Appendix A. Key Terms

By Noah Gift

This section contains select key terms that crop up often in teaching cloud computing, MLOps, and machine learning engineering:

Alerts

Alerts are health metrics that have actions associated with them. An example would be an alert that sends a text message to a software engineer when a web service returns multiple error status codes.

Amazon ECR

Amazon ECR is a container registry that stores Docker format containers.

Amazon EKS

Amazon EKS is a managed Kubernetes service created by Amazon.

Autoscaling

Autoscaling is the process of scaling load up or down automatically based on how many resources the nodes are using.

AWS Cloud9

AWS Cloud9 is a cloud-based development environment running in AWS. It has special hooks for developing serverless applications.

AWS Lambda

A serverless compute platform by AWS that has FaaS capability.

Azure Container Instances (ACI)

Azure Container Instances is a managed service from Microsoft that allows you to run container images without managing servers to host them.

Azure Kubernetes Service (AKS)

Azure Kubernetes Service is a managed Kubernetes service created by Microsoft.

`black`

The `black` tool formats the text of Python source code automatically.

Build server

A build server is an application that works in both the testing and deployment of software. Popular build servers can be both SaaS or open source. Here are a few popular options:

- Jenkins is an open source build server that can run anywhere including AWS, GCP, Azure, or a docker container or on your laptop.
- CircleCI is a SaaS build service that integrates with a popular Git hosting provider like GitHub.

CircleCI

A popular SaaS (software as a service) build system used in DevOps workflows.

Cloud native applications

Cloud native applications are services that utilize the unique capabilities of the cloud, like serverless.

Container

A container is a set of processes that are isolated from the rest of the operating system. They are often megabytes in size.

Continuous delivery

Continuous delivery is the process of delivering tested software automatically to any environment.

Continuous integration

Continuous integration is the process of automatically testing software upon check-in to the source control system.

Data engineering

Data engineering is the process of automating the flow of data.

Disaster recovery

Disaster recovery is the process of designing a software system to recover despite a disaster. This process could include archiving data to another location.

Docker format container

There are several formats for containers. An emerging form is Docker, which involves the definition of a *Dockerfile*.

Docker

Docker is a company that creates container technology, including an execution engine, collaboration platform via DockerHub, and a container format called *Dockerfile*.

FaaS (function as a service)

A type of cloud computing that facilitates functions that respond to events.

Google GKE

Google GKE is a managed Kubernetes service created by Google.

IPython

The `ipython` interpreter is an interactive terminal for Python. It is the core of the Jupyter notebook.

JSON

JSON stands for JavaScript Object Notation, and it is a lightweight, human-readable data format used heavily in web services.

Kubernetes clusters

A Kubernetes cluster is a deployment of Kubernetes that contains an entire ecosystem of Kubernetes components, including nodes, pods, the API, and containers.

Kubernetes containers

A Kubernetes container is a Docker image that deploys into a Kubernetes cluster.

Kubernetes pods

A Kubernetes pod is a group of one or more containers.

Kubernetes

Kubernetes is an open source system for automating the operations of containerized applications. Google created it and open-sourced in 2014.

Load testing

Load testing is the process of verifying the scale characteristics of a software system.

Locust

Locust is a load-testing framework that accepts Python-formatted load test scenarios.

Logging

Logging is a process of creating messages about the running state of a software application.

Makefile

A `Makefile` is a file that contains a set of directives used to build software. Most Unix and Linux operating systems have built-in support for this file format.

Metrics

Metrics are the creation of KPIs (Key Performance Indicators) for a software application. An example of a parameter is the percentage of CPU used by a server.

Microservice

A microservice is a lightweight, loosely coupled service. It can be as small as a function.

Migrate

Migrate is the ability to move an application from one environment to another.

Moore's Law

The perception that for some time, the number of transistors on a microchip doubles every two years.

Operationalization

The process of making an application ready for production deployment. These actions could include monitoring, load testing, and setting up alerts.

pip

The `pip` tool installs Python packages.

Ports

A port is a network communication endpoint. An example of a port is a web service running on port 80 via the protocol HTTP.

Prometheus

Prometheus is an open source monitoring system with an efficient time-series database.

pylint

The `pylint` tool checks the Python source code for syntax errors.

PyPI

The Python Package Index, where published packages are available to install with tools like `pip`.

pytest

The `pytest` tool is a framework for running tests on Python source code.

Python virtual environment

A Python virtual environment is created by isolating a Python interpreter to a directory and installing packages in that directory. The Python interpreter can perform this action via `python -m venv yournewenv`.

Serverless

Serverless is a technique of building applications based on functions and events.

SQS queue

A distributed messaging queue built by Amazon with near-infinite reads and writes.

Swagger

A swagger tool is an open source framework that simplifies the creation of API documentation.

Virtual machine

A virtual machine is the emulation of a physical operating system. It can be gigabytes in size.

YAML

YAML is a human-readable serialization format often used in configuration systems. It is easily portable to JSON format.