**Containers in the cloud**

- (Introduction to containers) It explains that Infrastructure as a Service (IaaS) allows sharing compute resources using virtual machines (VMs) with individual OS environments, but this can be slow and costly as demand increases. Containers offer a solution by providing a lightweight, self-contained environment with limited access to resources. They virtualize the OS and scale like Platform as a Service (PaaS) while retaining the flexibility of IaaS. Containers make applications highly portable and scalable, allowing seamless deployment across various environments. They enable rapid scaling of applications, and when built as microservices, they can be easily connected and scaled independently across multiple hosts.

- (Kurbernetes) Kubernetes is an open-source platform for managing containerized workloads and services. It allows you to orchestrate containers on multiple hosts, scale them as microservices, and deploy rollouts and rollbacks easily. At its core, Kubernetes consists of a set of APIs to deploy containers on a cluster of nodes. Nodes represent computing instances, and Pods are the smallest units in Kubernetes, representing running processes on the cluster. Typically, each Pod contains one container, but multiple containers can be packaged together if they have a hard dependency.

Deployments represent groups of replica Pods and ensure that Pods remain running even if nodes fail. Services provide a fixed IP address for Pods and act as a load balancer to route external requests to them. Kubernetes supports declarative configuration, allowing you to specify the desired state of your applications using Deployment configuration files. To scale or update your applications, you can modify the configuration file and use the **kubectl apply** command. Kubernetes will handle the necessary changes in the desired state and update the application accordingly, providing a seamless way to manage and scale containerized applications.

( Kubernetes is an open-source platform for managing containerized workloads and services. It allows you to scale, deploy, and manage containers easily. Kubernetes uses Pods, Deployments, and Services to organize and scale applications across multiple nodes, providing seamless management and updates using declarative configuration files. )

- (Google Kubernetes Engine) Google Kubernetes Engine (GKE) is a managed Kubernetes service provided by Google Cloud. It allows you to create Kubernetes clusters with multiple Compute Engine instances. GKE offers advanced cluster management features such as load balancing, node pools for flexibility, automatic scaling and upgrades, node auto repair, and monitoring with Google Cloud's operation suite. Running applications on GKE clusters is beneficial for bridging on-premises and cloud resources. To start a Kubernetes cluster on GKE, you can use the command **gcloud container clusters create k1**.

- (Hybrid and multi-cloud) Modern hybrid or multi-cloud architectures have emerged. They allow companies to keep some parts of their infrastructure on-premises while moving other workloads to the cloud. This approach offers several benefits:

1. Flexibility: You can move specific workloads to the cloud at your own pace, without requiring a full-scale migration.
2. Scalability: Cloud services provide flexibility and scalability, allowing you to adjust resources as needed without the delays associated with on-premises upgrades.
3. Cost Savings: Utilizing cloud services for certain workloads can lead to lower computing costs and higher availability.
4. Access to Specialized Services: You can add specialized cloud services, such as machine learning, content caching, data analysis, long-term storage, and IoT, to enhance your computing resources.

- (Anthos) Anthos is a hybrid and multi-cloud solution offered by Google Cloud. It is designed to manage distributed systems and services across on-premises and multiple cloud environments. Anthos is built on top of Kubernetes and Google Kubernetes Engine (GKE) On-Prem, providing a centralized control plane for policy-based application lifecycle management.

Here's an overview of the key components and features of Anthos:

1. Google Kubernetes Engine (GKE): A managed and production-ready environment for deploying containerized applications in the cloud. It ensures high availability, scalability, and supports Certified Kubernetes for portability across different environments.
2. GKE On-Prem: The on-premises counterpart of GKE, providing a conformant version of Kubernetes with pre-loaded best-practice configurations. It offers easy upgrades to validated Kubernetes releases and integrates with various Google Cloud container services.
3. Anthos Service Mesh and Istio: Service Mesh layers that manage and secure microservices across the hybrid network. They communicate using Cloud Interconnect to synchronize and pass data.
4. Cloud Logging and Cloud Monitoring: Built-in logging and monitoring solutions for Google Cloud, offering fully managed logging, metrics collection, dashboarding, and alerting for comprehensive observability.
5. Anthos Configuration Management: Provides a single authoritative source of truth for cluster configurations through a Policy Repository (a git repository). It allows administrators and developers to enforce configurations locally in each environment and enables code changes with a single repository commit.