ICS 421 Spring 2010
Parallel & Distributed Databases

Asst. Prof. Lipyeow Lim
Information & Computer Science Department
University of Hawaii at Manoa
Why Parallel Data Access?

SELECT *
FROM mydata

1 Terabyte (= 1024 GB)

10 MB/s

1.2 days to scan

1000 x parallel 10 MB/s

1.5 mins to scan!
How large is a petabyte?
Parallel DBMS

- eBay’s main Teradata data warehouse (DW):
  - > 2 petabytes of user data
  - 10s of 1000s of users
  - Millions of queries per day
  - 72 nodes
  - >140 GB/sec of I/O, or 2 GB/node/sec

- eBay’s Greenplum DW
  - 6 1/2 petabytes of user data
  - 96 nodes
  - 200 MB/node/sec of I/O

- Walmart – 2.5 petabytes

- Bank of America – 1.5 petabytes

- Some parallel DBMSs besides the usual Oracle-IBM-MS trio:
  - Teradata
  - Netezza
  - Vertica
  - DATAllegro
  - Greenplum
  - Aster Data
  - Infobright
  - Kognitio, Kickfire, Dataupia, ParAccel, Exasol, ...

2/2/2010
Lipyeow Lim -- University of Hawaii at Manoa
Parallelism

Pipeline parallelism
• many machines each doing one step in a multi-step process.

Partition parallelism
• many machines doing the same thing to different pieces of data.

Parallelism is natural to DBMS processing
Parallelism Terminology

- **Speed-up**
  - Same job + more resources = less time

- **Scale-up**
  - Bigger job + more resources = same time

- **Transaction scale-up**
  - More clients + more resources = same time
Parallel Architecture: Share What?

Shared Memory (SMP)
- Easy to program
- Expensive to build
- Difficult to scaleup

CLIENTS
Processors
Memory

Sequent, SGI, Sun

Shared Disk
- Hard to program
- Cheap to build
- Easy to scaleup

CLIENTS

VMScluster, Sysplex

Shared Nothing (network)

CLIENTS

Tandem, Teradata, SP2

2/2/2010
Lipyeow Lim -- University of Hawaii at Manoa
Different Types of DBMS Parallelism

• Intra-operator parallelism
  – get all machines working to compute a given operation (scan, sort, join)

• Inter-operator parallelism
  – each operator may run concurrently on a different site (exploits pipelining)

• Inter-query parallelism
  – different queries run on different sites

• We’ll focus on intra-query parallelism
Parallel vs Distributed DBMS

• A parallel database system
  – Improve performance via parallelization of various operations such as loading data, building indexes, evaluating queries

• A distributed database system
  – Data is physically stored across several (geographical) sites
  – Each site is managed by an independent DBMS
  – Distribution governed by factors like local ownership & increased availability

• The boundaries of these traditional definitions are blurring.
Types of Distributed DBMS

- **Homogeneous**: Every site runs the same type of DBMS.
  - Parallel DBMSs are usually homogeneous

- **Heterogeneous**: Different sites run different DBMSs (different RDBMSs or even non-relational DBMSs).
Data Partitioning & Fragmentation

- Parallel DB
  - Data partitioning

- Distributed DB
  - Fragmentation

- Same basic problem: How do we break up the data (tables) and spread them amongst the “nodes”
  - Horizontal vs Vertical
  - Range vs Hash
  - Replication

- DB user’s view should be one single table.
Automatic Data Partitioning

Partitioning a table:

**Range**

- Good for equijoins, range queries, group-by

**Hash**

- Good for equijoins

**Round Robin**

- Good to spread load

- Shared disk and memory less sensitive to partitioning,
- Shared nothing benefits from "good" partitioning