

CS 347:
 Distributed Databases and
 Transaction Processing
Notes 01: Introduction
 Hector Garcia-Molina

CS 347
Lecture 1
1

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, ...

CS 347
Lecture 1
2

In CS347

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

CS 347
Lecture 1
3

Multiple processors

- Opportunity for parallelism
- Opportunity for reliability
- Synchronization issues

⇒ To illustrate synchronization problems:
 Two Generals Problem

CS 347
Lecture 1
4

The one general problem (Trivial!)

CS 347
Lecture 1
5

The two general problem:

CS 347
Lecture 1
6

Rules:

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

CS 347 Lecture 1 7

How Many Messages Do We Need?

assume blue starts...

```

    graph LR
      BG((BG)) -- "attack at 9am" --> RG((RG))
  
```

Is this enough??

CS 347 Lecture 1 8

How Many Messages Do We Need?

assume blue starts...

```

    graph LR
      BG((BG)) -- "attack at 9am" --> RG((RG))
      RG -- "ack (red goes at 9am)" --> BG
  
```

Is this enough??

CS 347 Lecture 1 9

How Many Messages Do We Need?

assume blue starts...

```

    graph LR
      BG((BG)) -- "attack at 9am" --> RG((RG))
      RG -- "ack (red goes at 9am)" --> BG
      BG -- "got ack" --> RG
  
```

Is this enough??

CS 347 Lecture 1 10

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??

CS 347 Lecture 1 11

Probabilistic Approach?

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

assume blue starts...

```

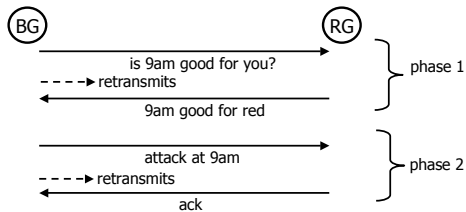
    graph LR
      BG((BG)) -- "attack at 9am" --> RG((RG))
      BG -- "attack at 9am" --> RG
      BG -- "attack at 9am" --> RG
      BG -- "attack at 9am" --> RG
  
```

CS 347 Lecture 1 12

Eventual Commit

- Eventually both sides attack...

assume blue starts...



CS 347

Lecture 1

13

Commit Protocols

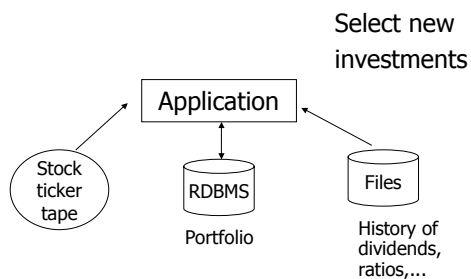
- Will study commit protocols like these...

CS 347

Lecture 1

14

Heterogeneity



CS 347

Lecture 1

15

Autonomy

Example: unable to get statistics for query optimization

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

CS 347

Lecture 1

16

- 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

CS 347

Lecture 1

17

- Renewed Interest in Distributed/Parallel Data Processing!

- Web data is massive, need to manage with many computers
- Peer-to-peer systems manage huge amounts of data
- How to crawl and search the web?
- Data from many sources (e.g., comparison shopping): how to integrate?
- Sensor Networks: data generated an many sensors/devices, need to analyze

CS 347

Lecture 1

18

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; Office: Gates 424; Email: zoltan@cs.stanford.edu.
- TEACHING ASSISTANT: David Menestrina; Office: Gates 432; Email: cs347-spr0708-staff@lists.stanford.edu; Office Hours: Fridays 2-4pm.
- SECRETARY: Marianne Siroker; Office: Gates Hall 436; Email: siroker@cs.stanford.edu; Phone: (650) 723-0872

CS 347

Lecture 1

19

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%.

CS 347

Lecture 1

20

Tentative Syllabus (Part I)

<u>DATE</u>	<u>TOPIC</u>
• Wednesday April 2	Introduction [N01]
• Monday April 7	Data Fragmentation (shorten?) [N02] Z
• Wednesday April 9	Query processing [N03] Z
• Monday April 14	Query processing [N03] Z
• Wednesday April 16	Query Processing (shorten?) [N04] Z
• Monday April 21	Peer to Peer Systems [N05]
• Wednesday April 23	Peer to Peer Systems [N05]
• Monday April 28	Map-Reduce [new]
• Wednesday April 30	Distributed IR [new] Z

CS 347

Lecture 1

21

Tentative Syllabus (Part II)

<u>DATE</u>	<u>TOPIC</u>
• Monday May 5	Midterm
• Wednesday May 7	Concurrency Control, Failures [N06]
• Monday May 12	Reliable Data Management [N07]
• Wednesday May 14	Reliable Data Management [N07] Rec
• Monday May 19	Replicated Data Management [N08]
• Wednesday May 21	Network Partitions [N09]
• Wednesday May 28	Distributed Entity Resolution [new]
• Monday June 2	Time [N10]
• Wednesday June 4	Heterogeneous Systems, Source Capabilities [N11, 12, 13] Z

CS 347

Lecture 1

22

Concepts you should be familiar with:

- CS245: query plan, cost estimation, join algorithms, recovery, logging,...
- Interconnection networks (bus, mesh, hypercube,...)
- Computer networks (LAN, WAN,...)

CS 347

Lecture 1

23

Introductory topics

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

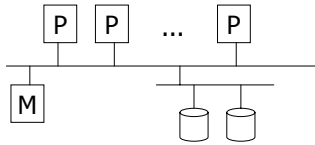
CS 347

Lecture 1

24

DB architectures

(1) Shared memory



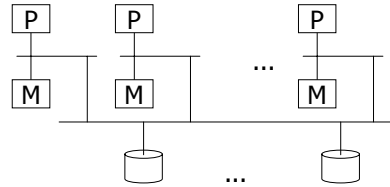
CS 347

Lecture 1

25

DB architectures

(2) Shared disk



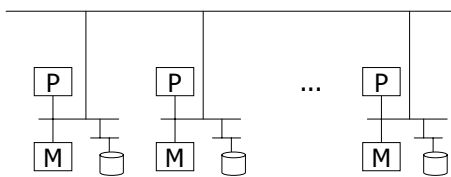
CS 347

Lecture 1

26

DB architectures

(3) Shared nothing



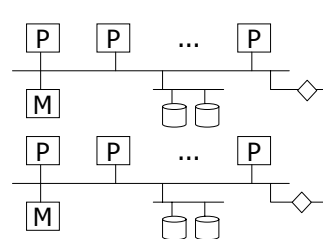
CS 347

Lecture 1

27

DB architectures

(4) Hybrid example



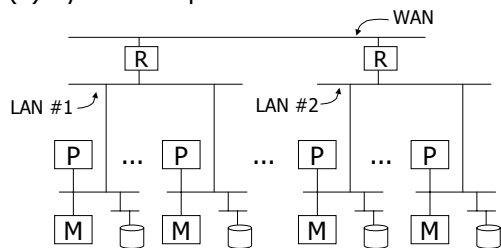
CS 347

Lecture 1

28

DB architectures

(4) Hybrid example 2



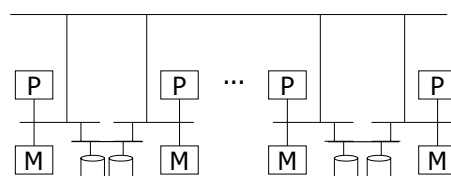
CS 347

Lecture 1

29

DB architectures

(4) Hybrid Tandem-like



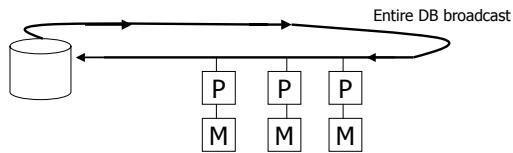
CS 347

Lecture 1

30

DB architectures

(5) Unusual?
Datacycle (Broadcast disks)

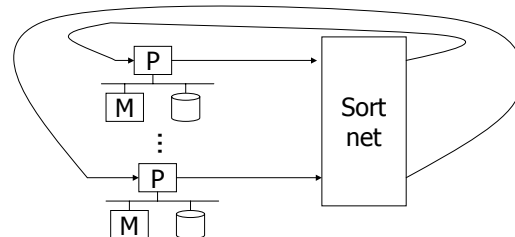


CS 347

Lecture 1

31

(5) Unusual Sorting network

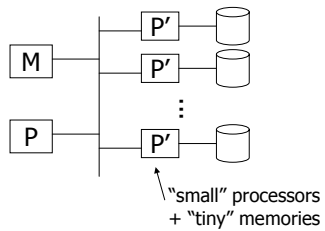


CS 347

Lecture 1

32

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

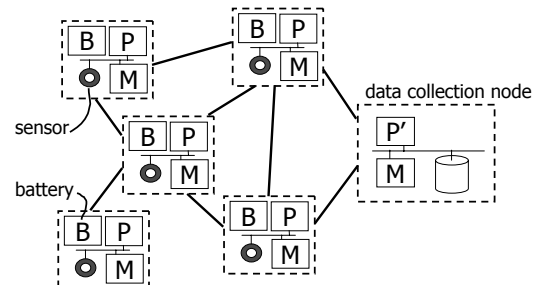


CS 347

Lecture 1

33

(6) Unusual — sensor networks



CS 347

Lecture 1

34

Issues for selecting architecture

- Reliability
- Scalability
- Geographic distribution of data
- Data "clusters"
- Performance
- Cost

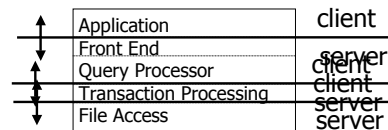
CS 347

Lecture 1

35

Client-Server Systems

(or how to partition software)



CS 347

Lecture 1

36

Transaction Servers

- Clients ship transactions consisting of 1 or more SQL commands

E.g., Open DataBase Connectivity (ODBC)
(standard API)

CS 347

Lecture 1

37

Data Servers

- Client requests pages or records
- Popular for OODB systems

CS 347

Lecture 1

38

Issues

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

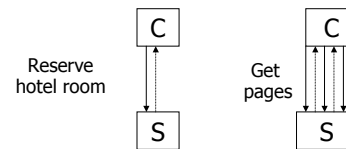
CS 347

Lecture 1

39

Basic Tradeoff

- Offloading work to clients
- Data transmitted

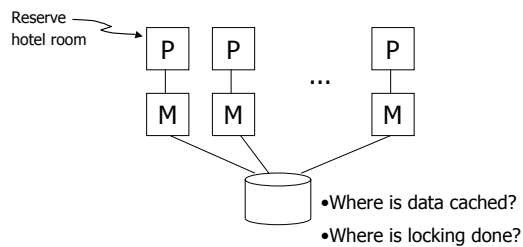


CS 347

Lecture 1

40

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



CS 347

Lecture 1

41

Parallel or distributed DB system?

- More similarities than differences!

CS 347

Lecture 1

42

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

CS 347

Lecture 1

43

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

CS 347

Lecture 1

44

Next

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

CS 347

Lecture 1

45

Query processing in parallel DBs:

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

CS 347

Lecture 1

46

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)

CS 347

Lecture 1

47