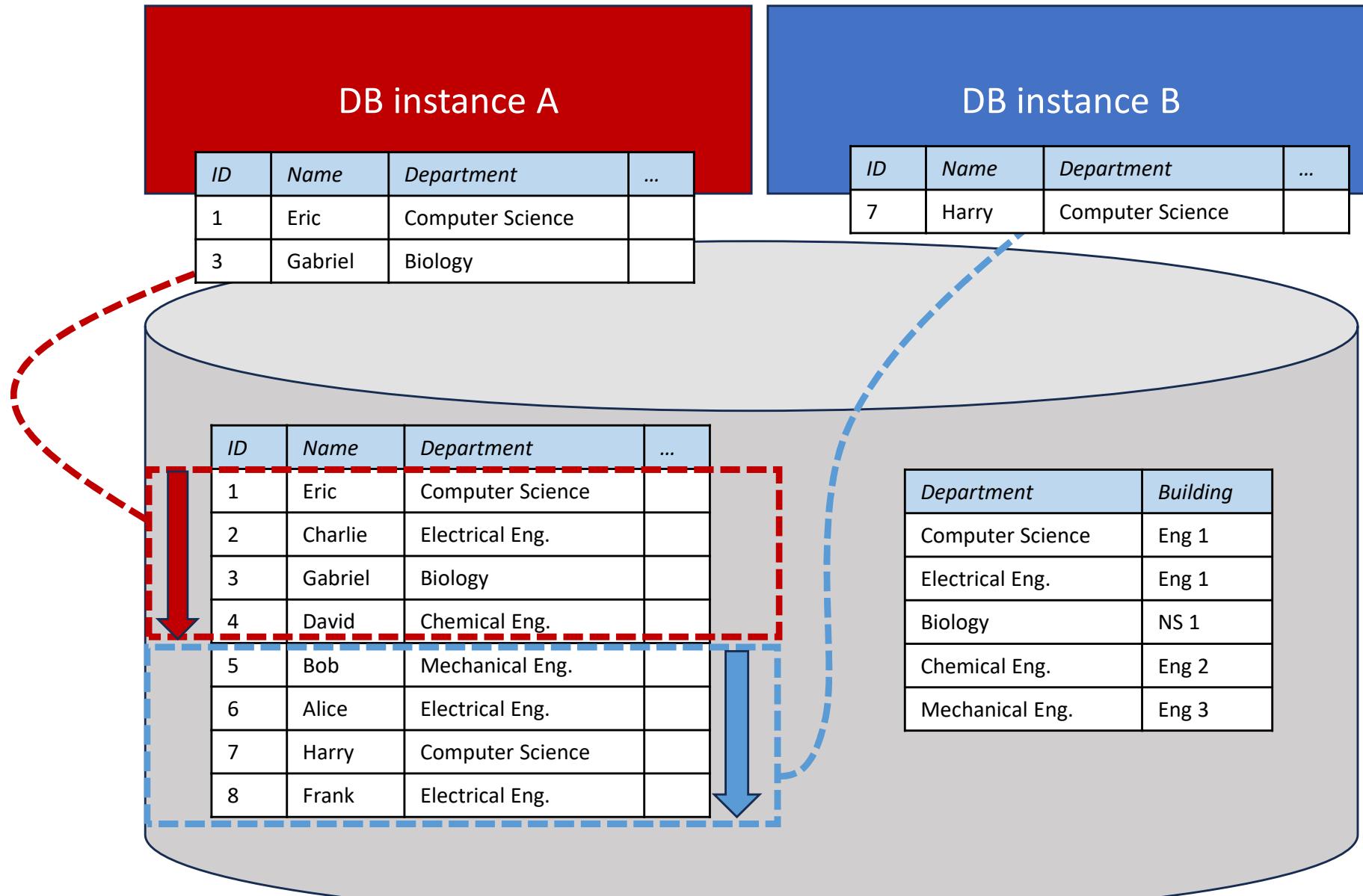
# Parallel Query Planning

## ■ Parallel Scan



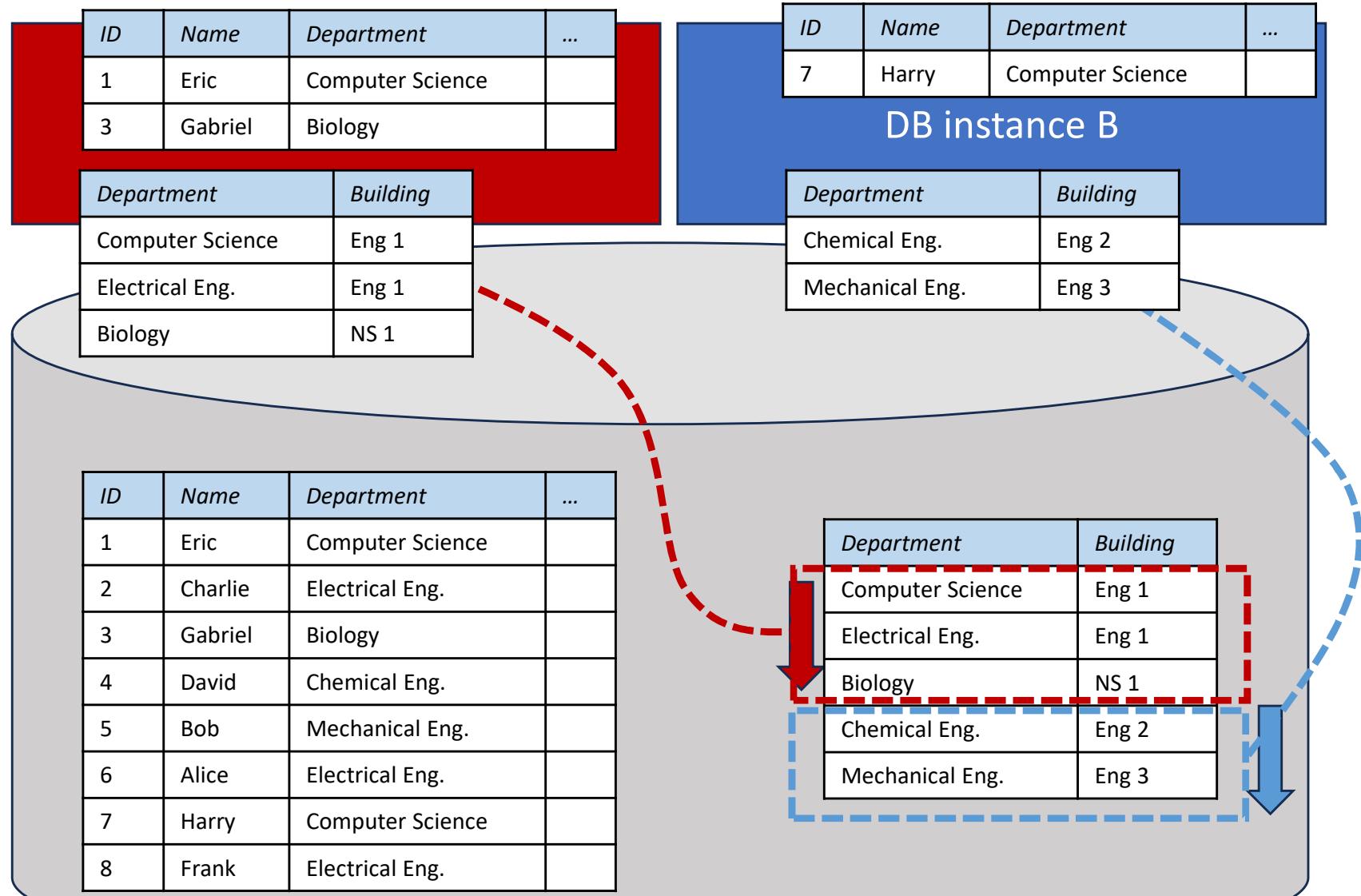
# Parallel Query Planning

## ■ Parallel Filter



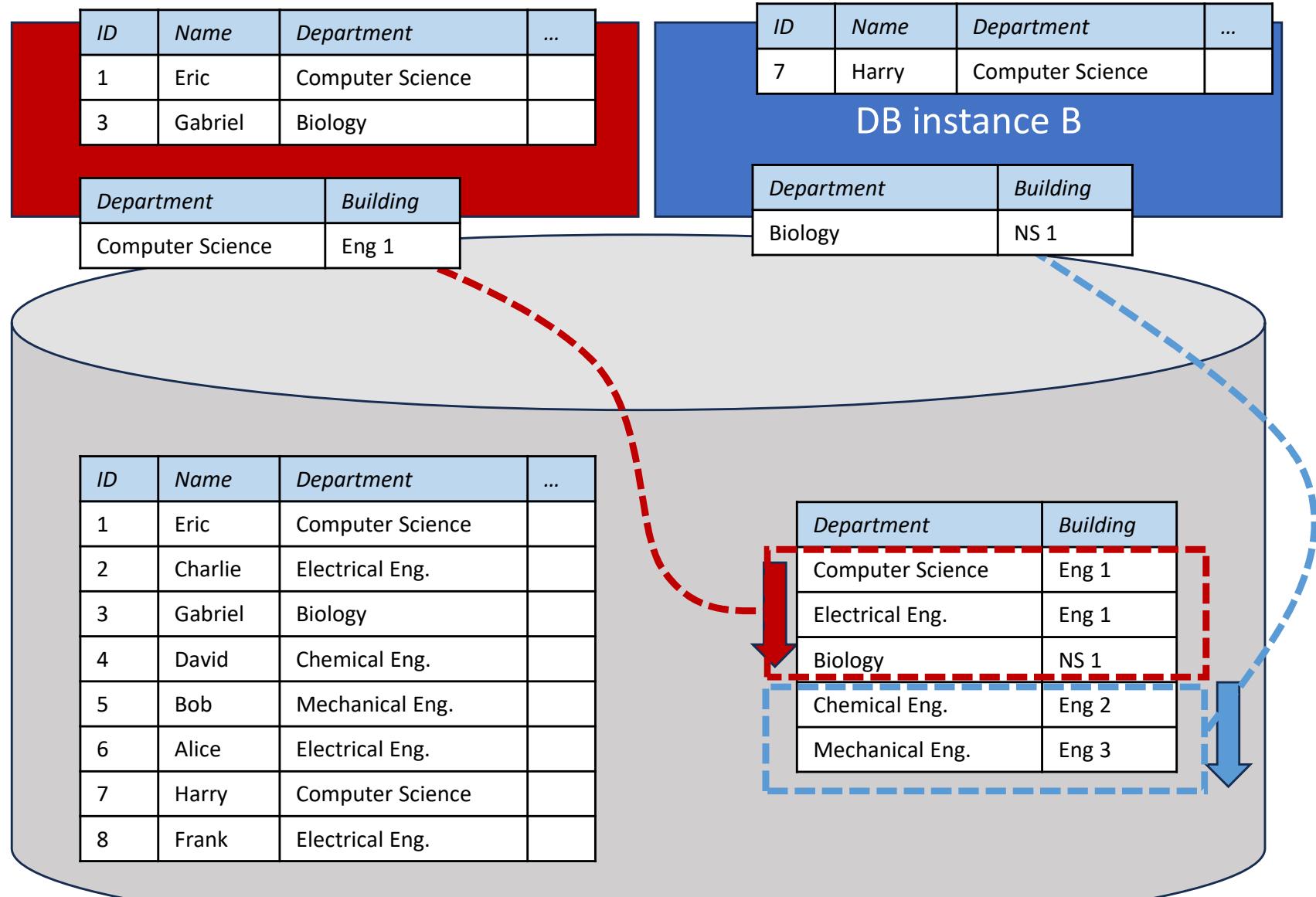
# Parallel Query Planning

## ■ Parallel Scan



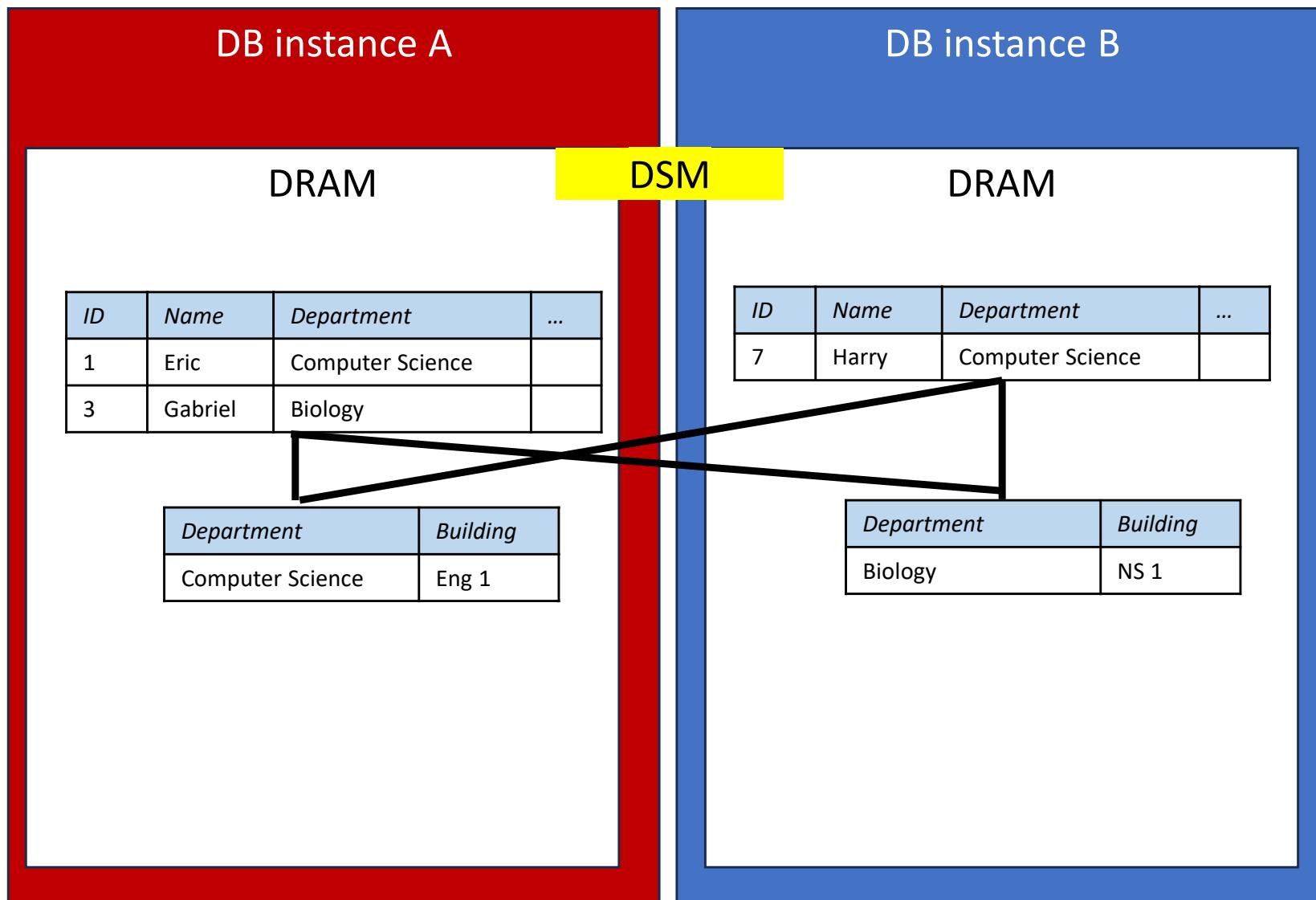
# Parallel Query Planning

## ■ Parallel Filter



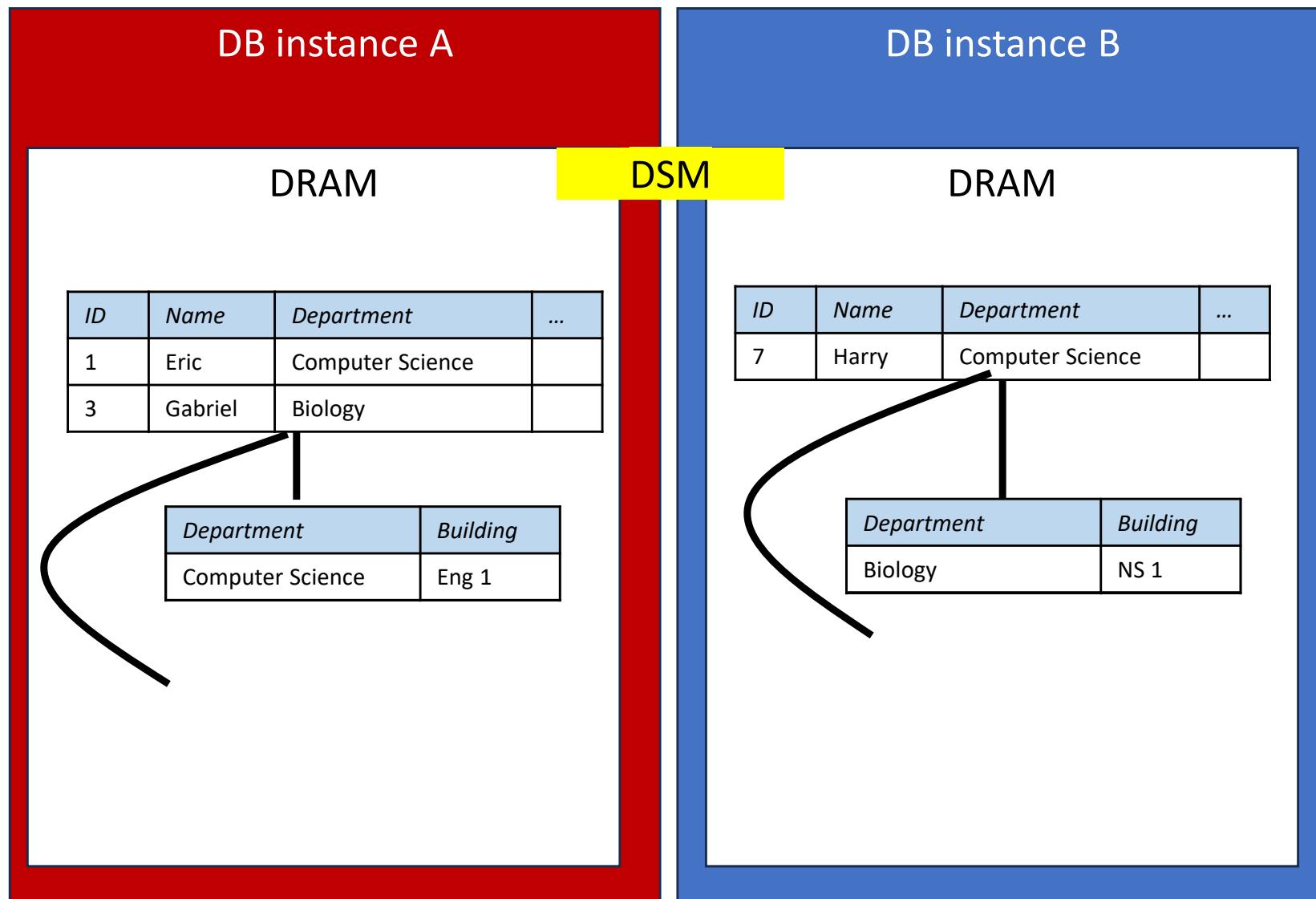
# Parallel Query Planning

## ■ Parallel Join



# Parallel Query Planning

## ■ Parallel Join



# Parallel Query Planning

- Access to shared cache requires ***Coherency Mechanism***

