CLOUD COMPUTING CONCEPTS

with Indranil Gupta (Indy)

CLOUDS ARE DISTRIBUTED SYSTEMS

Lecture A

A CLOUD IS A DISTRIBUTED SYSTEM

A Cloud...

- A cloud consists of
  - Hundreds to thousands of machines in a datacenter (server side)
  - Thousands to millions of machines accessing these services (client side)
- Servers communicate amongst one another
- Clients communicate with servers
- Clients also communicate with each other

A Cloud... IS a Distributed System

- Servers communicate amongst one another → Distributed System
  - Essentially a cluster!
- Clients communicate with servers
  - Also a distributed system!
- Clients may also communicate with each other
  - In peer-to-peer systems like BitTorrent
  - Also a distributed system!

Four Features of Clouds = All Distributed Systems Features!

I. Massive Scale: many servers

II. On-demand nature
   - access (multiple) servers anywhere

III. Data-Intensive Nature
   - lots of data => need a cluster (multiple machines) to store

IV. New Cloud Programming Paradigms
   - Hadoop/Mapreduce, NoSQL all need clusters

Cloud = A Fancy Word for a Distributed System

- A “cloud” is the latest nickname for a distributed system
- Previous nicknames for “distributed system” have included
  - Peer-to-peer systems
  - Grids
  - Clusters
  - Timeshared computers (Data Processing Industry)

(See Lecture Video on History of Clouds)

Cloud = A Fancy Word for a Distributed System (2)

- Nicknames come and go, but the core concepts underlying distributed systems stay the same
  - And they are used decade after decade
    - E.g., Lamport Timestamps were invented in the 1970s, and they are used in almost all distributed/cloud systems today
  - This course is about these distributed systems concepts
- A few years from now, there may be a new nickname for distributed systems
  - The core concepts will remain the same, and they will continue to be used in real systems

So What IS a "Distributed System"?

- Let's try to define the term!