Google Cloud Platform: Compute Engine

- Google Compute Engine is Google's **Infrastructure-as-a-Service (IaaS)** virtual machine offering.
- Compute engine don’t have any upfront cost.
- Compute Engine allows customers to use powerful virtual machines in the cloud as server resources instead of acquiring and managing server hardware.
- Customers can configure and run a wide variety of virtual machine configurations.
- VM Instances are comprised of on **Operating System** and infrastructure resources such as **CPU, Memory, Disk**, and **Networking**.

- **Machine Types**- Machines types are templates of virtualized hardware that will be available to the VM instance. These resources include the CPU, Memory, Disk capabilities, and so on.
- Standard machine type -
  - Ideal for typical balanced instances with respect to RAM and CPU
  - Have 3.75GB of RAM per virtual CPU
- High-memory machine types -
  - Ideal for applications that require more memory
  - Have 6.5GB of RAM per virtual CPU

- Shared-core machine types -
  - These machines have one virtual CPU on a single hyper-thread of a single host CPU that is running the instance. Ideal for non-resource intensive applications.
  - Very cost effective
- Large machine types -
  - Ideal for resource-intensive workloads
  - Up to 1TB of memory
- Custom machine types - This is ideal if you have a workload that maybe requires more processing power or memory than what is offered by the Google-provided types, or if you need GPUs.

- Disks - The disk you choose will be your single root disk in which your image is loaded during the boot process. Do you choose a persistent disk or a local disk?
- Persistent Disks - Persistent disks are network-based “disks” abstracted to appear as a block device. Data is durable, meaning the data will remain as you left it after reboots and shutdowns.
- Available as either a standard hard disk drive or as a solid state drive (SSD), persistent disks are located independently of the VM instances, which means they can be detached and reattached to other instances.
- Standard persistent disks
  - Ideal for efficient and reliable block storage
  - Max 64TB per instance
  - Only available within a single zone
- SSD persistent disks
  - Ideal for fast and reliable block storage
  - Max 64TB per instance
  - Only available within a single zone

- Images - Images contain a bootloader, Operating System, file system structure, and any software customizations needed for your deployment.
- The image describes what actually gets loaded onto the root disk.
- Tons of public images are available from Google and other authorized third-party vendors.
- Google Compute Engine (GCE) uses the selected image to create a persistent boot disk for each instance.

- What if I choose a zone and want it changed afterward?
- You can absolutely move VM instances to another zone, but it will require a short outage.
- To do this manually, you'll snapshot the disk on the instance you wish to move.
- Next, create a new disk in the desired zone from the snapshot.
- Create a new VM instance in the new zone and attach the new disk.

- Preemptibility - Preemptible VM is an affordable, short-lived instance ideal for batch jobs or fault-tolerant workloads.
- They're up to 80% cheaper than regular instances, so if your application can handle random the termination of VMs at any time, then this is best option.
- Some common applications that use preemptible VMs are modeling or simulations, rendering, media transcending, big data, continuous integration, and web crawling.

Will see you in Next Lecture…

See you in next lecture …