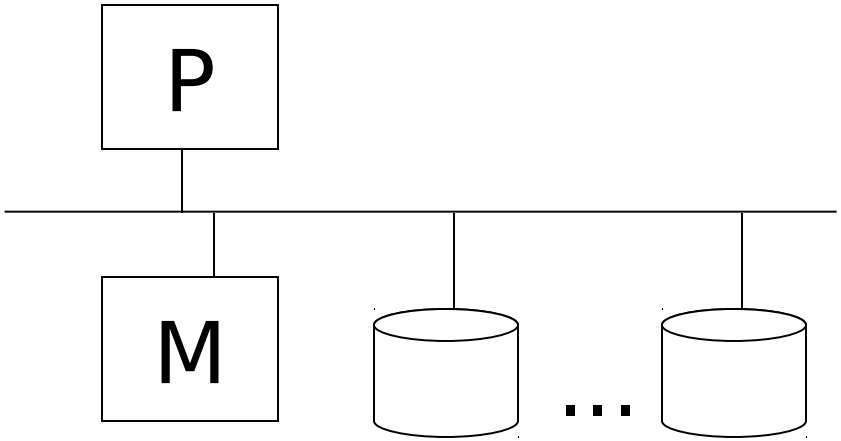
CS 347: Distributed Databases and Transaction Processing

Notes 01: Introduction

Hector Garcia-Molina

In CS245: Centralized DB system



Software:

Application
SQL Front End
Query Processor
Transaction Proc.
File Access

- Simplifications:
  - single front end
  - one place to keep locks
  - if processor fails, system fails, ...

In CS347

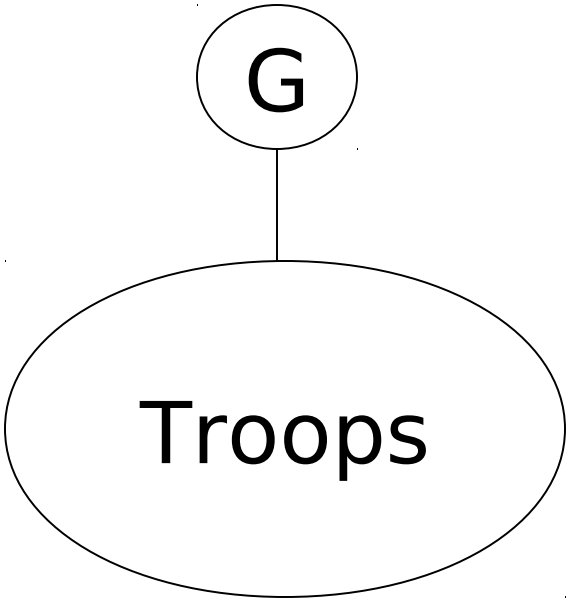
- Multiple processors (+ memories)
- Heterogeneity and autonomy of "components"

Multiple processors

- Opportunity for parallelism
- Opportunity for reliability
- Synchronization issues

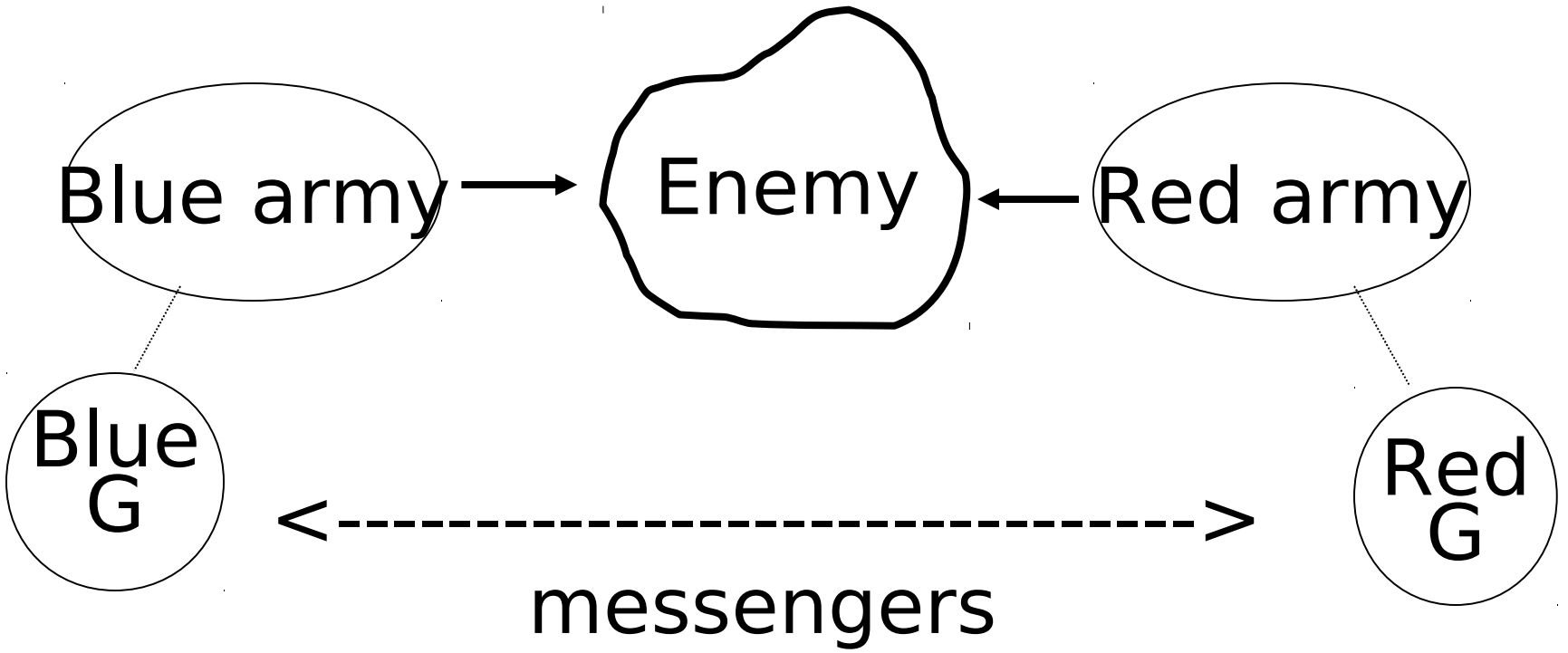
⇨ To illustrate synchronization problems:
Two Generals Problem

The <u>one</u> general problem (Trivial!)



➡ Battlefield

The two general problem:



messengers

Rules:

- Blue and red army must attack at same time
- Blue and red generals synchronize through messengers
- Messengers can be lost

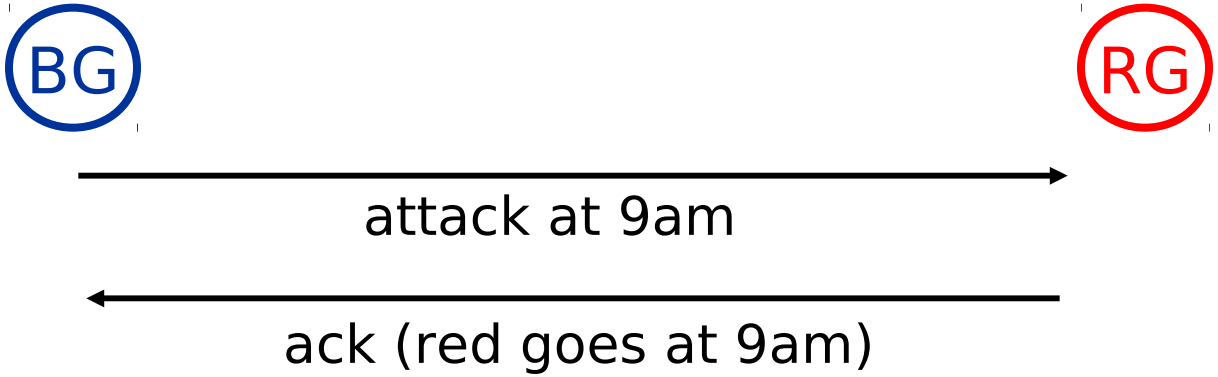
How Many Messages Do We Need?

assume blue starts...

Is this enough??

How Many Messages Do We Need?

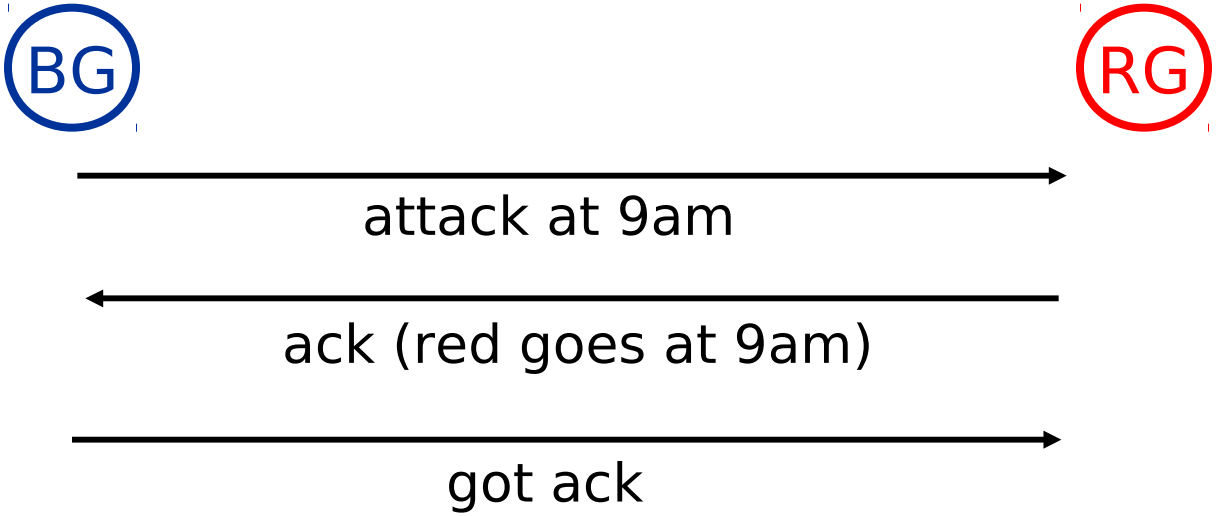
assume blue starts...



Is this enough??

How Many Messages Do We Need?

assume blue starts...



Is this enough??

Stated problem is Impossible!

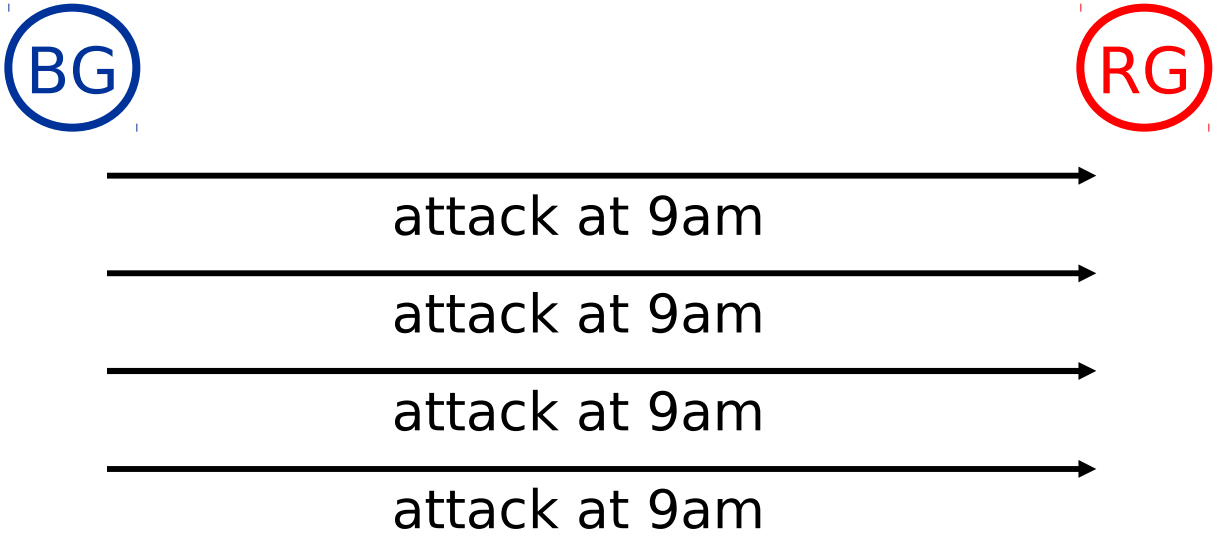
- **Theorem:** There is no protocol that uses a finite number of messages that solves the two-generals problem (as stated here)

Alternatives??

Probabilistic Approach?

- Send as many messages as possible, hope one gets through...

assume blue starts...

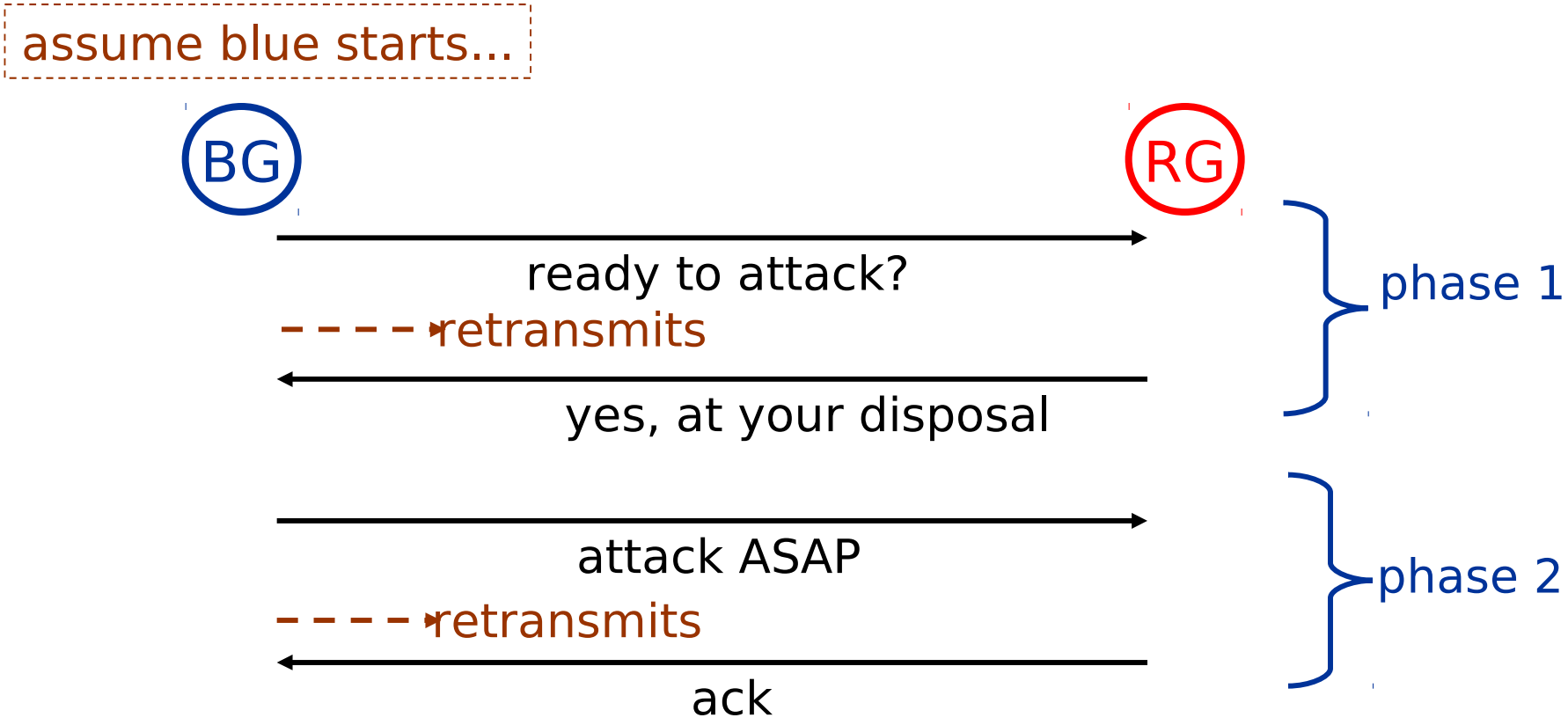


Eventual Commit

- Eventually both sides attack...

2-Phase Eventual Commit

- Eventually both sides attack...

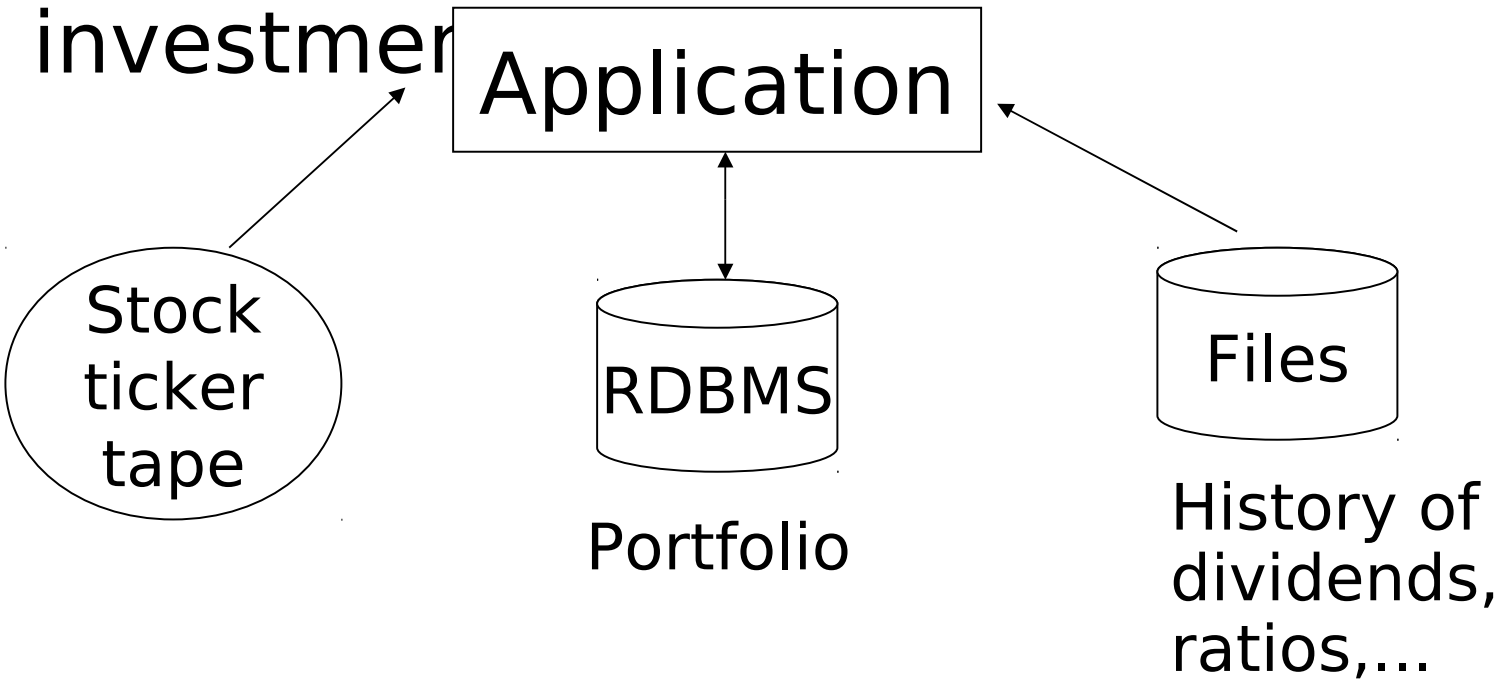


Commit Protocols

- Will study commit protocols like these...

Heterogeneity

Select new



Autonomy

Example: unable to get statistics for query optimization

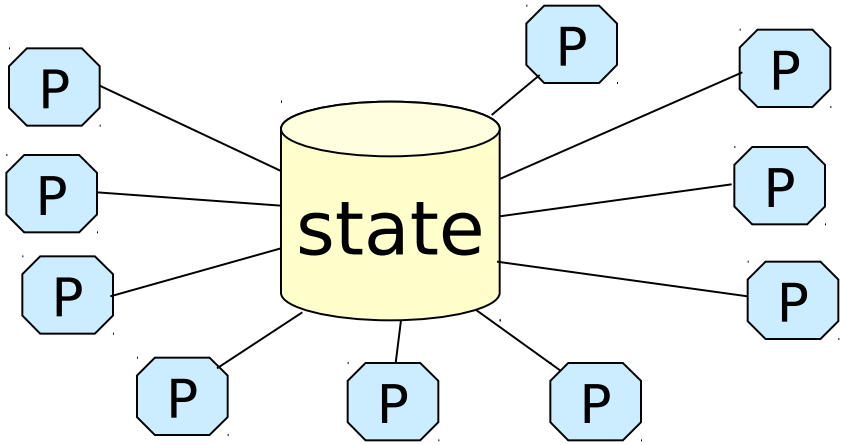
Example: blue general may have mind of his (or her) own!

- So, in CS347 we study data management
  with multiple processors and possible
  autonomy, heterogeneity
  - Impact on:
    - Data organization
    - Query processing
    - Access structures
    - Concurrency control
    - Recovery

- Renewed Interest in Distributed/Parallel Data Processing!
  - Massive web data, manage with many computers
  - How to crawl and search the web?
  - Peer-to-peer systems manage huge amounts of data
  - Data from many sources (e.g., comparison shopping): how to integrate?
  - Sensor Networks: data generated an many sensors/devices, need to analyze
  - Multi-player games (e.g., Second Life): tons of distributed data

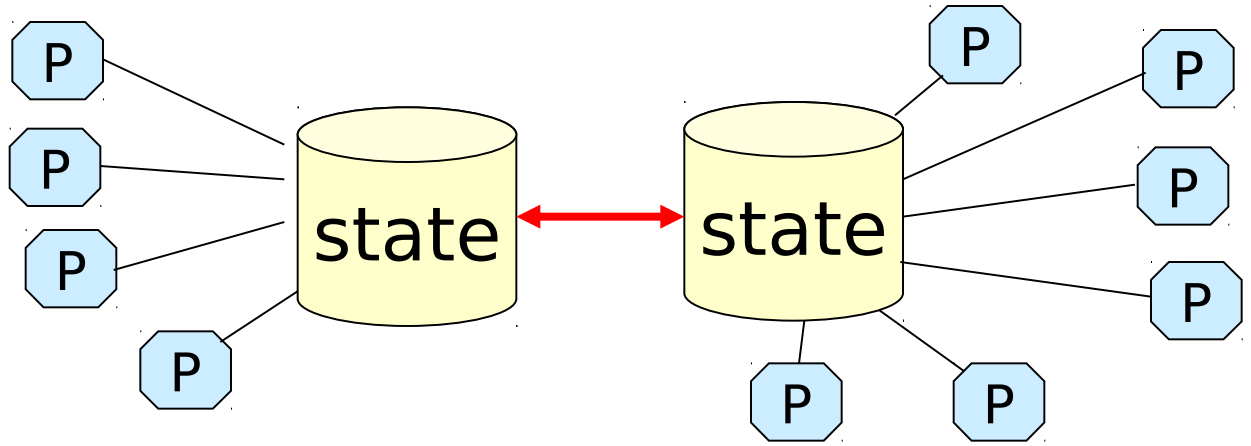
It's the ~~Economy~~ Data, Stupid!

- Example: Multi-player games



It's the ~~Economy~~ Data, Stupid!

- Example: Multi-player games



Logistics

- LECTURES: Mondays and Wednesdays 12:50pm to 2:05pm, Skilling 193
- INSTRUCTOR: Hector Garcia-Molina; Office: Gates Hall 434 Email: hector@cs.stanford.edu; Office Hours: Mondays, Wednesdays 11am to 12noon.
- ADDITIONAL INSTRUCTOR: Zoltan Gyongyi; Email: zoltan@cs.stanford.edu.
- TEACHING ASSISTANT: Petros Venetis; Office: Gates 424; Email: venetis@cs.stanford.edu; News Group: su.class.cs347; Office Hours: TBD
- SECRETARY: Marianne Siroker; Office: Gates Hall 436; Email: siroker@cs.stanford.edu; Phone: (650) 723-0872

Logistics

- TEXTBOOK: No required textbook. Some material for the lectures will be drawn from the following book:
  - M. Tamer Ozsu and Patrick Valduriez, "Principles of Distributed Database Systems," Second Edition, Prentice Hall 1999.
- CLASS WEB PAGE: http://www.stanford.edu/class/cs347 Will contain homework assignments, course news, etc. Be sure to check it periodically.
- ASSIGNMENTS: about 5 homeworks
- GRADING: Homeworks: 20%, Midterm 30%, Final: 50%.

Tentative Syllabus 2010 (Part I)

DATE	TOPIC
• Monday March 29	Introduction [N01]
• Wednesday March 31	Data Fragmentation [N02] Z
• Monday April 5	Query processing [N03] Z
• Wednesday April 7	Query processing [N03] Z
• Monday April 12	Query Optimization [N04] Z
• Wednesday April 14	Concurrency Control, Failures [N06]
• Monday April 19	Reliable Data Management [N07]
• Wednesday April 21	Reliable Data Management [N07]
• Monday April 26	Replicated Data Management [N08]
• Wednesday April 28	Midterm

Tentative Syllabus 2010 (Part II)

DATE TOPIC

- Monday May 3 Network Partitions [N09]
- Wednesday May 5 Peer to Peer Systems [N05]
- Monday May 10 Peer to Peer Systems [N05]
- Wednesday May 12 Map-Reduce Z
- Monday May 17 Distributed IR Z
- Wednesday May 19 Distributed Entity Resolution [new]
- Monday May 24 Time [N10]
- Wednesday May 26 Heterogeneous Systems, Source Capabilities [N11, 12, 13] Z
- Wednesday June 2 Extra Lecture (have one more this year!)

Concepts you should be familiar with:

- CS245: query plan, cost estimation, join algorithms, recovery, logging,...
- Interconnection networks

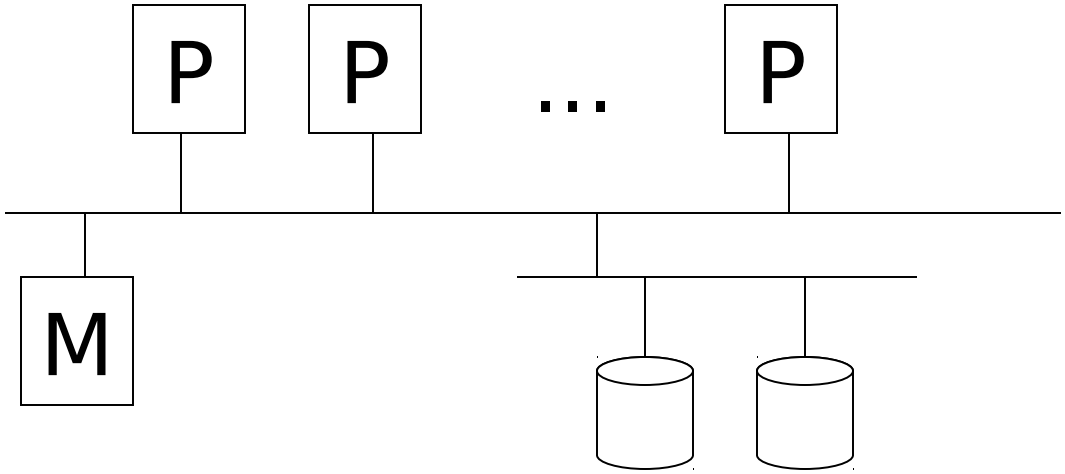
  (bus, mesh, hypercube,...)
- Computer networks (LAN, WAN,...)

Introductory topics

- Database architectures
- Client-server systems
- Distributed vs. parallel DB systems
- Cloud Computing

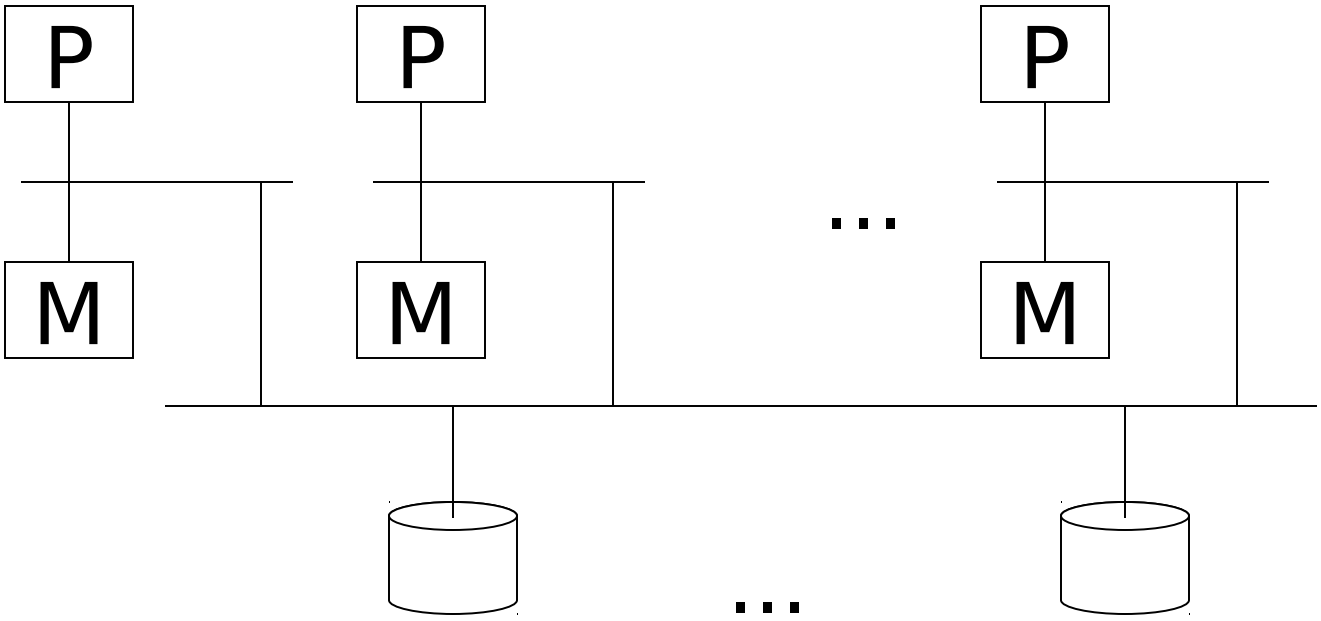
DB architectures

(1) Shared memory



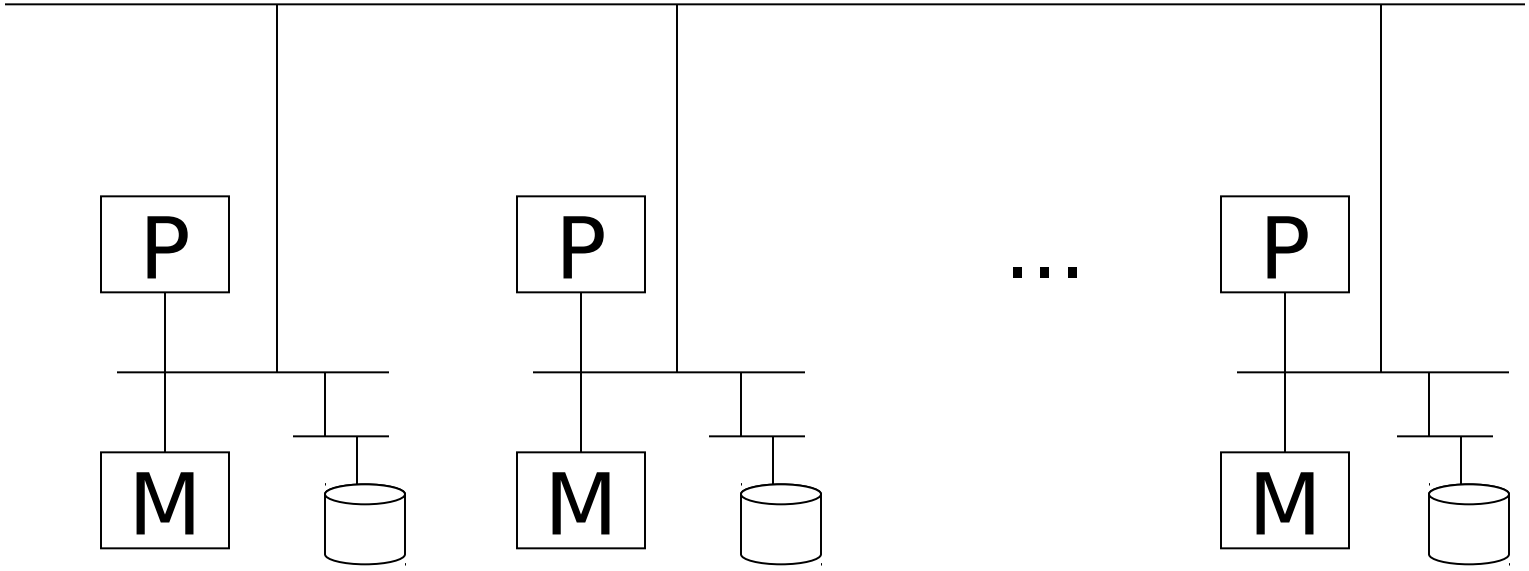
DB architectures

(2) Shared disk



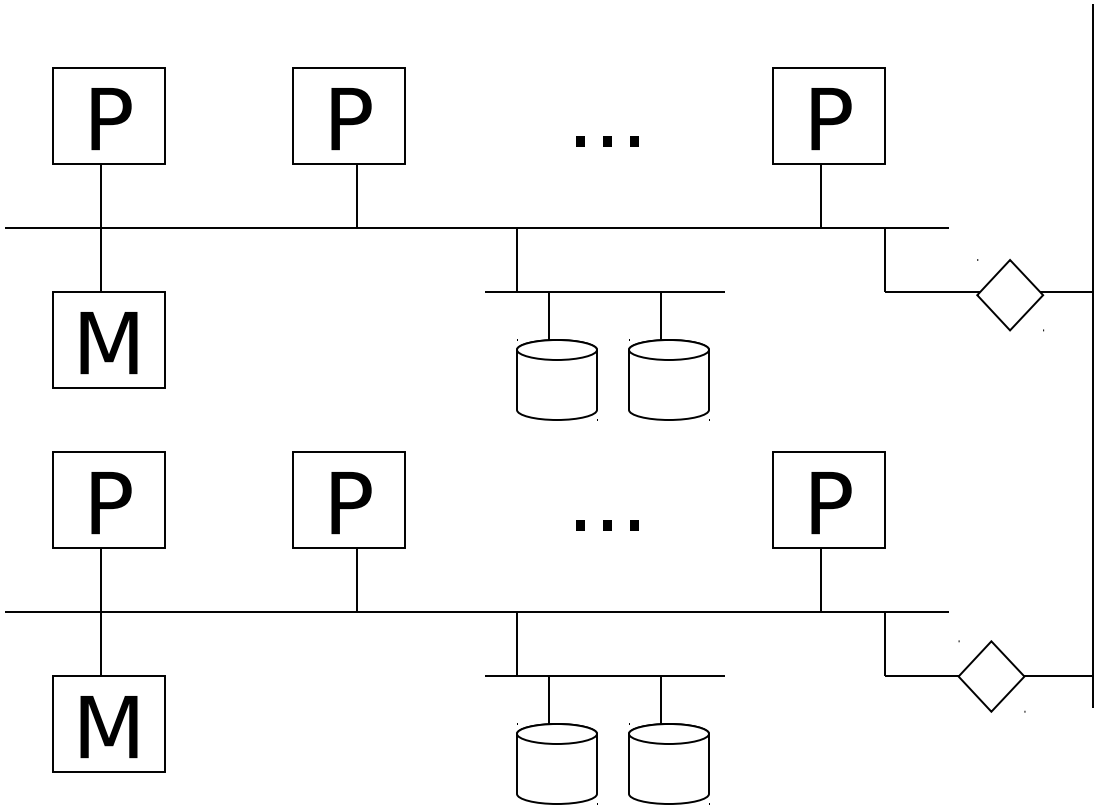
DB architectures

(3) Shared nothing



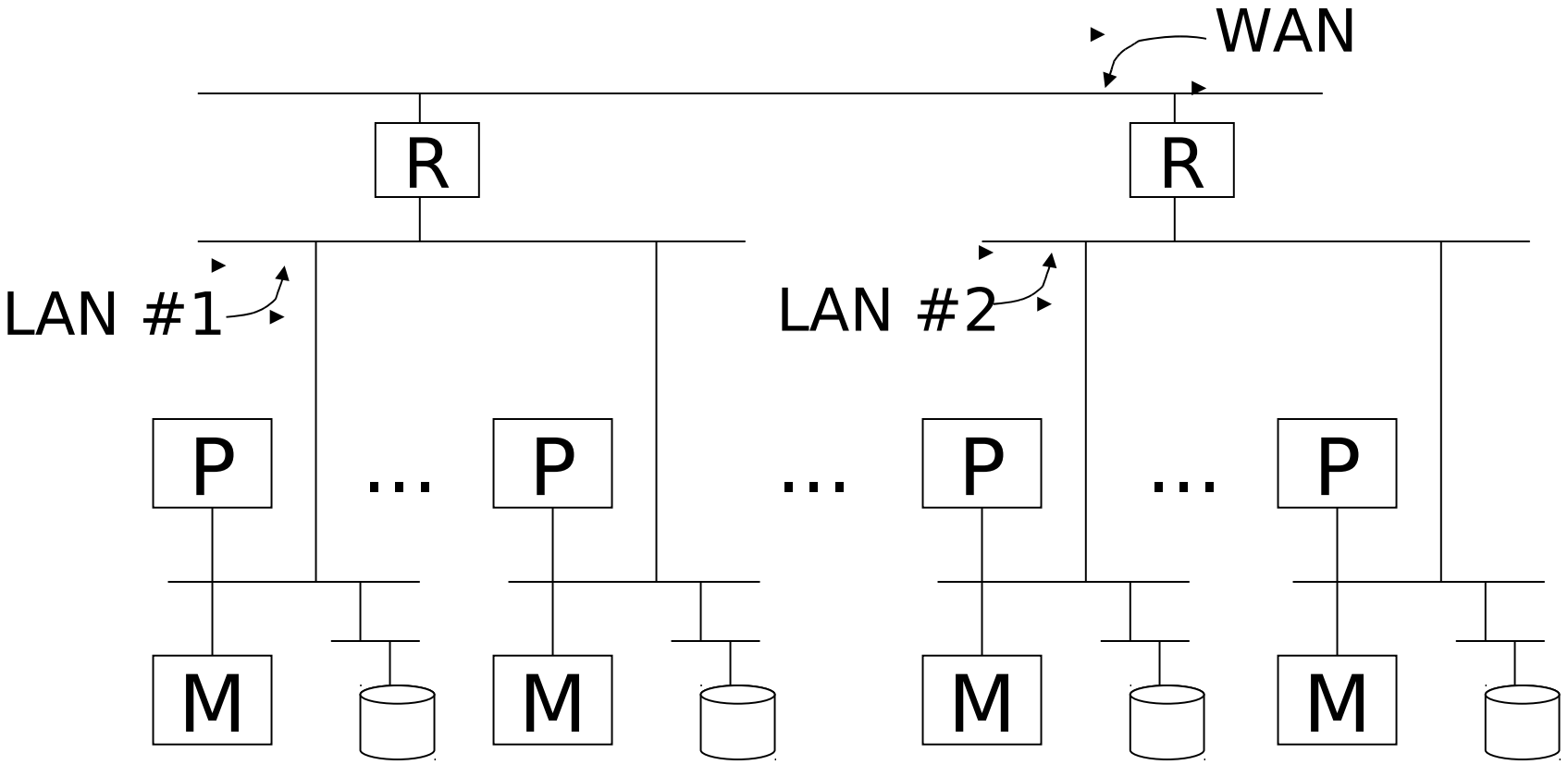
DB architectures

(4) Hybrid example



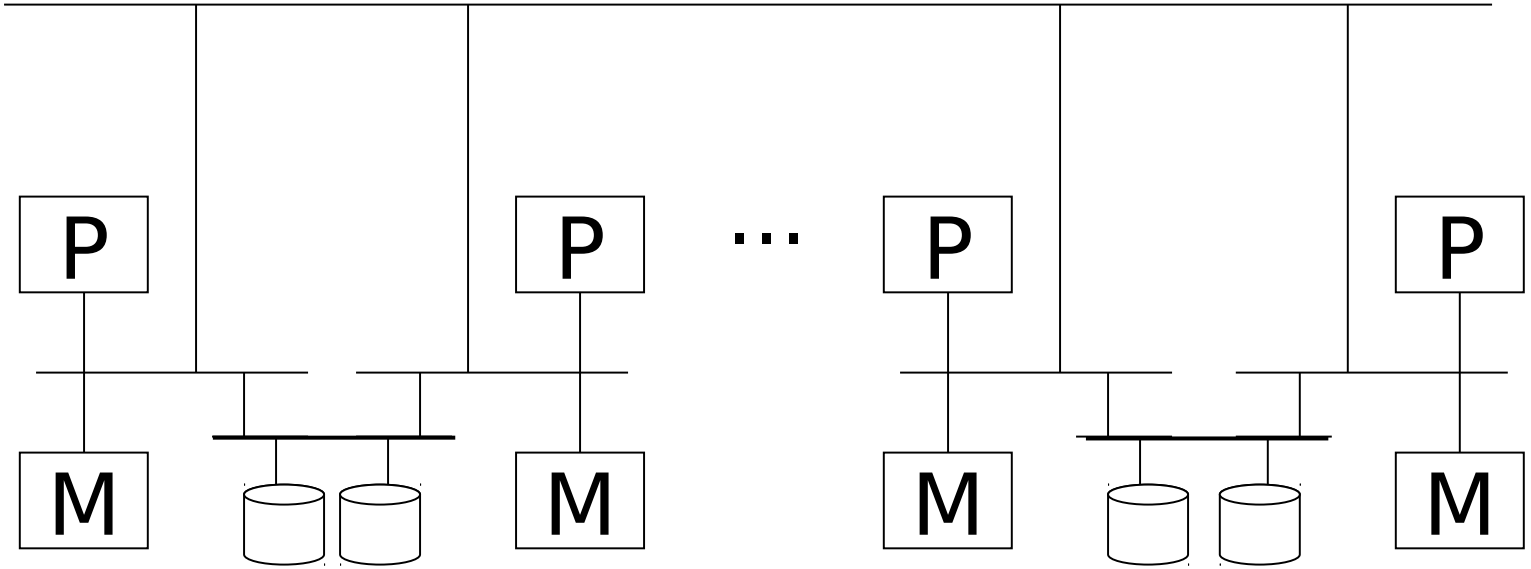
DB architectures

(4) Hybrid example 2



DB architectures

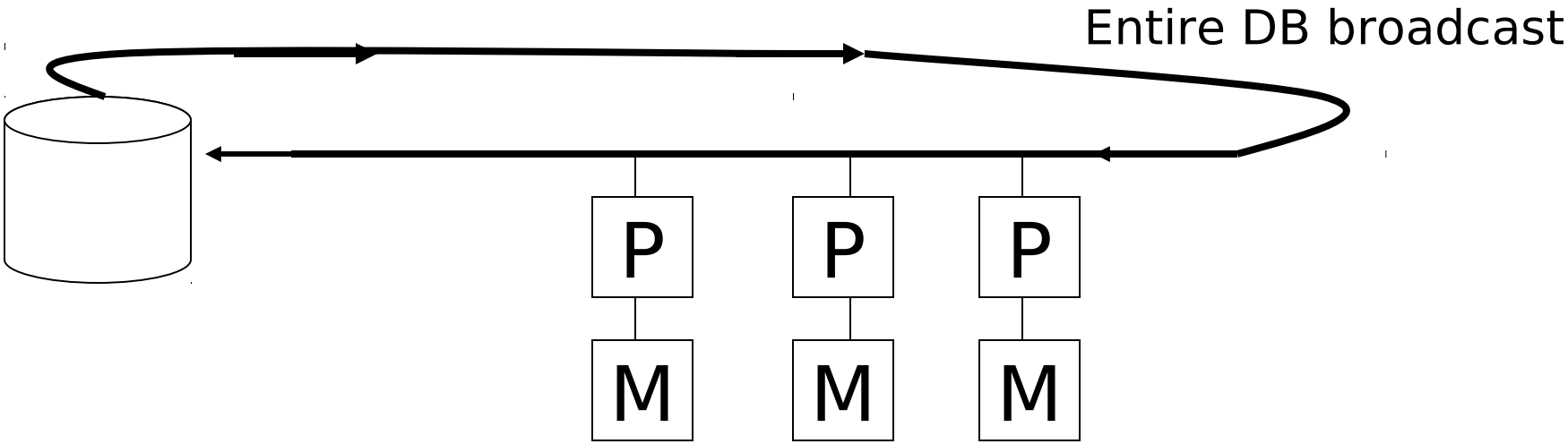
(4) Hybrid Tandem-like



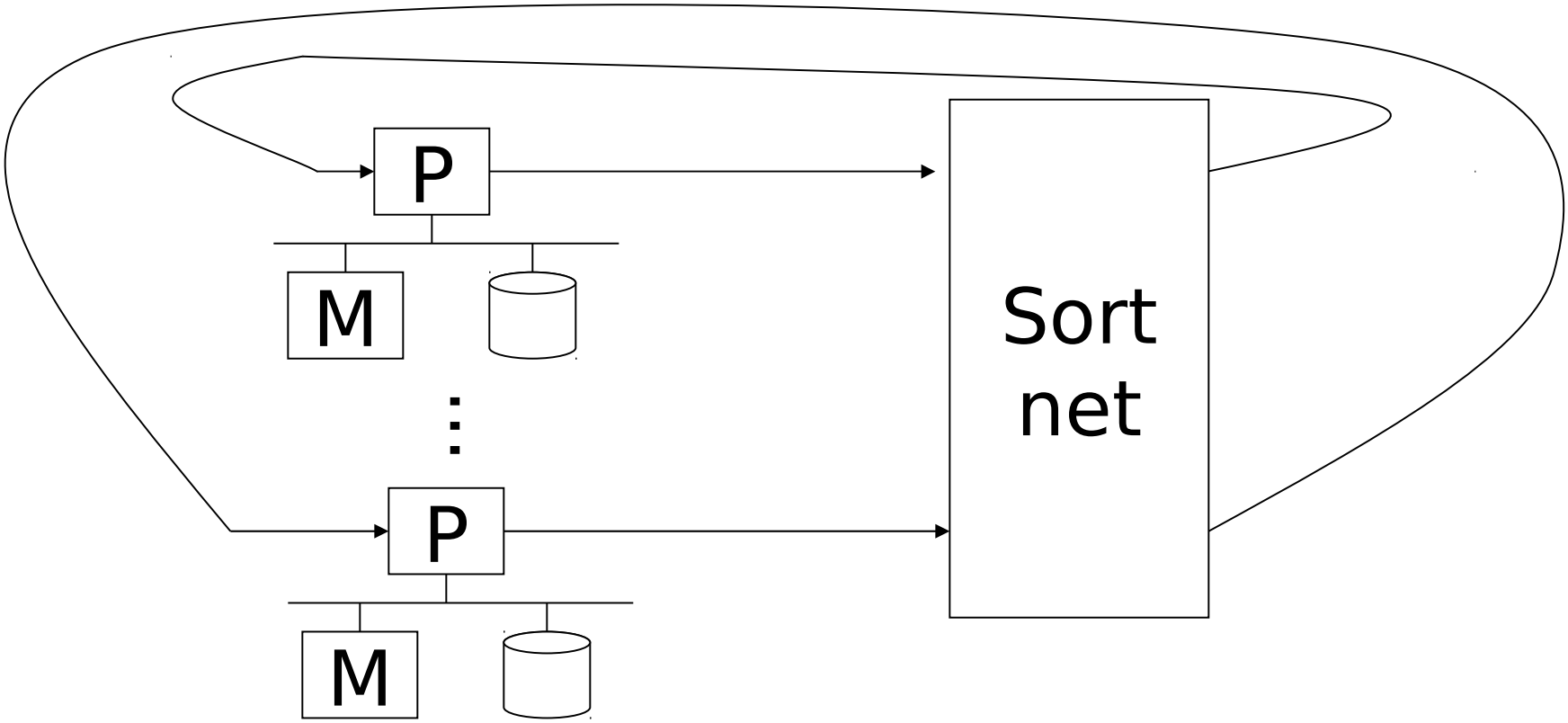
DB architectures

(5) Unusual?

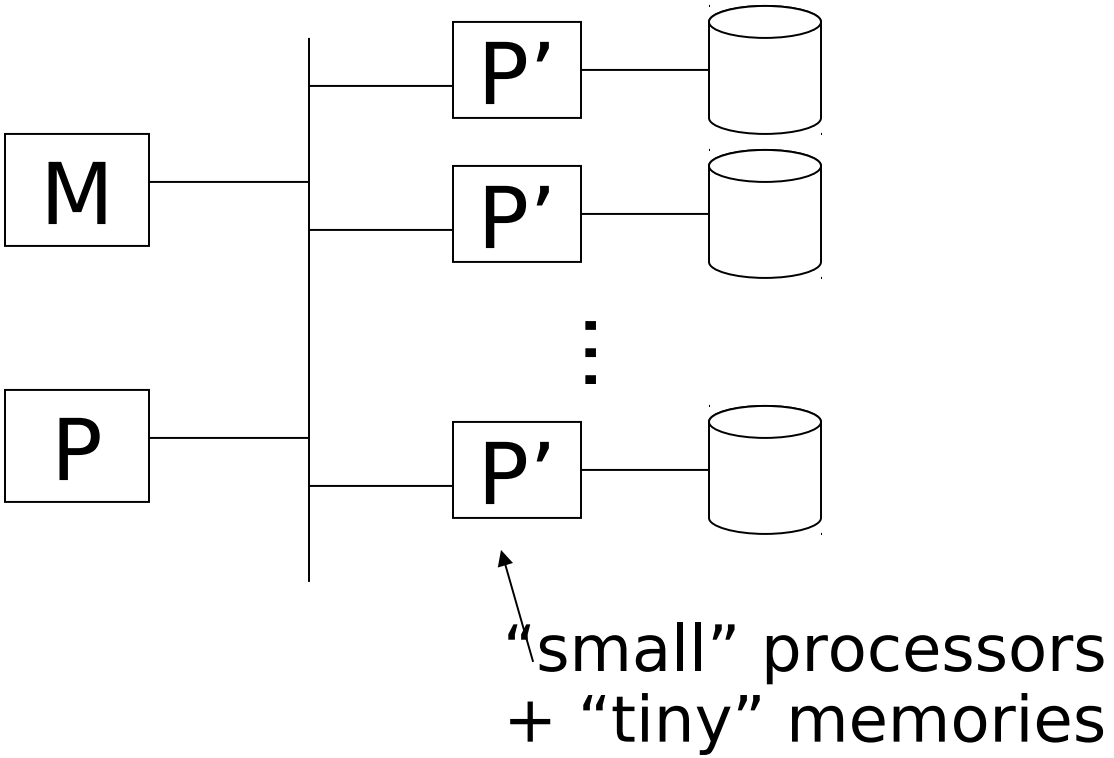
Datacycle (Broadcast disks)



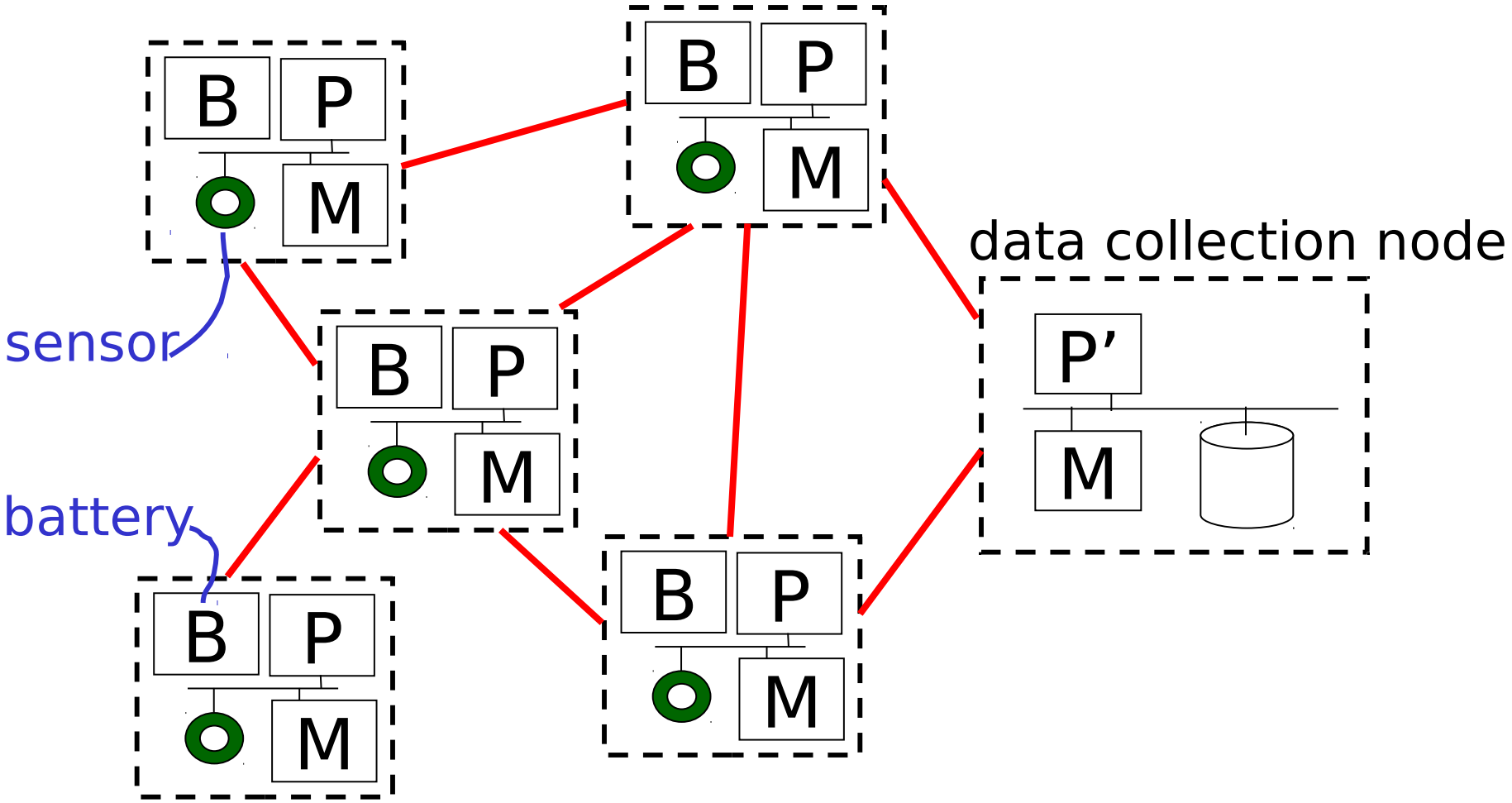
(5) Unusual Sorting network



(5) Unusual — processor per track or processor per disk



(6) Unusual — sensor networks

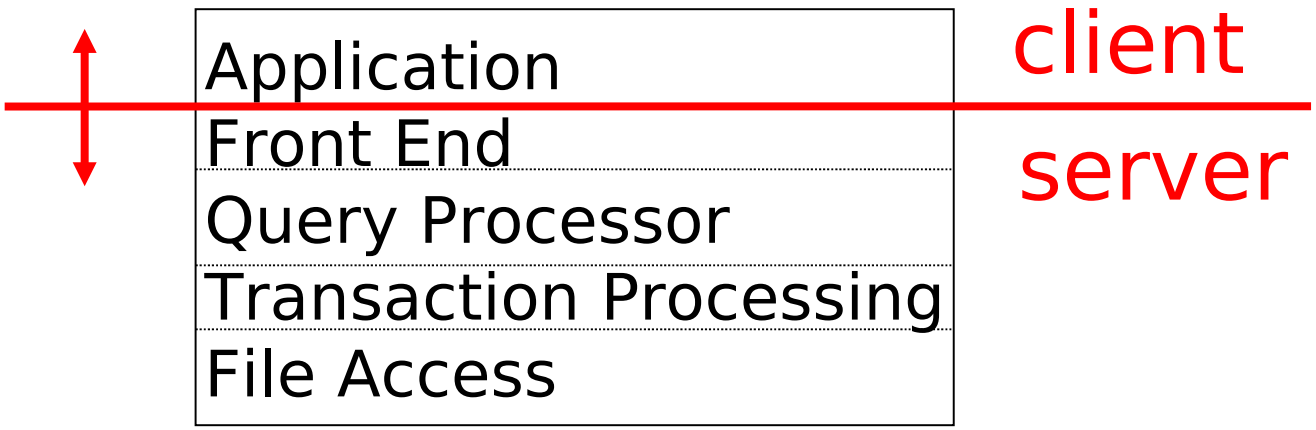


Issues for selecting architecture

- Reliability
- Scalability
- Geographic distribution of data
- Data “clusters”
- Performance
- Cost

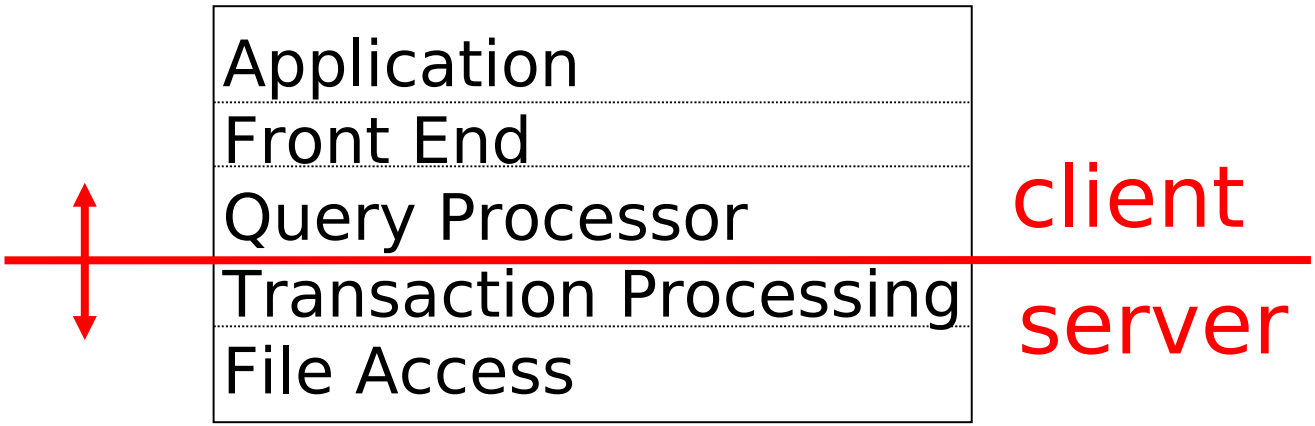
Client-Server Systems

(or how to partition software)



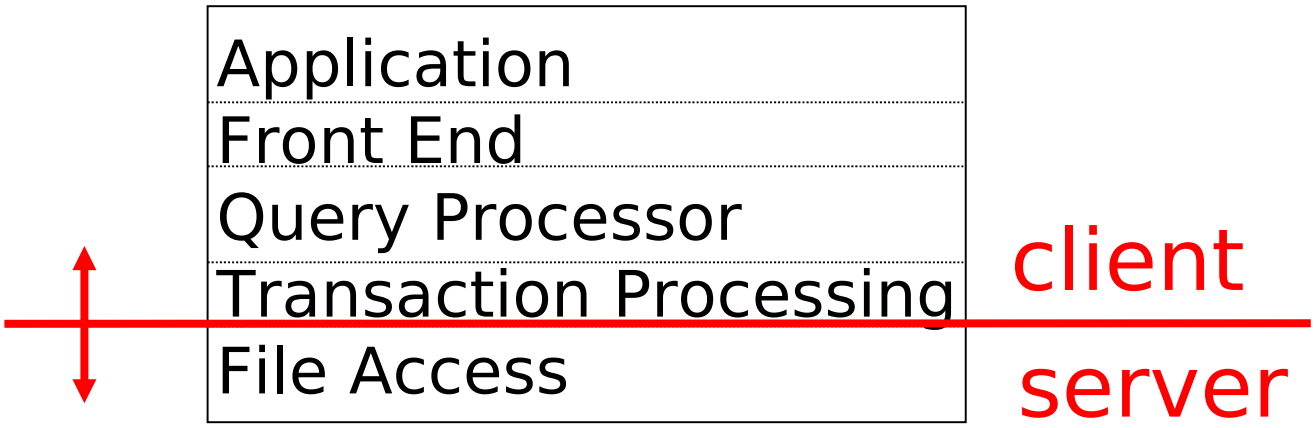
Client-Server Systems

(or how to partition software)



Client-Server Systems

(or how to partition software)



Transaction Servers

- Clients ship transactions consisting of 1 or more SQL commands

E.g., Open DataBase Connectivity (ODBC)

(standard API)

Data Servers

- Client requests pages or records
- Popular for OODB systems

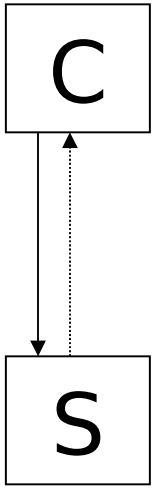
Issues

- Object granularity
- Where is data cached?
- Where is locking done?

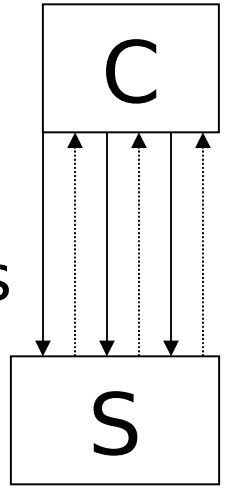
Basic Tradeoff

- Offloading work to clients
- Data transmitted

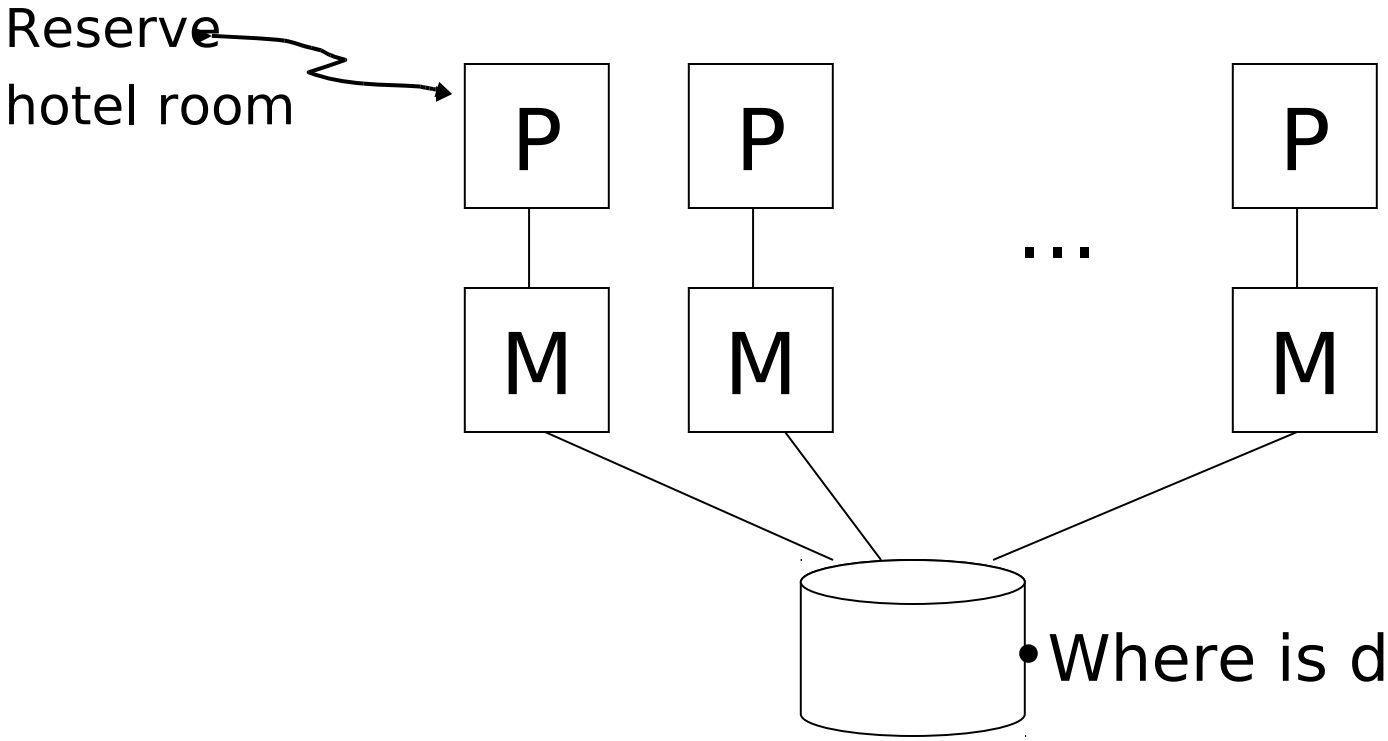
Reserve hotel room



Get pages



Note: Similar issues arise when we partition software/functionality within server



- Where is data cached?
- Where is locking done?

Parallel or distributed DB system?

- More similarities than differences!

- Typically, parallel DBs:
  - Fast interconnect
  - Homogeneous software
  - High performance is goal
  - Transparency is goal

- Typically, distributed DBs:
  - Geographically distributed
  - Data sharing is goal (may run into heterogeneity, autonomy)
  - Disconnected operation possible

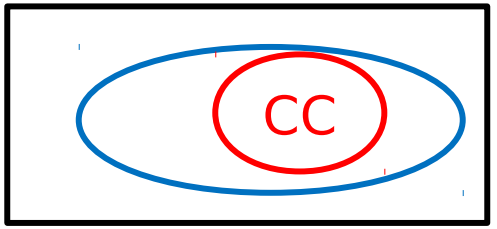
Cloud Computing

- Is CC just a marketing term??
  - utility (like power)
  - data or CPU cycles?
  - many processors, many storage units
  - business model

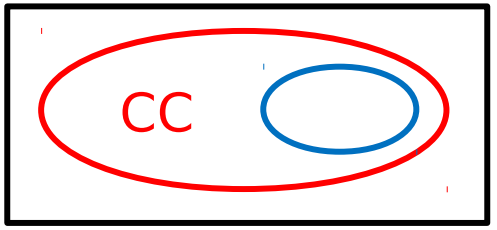
Is CC a subset, superset, disjoint from, or overlaps with:

- grid computing
- distributed computing
- Web 2.0
- Cluster Computing
- Peer-to-peer computing
- software as a service
- client-server computing
- data center as a computer
- massively parallel computing

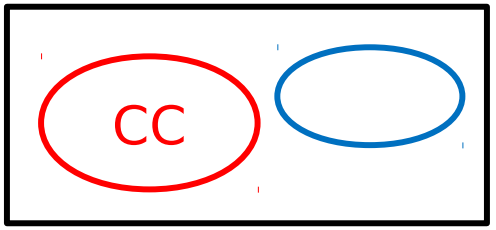
(A)



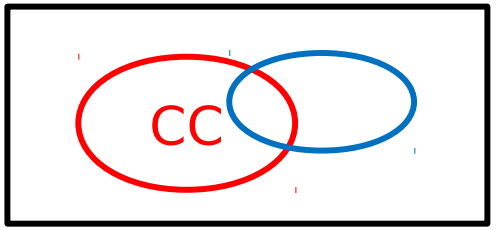
(B)



(C)



(D)



Clash of the Clouds (Economist April 4, 2009)

CC Issues

- Customer lock-in
- Privacy
- Standards
- Software licensing

Next

- How to describe distributed data
- Query processing in parallel DBs
- Query processing in distributed DBs

Query processing in parallel DBs:

- Typically: we can distribute/ partition/
  sort…. data to make certain DB operations (e.g., Join) fast

Query processing in distributed DBs:

- Typically: we are given data distribution; we need to find query processing strategy to minimize cost (e.g., communication cost)