Large-scale Data Processing

Yang Wang
Model

- A number of machines connected by network.
How is it different?

• Data/Computation must be partitioned
  – Coordination, locality, load balance, ...

• Dynamic and decentralized environment
  – Nodes can fail or join
  – Network can be unreliable
  – No global clock
  – ...

Goal of this course

Coordinate machines to complete tasks in a dynamic and decentralized environment
So many works ......

• Functionalities
  – What interface does it provide?

• Guarantees
  – What is the correct behavior?

• Performance
  – What is its throughput, latency, and scalability?
Functionalities

What does the user need to do?

What does the system provide?
**Functionalities**

**Computing platform**

- MPI
- MapReduce
- FortranMP

Users do more → Systems do more

**Storage platform**

- Block
- Key-value
- File system
- Database

Users do more → Systems do more
Functionalities

- Metadata services:
  - Lease management: Chubby/Zookeeper
  - Failure detection: Falcon
  - Load balancing: ...
  - Authentication
  - Access control
  - ....
Guarantees

• Consistency (safety, correctness):
  – Linearizable, sequential, casual, eventual, ...

• Liveness

• Availability

• Fault tolerance:
  – Crash only; crash and timing; arbitrary
Performance

- Latency
- Throughput
- Scalability

What is their relationship?
Latency

Mem | LAN | Disk | Geo-replication

Low | High
How to measure performance?

• Throughput:
  - We are interested in a system’s max throughput when it is saturated
  - How to saturate a system? Keep increasing load until system’s throughput does not change.
  - How to increase load:
    • Blocking mode: add number of clients
    • Non-blocking mode: add number of outstanding requests
    • Note you also need to ensure client machines are not saturated.

• Latency:
  - Can increase if load is increasing (queuing theory)
How to measure performance?

• Throughput-latency graph:

  When throughput is close to max, latency grows arbitrarily.

  We often report latency when reaching 80% max throughput.
How to measure performance?

- Scalability graph:

  - **Super-linear growth:** can happen
  - **Linear growth:** perfect scalability
  - **Sub-linear growth:** bottleneck somewhere

Max throughput

Number of servers
How to measure performance?

• Notes:
  • If you are using multiple clients, make sure they all start and stop at relatively the same time. Why?
  • Give the system enough warm up time. Your results should not include warm up time.
  • Allocate sufficient amount of memory
  • If you are testing a disk-based system, the dataset of your test should be big enough so that it cannot be all cached in memory. Rule of thumb: the size of your dataset should be at least twice of your total memory size
  • If you share the network with other users, make sure that does not severely affect your experiments
Warm-up project

• Draw the throughput-latency graph

• Choose a client-server system:
  • Examples: web server, database, etc

• Choose a workload
  • Examples: download file, update a row, etc

• Use the previous methods to draw the graph