Introduction

COSC349—Cloud Computing Architecture

David Eyers

Instructors

- Teaching team
  - Lectures: David Eyers dme@cs.otago.ac.nz Owheo 1.25
  - Labs: TBC (but also David!)

- COSC349 resources are available on the public web:
  - https://cosc349.cspages.otago.ac.nz/

- (There is a Blackboard section too, but it mostly links to the cspages site so that material can be made public.)

Class reps

- Please make sure that we have some
  - It's the fifth time we're running the paper
  - … but I am always really keen to get lots of feedback anyway

Schedule

- **Lectures**—all will be recorded
  - Tuesday 13:00–13:50 (OBS117)
  - Thursday 13:00–13:50 (Biochemistry G13)
- **Labs** (separate streams)—there **is a lab** in the first week
  - Wednesday 10:00–11:50, Owheo G.38 (Lab F)
  - Wednesday 12:00–13:50, Owheo G.38 (Lab F)
- **Tutorials** (separate streams)—there **is no tutorial** in the first week
  - Tuesday 10:00–10:50
  - Thursday 10:00–10:50

Assessment

- 20% Assignment 1, due Monday 4th September
- 20% Assignment 2, due Monday 2nd October
- 60% Final Exam

- Assignments can be done individually or in pairs

Course overview: probable lecture schedule

- Intro. & overview
- emulation
- virtualisation
- paravirtualisation
- CPU support
- hypervisor
- Linux VServer and Solaris Zones
- containers
- unikernels
- IaaS
- PaaS
- SaaS
- XaaS
- elasticity
- storage
- security
- middleware
- DC networking architecture
- Software Defined Networking (SDN)
- orchestration
- legal issues & failures
- Reengineering software for cloud use
- message queues
- distributed consensus
- trusted hardware
- emerging technology

Course overview: labs

- Linked from the course webpage
  - … although many will make use of external resources
- Not assessed
- First lab tomorrow: revision of Unix shell and Git
  - You will need to use Git for your assignment work
  - You will also need to be comfortable with Unix shell script
  - If you've done COSC202, you've seen this material…
    - … however very few completed all of it: talk to me if you're bored!

Course overview: tutorials

- These will be **held as needed**
  - In particular let us know if topics have not been explained such that you are confident with them

- A good reality check to apply in terms of learning is whether or not you are comfortable with the learning objectives documented for that lecture

Reading

- There is **no specific textbook** for this paper

- The online resources from cloud providers are generally extremely good…
  - … after all, they want you to use the services that they offer!

Aim of this paper

- **Non-aim**: that you understand how to *use* the cloud
- **Aim**: you understand how the cloud is *built*
  - If you understand how to build cloud computing services, you should be able to pick up and use cloud services on offer

- Why this approach?
  - Understanding the **underlying cloud technologies** is CS
  - Also good insurance for you—if you get the fundamentals, then you will be **able to adapt** more easily to new offerings
    - The cloud computing space continues to change very rapidly!

Learning objectives for lecture 1

- You can define **cloud computing**
- You can explain essential **cloud characteristics**
- You can contrast between the **three service models**
- You understand the benefits of the different **cloud deployment models**

For starters, let's define cloud computing

- USA NIST's cloud computing definition:
- *Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models.*

NIST's five essential cloud characteristics1/3

- 1—**On-demand self-service**
  - Consumers can unilaterally provision capabilities
  - No human action is required by the service provider

- 2—**Broad network access**
  - Capabilities are available over the network
  - Capabilities are accessed using standard mechanisms
  - Multi-device support, thick or thin clients

NIST's five essential cloud characteristics2/3

- 3—**Resource pooling**
  - Cloud provider's resources use a multi-tenant model
    - Typically dedicated physical infrastructure is not provisioned
  - Location independence
    - … at least cloud tenant may not know where resources are
  - Types of resources include
    - Data storage—usually block or object storage
    - Processing—usually CPU
    - Memory—usually RAM
    - Network bandwidth

NIST's five essential cloud characteristics3/3

- 4—**Rapid elasticity**
  - Easy to provision and release capabilities
  - Indeed may happen automatically
  - Resources should appear to be unlimited to the consumer

- 5—**Measured service**
  - Resources are metered transparently in some reasonable way
  - Cloud provider can take advantage of economies of scale
  - Cloud tenant can monitor and control their resource use

NIST's three service models

- **Software as a Service** (SaaS)
  - Cloud provider hosts complete software+hardware stack
- **Platform as a Service** (PaaS)
  - Cloud provider provides programming interface to consumer
- **Infrastructure as a Service** (IaaS)
  - Cloud provider just manages underlying hardware

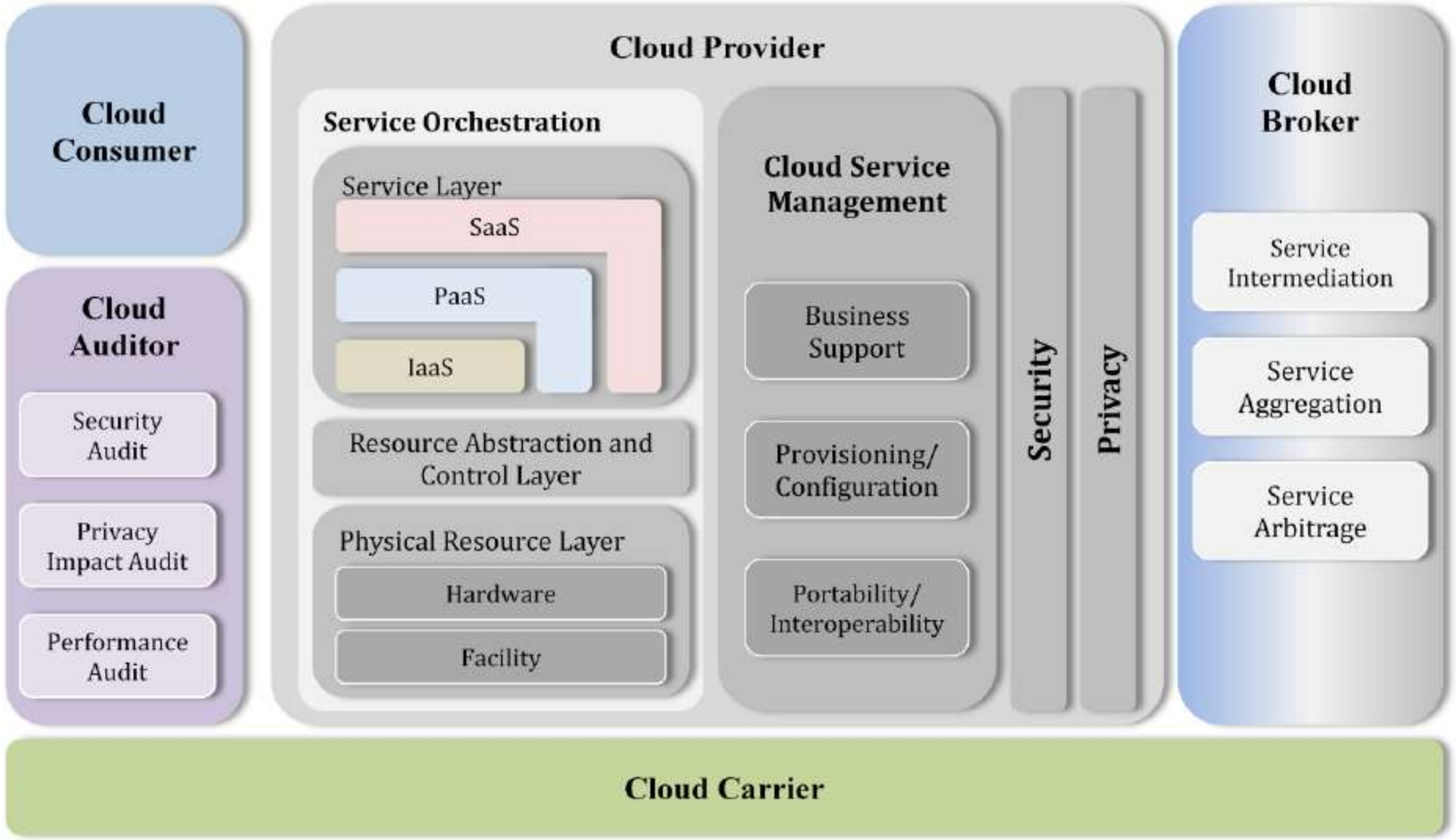
- We will return to these service models in later lectures

NIST's four deployment models

- **Private cloud**
  - Provisioned for exclusive use by one organisation
- **Community cloud**
  - Exclusive use by organisations with shared concern
- **Public cloud**
  - Open for use by the general public
  - Hosted by business and/or academic and/or government org.

- **Hybrid cloud**—some combination of the above

NIST cloud conceptual reference model



From NIST Special Publication 500-292 republished courtesy of the National Institute of Standards and Technology.