

CS 294-42: Cloud Computing, Infrastructure, Services, and Applications

Spring 2009 (MW 2:30-4:00, 310 Soda Hall)

Ion Stoica

(<http://www.cs.berkeley.edu/~istoica/classes/cs294/09/>)

1

Cloud Computing in Numbers

- Datacenter instance:
 - Costs in billion range
 - > 100,000 servers
- By 2010
 - Over 15mil of servers installed in US
 - Over 45B to power and cool servers

2

Why does Cloud Computing Matter?

- Fundamental change
 - The way applications are written and deployed
 - Internet traffic: Internet becomes last-hop between hosts and datacenter
 - Economics
- Opportunity to rethink:
 - Large scale distributed systems
 - Network architectures
 - Tradeoffs in computer systems

3

How is Cloud Computing Different?

- How is different from distributed systems?
- How is different from parallel systems?
- Axis:
 - Environment constraints
 - Scale
 - Type of failures
 - Application requirements
 - ...

4

Academia Challenges

- Rapid evolving field
 - Need to be in the avantgarde of understanding challenges and trends
 - Unfortunately, academia trails industry
 - Very hard to achieve the scale and generate the workload that reveal main challenges
- How to address above challenges?

5

Grading

- Project: 60%
- Class presentations: 40%
 - See Randy's guidelines for leading discussion on papers
 - <http://bnrg.eecs.berkeley.edu/~randy/Courses/CS294.F07/LeadingPapers.pdf>

6

Papers

- Is the problem real?
- What is the solution's main idea (nudget)?
- Why is solution different from previous work?
 - Are system assumptions different?
 - Is workload different?
 - Is problem new?
- Does the paper (or do **you**) identify any fundamental/hard trade-offs?

7

Papers (cont'd)

- Do you think the work will be influential in 10 years?
 - Why or why not?
- Predicting the future always hard, but worth a try
 - Look at past examples for inspiration

8

Streaming Over TCP

- Countless papers:
 - Why cannot be done...
 - New protocols to do it...
- Today
 - Virtually all streaming over TCP
 - During inauguration, Akamai served "...over 800 Gbps of Flash streaming"

9

Why did it Succeed?

10

Multicast

- Countless papers:
 - Why world will come to a standstill without multicast...
 - New protocols to do it...
- Today
 - Multicast is used only in enterprise settings at best
 - Overlay multicast widely used in the Internet
 - During inauguration Akamai served "...over 800 Gbps of Flash streaming"

11

Why Did it Fail?

12

Consistency Everywhere

- Many papers & systems:
 - Group synchronous communication
 - Causally ordered message delivery
 - ...
- Today:
 - Almost never used in WANs, and rarely used in LANs

13

Why Did it Fail?

14

Shared Memory

- Countless papers:
 - How shared memory simplifies programming parallel computers
 - Many, many systems proposed and build
- Today:
 - Message passing (MPI) took over as the de facto standard for writing parallel applications

15

Why Did it Fail?

16

Network Computer

- Big in 90s
 - Promoted by an alliance of Sun, Oracle, Acorn
- Promise: many of advantages of cloud computing
 - Easy to manage
 - Application sharing
 - ...
- Failed miserably

17

Why Did it Fail?

18

What are Hard/Fundamental Tradeoffs?

- Brewer's CAP conjecture: "Capability, Availability, Partition-tolerance", you can have only two in a distributed system
- In a *in-order, reliable* communication protocol cannot minimize overhead and latency simultaneously
- Hard to simultaneously maximize evolvability and performance

19