

ICS 421 Spring 2010

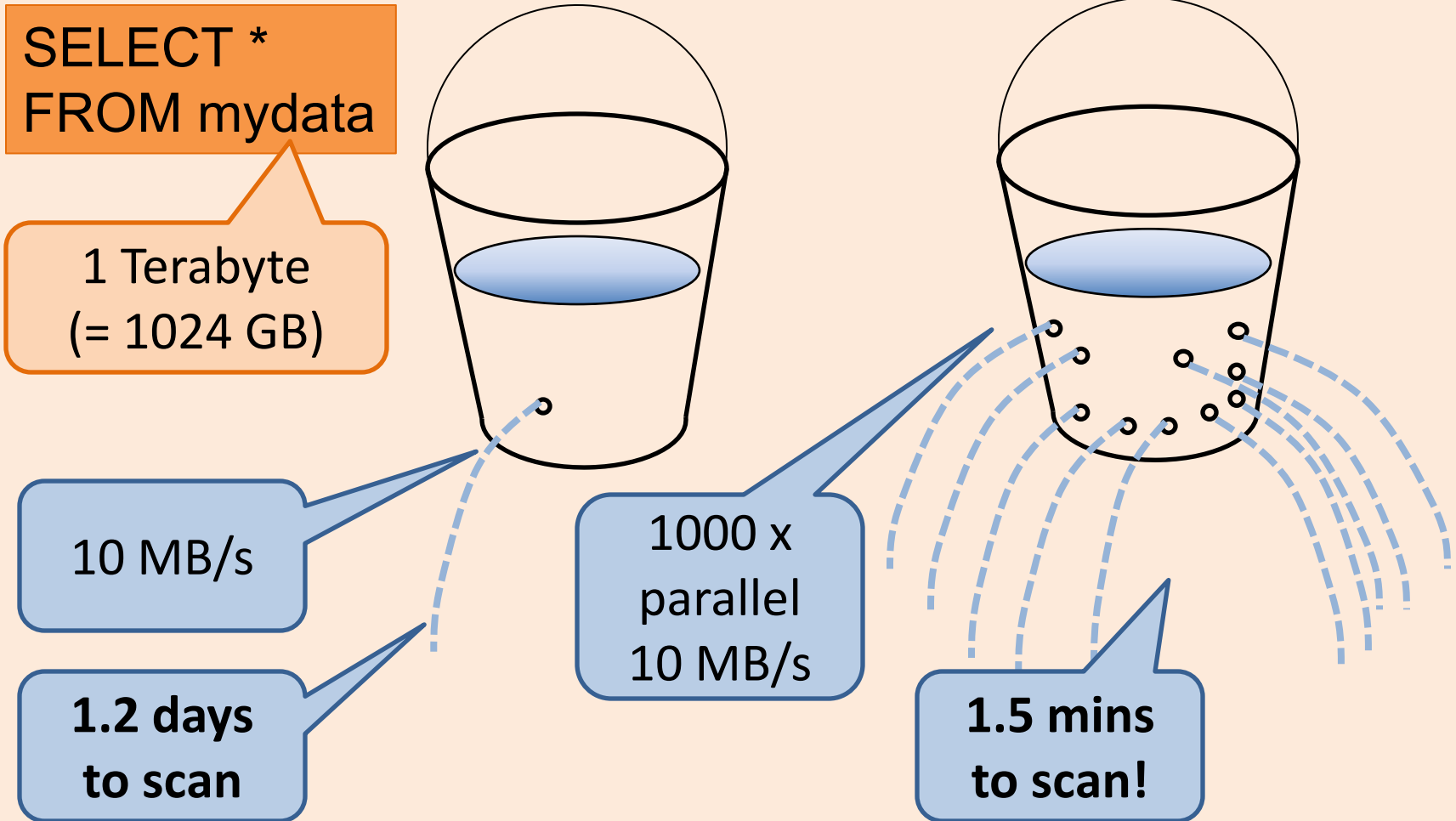
Parallel & Distributed Databases

Asst. Prof. Lipyeow Lim

Information & Computer Science Department

University of Hawaii at Manoa

Why Parallel Data Access ?



Multi-Petabyte Databases



[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
 - DATAlegro
 - Greenplum
 - Aster Data
 - Infobright
 - Kognitio, Kickfire, Dataupia, ParAccel, Exasol, ...

Parallelism

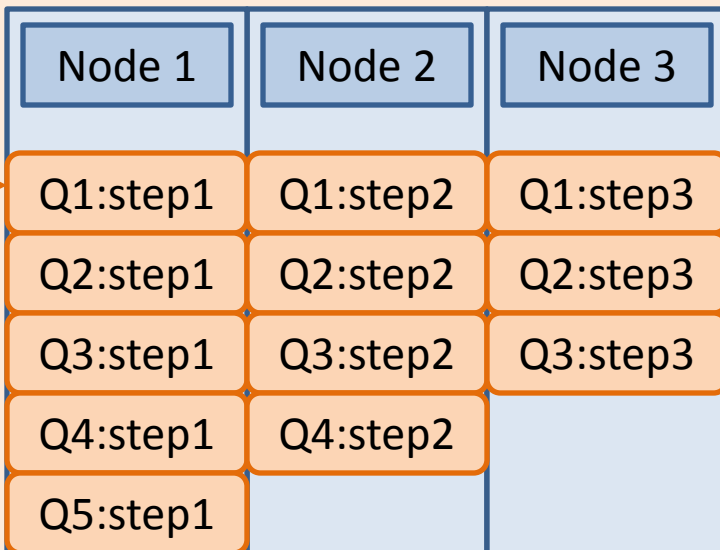
Pipeline parallelism

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

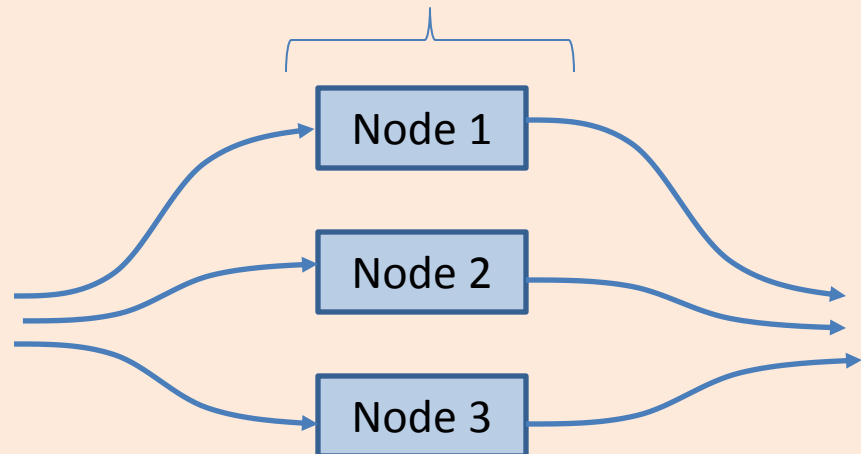
Partition parallelism

- many machines doing the same thing to different pieces of data.

Parallelism is natural to DBMS processing

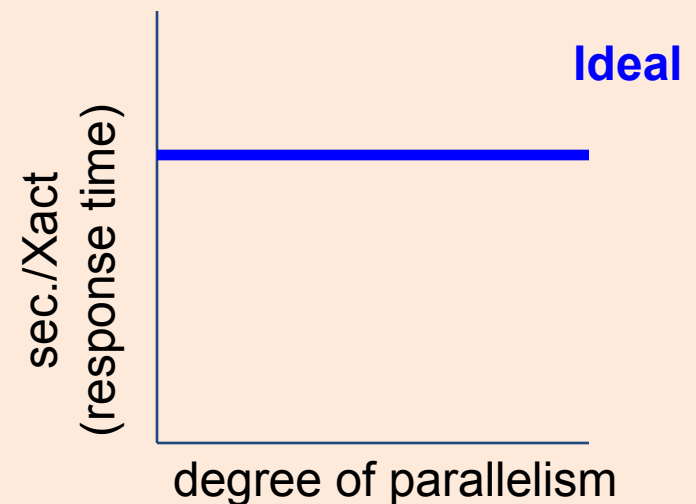
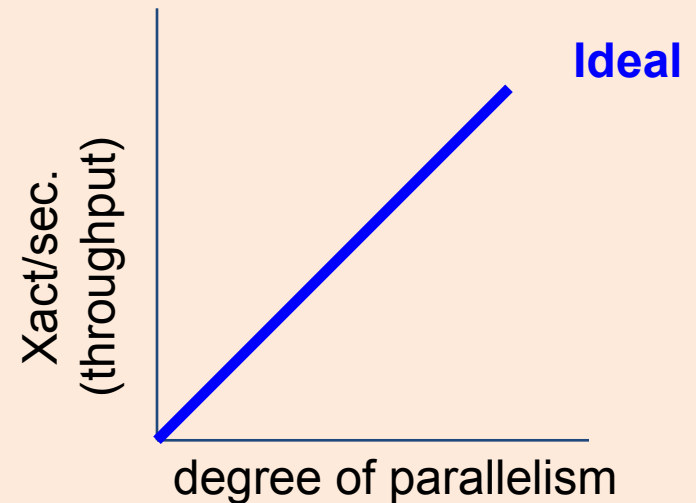


Processing Query 1



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

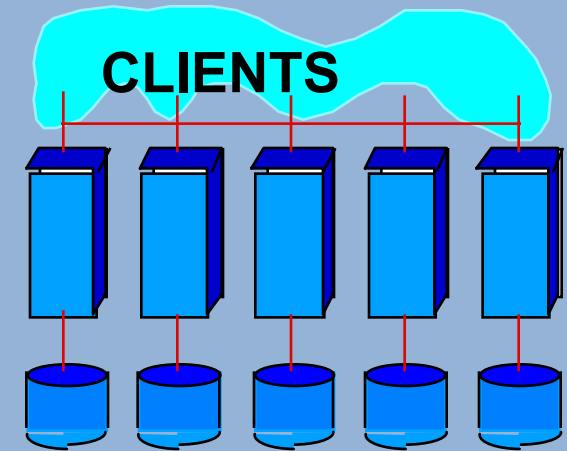
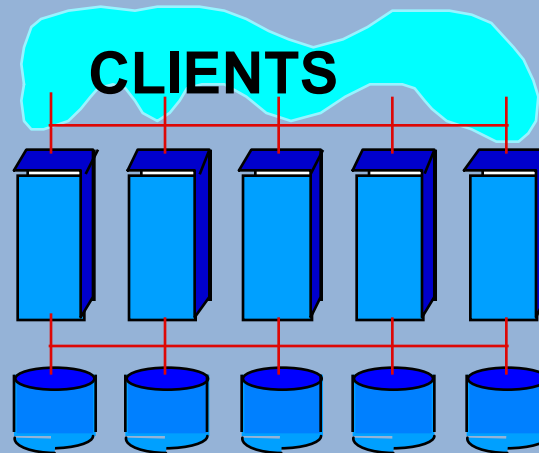
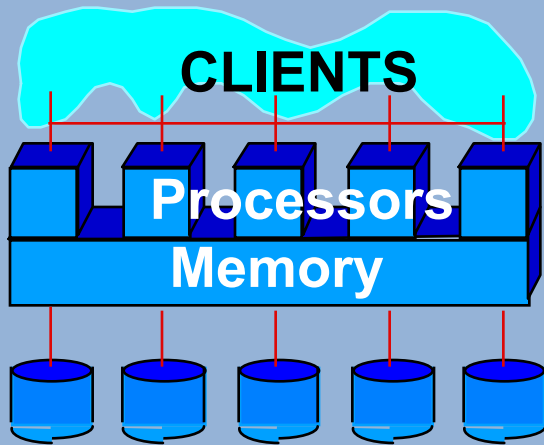


Parallel Architecture : Share What?

Shared Memory
(SMP)

Shared Disk

Shared Nothing
(network)



Easy to program
Expensive to build
Difficult to scaleup

Hard to program
Cheap to build
Easy to scaleup

Sequent, SGI, Sun

VMScuser, Sysplex

Tandem, Teradata, SP2

Different Types of DBMS Parallelism

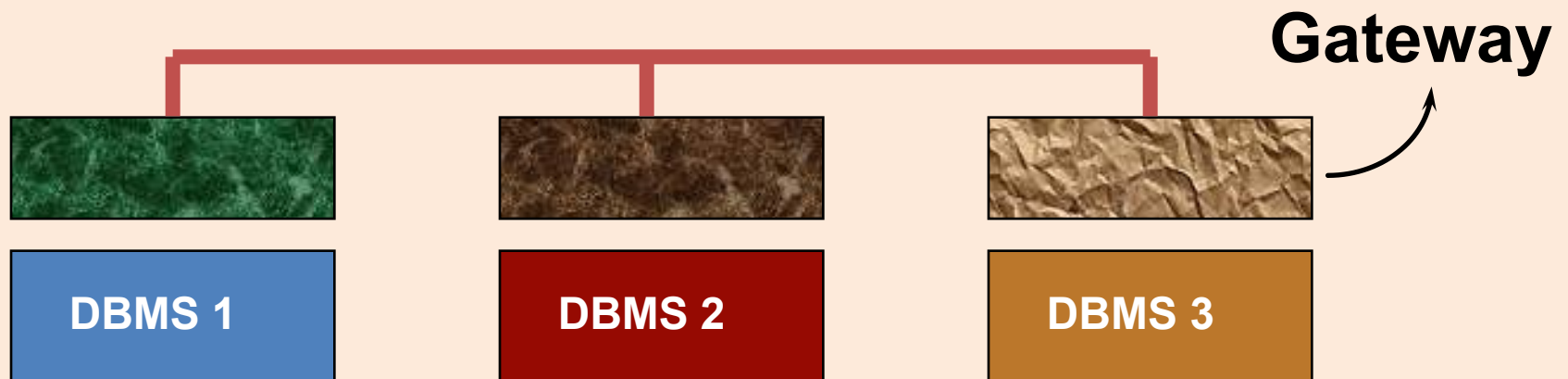
- Intra-operator parallelism
 - get all machines working to compute a given operation (scan, sort, join)
- Inter-operator parallelism
 - each operator may run concurrently on a different site (exploits pipelining)
- Inter-query parallelism
 - different queries run on different sites
- We'll focus on 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 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

TID

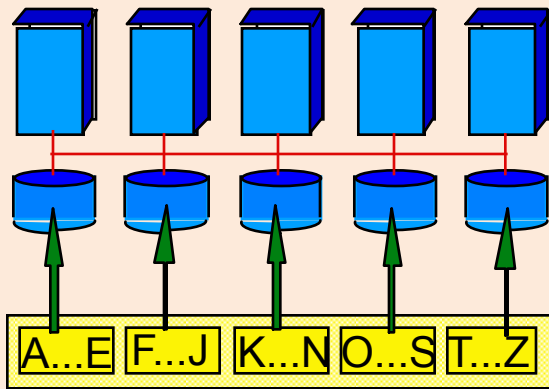
t1				
t2				
t3				
t4				

- 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

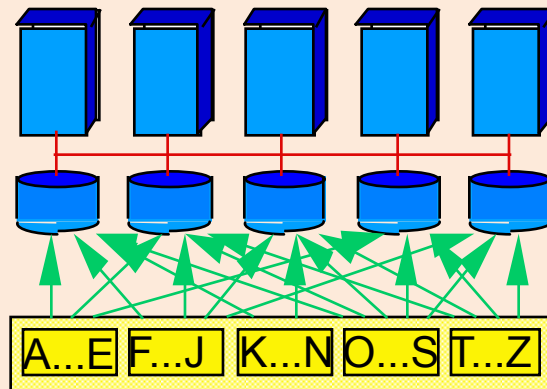
Partitioning a table:

Range



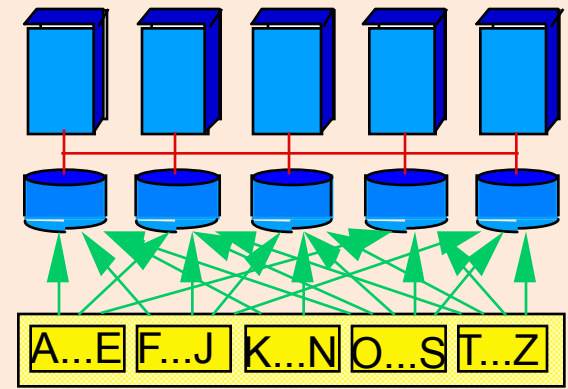
Good for equijoins,
range queries
group-by

Hash



Good for equijoins

Round Robin



Good to spread load

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