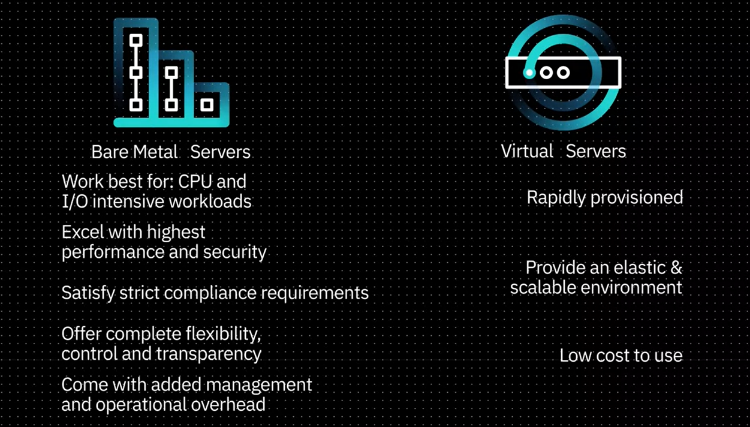
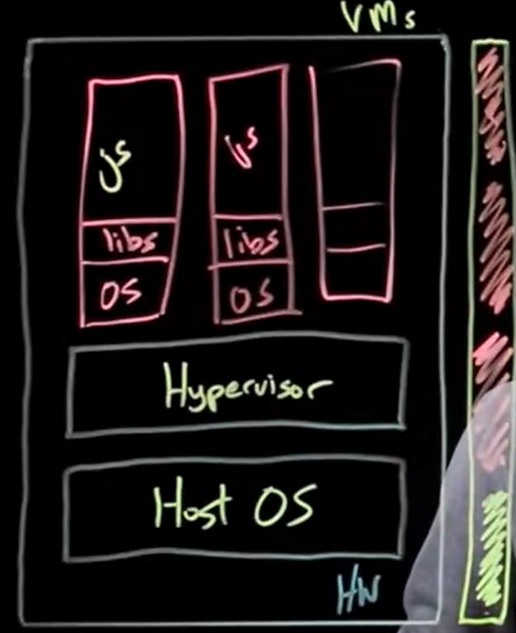
Module 3 (Components of Cloud Computing)

1. **Overview of Cloud Infrastructure:**
   1. A cloud service provider provides its cloud infrastructures/services **in different regions** (e.g. AWS has regions in US East, US Ohio, Asia East etc). This is done such that if there’s disruption in one of the regions (e.g. due to earthquake/natural disasters), the other regions won’t be affected.
   2. Each region can have **multiple Availability Zones**, each AZ has their own computing resources (e.g. in AWS -> you have US East-1 and US-East-2 zone)
   3. Cloud providers in general provide the following computing resources – **Virtual servers** (VMs – software based), **Bare metal servers** (Physical servers, not virtualized software), **Serverless computing resources** (Abstraction layer on top of VMs).
   4. Each Virtual Server or Bare metal server has their own **local storage** (which will be **deleted**, when the VS or BMS is **deleted/decommissioned**). Hence there are other types of storage options available, such as **Block storage, File storage, Object storage** (to **ensure more persistent data storage**). **Object storage** is the **most commonly used type** in the cloud… as it is resilient & highly distributed.
2. **Virtualization & Virtual Machines Explained:**
   1. Virtualization is the **process of creating a virtualized/virtual environment**
   2. This is done by using a **hypervisor**, which is a piece of software, that runs on top of the **physical server/host.**
   3. The hypervisor pulls **the resources from the physical server (e.g. processing power, storage space, memory etc) and allocates it to your virtualized environment.**
   4. **Two main types of hypervisors – type 1** (a hypervisor installed directly on the physical server – **aka bare metal hypervisors**) **and type 2 (a host OS sits between the hypervisor and host – aka hosted hypervisor**). Type 1 more common than type 2. Type 1 is more secure and lower (better) latency than type 2
   5. **You can run multiple VMs on the hypervisor.**
   6. **Benefits of Virtualization -> Cost savings, Decrease downtime, agility + speed**
   7. **Types of Virtual Machines (VMs): VMs** are also known as **Virtual Servers/Virtual Instances/Instances. VMs can be single tenant or multi tenant (multi users).**
3. **Bare Metal Servers:**
   1. A BMS is a **single tenant & dedicated physical server.**
   2. Needs to be preconfigured with certain settings first beforehand -> can take longer to provision than virtual machines/servers.
4. **Virtual Servers VS Bare Metal Servers**:
5. **Containers:**
   1. Containerization essentially eliminates the need to have guest OS -> reduce bloatware/bloating -> frees up available computing resources for other purposes
   2. See diagram below for more explanation:



VMs vs Containers (bottom the 3 -> is to create settings, then image for the container, then create the container itself with the necessary libs/dependencies for your app code to run in)

1. **Cloud Storage and Content Delivery Networks (CDNs):**
   1. Cloud storage is basically the storage of files or data in the cloud
   2. Several types - Direct-attached (Local storage), File storage, Block storage, Object storage; each have their pros and cons
   3. Understand the concept of Ephemeral (Temporary) vs Persistent storage
2. **File storage:**
   1. File storage is commonly used as a storage centre for files (e.g. departmental or centralized file storage system). It can be used as a form of data storage, but not advisable as it has poor speed -> application runs slowly.
   2. Must be attached to compute node before it can be used VIA **ETHERNET NETWORK**
   3. **SPEED VARIES; LOW SPEED if HIGH LOAD**
   4. **Can attach to MULTIPLE compute nodes at once.**
   5. **Good for workloads that don’t need fast speed (e.g. department file system)**
3. **Block storage:**
   1. Block storage breaks files into blocks/chunks of data, and stores each block separately in a unique address.
   2. Must be attached to compute node before it can be used **VIA HIGH-SPEED FIBRE NETWORK.**
   3. **HIGH SPEED**,but not suitable for shared storage between multiple servers
   4. **Only attach to ONE compute node at a time**
   5. **Good for workloads that need fast speed (e.g. databases)**
4. **Object storage:**
   1. Object storage is used to store files/objects which are static (e.g. image/video/text)
   2. **Hence, not suitable for uses in OS, databases etc, where the data is DYNAMIC**
   3. **Objects are stored in BUCKETS.**
   4. **Buckets can be subdivided into resilience options: 1) store data in ONE data centre VS 2) store data in MULTIPLE data centre (more resilient)**
   5. Does not need to be attached to a compute node before using it
   6. Instead, in object storage, we use API to upload or download the data we need
   7. Less expensive than other storage options (e.g. file/block storages)
   8. The most important thing to note about Object storage is that it’s effectively infinite:
      1. For file or block storage, you first decide how much storage space you want to use (e.g. 50GB), then pay a sum of money to use that storage space 50GB.
      2. For object storage, you operate on a pay-as-you-use basis (GB/cost)
   9. Good for storing unstructured data (data that has no hierarchical order/directory structure)
   10. **Slower than file/block storage! But much CHEAPER than file/block storage**
5. **Object storage tiers and APIs:**
   1. **Different tiers of object storage bucket tiers** – Standard tier, Vault/archive tier, Cold Vault tier (this is ordered from most frequently used to least frequently used & from most expensive to least expensive)
   2. **Uses S3 API (AWS) or S3 compatible API -> it’s a HTTP based API to connect to the bucket**
6. **Content Delivery Network (CDN):**
   1. **A CDN is a distributed server network that distributes stored/cache copies of files to users based on their geographical location.**
   2. **Originally there will a one request to the main server (e.g. in Asia) -> which sends the content/files to the CDN endpoints**
   3. **After this, if a user e.g. in USA, wants to access the content, it will send a request to the CDN endpoint in USA, instead of sending a request to the main server in Asia.**
   4. **This increases speed and reduces load on the main server (USA connect to USA CDN endpoint is much faster than USA send a request to main server in Asia)**