Cloud Hosting Intro

Objectives

- Explain the business and technical pros and cons of cloud computing
- Differentiate between the service models
- Explain the shared responsibility model
- Explain the difference between CapEx and OpEx
- Explain the concept of a virtual machine
- Compare and contrast a physical machine to a virtual machine
- Explain the differences between cloud deployment models
- Create a virtual machine on the cloud and deploy an application and runtime to it
- Use basic Linux commands
- Create and use an asymmetric key pair for an SSH session
- Create a container registry
- install Docker on a remote machine
- Push your image to a registry, pull the image and run a container on a remote machine

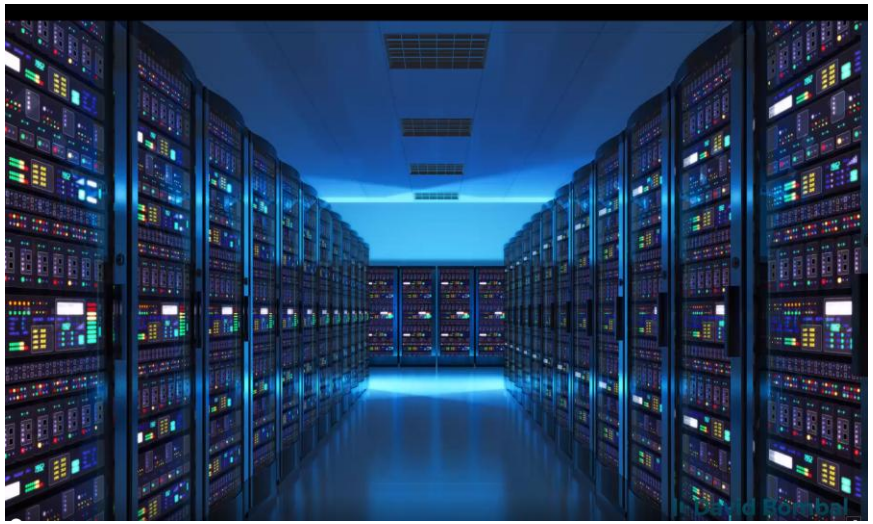
Introduction into Cloud Hosting

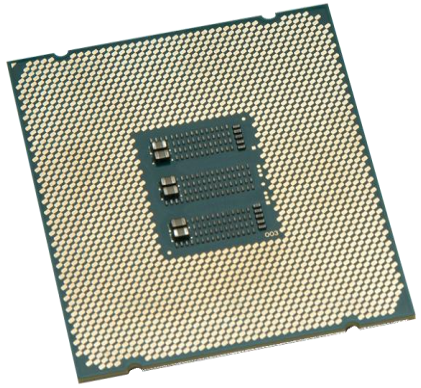
- You are now familiar with full stack development on your local machine. We now need to deploy remotely so potentially anyone in the world can access your application
- The backend can be implemented on a server in your home, in an office, in a data centre or in the cloud - as long as it is accessible from the internet
- Cloud hosting is just a load of data centres hosting lots of servers that you rent rather than buy
- An Azure and AWS data centre is generally referred to as an availability zone (AZ), there are exceptions
  - Google refers to them just as "zones"

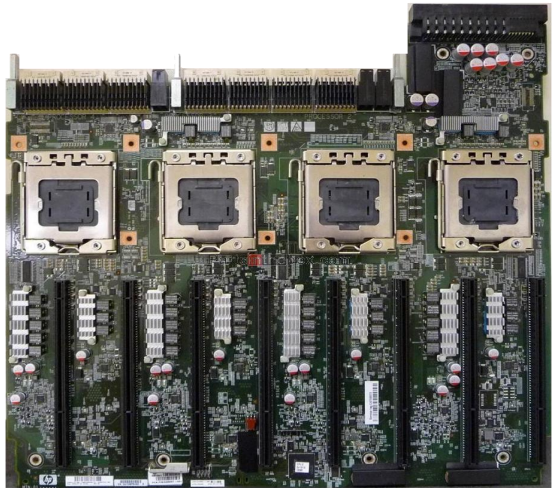
Cloud Business Model

- For an enterprise, you need a lot of compute and storage. You need premises with strong physical security, redundant power, high speed networking, racks of servers with redundant power and network cards, lots of storage for files and databases, infrastructure security and the list goes on. All this comes as a sunk cost referred to as Capital Expenditure. You pay for it then the assets depreciate over time, about 5 years. You need specialised staff to install, protect and maintain all this equipment
- If you want to scale your application, i.e. give it more CPU or Storage, you need to buy and install more stuff which can take a long time. Once you have scaled up, if the traffic for your service reduces, you are stuck with a load of kit that's hardly being used
- Amazon found themselves in this position and at times had a lot of spare capacity so why not rent it out? At least until you need it back. That idea grew into what we have today where the data centres are specifically used to rent these services
- This means businesses can replace CapEx with OpEx – i.e. operating costs which are service costs rather than depreciable asset costs. The rental model is basically pay for what you use. If you want more then pay more. If you want to scale down then do so and pay less
- The cost and hiring difficulties and permeant salary costs of all the specialist staff required to maintain and operate the infrastructure moves to cloud service provider so this is seen as a good business model
- Businesses can capitalise on economies of scale

What does "the cloud" & "on-prem" look like?





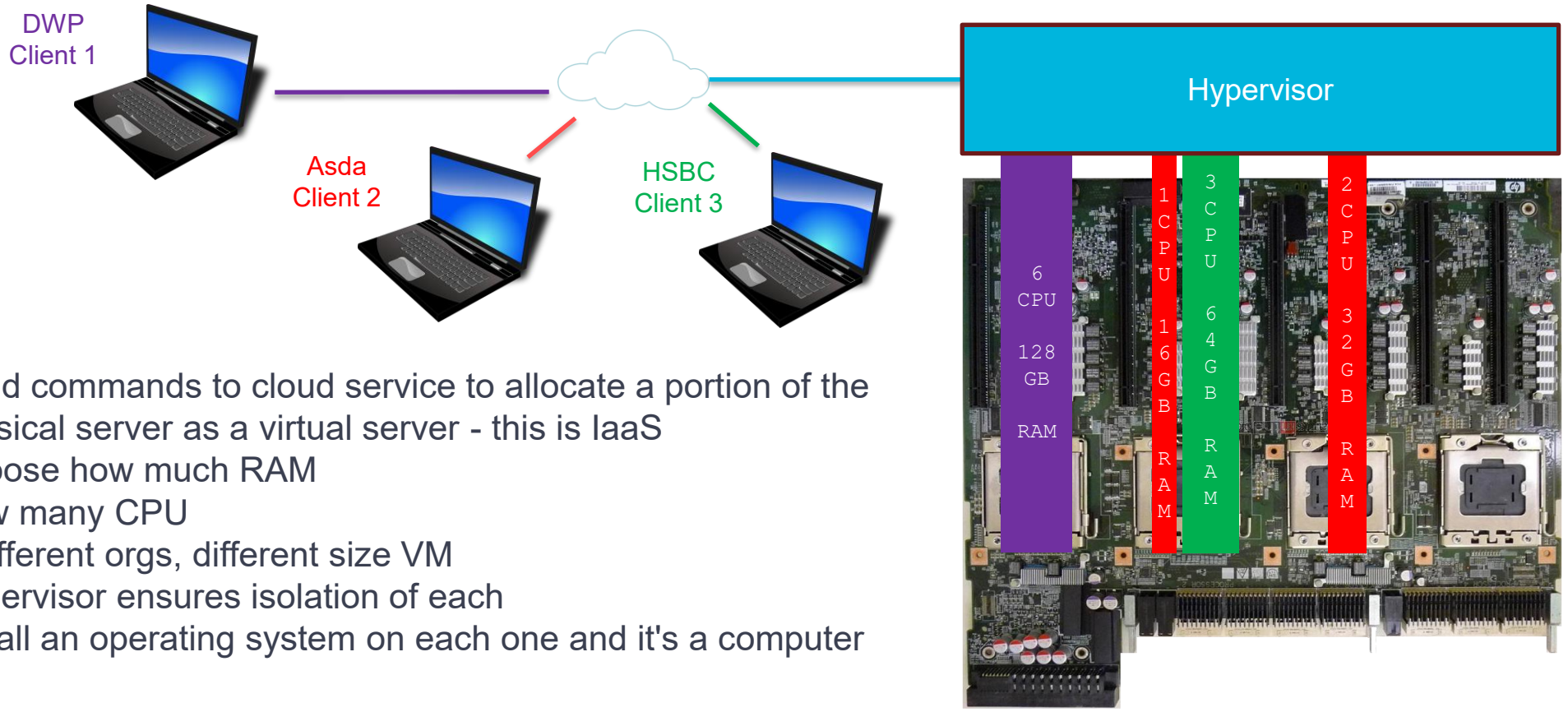


Intel Xeon (Xeon Sierra Forest CPU) 288 cores per chip - 1,152 in this server

Virtual Machines

- It would be very inconvenient for the cloud provider if we wanted to rent 10 physical servers one day then reduce to 5 the next day. Having to re-route networking and security – pulling cables and plugging them in elsewhere is infeasible – but possible (and is one model)
- You may find that the extremely powerful physical machines are too big for your application, and you can't get the most out of them and a whole machine is expensive.
- Businesses want to be able to scale their infrastructure regularly so a software approach to infrastructure architecture is required so the cloud providers can allocate machines that have been released, for immediate use for someone else
- A virtual machine is a software emulation of a physical machine hence the name virtual machine
- A physical machine can host multiple virtual machines which to the user, look just like a physical machine and although there is some performance overhead, virtual machines are now very mature and performant

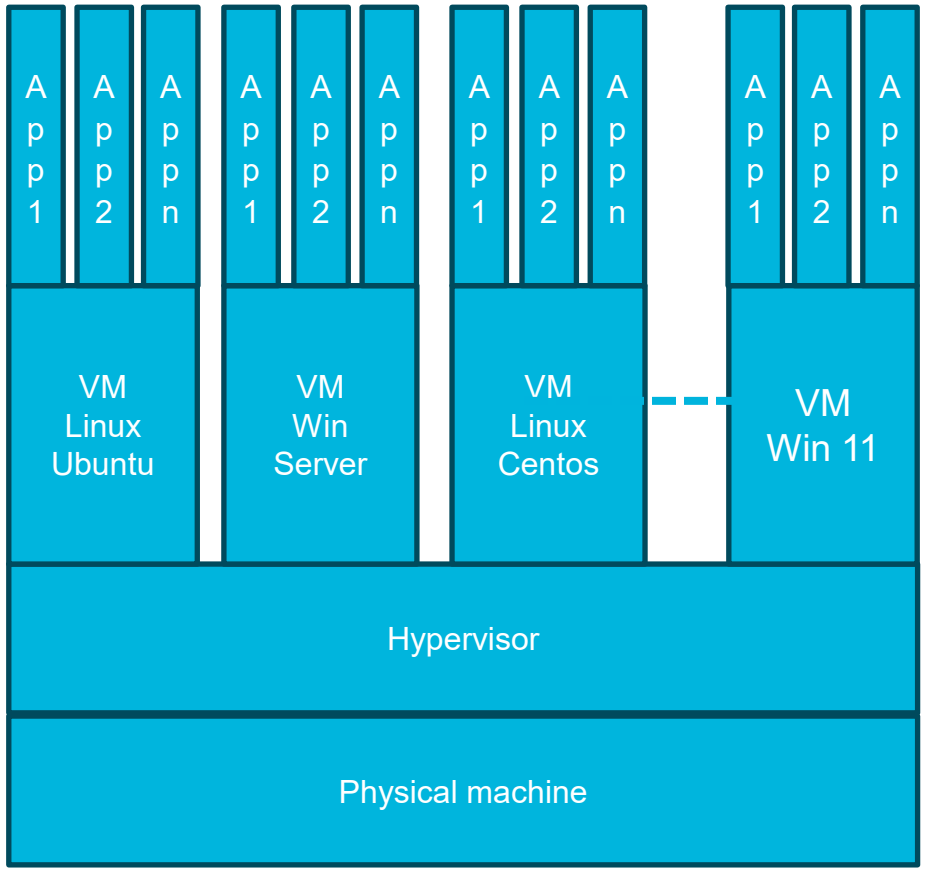
What are: Virtual Machine, Public Cloud, IaaS?



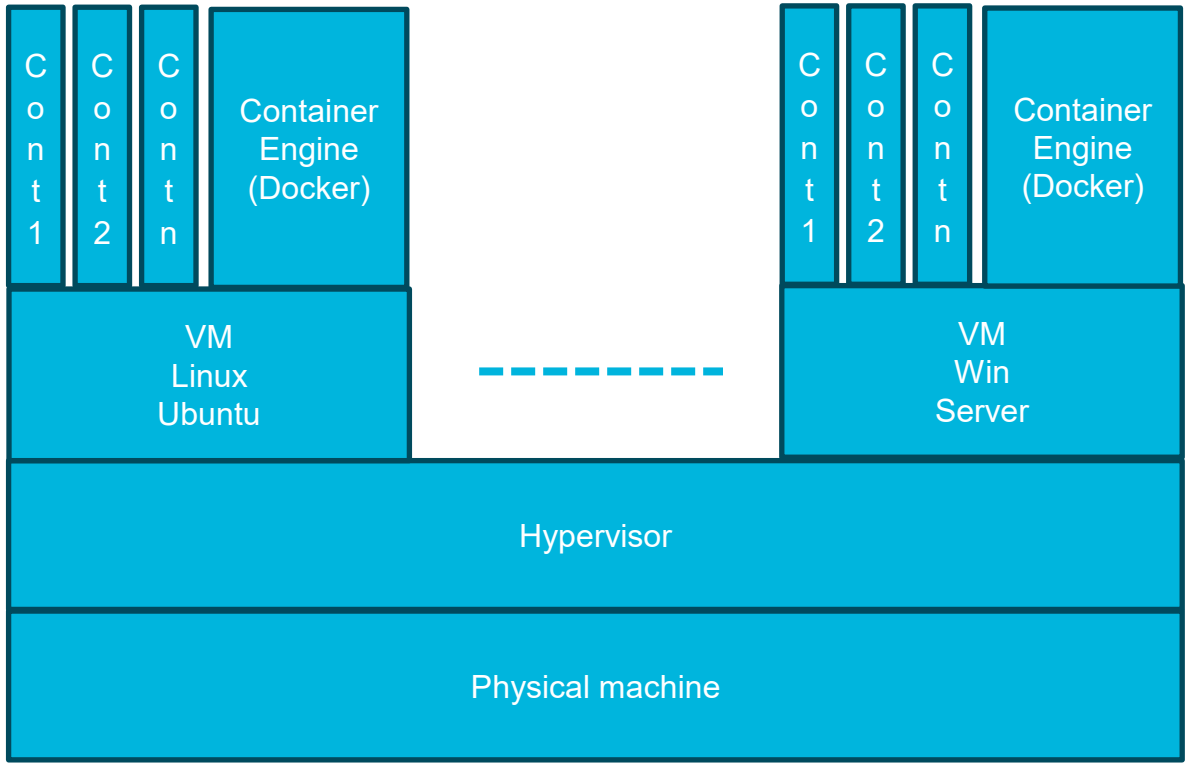
- Send commands to cloud service to allocate a portion of the physical server as a virtual server - this is IaaS
- Choose how much RAM
- How many CPU
- 3 different orgs, different size VM
- Hypervisor ensures isolation of each
- Install an operating system on each one and it's a computer

https://azure.microsoft.com/en-us/Free
https://aws.amazon.com/free

VM vs Container structure



Virtual Machines



Containers

- A VM is an isolated whole machine with its whole operating system and applications – like individual servers or PCs. They are isolated at the hypervisor level when using a type1 hypervisor as would be used in a data centre
- Because of the size of some of the operating systems – Win server 2025 requires 32GB of disk space (Ubuntu server is about 1.9GB), it can take minutes to load and start, and you often don't need most of the stuff in it
- If you only want to run a single app you have written, you don’t need all the OS tools, hundreds of drivers and bloat
- You can write your program on a laptop say, then, using a tool such as Docker, create an image and run as a container
- A container engine, e.g. Docker, runs on a single shared OS and creates a runtime environment for the containers
- The container is isolated from other containers at a process level – they are only accessible via a network IP

Cloud Deployment Models

Private cloud

- Can be on-prem to provide complete control of infrastructure
- Can be in the cloud but completely private from others. No one else shares your physical servers
- Better security and privacy. Hosted in own data centre or cloud private managed servers to control sensitive data and data sovereignty
- Downside: all admin is provided by the customer

Public cloud

- Share infrastructure with unknown others – relies on good security configuration for segregation
- Costs less as when you switch something off, someone else can use that resource
- Access from anywhere in the world
- Security and privacy is more complex and important to get right

Hybrid

- Some systems on prem and some in the cloud
- Maybe for security - or as part of a phased migration – or to keep sensitive data on prem
- Applications and data can span both
- Much more complex

Community Cloud

- A cloud environment shared by several organisations with common goals or requirements, such as security, compliance, or policy. e.g. Government, Education, healthcare

Cloud Service Models

- **IaaS**: Infrastructure as a service – you deploy and manage architecture, operating system, apps, databases etc
- Basically rent servers / memory / disk and related power, networking etc
- Fast server scaling
- No ownership of hardware - no more CapEx, move to OpEx
- Most control but also need more skilled staff such as architects, network specialists, dev-ops, maintenance team

- **PaaS**: Platform as a Service
- IaaS plus OS, tools, database etc. You just need to create and deploy the apps onto it and it will scale for you
- No need to worry about OS licenses, patching, upgrades etc.
- Less skilled resources needed – once architected, only need mainly developers. All patching etc is part of the service

- **SaaS**: Software as a Service
- Infrastructure, resilience, scaling, DR, elasticity etc is a managed service and so is the software and licenses
- Examples include: O365, SAP applications, Seibel CRM etc
- Unlike desktop apps, you can access SaaS applications from anywhere with any device that can access the internet

- **DBaaS**: Database as a service - full service between PaaS and SaaS in the cloud - e.g. MongoDB Atlas

Cloud Shared Responsibility Model