Parallel Data Processing

Prof. Yanlei Diao

Slides Courtesy of R. Ramakrishnan and J. Gehrke
In 50 years, computers will be intelligent, with a storage capacity of about $10^9$.

--- Alan Turing, 1950

Moore’s Law predicts the CPU and memory capacities to double about every 18 months.
If exponential growth continues...
Limitations of Moore’s Law

• Concerns regarding if Moore’s law can continue…
• Disk IO bandwidth did not increase much (7200 - 15000 rpm)
• In some domains data grows faster than Moore’s law
• New data processing needs grow fast (“democratization of data”)
Another technology breakthrough is...

Parallel Data Processing
Data Analytics for A Social Network

**User profiles:**
Background, pics, postings, friends…

**Click Streams:**
many billion rows/day
many TB/day

**Data Loading:**
High Volume + Transformation

**Data Processing Backend**

**Quick lookups and updates:**
Update your own profile, read friends’ profiles, write msgs,…

**Analysis Queries:**
Ad targeting, fraud detection, resource provisioning…
Some (Old) Numbers about Facebook

- 1.5 billion daily active users
- 22% Internet time of Americans
- 4 petabyte of new data each day
- >250 billion photos; 350 million photos and 5.8 billion likes a day.

Initial software: PHP + MySQL cluster + Memcached

One of the largest MySQL cluster (OLTP)
- serve images/profiles

Hive, FB’s data warehouse (OLAP), has 300 petabytes of data
- >30,000 servers
- ~180,000 severs or more
Google AlphaGo
Three Forms of Parallelism

Data Parallelism

Model Parallelism

Pipeline Parallelism
Topics

1. Parallel databases (80’s - 90’s)

2. MapReduce (2004 - present)

3. Relational processing on MapReduce
Parallel Databases 101

• Rise of parallel databases: late 80’s

• Architecture: **shared-nothing** systems
  - A number of nodes connected by fast Ethernet switches
  - Inter-machine messages are the only way for communication
  - But used special-purpose hardware (costly, slow to evolve)
  - Small scale (hence did not focus on fault tolerance)

• Typical systems
  - Gamma: U. of Wisconsin Madison
  - TeraData: Wal-Mart’s 7.5TB sales data in hundreds of machines
  - Tandem
  - IBM / DB2
  - Informix…
Some Parallel (||) Terminology

- **Speed-Up**
  - Holds the problem size, grows the system
  - Reports *serial time*/||-time
  - Ideally, linear

- **Scale-Up**
  - Grows both the system and the problem, reports running time
  - Ideally, constant
Some Parallel (||) Terminology

- **Speed-Up**
  - Holds the problem size, grows the system
  - Reports serial time/||-time
  - Ideally, linear

- **Utilization**
  - Speed-up / degree of ||-ism
  - Ideally, constant
Data Partitioning Schemes

Partitioning a table:

**Range by R.a**
- sequential scan
- associative search
- sorting
- may have data skew

**Hash by R.a**
- sequential scan
- equality search
- equijoins if matching the hash attribute
- range search,
- operations that do not match the hash attr.
- can have data skew

**Round Robin**
- sequential scan
- useless for other query operations