Overview of distributed query processing

Dario Della Monica

These slides are a modified version of the slides provided with the book

Özsu and Valduriez, *Principles of Distributed Database Systems* (3rd Ed.), 2011

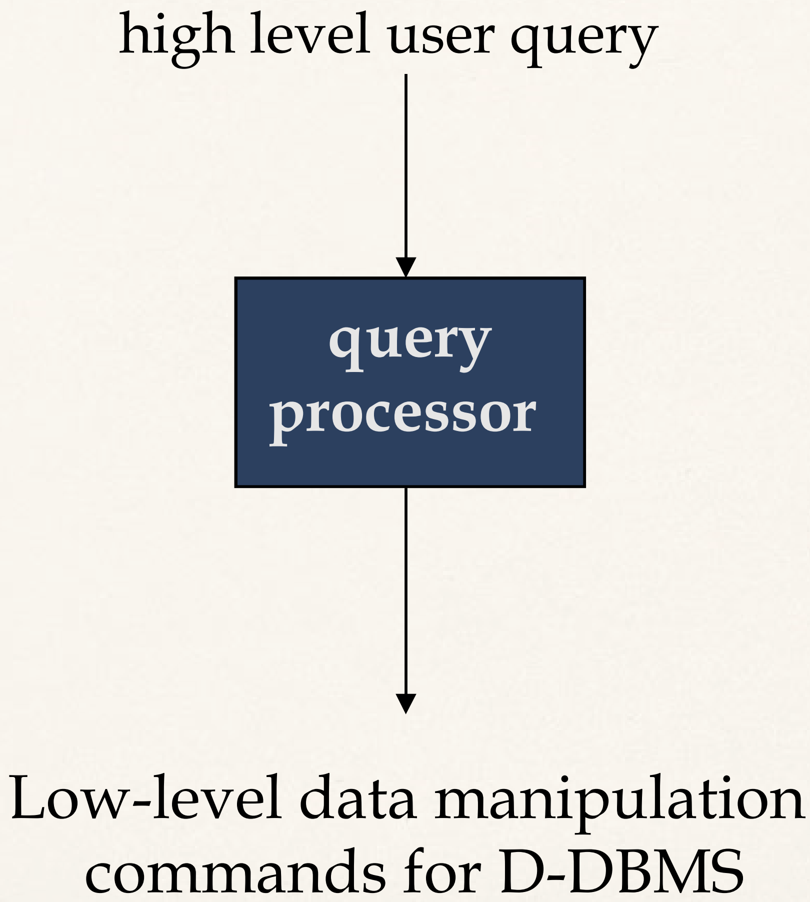
The original version of the slides is available at: extras.springer.com

Outline (distributed DB)

- Introduction (Ch. 1) *
- Distributed Database Design (Ch. 3) *
- Distributed Query Processing (Ch. 6-8) *
  - → **Overview (Ch. 6)** *
  - → Query decomposition and data localization (Ch. 7) *
  - → Distributed query optimization (Ch. 8) *
- Distributed Transaction Management (Ch. 10-12) *

* Özsu and Valduriez, *Principles of Distributed Database Systems* (3rd Ed.), 2011

Query Processing in a D-DBMS



Selecting Alternatives

```
SELECT    *
FROM      EMP,ASG
WHERE     EMP.ENO = ASG.ENO
AND       RESP = "Manager"
```

$$EMP \bowtie_{ENO} (\sigma_{RESP=\text{“Manager”}} (ASG))$$

$$\sigma_{RESP=\text{“Manager”}} (EMP \bowtie_{ENO} (ASG))$$

What are the Additional Problems?

- More parameters
  - Fragmentation
  - Replication
  - Data exchange alternatives/multiple sites
- To transform a **global query on relations** of a distributed DB (seen as a single DB by the user) into **local queries on fragments** stored on several local DB's (**data localization**)
- QEP must include information on communications (data transfers among sites) and on which sites operations are performed
- Use of semijoins to reduce the amount of data transferred among sites
  - Focus of the optimizer is selecting optimal order for join and semijoin operations
- Centralized vs. distributed optimization
- Cost to minimize
  - Centralized DB: CPU and I/O cost only (actually, only I/O)
  - Distributed DB: also communication costs
  - Communication costs are the dominating ones (even though this might not be the case with increased network speed, especially within Local Area Network)

What are the Additional Problems? – Example

- Global query: EMP $\bowtie_{ENO}$ ($\sigma_{RESP=\text{“Manager”}}$ (ASG))
- Fragmentation and allocation

$ASG_1 = \sigma_{ENO \leq \text{“E3”}}(ASG)$ (site 1)
$ASG_2 = \sigma_{ENO > \text{“E3”}}(ASG)$ (site 2)
$EMP_1 = \sigma_{ENO \leq \text{“E3”}}(EMP)$ (site 3)
$EMP_2 = \sigma_{ENO > \text{“E3”}}(EMP)$ (site 4)
Query result (site 5)

What are the Additional Problems? – Example

- Global query: $EMP \bowtie_{ENO} (\sigma_{RESP=\text{“Manager”}} (ASG))$
- Fragmentation and allocation

$ASG_1 = \sigma_{ENO \leq \text{“E3”}}(ASG)$	(site 1)
$ASG_2 = \sigma_{ENO > \text{“E3”}}(ASG)$	(site 2)
$EMP_1 = \sigma_{ENO \leq \text{“E3”}}(EMP)$	(site 3)
$EMP_2 = \sigma_{ENO > \text{“E3”}}(EMP)$	(site 4)
Query result	(site 5)

Relational algebra must be extended to model exchanging data between sites

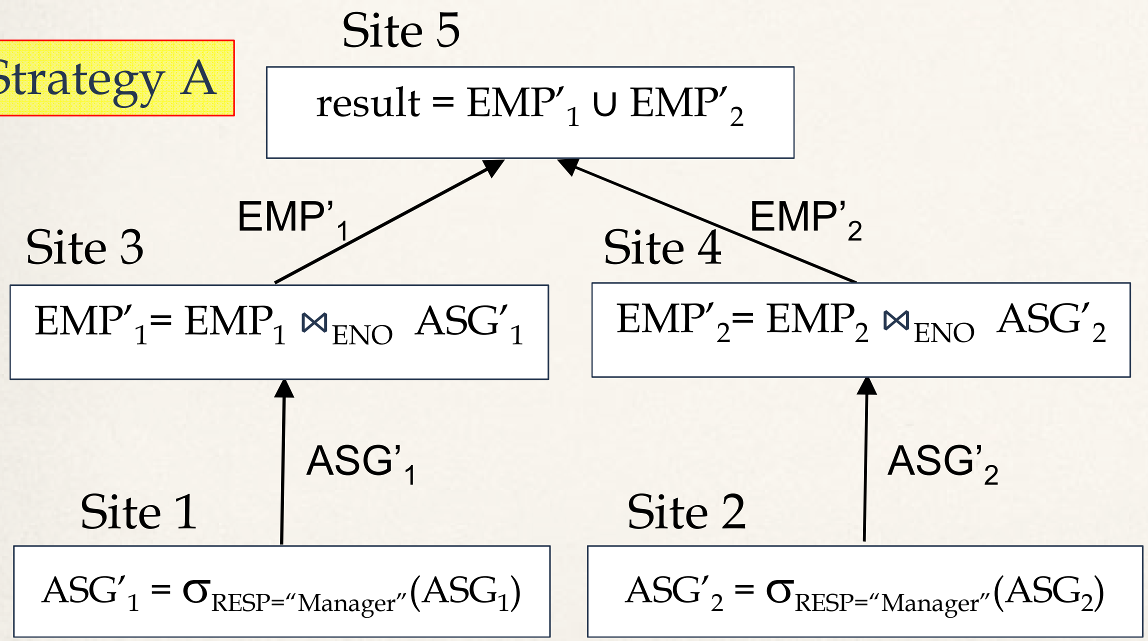
What are the Additional Problems? – Example

- Global query: $EMP \bowtie_{ENO} (\sigma_{RESP=\text{"Manager"}}(ASG))$
- Fragmentation and allocation

$ASG_1 = \sigma_{ENO \leq \text{"E3"}}(ASG)$	(site 1)
$ASG_2 = \sigma_{ENO > \text{"E3"}}(ASG)$	(site 2)
$EMP_1 = \sigma_{ENO \leq \text{"E3"}}(EMP)$	(site 3)
$EMP_2 = \sigma_{ENO > \text{"E3"}}(EMP)$	(site 4)
Query result	(site 5)

Relational algebra must be extended to model exchanging data between sites

Strategy A

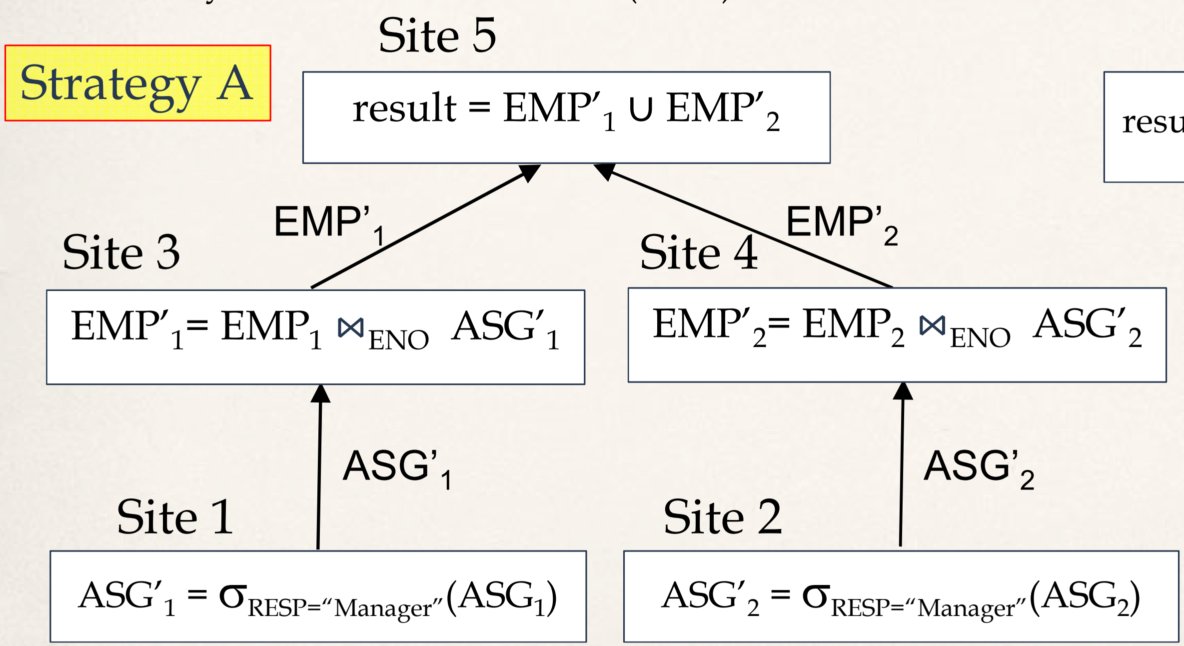


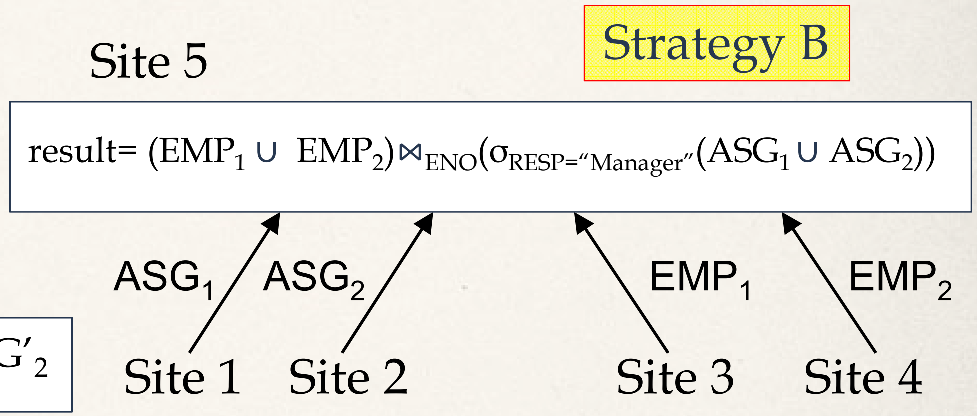
What are the Additional Problems? – Example

- Global query: $EMP \bowtie_{ENO} (\sigma_{RESP=\text{"Manager"}}(ASG))$
- Fragmentation and allocation

$ASG_1 = \sigma_{ENO \le \text{"E3"}}(ASG)$	(site 1)
$ASG_2 = \sigma_{ENO > \text{"E3"}}(ASG)$	(site 2)
$EMP_1 = \sigma_{ENO \le \text{"E3"}}(EMP)$	(site 3)
$EMP_2 = \sigma_{ENO > \text{"E3"}}(EMP)$	(site 4)
Query result	(site 5)

Relational algebra must be extended to model exchanging data between sites



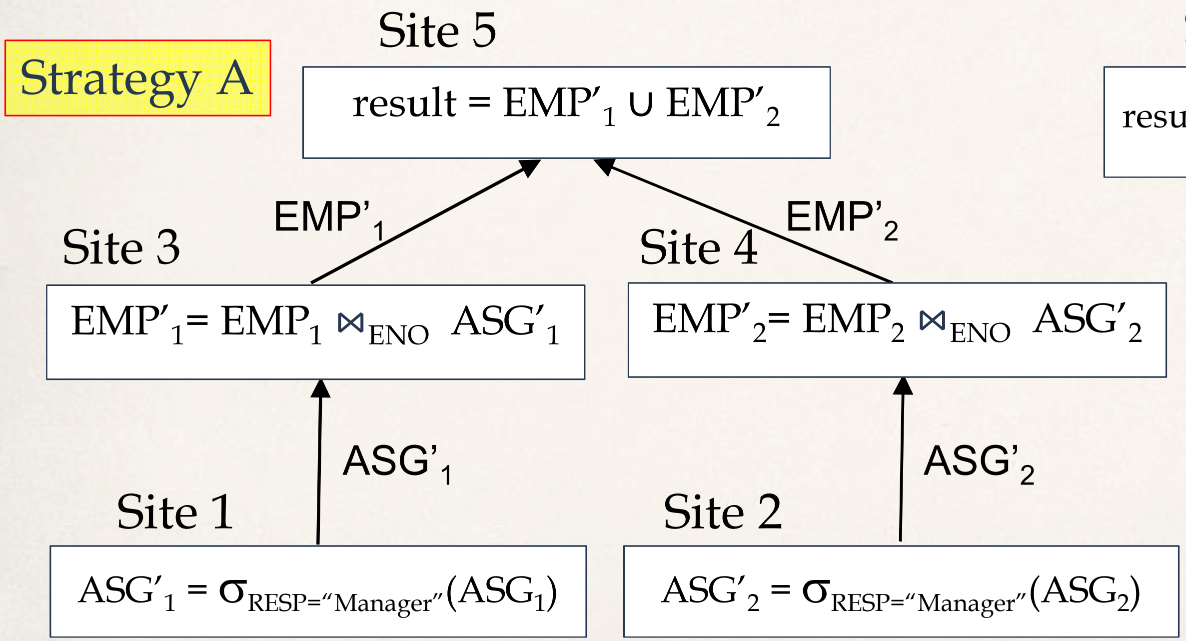


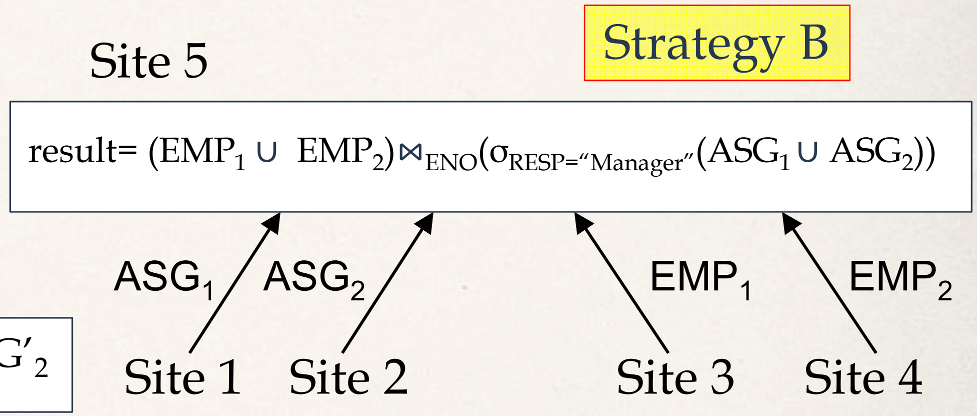
What are the Additional Problems? – Example

- Global query: $EMP \bowtie_{ENO} (\sigma_{RESP=\text{"Manager"}}(ASG))$
- Fragmentation and allocation

$ASG_1 = \sigma_{ENO \leq \text{"E3"}}(ASG)$	(site 1)
$ASG_2 = \sigma_{ENO > \text{"E3"}}(ASG)$	(site 2)
$EMP_1 = \sigma_{ENO \leq \text{"E3"}}(EMP)$	(site 3)
$EMP_2 = \sigma_{ENO > \text{"E3"}}(EMP)$	(site 4)
Query result	(site 5)

Relational algebra must be extended to model exchanging data between sites





Assume
- *card*(EMP) = 400
- *card*(ASG) = 1000
- 20 managers in ASG
- indexes on ASG.RESP and EMP.ENO
- access cost per tuple = 1 unit
- network transfer cost per tuple = 10 units

Cost of Alternatives

- Assume
  - ➡ *card* (EMP) = 400, *card*(ASG) = 1000, 20 managers in ASG
  - ➡ indexes on ASG.RESP and EMP.ENO
  - ➡ tuple access cost = 1 unit; tuple transfer cost = 10 units
- Strategy A

➡ produce ASG': (10+10) ∗ tuple access cost	20
➡ transfer ASG' to the sites of EMP: (10+10) ∗ tuple transfer cost	200
➡ produce EMP': (10+10) ∗ 2 ∗ tuple access cost	40
➡ transfer EMP' to result site: (10+10) ∗ tuple transfer cost	200
Total Cost	460

- Strategy B

➡ transfer EMP to site 5: 400 ∗ tuple transfer cost	4,000
➡ transfer ASG to site 5: 1000 ∗ tuple transfer cost	10,000
➡ produce ASG': 1000 ∗ tuple access cost	1,000
➡ join EMP and ASG': 400 ∗ 20 ∗ tuple access cost	8,000
Total Cost	23,000

Distributed Query Processing Methodology