Cloud Computing

Welcome
(Module Administrivia)

Dell Zhang
Birkbeck, University of London
2014/15
Who is this module for?

- Software Developers
- Hardware Engineers
- System Administrators
- IT Managers
- IT Salesmen
What is this module about?

- Data-intensive information processing
- Big-data (“Web-scale”) problems
- Focus on applications
- MapReduce... and beyond
What is MapReduce?

• Programming model for expressing distributed computations at a massive scale
• Execution framework for organizing and performing such computations
• Open-source implementation called Hadoop
On the job market

- Hadoop Jobs in London

[Graph showing the increase in Hadoop jobs in London from 2008 to 2013]
On the job market

- Hadoop Jobs in London
Powering London’s digital economy

Pre-requisites

• Good knowledge of Java programming
  – BSc: SP2; MSc: ISD or PiJ
  – But this module is not just about programming: we’ll expect you to pick up Hadoop (quickly) along the way
  – Focus on “thinking at scale” and algorithm design

• No previous experience necessary in
  – Parallel and distributed programming etc.
This module is not for you ...

• If you’re not genuinely interested in the topic
• If you can’t put in the time
• If you’re not ready to do a lot of work
• If you’re not open to thinking about computing in new ways
• If you can’t cope with the uncertainty and unpredictability, etc. that comes with bleeding edge software
Assessment

• Coursework (20%)
  – BSc: One essay of 2,000–3,000 words
  – MSc: Two programming assignments
• Examination (80%)
Resources

• Hadoop on your local machine
  – Linux: native installation
  – Windows/Mac: in a virtual machine (e.g., VirtualBox)
    • Hortonworks Sandbox

• Hadoop in the cloud
  – Amazon EMR
Teaching Assistant

Martyn Harris
Important Aside

• Amazon Web Services (AWS) in Education
  – Create your account
  – Redeem the free credits

• Stay tuned for more details from the module page (also in Moodle) ...