

GLOSSARY OF

Cloud Native Terms

...native ap
...container registry
...ops developer-ready
...Google Cloud Platform
...Infrastructure I image inf
...measured microservices micro
...PaaS PCF PKS Photon OS platform
...private cloud ...quality of service R RabbitMQ RBAC R
...service service ...e-defined data center (SDDC) software a
...ditional application V Vagrant Volume W workload X XML Y YARN YAN
...uild BOSH C Cassandra cloud computing cloud Infrastructure cloud na
...erized application containers as a service container host container reg
...infrastructure D day one day two desired state DevOps developer-rea
...ance Fluentd G GCP open service broker Gemfire Google Cloud Platfor
...od autoscaler hybrid cloud hyper-converged infrastructure I image inf
...let Kubo L LDAP Lightwave linkerd M Memcached microservices micro
...inationated P platform orchestration PaaS PCF PKS Photon OS platform
...private cloud Prometheus pull Q quality of service R RabbitMQ RBAC R
...service discovery software-defined data center (SDDC) software as a
...ditional application V Vagrant Volume W workload X XML Y YARN YAN
...uild BOSH C Cassandra cloud computing cloud Infrastructure cloud na

SH **C** Cassandra cloud computing cloud Infrastructure cloud native applications Cloud Spanner cluster CNCF CNI Concourse CoreDNS containerize containerized application containers as a service container host container registry controllers CI/CD continuous integration continuous delivery converged infrastructure **D** day one day two desired state DevOps developer-ready infrastructure digital transformation Docker Docker Swarm **E** elastic ELK stack etcd **F** fault tolerance Fluentd GCP open service broker Gemfire Google Cloud Platform Google Kubernetes Engine Greenplum Database GRPC **H** Hadoop Harbor Helm Chart horizontal pod autoscaler hybrid cloud hyper-converged infrastructure **I** image infrastructure as a service (IaaS) ingress **J** Jaeger JSON **K** K8s KaaS Kafka kubernetes kubectl kubelet Kubo **L** LDAP Lightwave linkerd **M** Memcached microservices microsegmentation Minikube MongoDB MySQL **N** namespace NodePort NoSQL NSX **O** OCI OpenTracing opinionated **P** platform orchestration PaaS PCF PKS Photon OS platforms platform as a service platform developer platform operator pod PostgreSQL prescriptive platform private cloud Prometheus pull **Q** quality of service **R** RabbitMQ RBAC Redis replica set repository refactoring repackaging replatforming rkt runC **S** service service discovery software-defined data center (SDDC) software as a service (SaaS) Spanner Spark spec Spring Cloud Data Flow StatefulSet **T** tag the cloud traditional application **V** Vagrant Volume **W** workload **X**ML **Y** YARN YAML **Z** ZooKeeper 12-factor app **A** ACID Active Directory API server Azure Container Registry **B** build BOSH **C** Cassandra cloud computing cloud Infrastructure cloud native applications Cloud Spanner CNCF CNI Concourse CoreDNS container containerize containerized application containers as a service container host container registry controllers CI/CD continuous integration continuous delivery continuous deployment converged infrastructure **D** day one day two desired state DevOps developer-ready infrastructure digital transformation Docker Docker Swarm **E** elastic ELK stack etcd **F** fault tolerance Fluentd **G** GCP open service broker Gemfire Google Cloud Platform Google Kubernetes Engine Greenplum Database PC **H** Hadoop Harbor Helm Chart horizontal pod autoscaler hybrid cloud

Introduction

This glossary presents definitions for terminology in the **cloud native space**. The definitions are not intended to be axiomatic, dictionary-style definitions but rather **plain-language descriptions** of **what a term means** and an explanation of why the **technology associated with it** matters. For some of the terms, meaning varies by usage, situation, perspective, or context.



List of TERMS



ACID, 8
Active Directory, 8
Admiral, 8
agile software
 development, 9
AKS, 9
API-first, 9
API server, 9
Azure Container Registry, 9

A

B

build, 10
BOSH, 10

Calico, 10
Cassandra, 10
Clarity, 10
cloud computing, 11
Cloud Foundry Container
 Runtime, 11
cloud infrastructure, 11
cloud native
 applications, 12
Cloud Spanner, 12
cluster, 12
CNCF, 13
CNI, 13
Concourse, 13
CoreDNS, 13
container, 13
containerize, 14
containerized
 application, 14
containers as
 a service, 14
container host, 14
container registry, 14
controllers, 14
CI/CD, 14
continuous integration, 14
continuous delivery, 15
continuous
 deployment, 15
converged
 infrastructure, 15
CredHub, 15

C

D

- day one, 15
- day two, 15
- desired state, 15
- DevOps, 15
- developer-ready
 - infrastructure, 16
- Diego, 16
- digital transformation, 16
- Dispatch, 16
- Docker, 17
- Docker Swarm, 17

E

- elastic, 17
- ELK stack, 17
- etcd, 17

F

- fault tolerance, 18
- Flannel, 18
- Fluentd, 18
- function as a service, 18

G

- GCP open service
 - broker, 19
- Gemfire, 19
- Go, 19
- Google Cloud
 - Platform, 19
- Google Kubernetes
 - Engine, 19
- Greenplum Database, 19
- GRPC, 19

H

- Hadoop, 20
- Harbor, 20
- Hatchway, 20
- Helm Chart, 20
- horizontal pod
 - autoscaler, 21
- hybrid cloud, 21
- hyper-converged
 - infrastructure, 21

I

- image, 21
- infrastructure as a service
 - (IaaS), 21
- ingress, 22
- Istio, 22

J

- Jaeger, 22
- JSON, 22

K

- K8s, 23
- KaaS, 23
- Kafka, 23
- kops, 23
- Kubernetes, 23
- kubectrl, 23
- kubelet, 24
- Kubo, 24

L

LDAP, 24
Lightwave, 24
linkerd, 25

M

MANO, 25
Memcached, 25
microservices, 25
microsegmentation, 26
Minikube, 26
MongoDB, 26
Multicloud, 26
MySQL, 26

N

namespace, 27
namespace isolation, 27
network functions
 virtualization, 27
Node.js, 27
NodePort, 27
NoSQL, 28
NSX, 28

O

OCI, 28
OpenStack, 28
OpenTracing, 28
opinionated platform, 29
orchestration, 29
overlay network, 29

P

PaaS, 29
PAS, 29
PCF, 29
Photon OS, 29
PKS, 29
platforms, 30
platform as a service, 30
platform developer, 30
platform operator, 31
platform reliability
 engineer, 31
pod, 31
PostgreSQL, 31
prescriptive platform, 32
private cloud, 32
Prometheus, 32
pull, 32

Q

quality of service, 33

R

RabbitMQ, 33
RBAC, 33
Redis, 33
registry, 34
replica set, 34
repository, 34
refactoring, 34
repackaging, 34
replatforming, 34
rkt, 34
runC, 34

scheduler, 35
service, 35
service discovery, 35
service mesh, 35
sidecar, 36
software-defined data center (SDDC), 36
software as a service (SaaS), 36
Spanner, 36
Spark, 36
spec, 37
Spring, 37
Spring Cloud
 Data Flow, 37
StatefulSet, 37
swarm, 37

S

tag, 38
Tern, 38
Terraform, 38
the cloud, 38
traditional application, 38

T

UID, 39
underlay network, 39

U

Vagrant, 39
virtual infrastructure manager, 39
VMware Integrated OpenStack, 39
Volume, 39
vSphere Cloud Provider, 40
vSphere Integrated Containers, 40

V

workload, 41

W

XML, 41

X

YARN, 42
YAML, 42

Y

ZooKeeper, 42

Z

12-factor app, 43

#

A

ACID

ACID stands for Atomicity, Consistency, Isolation, and Durability—properties of database transactions that, taken together, guarantee the validity of data in the face of power failures or system errors.

Active Directory

Microsoft Active Directory (AD) is a directory service that authenticates users and controls access to personal computers, servers, storage systems, applications, and other resources. An Active Directory domain controller combines a Kerberos key distribution center (KDC) with an LDAP server to provide authentication and authorization. To authenticate the identity of users, AD uses the highly secure Kerberos protocol or the legacy NT LAN Manager (NTLM). To authorize access to resources, AD typically uses a Privilege Attribute

Certificate (PAC), which is a data structure in a Kerberos ticket that contains group memberships, security identifiers, and other information about a user's profile. See *LDAP*.

Admiral

An open source project from VMware, Admiral is a container management portal. It supplies a user interface for DevOps teams and others to provision and manage containers. With Admiral, you can use Docker compose to combine different containers into an application. On vSphere Integrated Containers, you can use Admiral to manage container hosts and apply governance to their usage, including capacity quotas. Admiral can also show metrics and other information about container instances. See *vSphere Integrated Containers*.

agile software development

A methodology that emphasizes iterative, incremental, collaborative development using cross-functional teams optimized for rapidly responding to changing requirements derived from feedback, learning, and new information.

AKS

Azure Container Service (AKS) is Microsoft's managed Kubernetes service that runs in Azure.

API-first

A guideline for developing software that makes the API the cornerstone of the application. From the beginning of the software-development project, the emphasis is on building an API to be consumed by client applications and services, and every functional requirement is to be fulfilled through the API.

API server

In Kubernetes, the API server provides a frontend that handles REST requests and processes data for API "objects," such as pods, services, and replication controllers.

Azure Container Registry

ACR is a private image registry from Microsoft that includes geo-replication.

B

build

With Docker, it is the process of building Docker images by using a Dockerfile. In the context of the CI/CD pipeline, the build process generates an artifact, such as a set of binary files that contain an application.

BOSH

An open source system that unifies release engineering, deployment, and lifecycle management for large distributed systems. BOSH performs monitoring, failure recovery, and software updates with zero-to-minimal downtime. Just as Kubernetes maintains the desired state of a containerized application, BOSH maintains the desired state of the underlying infrastructure, including Kubernetes itself, on which the application runs.

C

Calico

Formally known as Project Calico, this open source project provides a software-defined Layer 3 network fabric for containers. It integrates with Kubernetes, OpenStack, and other cloud platforms.

Cassandra

A NoSQL database, Apache Cassandra manages structured data distributed across commodity hardware. Common use cases include recommendation and personalization engines, product catalogs, play lists, fraud detection, and message analysis.

Clarity

An open source framework from VMware formally known as Project Clarity, it is a design system for crafting cohesive, interactive user experiences in web applications. Clarity includes an HTML and CSS framework, user experience guidelines, and Angular components.

cloud computing

Cloud computing is an umbrella term for elastic, on-demand, shared computing resources and services—such as computational power, storage capacity, database usage, analytics, and software applications—delivered as a service over the Internet, typically with metered pricing. The organizations that provide cloud computing are frequently referred to as *cloud providers*. See also the definitions of the three cloud-computing service models: *IaaS*, *PaaS*, and *SaaS*. For a formal definition and taxonomy of cloud computing, see the *The NIST Definition of Cloud Computing*, NIST Special Publication 800-145, at <https://csrc.nist.gov/publications/detail/sp/800-145/final>.

Cloud Foundry Container Runtime

Formerly called Kubo, for Kubernetes on BOSH, CF CR is an open source project for deploying and managing Kubernetes by using BOSH. For more information on CF CR, see [CloudFoundry.org](https://www.cloudfoundry.org). See also: *BOSH*.

cloud infrastructure

Encompasses the servers, virtual machines, storage systems, networking, and other components required for cloud computing and infrastructure as a service. Cloud infrastructure provides the building blocks, or primitives, for creating hybrid and private clouds that deliver cloud computing services.



cloud native applications

Generally speaking, cloud native applications are apps that are developed and optimized to run in a cloud as distributed applications. More specifically, according to the Cloud Native Computing Foundation, cloud native applications, which are also generally referred to as “modern” applications, are marked by the following characteristics:

- Containerized for reproducibility, transparency, and resource isolation.
- Orchestrated to optimize resource utilization.
- Segmented into microservices to ease modification, maintenance, and scalability.

Different organizations, however, have different definitions. Dell EMC, for example, defines *cloud native application* as a highly scalable next-generation distributed application architecture that uses open standards and is dynamic in nature.

Cloud native applications are typically developed and deployed on a containers as a service platform (CaaS) or a platform as a service (PaaS). See also: *12-factor app*.

Cloud Spanner

A globally distributed, strongly consistent database service that combines the benefits of a relational database structure with non-relational horizontal scale.

cluster

Three or more interconnected virtual machines or physical computers that, in effect, form a single system. A computer in a cluster is referred to as a node. An application running on a cluster is typically a distributed application because it runs on multiple nodes. By inherently providing high availability, fault tolerance, and scalability, clusters are a key part of cloud computing.

CNCF

Cloud Native Computing Foundation. An open source project hosted by the Linux Foundation, the CNCF hosts Kubernetes and other key open source projects, including Prometheus, OpenTracing, Fluentd, and linkerd. VMware is a member of the Linux Foundation and the Cloud Native Computing Foundation.

CNI

Container Network Interface. It is an open source project hosted by the CNCF to provide a specification and libraries for configuring network interfaces in Linux containers.

Concourse

Concourse is a system for continuous integration and continuous delivery that works with Pivotal Cloud Foundry and other platforms to help enterprise development teams release software early and often. Note that in the context of Concourse, the *D* in *CI/CD* stands for delivery,

not deployment. Concourse automates the testing and packaging of frequent code commits. See *CI/CD*.

CoreDNS

An open source project, CoreDNS can integrate with Kubernetes, etcd, Prometheus, and other software to provide DNS and service discovery with plugins. CoreDNS is hosted by the CNCF.

container

A portable, executable format, known as an image, for packaging an application with all its dependencies and instructions on how to run it. When the container image is executed, it runs as a process on a computer or virtual machine with its own isolated, self-described application, file system, and networking. The use of containers is increasing because they provide a portable, flexible, and predictable way of packaging, distributing, modifying, testing, and running applications. Containers speed up software development and deployment.



containerize

To package an application in a container.

containerized application

An application that has been packaged to run in one or more containers.

containers as a service

A container-as-a service platform helps developers build, deploy, and manage containerized applications, typically by using Kubernetes or another orchestration framework, such as Mesos or Docker Swarm.

container host

A Linux operating system optimized for running containers. Examples include CoreOS and Project Photon OS by VMware.

container registry

See *registry*.

controllers

In Kubernetes, controllers are processes started by the Kubernetes Controller Manager to perform the routine tasks associated with managing a cluster.

CI/CD

Refers to either the continuous integration and continuous delivery pipeline or the continuous integration and continuous deployment pipeline. Context often, but not always, disambiguates the abbreviation. See *continuous integration*, *continuous deliver*, and *continuous deployment*.

continuous integration

Constantly combines source code from different developers or teams into an app and then tests it.

D

continuous delivery

Readies an application or part of an application for production by packaging and validating it.

continuous deployment

Automatically deploys an application or part of an application into production.

converged infrastructure

Technology that brings together the disparate infrastructure elements powering IT, including servers, data storage devices, networking functions, virtualization, management software, orchestration, and applications. See *hyper-converged infrastructure*.

CredHub

A component of Cloud Foundry that manages and securely stores credentials like passwords, certificates, certificate authorities, and keys. CredHub is deployed by BOSH. See *Pivotal Cloud Foundry* and *BOSH*.

day one

Refers to deployment.

day two

Refers to post-deployment operations.

desired state

A key benefit of Kubernetes is that it automatically maintains the *desired state*—the state that an administrator or platform operator specifies an application should be in.

DevOps

Delivering software in an expedient, reliable, sustainable way requires collaboration between IT teams and developers. DevOps takes place when developers and IT come together to focus on operations in the name of streamlining and automating development and deployment. DevOps is a key practice driving the development and deployment of cloud native applications.

D

developer-ready infrastructure

VMware vSphere, VMware NSX, VMware vSAN, and VMware vRealize Operations lays the foundation for a software-defined data center (SDDC). Running VMware Pivotal Container Service or Pivotal Cloud Foundry on top of a VMware SDDC, for example, produces developer-ready infrastructure—agile, self-service infrastructure that is ready to use to build and run cloud native applications.

Diego

The container management system for Pivotal Cloud Foundry.

digital transformation

Optimizing the use of your computing resources, organizational processes, and software development practices to extend your enterprise’s adaptability, productivity, innovation, competitive advantage, and global reach. At a high level, digital transformation often entails

the adoption of new technologies, including cloud computing, mobile devices, social media, and big data analytics. At a lower level, cloud native technologies and practices—such as containers, Kubernetes, microservices, container platforms, DevOps, and the CI/CD pipeline—converge into a powerful recipe for digital transformation.

Dispatch

An open source framework from VMware, Dispatch manages serverless applications and services built using functions. As a serverless solution, Dispatch uses a flexible driver interface to integrate with one or more functions-as-a-service implementations. Dispatch manages function runtimes and containers built on Photon OS. See *function as a service*.

E

Docker

Docker is a widely used container format. Docker defines a standard format for packaging and porting software, much like ISO containers define a standard for shipping freight. As a runtime instance of a Docker image, a container consists of three parts:

- A Docker image
- An environment in which the image is executed
- A set of instructions for running the image

Docker Swarm

Is the name of a standalone native clustering tool for Docker. Docker Swarm combines several Docker hosts and exposes them as a single virtual Docker host. It serves the standard Docker API, so any tool that already works with Docker can transparently scale up to multiple hosts.

elastic

A resource or service that can dynamically expand or contract to meet fluctuations in demand.

ELK stack

Elasticsearch, Logstash, and Kibana combine to form the ELK stack. Taken together, these three open source projects provide a platform to collect, search, analyze, and visualize data. Elasticsearch is a distributed search and analytics engine that lets data engineers query unstructured, structured, and time-series data. Logstash lets you collect unstructured data, enrich it, and route it to another application, such as Elasticsearch. Kibana is a visualization engine to display data in dashboards as graphics and maps.

etcd

A distributed key-value store that Kubernetes uses to store data about its state and configuration.

F

fault tolerance

Fault tolerance is the property that lets a system continue to function properly in the event of component failure.

Flannel

Flannel furnishes an overlay network for containers running in Kubernetes. See *overlay*.

Fluentd

A data collector for unified logging. Fluentd, which works with cloud native applications, is hosted by the CNCF.

function as a service

FaaS is a cloud computing model that lets you run and manage application functions without managing a traditional server as the application's backend—giving rise to the nomenclature “serverless.” The functions typically respond to events, making FaaS a useful method of processing events in

the context of the Internet of things. Here's an example: When you wake up in the morning, your smart watch monitors your blood sugar. When it reaches a certain level, the watch sends an event to a remote, vendor-managed server in the cloud, which uses a FaaS framework to process the event with a function and send a request to your networked microwave oven, triggering it to start cooking your oatmeal. AWS Lambda is an example of a commercial serverless platform. OpenFaaS is an example of a FaaS framework for building serverless functions with Docker and Kubernetes. See *Dispatch*.

G

GCP open service broker

It lets apps access Google cloud APIs from anywhere.

Gemfire

Pivotal Gemfire is a distributed data management platform that compresses operational data and holds it in memory to provide real-time, consistent, and scalable access to data-intensive NoSQL applications.

Go

Created at Google, Go is a programming language popular with developers who build cloud native technology. Docker and Kubernetes are written in Go, which is sometimes referred to as golang.

Google Cloud Platform

GCP.

Google Kubernetes Engine

It is a managed environment to deploy and scale containerized applications that are orchestrated by Kubernetes.

Greenplum Database

An ACID-compliant transactional database that employs a shared-nothing, massively parallel processing architecture, Pivotal Greenplum complies with SQL standards. It interoperates with industry-standard business intelligence and ETL tools as well as Hadoop. With a library of analytics functions and a framework for building custom functions, Greenplum addresses data warehousing use cases for big data.

GRPC

A project of the CNCF, GRPC is a open-source universal remote procedure call (RPC) framework for distributed systems. You can use it to define a service by using Protocol Buffers, a binary serialization language. GRPC also lets you automatically generate client and server stubs for a service in various languages.

H

Hadoop

Hadoop comprises the Hadoop Distributed File System (HDFS) and MapReduce. HDFS is a scalable storage system built for Hadoop and big data. MapReduce is a processing framework for data-intensive computational analysis of files stored in a Hadoop Distributed File System. Apache Hadoop is the free, open-source version of Hadoop that is managed by the Apache Software Foundation. The open-source version provides the foundation for several commercial distributions, including Hortonworks, IBM Open Platform, and Cloudera. There are also Hadoop platforms as a service. Microsoft offers HDInsight as part of its public cloud, Azure. Amazon Elastic MapReduce, or EMR, delivers Hadoop as a web service through AWS.

Harbor

An open source project donated by VMware to the CNCF, Harbor is a secure registry that hosts repositories of container images.

Hatchway

An open source storage project from VMware formally known as Project Hatchway, it provides storage infrastructure options for containers in vSphere environments, including hyper-converged infrastructure (HCI) with VMware vSAN. Hatchway integrates with Kubernetes to let you consume storage infrastructure as code. Data services like snapshot, cloning, encryption, deduplication, and compression can be made available at the level of a container volume. See <https://vmware.github.io/hatchway/>. See also *vSphere Cloud Provider*.

Helm Chart

A package of Kubernetes resources that are pre-configured, customized, and reproducible; you can then manage a chart with the Helm tool. The charts help improve the portability of Kubernetes applications. A single chart can contain an entire web application, including databases, caches, HTTP servers, and other resources.



horizontal pod autoscaler

In Kubernetes, a horizontal pod autoscaler is a controller that adds resources to handle an increase in demand when the requests to a service exceed the threshold set by the administrator.

hybrid cloud

Any modernized infrastructure that involves two or more delivery models, such as private cloud and public cloud resources.

hyper-converged infrastructure

Integrates the same key types of IT components that converged infrastructure does, but in a scalable rack or appliance that simplifies management, improves performance, and adds elastic scalability. See *converged infrastructure*.

image

With Docker, an image is the basis of a container. An image specifies changes to the root file system and the corresponding execution parameters that are to be used in the container runtime. An image typically contains a union of layered files systems stacked on top of each other. An image does not have state and it never changes.

infrastructure as a service (IaaS)

Infrastructure-as-a-service (IaaS) provides on-demand access to underlying IT infrastructure, including resources for storage, networking, and compute. With IaaS, a user can provision IT services when they need them to deploy and run arbitrary software. Users typically pay only for the resources they consume. The user, however, does not manage or control the underlying cloud infrastructure. See *cloud computing*.

I

ingress

In Kubernetes, ingress refers to an API object that controls external access to the services in a Kubernetes clusters, such as HTTP and HTTPS. Ingress can perform load balancing.

Istio

A platform that deploys a service mesh to connect, manage, and secure microservices on Kubernetes. Istio intercepts network communications among the microservices that make up a containerized application deployed on Kubernetes to manage the microservices as they interact. See also *microservices*, *sidecar* and *service mesh*.

J

Jaeger

A distributed tracing system released as open source software by Uber Technologies, Jaeger can monitor microservice-based architectures. Use cases include distributed transaction monitoring, root cause analysis, service dependency analysis, and performance optimization. Jaeger is hosted by the CNCF.

JSON

JavaScript Object Notation is a minimalist data-interchange format commonly used to annotate data, such as API output.

K

K8s

An abbreviation of sorts for Kubernetes.

KaaS

Kubernetes as a service.

Kafka

Apache Kafka partitions data streams and spreads them over a distributed cluster of machines to coordinate the ingestion of vast amounts of data for analysis. More formally, Kafka is a distributed publish-subscribe messaging system. A key use of Kafka is to help Spark or a similar application process streams of data. In such a use case, Kafka aggregates the data stream—for example, log files from different servers—into “topics” and presents them to Spark Streaming, which analyzes the data in real time.

kops

This term stands for Kubernetes Operations, a command-line tool to help you install, maintain, and upgrade Kubernetes clusters.

Kubernetes

An orchestration system that automates the deployment and management of containerized applications. As an application and its services run in containers on a distributed cluster of virtual or physical machines, Kubernetes orchestrates all the moving pieces to optimize the use of computing resources, to maintain the desired state, and to scale on demand. Kubernetes is also referred to as an orchestration framework or an orchestration engine. See *desired state* and *orchestration*.

kubectl

A command-line interface that you install on your computer and use to run commands that control and manage Kubernetes clusters.

K

kubelet

The agent that runs on each node in a Kubernetes cluster to manage pods. A PodSpec specifies how kubelet is to work. A PodSpec is a YAML or JSON object that describes a pod. The kubelet takes a set of PodSpecs that are provided through various mechanisms (primarily through the API server) and ensures that the containers described in those PodSpecs are running and healthy.

Kubo

See *Cloud Foundry Container Runtime*.

L

LDAP

Lightweight Directory Access Protocol. It is a standard protocol for storing and accessing directory service information, especially usernames and passwords. Applications can connect to an LDAP server to verify users and groups.

Lightwave

An open source security platform from VMware, Project Lightwave secures cloud platforms by providing a directory service, Active Directory interoperability, Kerberos authentication, and certificate services. Lightwave empowers IT security managers to impose the proven security policies and best practices of on-premises computing systems on their cloud computing environment. More specifically, Lightwave includes the following services:

- Directory services and identity management with LDAP and Active Directory interoperability

M

- Authentication services with Kerberos, SRP, WS-Trust (SOAP), SAML WebSSO (browser-based SSO), OAuth/OpenID Connect (REST APIs), and other protocols
- Certificate services with a certificate authority and a certificate store

linkerd

A service mesh that adds service discovery, routing, failure handling, and visibility to cloud native applications. linkerd is hosted by the CNCF.

MANO

In the telecommunications industry and others, MANO stands for management and orchestration of telco networking resources. In this context, MANO includes vRealize Suite: vRealize Operations, vRealize Automation, vRealize Network Insight, and vRealize Log Insight.

Memcached

As a system that caches data in the distributed memory of a cluster of computers, Memcached accelerates the performance of web applications by holding the results of recent database calls in random-access memory (RAM).

microservices

A “modern” architectural pattern for building an application. A microservices architecture breaks up the functions of an application into a set of small, discrete, decentralized, goal-oriented processes, each of which can be independently developed, tested, deployed, replaced, and scaled. See *cloud native application*.

M

microsegmentation

With VMware NSX, microsegmentation policies can specify granular traffic flow patterns among, for instance, the Kubernetes namespaces in which containerized applications are running. With microsegmentation, you can craft rules that impose security requirements on workloads and isolate resources at the level of microservices.

Minikube

A tool that lets you run a single-node Kubernetes cluster inside a virtual machine or locally on a personal computer.

MongoDB

A distributed NoSQL document database, MongoDB stores data with a flexible, schema-free data model that can adapt to change. MongoDB includes secondary indexes, geospatial search, and text search. Common use cases include serving data to mobile applications and performing real-time analytics.

Multicloud

This approach to cloud computing combines several cloud providers, platforms, or services in one environment or architecture. A multicloud strategy reduces reliance on a single vendor, protects your cloud services from outages, lets you tailor your architecture to your needs, and gives you the flexibility to switch solutions if your needs change. Using multiple cloud platforms and services, however, can complicate security, governance, and compliance. The portability of containers facilitates a multicloud strategy.

MySQL

It is an open source relational database management system (RDMS) that is commonly used in various types of applications, especially web apps. It is also widely embedded in the solutions distributed by independent software vendors (ISV) and original equipment manufacturers (OEM). In the name, SQL stands for Structured Query Language.

N

namespace

In the context of a Linux computer, a namespace is a feature of the kernel that isolates and virtualizes system resources. Processes that are restricted to a namespace can interact only with other resources and processes in the same namespace.

In Docker, namespaces isolate system resources like networking and storage.

In Kubernetes, when many virtual clusters are backed by the same underlying physical cluster, the virtual clusters are called namespaces.

namespace isolation

Limits the resources that a container can connect to or interact with.

network functions virtualization

NFV refers to the concept of virtualizing network node functions for creating communication services. In

the telecommunications industry, it is part of a cloud native design strategy. NFV-I stands for network functions virtualization infrastructure. In a telecommunications environment in which NFV is being used, a virtual network function, or VNF, is a specific function that can be combined with other functions to provide a communication service. A VNF manager controls a collection of VNF functions, including integrating a VNF with NFV infrastructure and assigning resources to it.

Node.js

An event-driven JavaScript runtime environment that executes JavaScript code on a server to produce dynamic content in web applications.

NodePort

In Kubernetes, a NodePort presents a service, such as a web server, on a port on the nodes in a Kubernetes cluster for external access.

N

NoSQL

A NoSQL database stores data that is structured in a way other than the tabular relationships of traditional relational databases. NoSQL is also known as non-SQL, non-relational, and not-only SQL. NoSQL databases are commonly used for big data and real-time data processing. Popular examples of NoSQL databases include MongoDB, Cassandra, and Pivotal Gemfire.

NSX

VMware NSX is a product that provides software-defined network virtualization.

O

OCI

Stands for Open Container Initiative, an organization dedicated to setting industry-wide container standards. OCI was formed under the auspices of the Linux Foundation for the express purpose of creating open industry standards around container formats and runtime. The OCI contains two specifications: the Runtime Specification (runtime-spec) and the Image Specification (image-spec). VMware is a member of OCI. See <https://www.opencontainers.org/>.

OpenStack

An open source platform for cloud computing that gives you control over pools of compute, storage, and networking resources through APIs and a dashboard. See *VMware Integrated OpenStack*.

OpenTracing

A vendor-neutral standard for distributed tracing. It is hosted by the CNCF.

P

opinionated platform

See *prescriptive platform*.

orchestration

Because it can automatically deploy, manage, and scale a containerized application, Kubernetes is often referred to as an orchestration framework or an orchestration engine. It orchestrates resource utilization, failure handling, availability, configuration, desired state, and scalability.

overlay network

Most orchestration systems include a software-defined networking component known as an overlay network. The overlay rides on top of the underlay to provide networking, such as IP addresses and ports, for the lifecycle of containers and hosts. The overlay can isolate communication among applications that use the same physical network. Overlay technologies include Flannel, Calico, and VMware NSX. See also *NSX*, *underlay network*, and *orchestration*.

PaaS

Platform as a service.

PAS

Pivotal Application Service. Formerly known as Elastic Runtime, PAS runs Java, .NET, and Node apps on Pivotal Cloud Foundry.

PCF

Pivotal Cloud Foundry, a private platform as a service for developing and deploying cloud native applications.

Photon OS

An open source project from VMware, Project Photon OS is a Linux operating system optimized for running containers.

PKS (VMware Enterprise PKS)

VMware Enterprise PKS is an enterprise Kubernetes platform.

P

platforms

The overarching business objective of using a container platform is to accelerate the development and deployment of scalable, enterprise-grade software that is easy to modify, extend, operate, and maintain. Three types of platforms provide varying degrees of support for container technology:

- A platform for running individual container instances. A platform for running container instances helps developers build and test a containerized application. It does not, however, orchestrate the containerized application with Kubernetes, nor does it provide a service broker so that developers can integrate tools, databases, and services with an app. An example of a container instance platform is VMware vSphere Integrated Containers.
- Containers as a service.
- Platform as a service.

platform as a service

Platform-as-a-service (PaaS) is a cloud-based environment for developing, testing, and running applications using programming languages, libraries, services, and tools supported or offered by the platform's provider. A platform as a service is sometimes referred to simply as an application platform. In this context, an application platform helps developers not only write code but also integrate tools and services, such as a database, with their application as, for instance, microservices. An example of a private platform as a service that is also referred to as an application platform is Pivotal Cloud Foundry. See *containers as a service*, *infrastructure as a service*, and *cloud computing*.

platform developer

An engineer who customizes a Kubernetes platform (or another modern platform) to fit the needs of their project or organization.

platform operator

An engineer who manages a platform like Kubernetes.

platform reliability engineer

An engineer who operates a platform, such as Pivotal Cloud Foundry. The work might include the management of disaster recovery, networking, storage, monitoring applications, and service brokers. A key part of a platform reliability engineer's job might be to establish and maintain a continuous integration and continuous delivery pipeline. See *DevOps*.

pod

On Kubernetes, a pod is the smallest deployable unit in which one or more containers can be managed—in other words, you run a container image in a pod. A set of pods typically wraps a container, its storage resources, IP address, and other options up

into an instance of an application that will run on Kubernetes. Docker is usually the container runtime used in a pod. A Kubernetes administrator or application developer specifies a pod by using a YAML file. Pods are commonly managed by a *deployment*, which see.

PostgreSQL

Also known as Postgres, it is an extensible object-relational database management system that securely stores data for large Internet-facing applications or data warehouses. Postgres is ACID-compliant; see *ACID*.

P

prescriptive platform

In the context of application platforms, a prescriptive platform hides the platform's complexity from developers by prescribing that developers use the system's formats, pipeline, and methods for building and running applications. For example, a prescriptive container platform might prescribe a scheduler, a runtime engine, integration with the underlying infrastructure, continuous delivery, and other aspects of the platform. A prescriptive platform is also referred to as an "opinionated" platform.

private cloud

A fully virtualized data center that includes two key capabilities that increase agility and are different from a virtualized data center: self-service and automation.

Prometheus

A open source monitoring system for Kubernetes. Prometheus is hosted by the CNCF.

pull

Downloading a container image from a registry into a local cache so that you can launch containers based on the image.

Q

quality of service

It is often abbreviated QoS.

R

RabbitMQ

An open source message broker, RabbitMQ implements the Advanced Method Queuing Protocol to give applications a common intermediate platform through which they can connect and exchange data.

RBAC

role-based access control. On Kubernetes, RBAC is a module that authorizes access to resources by role. RBAC empowers administrators to dynamically configure access policies through the Kubernetes API.

Redis

A key-value database, Redis can store a dataset in a networked, in-memory cache. Because keys in Redis can contain strings, hashes, lists, sets, sorted sets, bitmaps, and hyperlogs, Redis is often referred to as a data structure server. Data scientists, for instance, can perform operations on these data types to do things like compute set intersection, union and difference, and ranking.

R

registry

A hosted service that contains repositories of container images. Harbor, an open source project from VMware, is an example of a registry.

replica set

In Kubernetes, a replica set is a controller that manages the lifecycle of pods. See *controllers*.

repository

In the context of containers, a repository is a set of container images. The repository can be shared with other users through a registry server, and the images in the “repo” can be tagged with labels.

refactoring

Re-architecting an application or modifying its code to improve it. An application, for example, might be refactored by decomposing it into microservices.

repackaging

Placing a traditional application in a container format.

replatforming

Moving an application to another, more efficient platform. If the application being migrated is a traditional application and if the new platform uses containers, replatforming also involves repackaging.

rkt

Pronounced like *rocket*, rkt is a standards-based container engine from CoreOS.

runC

The code module that launches containers. It is part of containerd and managed by OCI, which stands for Open Container Initiative. See *OCI*.

S

scheduler

A module of a system or a software component that schedules and runs the deployment of containers, jobs, tasks, or another type of workload. Most public cloud services, such as Microsoft Azure, include a scheduler that lets you create jobs in the cloud. The jobs can, in turn, invoke services or tasks, such as backing up data or cleaning up logs.

service

The definition of *service* varies by context. In Kubernetes, it is an API object that describes how to access applications, such as a set of pods, by using methods like ports or load-balancers.

A service may also be a microservice within the context of some larger application. An HTTP server, for example, is a service.

service discovery

In the context of running cloud native applications built with microservices, service discovery automatically detects the dynamically assigned networking information of the services or the devices on which they are running.

service mesh

When a containerized application is built as a collection of services or microservices, it forms a mesh of services. A service mesh creates a layer above IP addresses and ports to connect the services and manage their interactions. A service mesh might deliver, for instance, load balancing, monitoring, and service-to-service authentication. Examples of technology that provides a service mesh are Istio and linkerd.

S

sidecar

A cloud native architectural pattern that isolates a component or service of an application by placing it in a separate but co-located container for independence and flexibility.

software-defined data center (SDDC)

A data center in which infrastructure is virtualized and delivered as a service. The infrastructure of an SDDC includes virtualized networking and software-defined data storage and management. An SDDC supports applications in a way that is more flexible, agile, efficient, and cost-effective than traditional approaches. In a SDDC, all the components of infrastructure—compute, networking, storage, security, and availability—are abstracted and delivered as automated, policy-driven software. An SDDC radically reduces manual processes, speeds up IT service delivery, reduces costs, and improves ROI.

software as a service (SaaS)

An application running on a cloud infrastructure that is used over a network, typically the Internet, instead of being downloaded and installed on local machines. The consumer of the service does not manage or control the underlying cloud infrastructure or the application's capabilities. Also known as a web app.

Spanner

See *Cloud Spanner*.

Spark

Apache Spark is an engine for large-scale data processing that can be used interactively from the Python shell. Spark combines streaming, SQL, and complex analytics by powering a stack of tools that can coexist in the same application. Spark can access diverse data sources, including not only the Hadoop File System (HDFS) but also Cassandra and MongoDB. Data scientists like Spark because they get access to Python's powerful numeric processing libraries.

spec

In Kubernetes, spec stands for specification. The specification is a description of a desired state, including the configuration supplied by a user.

Spring

The Spring Framework is an open-source Java framework by Pivotal that handles the infrastructure for developing Java applications. Spring includes Spring Boot, Spring Cloud, and Spring Cloud Data Flow.

Spring Cloud Data Flow

A toolkit for building data integration and real-time data processing pipelines. The Spring Cloud Data Flow server uses Spring Cloud Deployer to integrate pipelines with Pivotal Cloud Foundry, Mesos, or Kubernetes. Spring Cloud Data Flow helps engineers develop analytics pipelines by providing a distributed system that unifies ingestion, real-time analytics, batch processing, and data export.

StatefulSet

In Kubernetes, a StatefulSet manages the deployment and scaling of a set of pods according to your desired state. A stateful set can, for example, manage persistent storage and other resources for stateful pods.

swarm

With Docker, a swarm is a cluster of one or more Docker Engines running in swarm mode. Docker Swarm, however, is not the same thing as the swarm mode features in Docker Engine. See *Docker Swarm*.

T

tag

With Docker, a tag is a label that a user applies to a Docker image to distinguish it from other images in a repository.

Tern

An open source tool from VMware that identifies the sources, versions, and licenses for packages installed in a Docker container. The resulting manifest can help ensure that the packages in a container comply with licensing requirements. See <https://github.com/vmware/tern>.

Terraform

Software from HashiCorp for creating infrastructure as code in a cloud service provider's environment, such as Google Cloud Platform.

the cloud

Computing resources available over the Internet. See *cloud computing*.

traditional application

A traditional application is monolithic in design with an n-tier application architecture that generally consists of database, application, and web servers. These components are usually tightly coupled with the infrastructure and dependent on it for high availability.

U

UID

It can stand for user identifier, user ID, or unique identifier, depending on the context or the system. With Kubernetes, for example, a UID is a string that uniquely identifies an object.

underlay network

The underlay network connects machines, whether virtual or physical, by using either a traditional hardware-based approach to networking or a combination of hardware and software. See also *overlay network*, *service mesh*, and *NSX*.

V

Vagrant

HashiCorp's Vagrant turns a machine's configuration into a distributable template to produce a predictable development environment for applications.

virtual infrastructure manager

VIM manages the resources associated with network functions virtualization infrastructure, or NFV-I. See *network functions virtualization*.

VMware Integrated OpenStack

A distribution of OpenStack supported by VMware, it lets you run OpenStack on VMware infrastructure. See *OpenStack*.

volume

With Docker, a volume (or data volume) is a designated directory within one or more containers that bypasses the Union File System. Volumes are designed to persist data independent of the container's life cycle.



vSphere Cloud Provider

It is a persistent storage solution from VMware that supports stateful, data-intensive containerized applications, including those with databases. vSphere Cloud Provider, which is part of Kubernetes through Project Hatchway, enables VMware Enterprise PKS to support the following storage primitives on vSphere storage: volumes, persistent volumes (pv), persistent volumes claims (PVC), storage class, and stateful sets. The vSphere Cloud Provider also furnishes enterprise storage features like storage policy-based management. See *Pivotal Container Service* and *Hatchway*.

vSphere Integrated Containers

It is a product from VMware that creates container infrastructure within VMware vSphere so that you can run containerized applications on virtual machines. The solution includes an engine to run containers, a portal to manage containers, a registry to securely store container images, and a Docker-compatible interface. A key benefit of vSphere Integrated Containers is that it combines the portability of containers with the security, visibility, and management of vSphere.

W

workload

A workload is the computational or transactional burden of a set of computing, networking, and storage tasks associated with an application. Similar apps with the same technology and tools can have radically different workloads under different circumstances or during different times. Workloads can often be measured by CPU or memory consumption, network traffic, requests, database queries, transactions, and so forth. In very basic, broad terms, an application is a thing that processes something; a workload is the processing that's being done; and a use case is the reason that you do it. In the context of cloud computing and Kubernetes clusters, a workload can be seen as the amount of work that an instance of an app or part of an app performs during a certain time period.

X

XML

Extensible Markup Language. It is a flexible but verbose format for structuring and exchanging data. XML is often used in legacy applications, Java applications, and web applications for a variety of purposes, such as structuring configuration files or exchanging data. Although XML is sometimes used in cloud native applications, JSON or YAML (which see) are the preferred data formats.

Y

YARN

A sub-project of Apache Hadoop, YARN separates resource management from computational processing to expand interactional patterns beyond MapReduce for data stored in HDFS. YARN allocates resources for Hadoop applications such as MapReduce and Storm as they perform computations. YARN, in effect, stands at the center of a Hadoop environment by providing a data operating system and pluggable architecture for other applications.

YAML

A human-readable data serialization standard commonly used in configuration files to structure information and commands. In Kubernetes, specification files are written in YAML.

Z

ZooKeeper

Apache ZooKeeper coordinates distributed applications masquerading as animals. It provides a registry for their names. It configures and synchronizes them. It keeps them from running amok.



12-factor app

A methodology for developing a software-as-a-service (SaaS) application—that is, a web app—and typically deploying it on a platform as a service or a containers as a service.

- A
- B
- C
- D
- E
- F
- G
- H
- I
- J
- K
- L
- M
- N
- O
- P
- Q
- R
- S
- T
- U
- V
- W
- X
- Y
- Z

