

COMP2121: Principles of Distributed Systems and Networks

Introduction

Unit coordinator Dr. Vincent Gramoli
School of Information Technologies

The University of Sydney

Copyright

COMMONWEALTH OF AUSTRALIA Copyright Regulations 1969 WARNING

- This material has been reproduced and communicated to you by or on behalf of the University of Sydney pursuant to Part VB of the Copyright Act 1968 (**the Act**).
- The material in this communication may be subject to copyright under the Act. Any further copying or communication of this material by you may be the subject of copyright protection under the Act.
- **Do not remove this notice.**

The University of Sydney

Page 2

Outline

- Administration
Syllabus, goal, plagiarism, etc.
- Definition
What is a distributed system?
- Examples
Several distributed systems you might have heard of
- Challenges
Scalability, consistency, fault-tolerance

The University of Sydney

Page 3

Administration



Timetable

- Title: Distributed Systems & Network Principles
- UOS code: COMP2121
- Credit point: 6
- Lecture:
 - Monday 2PM-4PM, weeks 1-13
 - Merewether Lecture Theatre
- Labs: depend on your timetable (Wed, 12PM-1PM)
 - Go to the lab you are scheduled for
 - If for some reason you miss it, you can attend a later lab session if there is space and the tutor agrees, but ask the tutor before taking a seat
- Get help from staff if you feel you are falling behind

The University of Sydney

Page 5

Ourselves

- Lecturer **Vincent Gramoli**
 - Office: 417
 - Vincent.Gramoli@sydney.edu.au
 - <http://poseidon.it.usyd.edu.au/~gramoli/web/php/>
 - Office hour: Wednesday 16h-17h
 - Background: Distributed computing, concurrency
- **Teaching assistant:** Jian Guo jguo4890@uni.sydney.edu.au
- **Tutors:** depend on your timetable
 - Rabia Choudry rcha3704@uni.sydney.edu.au
 - Parinya Ekparinya pekp6601@uni.sydney.edu.au
 - Jian Guo jguo4890@uni.sydney.edu.au
 - Joshua Murray jmr9664@uni.sydney.edu.au
 - Patrick Nappa pnp4580@uni.sydney.edu.au
 - Michael John Spain mpg1382@uni.sydney.edu.au
 - Omid Tavallaei otav8458@uni.sydney.edu.au
 - Tyson Thomas tth6664@uni.sydney.edu.au
 - Julia Wong jwon5553@uni.sydney.edu.au

The University of Sydney

Page 6

Resources

eLearning: <https://elearning.sydney.edu.au/>

- Login using Unikey and password
- Submit official assignment work here or on PASTA
- Link to Unit of Study (UoS) outline
 - Official schedule, list of learning outcomes, etc.
- Copies of slides
- Lab instructions
- Assignment instructions
- Lecture recording
 - We intend to record the lectures, but the technology is not reliable
- See you grade

PASTA will be used for programming assignments:

<http://soft-app-pro-12.ucc.usyd.edu.au:8080/PASTA/>

Discussion forum is linked on the elearning website (invitations sent):

<https://edstem.org/courses/1441/>

- Do not send email to the lecturer, but post questions online (on this forum)
- Everyone is welcome to answer and rate answers

What you need to know

Distributed systems

- What is a distributed system?
- How it works?
- How to run yours?

Programming

- Most labs will require some programming
- All programming will be done in **Java**
- Your code will be required to run in the lab rooms
 - If you implement it at home, test it in the lab too!

Expectations

- Students attend scheduled classes, and devote an extra 6-9 hours per week
 - doing assessments
 - preparing and reviewing for classes
 - revising and integrating the ideas
 - practice and self-assessment
- Students are responsible learners
 - Participate in classes, constructively
 - Respect for one another (criticize ideas, not people)
 - Humility: none of us knows it all; each of us knows valuable things
 - Check eLearning, Ed, PASTA site at least once a week!
 - Notify academics whenever there are difficulties
 - Notify group partners honestly and promptly about difficulties

Advice

- Lectures notes are for help
- You should understand in-depth
- Practice your reasoning by re-doing the examples at home
- Think about implications, ask questions
- Re-read your notes or the lecture notes at home after the class to memorize easily

Overall Goal

Distributed systems

- W5: Programming assignment 1: 10%
- W7: Mid-semester exam: 10%
- W8: Programming assignment 2: 10%
- W12: Programming assignment 2: 20%
- Exam period: Final exam: 50%
- School of IT policy: You must have at least 40% of the exam total mark to pass the unit

Late assessments in COMP2121

Suppose you hand in work after the deadline:

- If you have not been granted special consideration or arrangements
 - A penalty of 20% of the available marks will be taken, per day (or part) late

Examples:

- Your work would have scored 60% and is 1 hour late
 - you get 40%
- Your work would have scored 70% and is 28 hours late
 - you get 30%

Warning: submission sites get very slow near deadlines

Submit early: you can resubmit if there is time before the deadline

The University of Sydney

Page 13

Special Consideration (University policy)

- If your performance on assessments is affected by illness or misadventure, follow proper bureaucratic procedures:
 - Have professional practitioner sign special USyd form
 - Submit application for special consideration online, upload scans
 - Note you have only a very short deadline for applying
 - http://sydney.edu.au/current_students/special_consideration/
 - notify coordinator by email as soon as anything starts to go wrong
- There is a similar process if you need special arrangements for religious observance, military service, representative sports, etc.

The University of Sydney

Page 14

Prerequisites

- Prerequisites: Java (INFO1103 or INFO1903) and (INFO1105 or INFO1905)
- Corequisites: Algorithms and complexity (COMP2007 OR COMP2907)
- You must have passed INFO1103 Introduction to Programming
- COMP2129 would help for concurrency

The University of Sydney

Page 15

Academic Integrity (University policy)

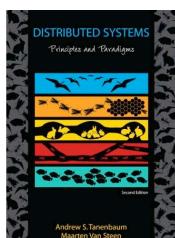
- "The University of Sydney is unequivocally opposed to, and intolerant of, plagiarism and academic dishonesty.
 - Academic dishonesty means seeking to obtain or obtaining academic advantage for oneself or for others (including in the assessment or publication of work) by dishonest or unfair means.
 - Plagiarism means presenting another person's work as one's own work by presenting, copying or reproducing it without appropriate acknowledgement of the source." [from site below]
 - <http://sydney.edu.au/elearning/student/EI/index.shtml>
- Submitted work is compared against other work (from students, the internet, etc.)
 - Turnitin for textual tasks (through eLearning), other systems for code
- Penalties for academic dishonesty or plagiarism can be severe

The University of Sydney

Page 16

Textbook

Distributed Systems – Principles and paradigms Tanenbaum and Van Steen. 2nd Edition.



This and other relevant works can be found in the university library

The University of Sydney

Page 17

Syllabus

Week	Lecture	Lab/Tutorial	Material
1	Introduction	None	Ch. 1, 2
2	Concurrency 1	Multithreading	Ch. 2, 3
3	Communication 1	Routing	Ch. 3, 4
4	Communication 2	Client-Server	Ch. 4
5	Synchronization 1	Physical and logical time	Ch. 6, http://www.theaustralian.com.au/australian-it/leap-second-change-sees-computer-systems-crash-and-story-ftrgakx1226414387145
6	Naming	Remote Method Invocation	Ch. 5, Sect. 10.3.4. (Section 16.3 of Silberschatz, Galvin and Gagne's book 8th edition.)
7	Consistency	Quiz	Ch. 7,
8	Synchronization 2	Consistency	(Hagit Attiya, Jennifer Welch. Distributed computing. Fundamentals, simulations, and advanced topics. 1
9	No lecture (abor day)	Transactional memory	http://sydney.edu.au/engineering/it/~gramoli/doc/pub2/wiley-42-preprint.pdf and Ch. 6
10	Failures	Consensus	Sect. 2.2.2, 4.5.2, Slicing
11	Security (guest lecture)	Security	Ch. 9
12	Blockchains	Demo for assignment 2	
13	Review	Review	

The University of Sydney

Page 18

Definition

What is a Distributed System?

A collection of independent computers that appears to its users as a single coherent system.



What is a Distributed System?

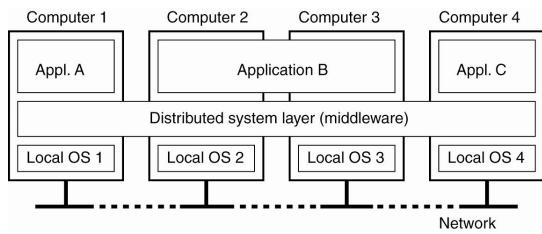
Transparency helps the users observe a single coherent system

	Description
Access	Hide differences in data representation and how a resource is accessed
Location	Hide where a resource is located
Migration	Hide that a resource may move to another location
Relocation	Hide that a resource may be moved to another location while in use
Replication	Hide that a resource is replicated
Concurrency	Hide that a resource may be shared by several competitive users
Failure	Hide the failure and recovery of a resource

- The different forms of transparency in a distributed systems

Middleware

A distributed systems can be viewed as a middleware



- The middleware layer extends over multiple machines and offers each application the same interface

Examples

Cloud computing



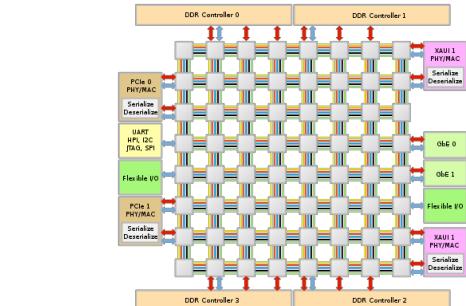
Cluster



The University of Sydney

Page 25

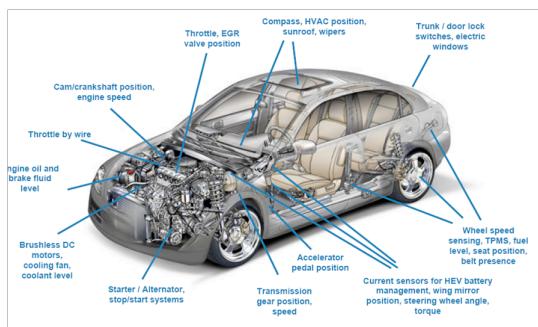
Chip Multiprocessors



The University of Sydney

Page 26

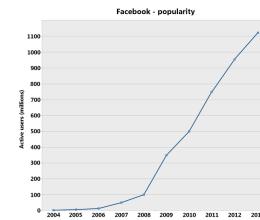
Sensor Networks



The University of Sydney

Page 27

Social Networks



SOURCE: http://en.wikipedia.org/wiki/File:Facebook_popularity.PNG 31 March 2013

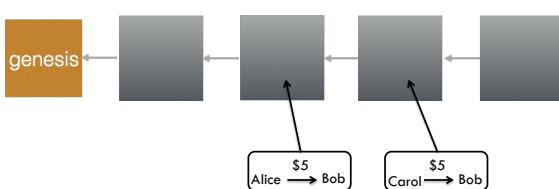


INTERNET USERS AND POPULATION STATISTICS FOR OCEANIA					
SOUTH PACIFIC REGION	Population (2012 Est.)	% Pop. of World	Internet Users, 30-June-2012	Penetration (% Population)	Internet % Users
Total for Oceania	35,903,569	0.5 %	24,287,919	67.6 %	1.0 %
Rest of the World	6,981,943,353	99.5 %	2,381,230,457	34.1 %	99.0 %
WORLD TOTAL	7,017,846,922	100.0 %	2,405,518,376	34.3 %	100.0 %

SOURCE: <http://www.internetworldstats.com/stats.htm> 30 June 2012

Page 28

Blockchain



The University of Sydney

Page 29

Challenges



Scalability

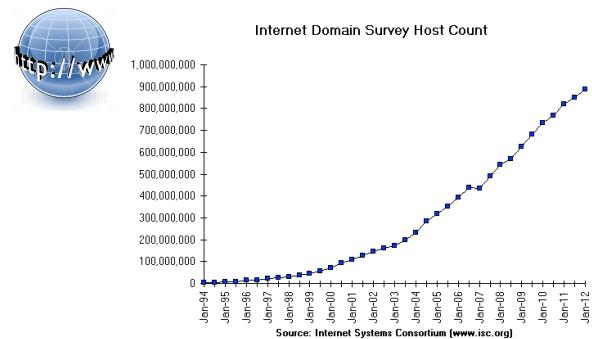
- **Scalability of a distributed system:** the ability for the system to preserve some properties as the system grows in
 - the number of requests or participants,
 - the distance between resources and users, or
 - the heterogeneity.

Concept	Example
Centralized services	A single server for all users
Centralized data	A single on-line telephone book
Centralized algorithms	Doing routing based on complete information

The University of Sydney

Page 31

Scalability



The University of Sydney

Page 32

Scalability



- **Burst of load:**
 - 456 tweets per second (TPS) when Michael Jackson died (June 25, 2009).
 - 6,939 TPS after midnight in Japan on 2011 New Year's day.
- **Increase in participation:**
 - +182%: Increase in number of mobile users in 2010.
 - >500,000 new accounts created on a single day.

Source: <http://blog.twitter.com/2011/03/numbers.html>

The University of Sydney

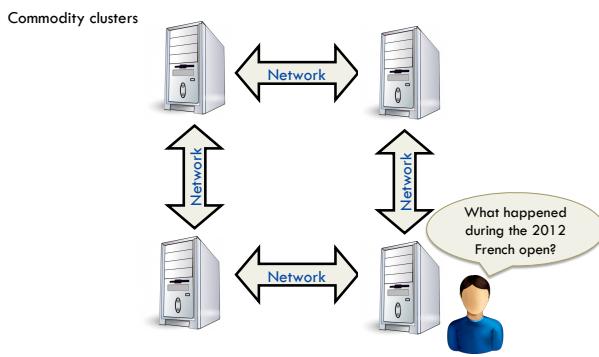
Page 33

Consistency

- **Consistency;** a property applying to a collection of data items that are accessed by distributed participants.
- **Examples of inconsistencies:** As a participant, I observe that Djokovic lost against Nadal but then Djokovic won against Federer in the 2012 French Open.

Page 34

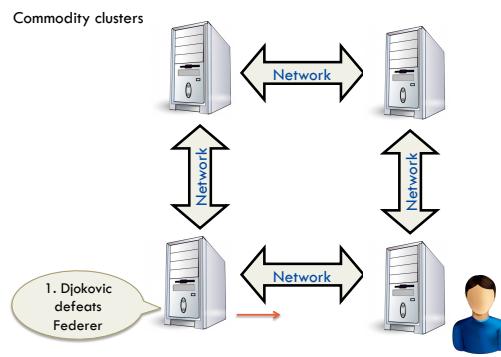
Consistency



The University of Sydney

Page 35

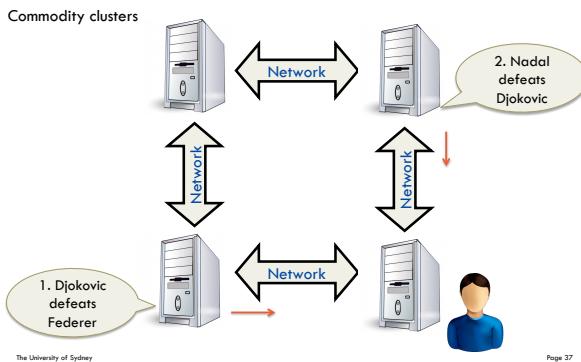
Consistency



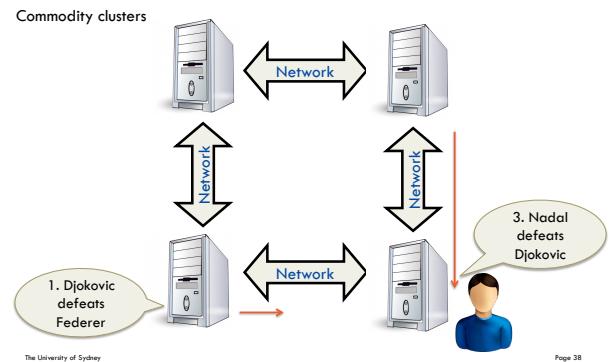
The University of Sydney

Page 36

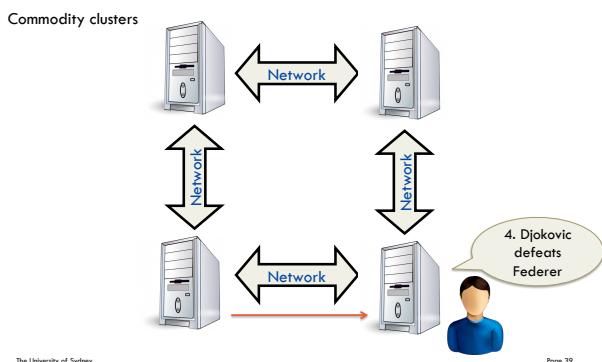
Consistency



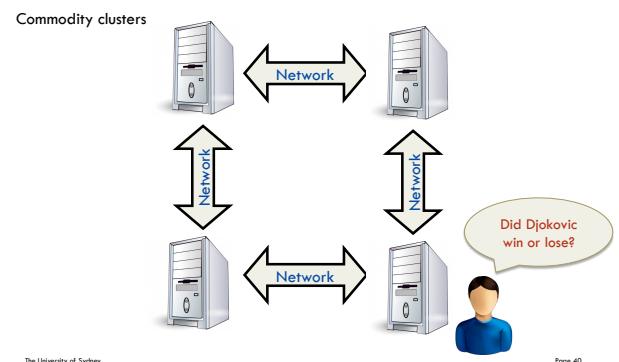
Consistency



Consistency



Consistency



Fault-Tolerance

- *Fault-tolerance of a distributed system:* the ability for the system to recover from partial failures.



Fault-Tolerance

- *Fault-tolerance of a distributed system:* the ability for the system to recover from partial failures.



Fault-Tolerance

- *Fault-tolerance of a distributed system:* the ability for the system to recover from partial failures.
- How to keep the distributed system up and running, thereby appearing as a single running system to its users?

