UNIVERSITY OF HULL

08346 Distributed Systems Programming Introduction

Dr John Dixon

Submit instant feedback on this lecture or module here: http://bit.ly/2vHICZJ

UNIVERSITY OF HULL

Feedback

- At the end of the module you will normally be asked to provide formal feedback (not hugely useful for you, but very useful for improving next year so you should definitely do this).
- Also submit informal feedback whenever you like here: http://bit.ly/2vHICZJ
  - You can remain anonymous if you wish (although I would recommend that you don't)
  - This will help me know if there are any problems or if I can make changes whilst you are still here to benefit from them!

UNIVERSITY OF HULL

Overview

We are looking at these topics today:

- What is this module?
- What should you expect to learn?
- What I expect from you.
- How the module is broken down (ACW/Exam)
- Where you can find resources, etc.
- An introduction, identifying the background of Distributed Computing

UNIVERSITY OF HULL

Administration

- This module takes place over the course of this trimester only.

- There will be two, 50min lectures per week, on:
  - Tues (5pm)
  - Fri (11am)

- And 1 lab session per week, for 8 weeks:
  - Starting NEXT WEEK
  - In weeks 25-27,29,30 and 33-34 (where weeks 31 and 32 are Easter)

- Remember to check/download your timetable
  - Make sure to turn up and sign in to your scheduled lab

UNIVERSITY OF HULL

Assessment

- Assessment is:
  - 50% Coursework
  - 50% Examination
- Coursework is:
  - A single distributed programming exercise
  - Marked out of 100
  - Released in the Week Commencing March 16th
  - Deadline is April 27th
  - Pass is 40% or above
- Exam is:
  - Two hours
  - Two questions
  - Covering both technical and design aspects of distributed systems
  - Marked out of 40 (20 marks per question)
  - Pass is 40% or above

UNIVERSITY OF HULL

Expectations

- The majority of you are third year students. This is a Level 6 module
- You will (hopefully) be in a job this time next year

- I expect you to…
- …behave like adults – not children
- …respect your peers and me (not disrupting lectures by talking, turning up late, etc.)
- …be able to manage your own time effectively
- …study outside lectures and read around the subject area to enhance your understanding of the principles being taught
- …ask questions, offer insight and contribute to lectures/labs/etc. where appropriate

UNIVERSITY OF HULL

Expectations

- You should expect me to:

- Answer your questions or direct you to resources (where appropriate)
- Teach/explain/outline/identify key concepts, as appropriate to your level of study and the module learning outcomes
- Set labs which are relevant to the module/exam/coursework
- Set coursework which is relevant to the module learning outcomes
- Mark and provide feedback for your coursework within a 4-week period

UNIVERSITY OF HULL

Module Learning Outcomes

By the end of this module you should be able to…

1. Explain, with comprehension, the key concepts and principles of research, selection and assessment of distributed system architectures.
2. Discuss and make judgements through critical analysis and evaluation in relation to the principal characteristics of distributed applications and their impact on design, implementation and deployment, justifying your arguments.
3. Critically evaluate a range of contemporary distributed computing technologies, integrating reference to literature effectively with own ideas.
4. Specify, design and implement a distributed software application, which is both appropriate and relevant for a suggested purpose.

UNIVERSITY OF HULL

Module Aims and Objectives

The aim of this module is to provide detailed coverage of the architectures, technologies and programming paradigms used in implementing and deploying distributed computing applications. Detailed consideration will be given to distributed system architectures. The primary focus of this module will not be on low-level network interactions but on higher-level distributed computing technologies (e.g. remote invocation, Web services, mobile computing) which allow the construction of secure and reliable systems.

On successful completion of this module, students will be able to demonstrate an understanding of:

- The major architectural models and programming paradigms used in distributed and networked computing.
- The design and deployment of distributed computing applications.
- Contemporary distributed computing technologies, their features and limitations.
- The relationship between design, deployment and operational requirements.

UNIVERSITY OF HULL

Indicative Content

- Distributed Systems and Applications
- Technologies used for distributed computing and the creation of distributed systems
- Patterns for architecting and creating distributed systems
- Security in distributed systems
- Concurrency in distributed systems
- Design and deployment of distributed systems

UNIVERSITY OF HULL

Resources

Resources supplied by me will be available on Canvas.

There will also likely be links to other resources on Canvas, in presentations, in the coursework specification and on the reading list at:

http://readinglists.hull.ac.uk/modules/08346/lists.html

You are expected to do your own reading and research too. Good resources are available in books and also online, especially through:

- MSDN – e.g. https://msdn.microsoft.com/en-us/library/windows/desktop/ee663286(v=vs.85).aspx
  https://msdn.microsoft.com/en-us/library/dd129906.aspx

UNIVERSITY OF HULL

Questions…?

Are there any questions on what I have explained so far?

UNIVERSITY OF HULL

Distributed Computing: A Primer

What is a Distributed System?

Definition: A system in which components, located at networked computers, communicate and coordinate their actions (only) by passing messages.

This leads to some typical characteristics / features:

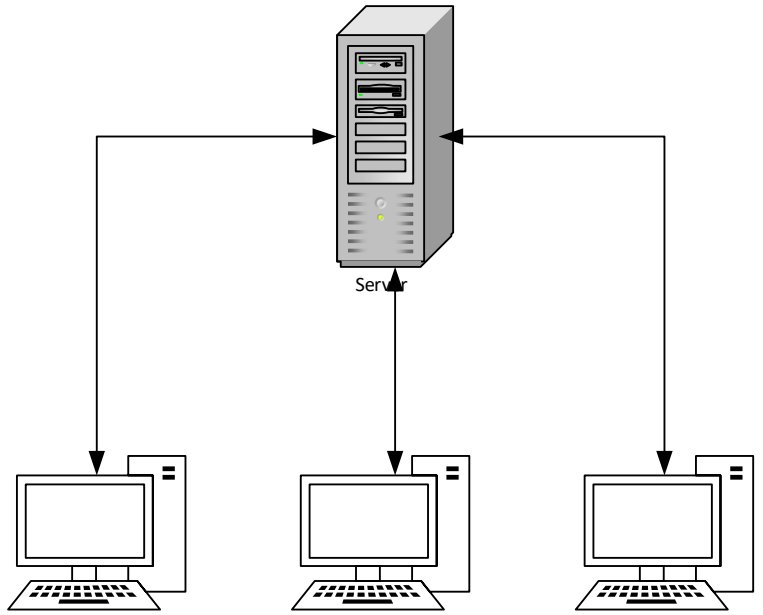
- Multiple networked components
- Components communicate via network / messages
- Resources shared between components
- Transparency - system is perceived as a whole, even across multiple devices
- Coordination is important to make sure the entire system can operate concurrently
- Failure and fault tolerance is key to keep the whole system operational
- Performance and scalability is very important

Distributed Computing: A Primer

Example of a Distributed System

Common Client/Server Model

Could be a server sending data back to clients on request (e.g. web) or perhaps server managing synchronisation and communication between clients (e.g. multiplayer gaming)



Where is execution happening?

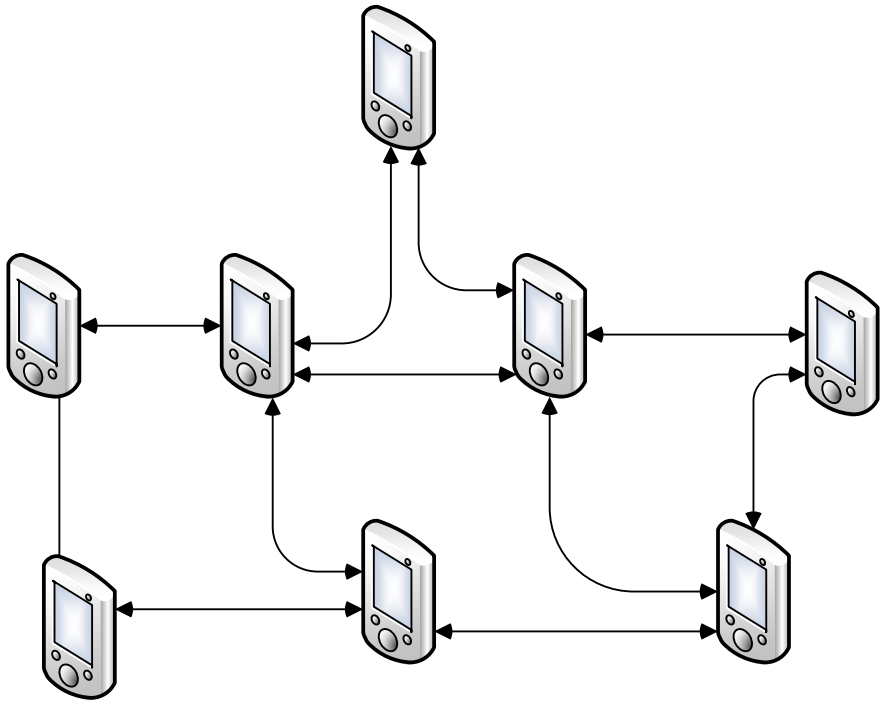
- In the server – perhaps the client is simply an interface.
- In the client – perhaps the server is just logging.
- Both – perhaps execution is split between server and client, depending on what is being executed!

UNIVERSITY OF HULL

Distributed Computing: A Primer

Another example of a Distributed System

Decentralised Peer-to-Peer Network (e.g. a sensor network)



- May not need a central point
- Processing is done across each device
- Global ‘knowledge’ increases
- E.g. A group of sensors are spread out across a room. One sensor can inform all other sensors that a high temperature has been sensed. Other sensors may identify that a spike in flammable gas was identified, whilst others pick up an increase in light, smoke in the atmosphere, an increase in CO2 and decrease in Oxygen, etc.

  Each sensor alone cannot determine that there is definitely a fire, but as a group, they can be very sure.

UNIVERSITY OF HULL

Distributed Computing: A Primer

What is a Distributed Application?

- An application that…
  - …spreads its execution over more than one computer or logical computer boundary
  - …uses networking to allow distributed application components to interact
  - …passes messages to communicate via the network
- For example:
  - SETI@Home : https://setiathome.berkeley.edu/
  - This is a distributed calculation of signals received from outer space
  - Everyone is sent a unique chunk of data for analysis
  - Results are stored for research purposes

UNIVERSITY OF HULL

Distributed Computing: A Primer

An physical example of a Distributed Application

Consider a problem…

I want you to take all zeros out of this number:

61237506127893815410235612035671205325472310521305210657123015025412542 3

How long might that take you as an individual? Quite some time!

But if we break the problem into 26 groups of 2 numbers and given each group to a different individual in the room – the problem is solved much quicker.

Of course there is the overhead of splitting, distributing and recompiling each number, but computers are pretty good at that and most distributed problems are far more complex!

UNIVERSITY OF HULL

Distributed Computing: A Primer

Four perspectives on Distributed Computing

Computer Science

- The study of distributed systems, architectural patterns, communication and concurrency
- Research into scalability, optimisation, security, etc.

Scientific Computing

- Analysing data
- Computing big data, linked data, etc.
- Large-scale, fast information processing and computation
- Data science

High-Performance Computing

- Parallel, distributed computation
- Incredible speed if problems can be broken down for concurrent execution
- GPUs

Real-World Distributed Computing

- Allows computation across a large physical space
- Save time/money (cloud computing)
- Reduces load on small, slow endpoint devices
- Allows common functionality or common content to be shared
- Useful for businesses and individuals

UNIVERSITY OF HULL

Distributed Computing: A Primer

The Real-World

- We are using distributed systems ever more in a world that is becoming increasingly connected
- Devices and appliances which were once thought of as standalone (e.g. lights, washing machines, central heating thermostats, refrigerators, etc.) are being increasingly relied on as part of a system of distributed computing.
- Almost all software applications written today are connected to the Internet and involve some form of distributed computation.
- Getting distributed computing right is key for us as professional developers.

Distributed Computing: A Primer

Real-World Challenges

- Distributed applications have to contend with a variety of
  - connectivity characteristics
    - bandwidth differences
    - intermittent connectivity (e.g. mobile devices)
  - usage patterns
  - hardware, operating systems and network platforms
- Applications must also
  - provide security of data and resources, particularly with mobile devices that are easily lost or stolen
  - ensure user privacy
  - be resilient to attack
- Distributed systems are inherently more complex
  - Unique problems specific to distributed computing
  - Difficult to model

UNIVERSITY OF HULL

Reflection

We've looked at:

- An introduction to the module
- A quick primer on Distributed Systems
- A quick primer on Distributed Applications
- Key concepts in Distributed Computing research and real-world use
- A quick primer on Distributed Computing in the real world

In the next session we will expand on Distributed Systems and Applications and examine design considerations