Kubernetes Architecture

Computer Systems and Networks - Midterm Research

Ahmed Nouralla
Security and Network Engineering Lab
Innopolis University

November 19, 2024

Overview

1. Introduction

- Kubernetes (aka. K8s or Kube) is an open-source system for automating the deployment, scaling, and management of **containerized** applications.
- Originally designed by **Google** and is currently the most used container orchestration platform.
- From a high-level perspective, K8s provides **abstractions** to manage workloads running on a cluster of machines.

2. Context

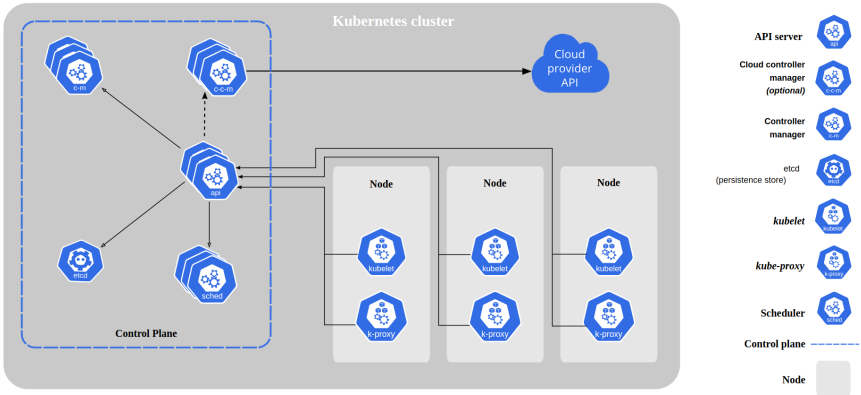

Figure: Kubernets Architecture

- The cluster is controlled and managed by a master node, storing cluster metadata in a key-value store (‘etcd‘) and exposing an API for interaction.
- A typical use case involves creating/querying objects via the API.

Examples

- ‘kubectl apply -f deployment.yaml‘ initiates application deployment
- ‘kubectl describe deployment‘ queries deployment status.

Security Consideration

Kubernetes API is the central point with full control over the cluster. If access control is improperly configured, it can give an attacker a full control over the infrastructure.

3. Main architectural drivers

Kubernetes was designed with a huge focus on

- **High Availability**: striving for no downtime through replication and fault tolerance.
- **Self-healing**: containers periodical LivenessProbes and automatic restarting.
- **Automatic Scaling**: through components such as the HorizontalPodAutoscaler.

4. Structure

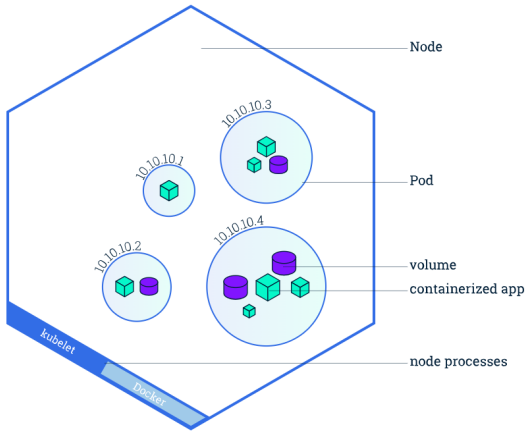

Figure: A single node in a K8s cluster may be responsible for multiple objects

Object	Overview
Pod	Represents a logical host that typically runs one containerized application but may run additional sidecar containers.
ReplicaSet	Ensures that a specified number of pod replicas are running at one time.
Deployment	Represents an application running in the cluster, provides declarative updates for Pods and ReplicaSets.
Service	Represents a network service that makes a set of pods accessible using a single DNS name and can load-balance between them.
ConfigMap	An API object used to store non-confidential data as key-value pairs that are accessible by pods (e.g., as env. vars or files).
Secret	Similar to ConfigMaps, but are specifically intended to hold confidential data (e.g., passwords and tokens).

Table: Popular k8s object kinds

Ingress	An API object that manages external access to the services in a cluster, typically HTTP.
StatefulSet	A deployment for stateful applications; provides guarantees about the ordering and uniqueness of deployed Pods.
DaemonSet	DaemonSet ensures that a copy of a certain pod (e.g., logs collector) is available on every node in the cluster.
PersistentVolume	Abstraction of a persistent storage that can use a local or remote (cloud) storage as a backend. Pods can acquire portions of that storage using a PersistentVolumeClaim
LimitRange	Enforces minimum and maximum resource usage limits per pod or container in a namespace.

Table: Popular k8s object kinds (cont.)

5. Behavior

- Kubernetes objects are managed through a **control loop**
- Essentially, the master node periodically compares object ‘status‘ vs. ‘spec‘, making changes as needed.
- For example, if a deployment specification requires ‘replicas: 3‘ and the controller sees that object status contains ‘replicas: 2‘, it will trigger the scheduling of an additional pod.

6. Rationale

- K8s started at Google as an internal cluster management tool called Borg.
- It was later open-sourced under the Cloud Native Computing Foundation (CNCF) to encourage contribution from the community as Linux containers got more popular.
- K8s current architecture was heavily influenced by its predecessors (Borg and Omega). In particular, the transition to a master-slave architecture with loosely-coupled components and the interaction only through an API.

7. Similar or competing systems

- AWS Elastic Container Service (ECS): a managed container orchestrator that integrates tightly with other AWS services, easier to setup and operate compared to K8s, but with less flexibility in multi-cloud or on-premises scenarios.

- RedHat OpenShift (Container Platform): a suite of components built around K8s with additional enterprise features (e.g., integration with popular cloud providers, build-related artifacts, and opinionated defaults).

- Hashicorp Nomad: a lightweight, flexible workload orchestrator for deploying and managing containers, virtual machines, and other workloads across data centers.

References

A. Verma, et al. (2015)
Large-scale cluster management at Google with Borg
Proceedings of the Tenth European Conference on Computer Systems. ACM. pp. 1–17.
Available: https://doi.org/10.1145/2741948.2741964

B. Burns, et al. (2016)
Borg, Omega, and Kubernetes
Commun. ACM 59, 5.
Available: https://doi.org/10.1145/2890784

Official Kubernetes Documentation
Available: https://kubernetes.io/docs/home/