Managing thousands of Kubernetes clusters with Gardener

https://gardener.cloud

About us

[@mvladev](https://github.com/mvladev)

- 5+ years experience with K8S
- Attending in several K8S SIGs
- Consulting-as-a-Service (CaaS) for a lager beer

[@ialidzhikov](https://github.com/ialidzhikov)

- Gardener project contributor for the last more that 1 year
- /area core
- /kind enhancement, /kind bug

Kubernetes Clusters as a Service

Gardener

The Kubernetes Botanist

What does Kubernetes not cover?

- **Install and manage multiple clusters**
- **Homogeneously** across **Multi-Cloud**
  - Public Cloud Providers
  - Private Cloud
- **Zero Ops**
  - Managed Nodes
  - Managed Control Plane
  - Day 2 Operations

The Gardener mission statement

Provide Kubernetes Clusters-as-a-Service homogeneously on hyper-scalers and on-premise fully managed and with minimal TCO.

https://gardener.cloud

The Gardener mission statement

https://gardener.cloud

In more details...

- Homogeneous Kubernetes clusters as a service
- On hyper-scalers and on-premise (OpenStack, …)
- Fleet management of thousands of clusters
- Fully managed / minimal TCO (Day-1 & 2)
- Multi OS support (CoreOS, SUSE, Ubuntu, Flatcar, …)
- Configurable control plane (OIDC, Runtime Configuration, …)
- Support for older and newer Kubernetes versions

Why do we need our own solution?

- SAP is going multi-cloud big time
- There is no open-source tool to match our needs (big scale, thousands of clusters and teams)
- Cost-effective and full control over the Kubernetes components
  - OIDC, Audit Logging, ...
- Support for on-premise IaaS
- ZeroOps
- Open Source is king

DEMO TIME

Gardener dashboard

Cluster creation

Cluster management

Monitoring & Logging

Kubernetes Conformance

- Gardener Conformance test results are available at the CNCF test grid - testgrid.k8s.io/gardener-all
- Kubernetes Conformance test coverage
  - we run the conformance tests on a regular basis for all supported Kubernetes versions (v1.10-v1.17) on almost all supported providers (GCP, AWS, OpenStack, Azure, Alicloud)

How does Gardener work?

Following the definition of Kubernetes...

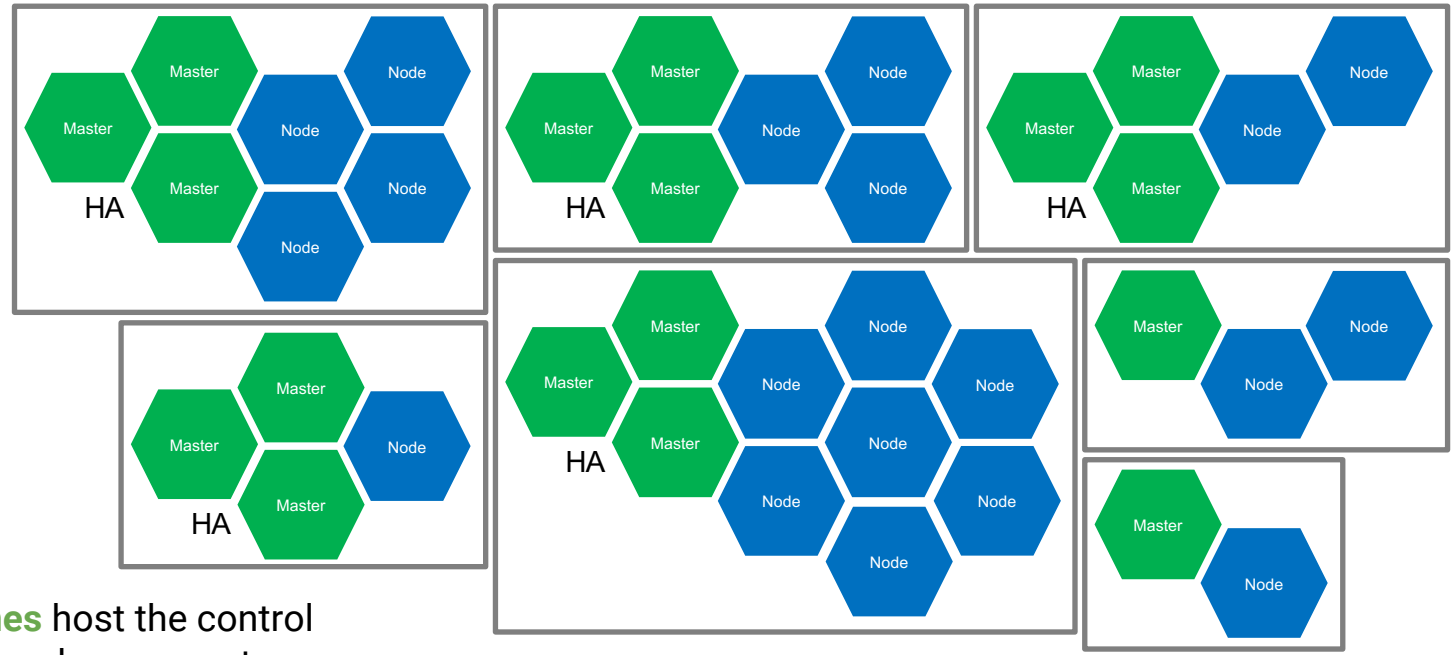
Kubernetes is an **open-source** system for **automating deployment**, **scaling**, and **management** of **containerized applications/software**.

...we do the following:

We use **Kubernetes** to **deploy**, **host** and **operate Kubernetes**.
Cluster control planes are **seeded** into already **existing clusters**.

aka the Inception, or Kubeception Model

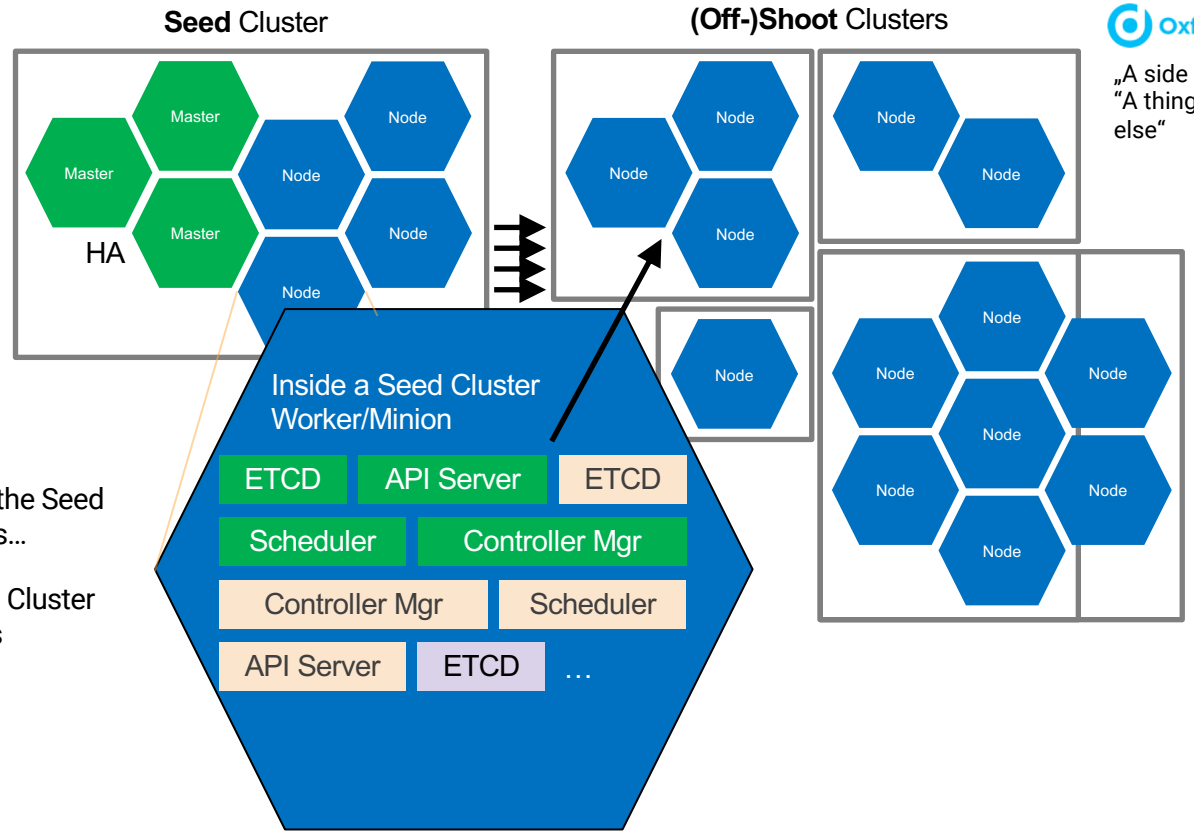
Common Kubernetes Cluster Setup



The **green machines** host the control plane, often in HA and on separate hardware (usually underutilized or, worse, overutilized)

The **blue machines** host the actual workload (usually pretty well utilized)

Gardener Kubernetes Cluster Setup



Oxford Dictionaries

„A side shoot or branch on a plant“ or “A thing that develops from something else“

Zooming into the Seed Cluster reveals…

Multiple Shoot Cluster Control Planes

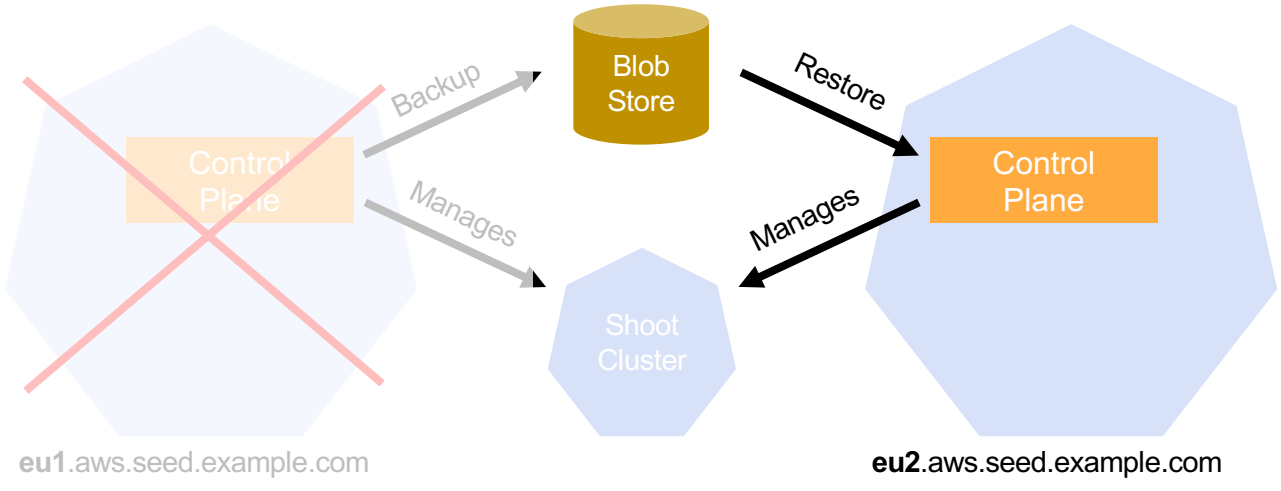
Machine De-/Provisioning, Self-Healing, Auto-Update and Auto-Scaling via our own Machine Controller Manager

Resilience, in case Shoot cluster has issues...

- **Kubernetes** (brings back the shoot cluster control plane / resources)
- **Machine Controller** (brings back machines (nodes))
- **ETCD Backup & Restore** (brings back the persistence)
- **Gardener** reconciliation (brings back infrastructure, configuration)

Resilience, in case Seed cluster has issues...

- Even though a seed cluster is set up as a shoot cluster, regional problems may take it offline longer than we like, so we can **move control planes** (not yet automatically)

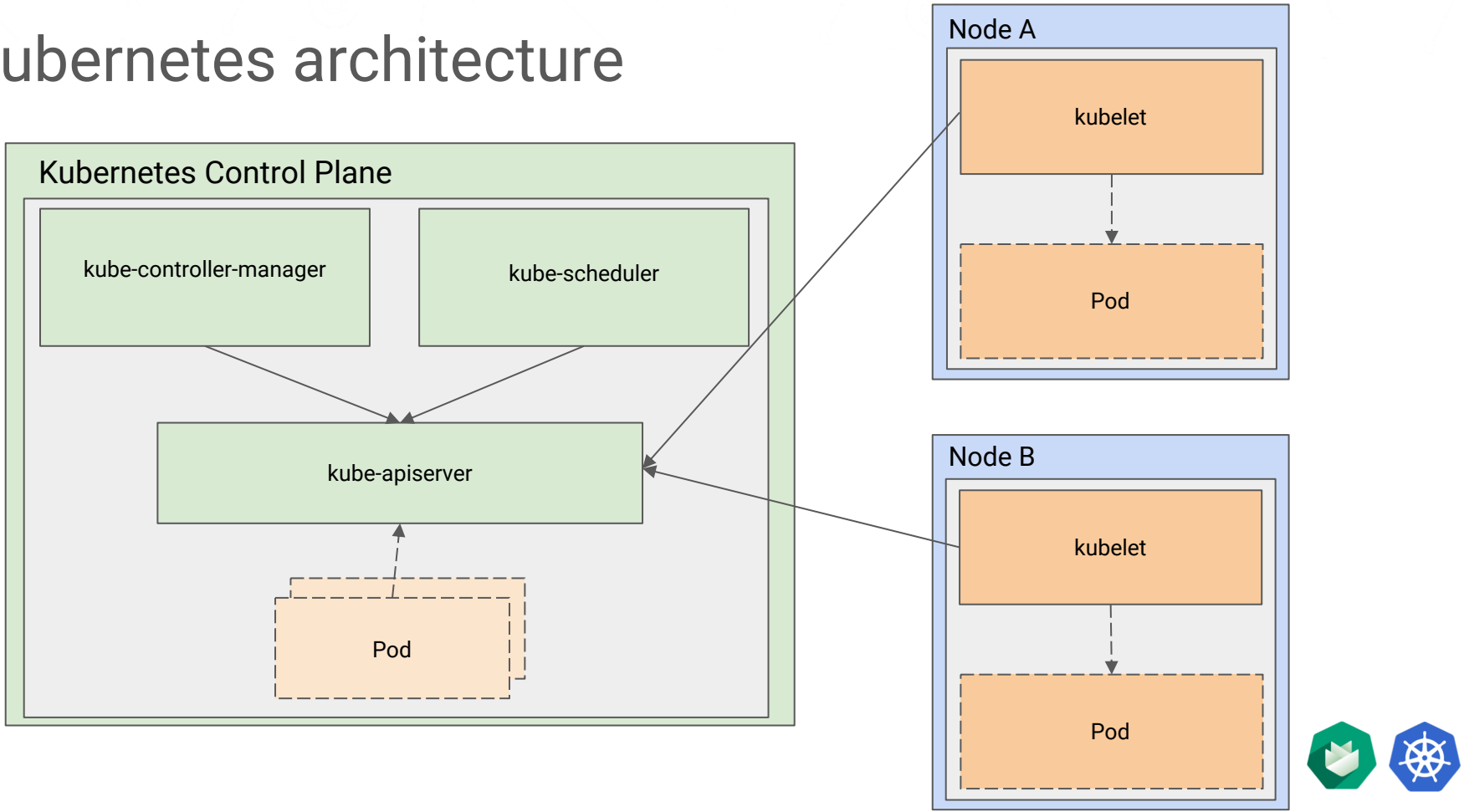


DEMO TIME

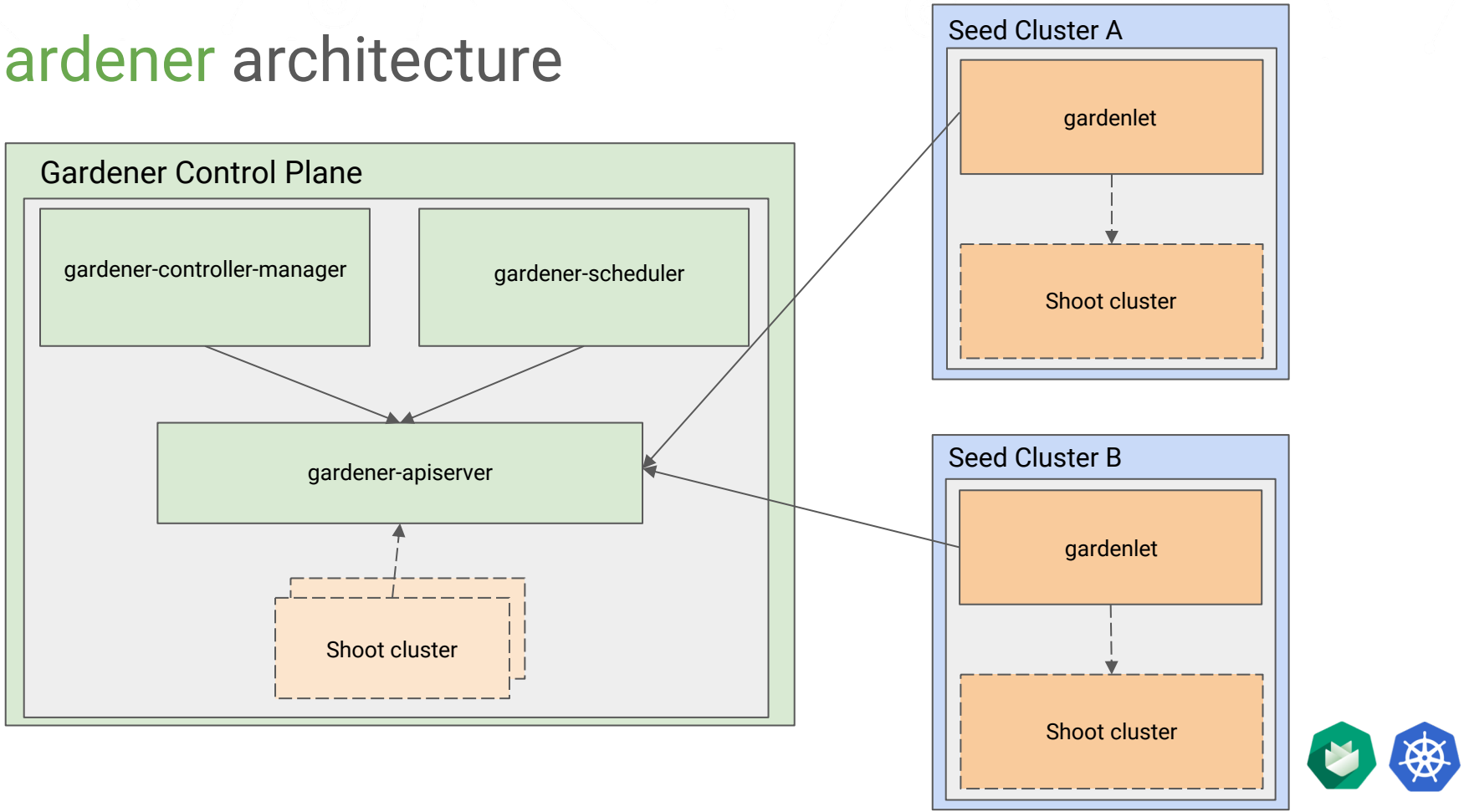
Cluster Inspection

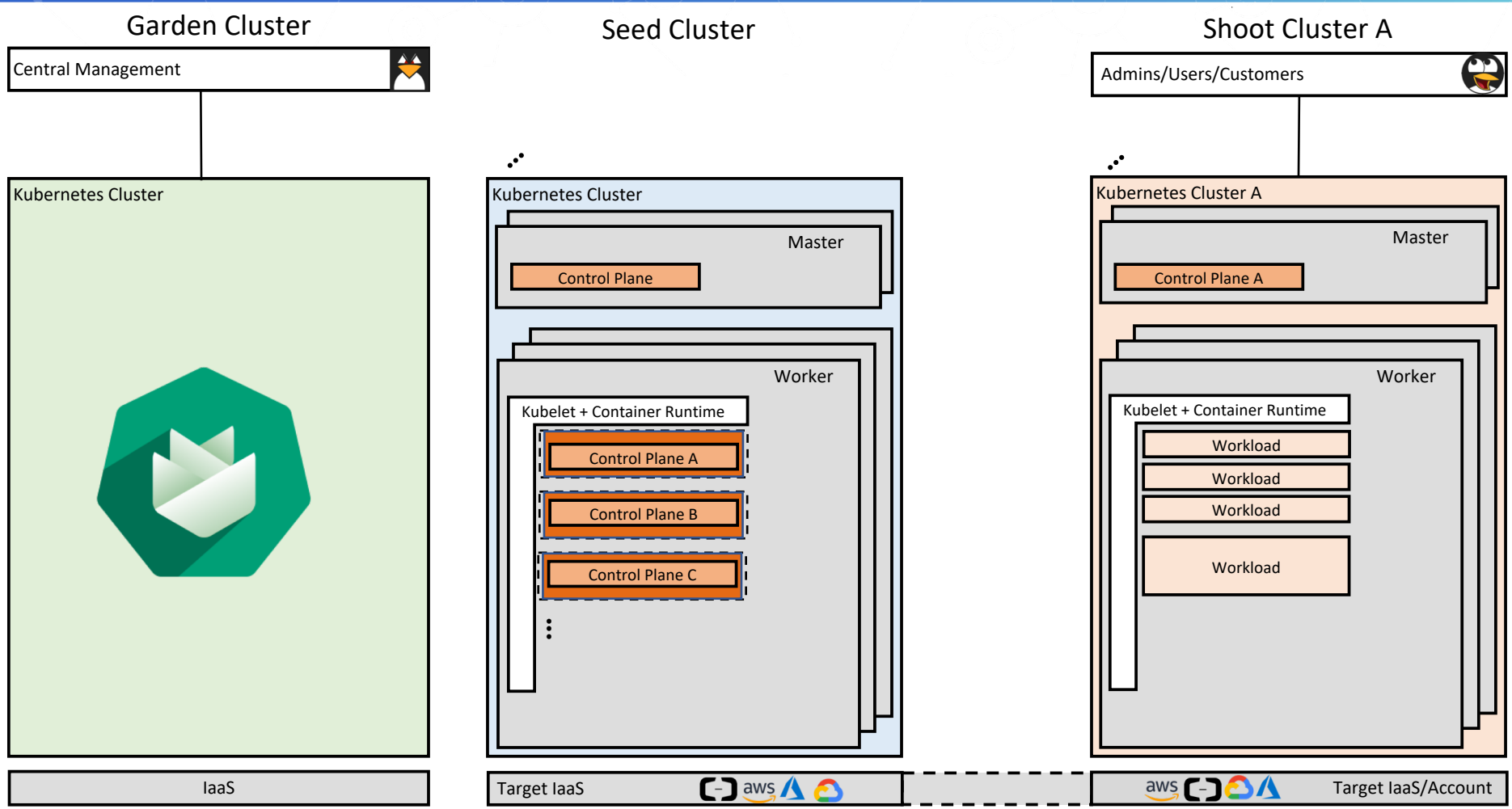
Self-healing

Kubernetes architecture



Gardener architecture




Garden Cluster
Central Management
Kubernetes Cluster
IaaS
Seed Cluster
Kubernetes Cluster
Master
Control Plane
Worker
Kubelet + Container Runtime
Control Plane A
Control Plane B
Control Plane C
Target IaaS
Shoot Cluster A
Admins/Users/Customers
Kubernetes Cluster A
Master
Control Plane A
Worker
Kubelet + Container Runtime
Workload
Workload
Workload
Workload
Target IaaS/Account

Back to the roots...

Lingua Franca - Gardener Shoot (Cluster) Resource

Native Kubernetes Resource →

Define your Infrastructure needs →

Specify (multiple) Worker pools →

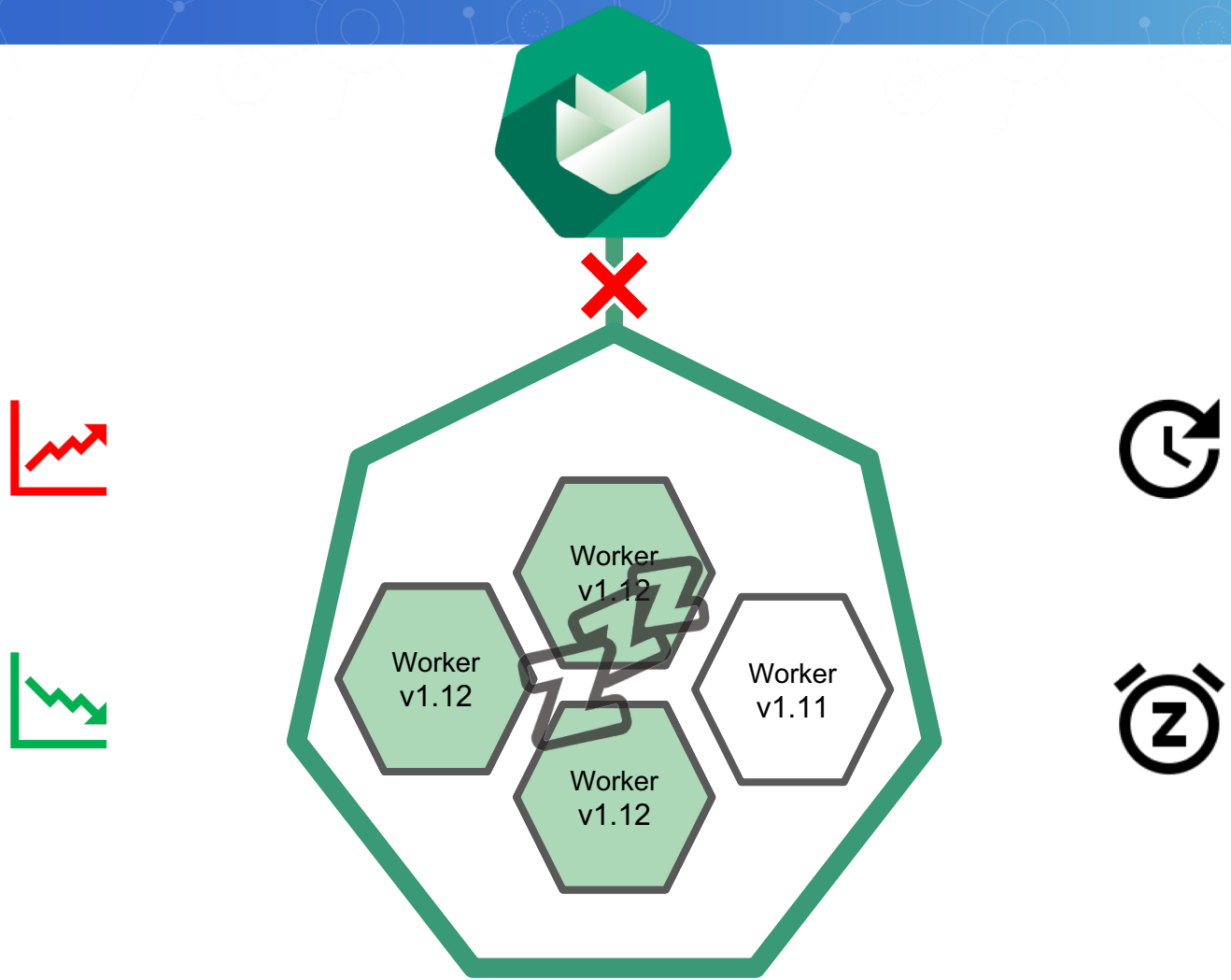
Set Kubernetes version →

Define when and what to update →

Gardener reported status →

```yaml
apiVersion: core.gardener.cloud/v1beta1
kind: Shoot
metadata:
  name: my-cluster
  namespace: garden-project
spec:
  provider:
    type: aws
    infrastructureConfig:
      apiVersion: aws.provider.extensions.gardener.cloud/v1alpha1
      kind: InfrastructureConfig
      networks:
        vpc:
                  cird: 10.222.0.0/16
        zones:
        - name: eu-west-1
          internal: 10.250.112.0/22
          public: 10.250.96.0/22
          workers: 10.250.0.0/19
    workers:
    - name: cpu-worker
      machine:
        type: m5.large
      minimum: 2
      maximum: 2
      volume:
        size: 50Gi
        type: gp2
      zones:
      - eu-west-1a
  kubernetes:
    version: 1.17.2
    kubeAPIServer: {}
  maintenance:
    autoUpdate:
      machineImageVersion: true
status:
  ...
```

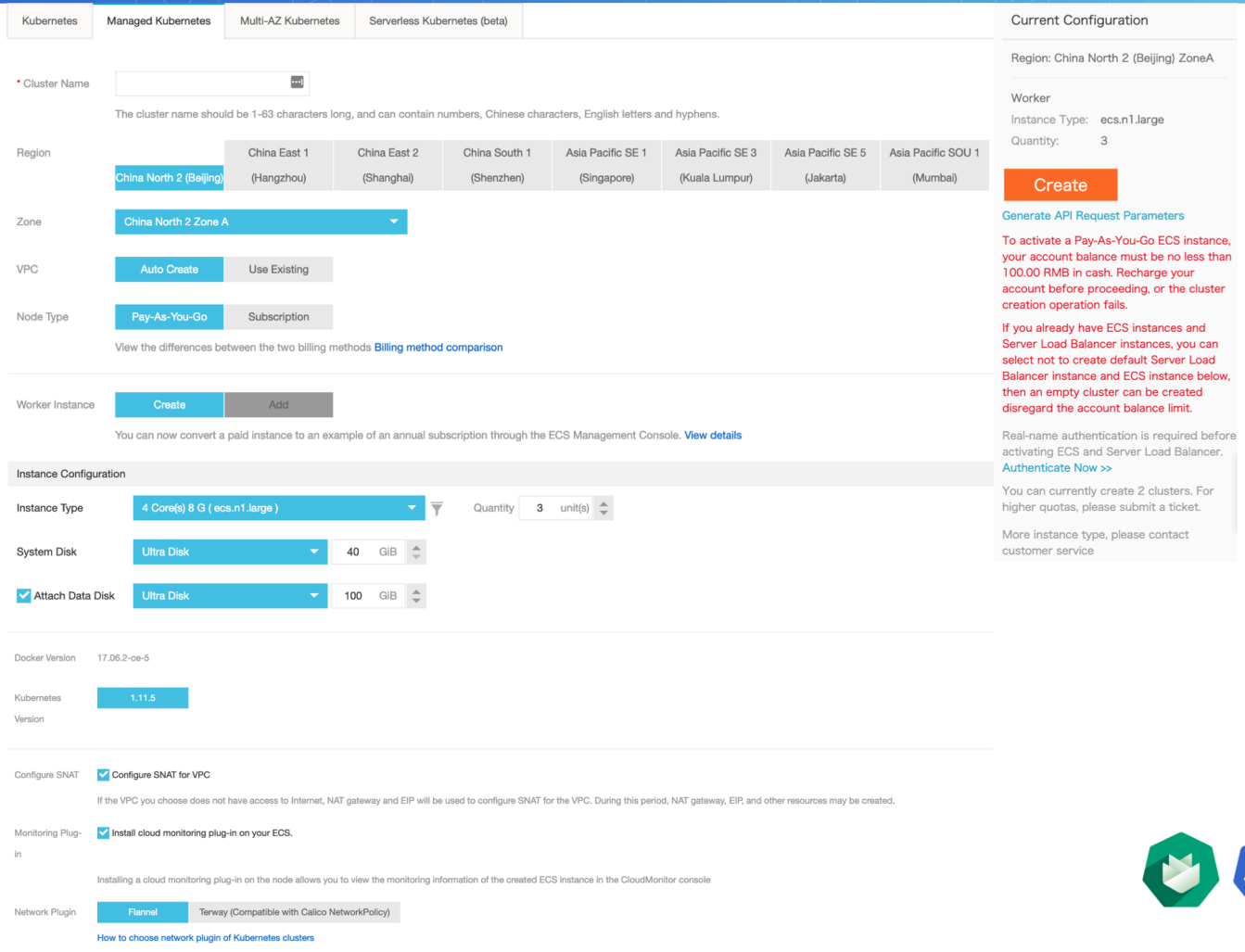
Circumvent Vendor Lock-in (Lingua Franca)


Worker
v1.12
Worker
v1.12
Worker
v1.11
Worker
v1.12

Why do we keep mentioning “homogenous clusters” so often?

Other offerings...

Aliyun ACK


Kubernetes
Managed Kubernetes
Multi-AZ Kubernetes
Serverless Kubernetes (beta)
* Cluster Name
The cluster name should be 1-63 characters long, and can contain numbers, Chinese characters, English letters and hyphens.
Region
China North 2 (Beijing)
China East 1 (Hangzhou)
China East 2 (Shanghai)
China South 1 (Shenzhen)
Asia Pacific SE 1 (Singapore)
Asia Pacific SE 3 (Kuala Lumpur)
Asia Pacific SE 5 (Jakarta)
Asia Pacific SOU 1 (Mumbai)
Zone
China North 2 Zone A
VPC
Auto Create
Use Existing
Node Type
Pay-As-You-Go
Subscription
View the differences between the two billing methods Billing method comparison
Worker Instance
Create
Add
You can now convert a paid instance to an example of an annual subscription through the ECS Management Console. View details
Instance Configuration
Instance Type
4 Core(s) 8 G (ecs.n1.large)
Quantity
3
unit(s)
System Disk
Ultra Disk
40
GiB
Attach Data Disk
Ultra Disk
100
GiB
Docker Version
17.06.2-ce-5
Kubernetes Version
1.11.5
Configure SNAT
Configure SNAT for VPC
If the VPC you choose does not have access to Internet, NAT gateway and EIP will be used to configure SNAT for the VPC. During this period, NAT gateway, EIP, and other resources may be created.
Monitoring Plug-in
Install cloud monitoring plug-in on your ECS.
Installing a cloud monitoring plug-in on the node allows you to view the monitoring information of the created ECS instance in the CloudMonitor console
Network Plugin
Flannel
Terway (Compatible with Calico NetworkPolicy)
How to choose network plugin of Kubernetes clusters
Current Configuration
Region: China North 2 (Beijing) ZoneA
Worker
Instance Type: ecs.n1.large
Quantity: 3
Create
Generate API Request Parameters
To activate a Pay-As-You-Go ECS instance, your account balance must be no less than 100.00 RMB in cash. Recharge your account before proceeding, or the cluster creation operation fails.
If you already have ECS instances and Server Load Balancer instances, you can select not to create default Server Load Balancer instance and ECS instance below, then an empty cluster can be created disregard the account balance limit.
Real-name authentication is required before activating ECS and Server Load Balancer.
Authenticate Now >>
You can currently create 2 clusters. For higher quotas, please submit a ticket.
More instance type, please contact customer service

AWS EKS

General configuration

Cluster name
Enter a unique name for your Amazon EKS cluster.

Kubernetes version
Select the Kubernetes version to install.

1.14

Role name Info
Select the IAM Role to allow Amazon EKS and the Kubernetes control plane to manage AWS resources on your behalf.

eks

Networking Info

VPC
Select a VPC to use for your EKS Cluster resources.

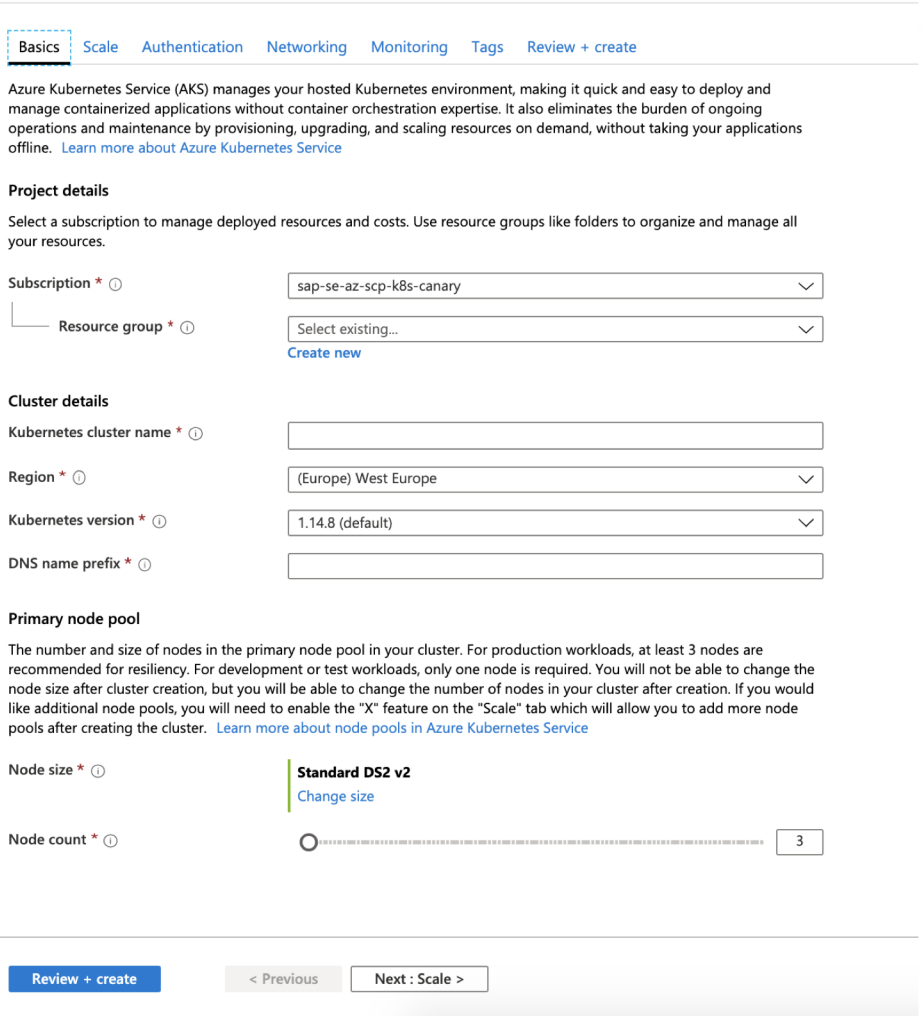
vpc-543bc332 - 172.31.0.0/16

Subnets
Choose the subnets in your VPC where your worker nodes will run.

Find subnet

Subnet	Name	Availability Zone	Subnet IPv4 CIDR
subnet-68866820		eu-west-1c	172.31.16.0/20
subnet-7bf1ea1c		eu-west-1b	172.31.0.0/20
subnet-3b4d6a60		eu-west-1a	172.31.32.0/20

Azure AKS


Basics
Scale
Authentication
Networking
Monitoring
Tags
Review + create
Azure Kubernetes Service (AKS) manages your hosted Kubernetes environment, making it quick and easy to deploy and manage containerized applications without container orchestration expertise. It also eliminates the burden of ongoing operations and maintenance by provisioning, upgrading, and scaling resources on demand, without taking your applications offline.
Learn more about Azure Kubernetes Service
Project details
Select a subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.
Subscription *
sap-se-az-scp-k8s-canary
Resource group *
Select existing...
Create new
Cluster details
Kubernetes cluster name *
Region *
(Europe) West Europe
Kubernetes version *
1.14.8 (default)
DNS name prefix *
Primary node pool
The number and size of nodes in the primary node pool in your cluster. For production workloads, at least 3 nodes are recommended for resiliency. For development or test workloads, only one node is required. You will not be able to change the node size after cluster creation, but you will be able to change the number of nodes in your cluster after creation. If you would like additional node pools, you will need to enable the "X" feature on the "Scale" tab which will allow you to add more node pools after creating the cluster.
Learn more about node pools in Azure Kubernetes Service
Node size *
Standard DS2 v2
Change size
Node count *
3
Review + create
< Previous
Next : Scale >

Google GKE

'Standard cluster' template

Continuous integration, web serving, backends. Best choice for further customization or if you are not sure what to choose.

Some fields can't be changed after the cluster is created. Hover over the help icons to learn more. Dismiss

Name

standard-cluster-1

Location type

- (•) Zonal
- ( ) Regional

Zone

us-central1-a

Master version

Try the new Release Channels feature instead of managing the master version directly. Use Release Channels

1.13.11-gke.23 (default)

Node pools

Node pools are separate instance groups running Kubernetes in a cluster. You may add node pools in different zones for higher availability, or add node pools of different type machines. To add a node pool, click Edit. Learn more

default-pool

Number of nodes: 3

Pod address range limits the maximum size of the cluster. Learn more

Machine configuration

Machine family: General-purpose | Memory-optimized

Machine types for common workloads, optimized for cost and flexibility

Series: N1

Powered by Intel Skylake CPU platform or one of its predecessors

Machine type: n1-standard-1 (1 vCPU, 3.75 GB memory)

vCPU	Memory
1	3.75 GB

CPU platform and GPU

Auto-upgrade: On

More options

Create Cancel Equivalent REST or command line

	ACK	EKS	AKS	GKE	Gardener
Kubernetes version (latest available 1.17)	1.11.x 1.12.x 1.14.x … outdated …	1.12.x 1.13.x 1.14.x … out of control …	1.13.x 1.14.x 1.15.x preview (1.16 & 1.17)	1.14.x (regular channel) 1.13.x (stable channel) 1.16.x (rapid channel)	1.10.x 1.14.x 1.11.x 1.15.x 1.12.x 1.16.x 1.13.x 1.17.x
Config. Control Plane	x	x	x	x	V
Custom DNS	x	x	x	x	V
Managed Certs	x	x	x	x	V
Worker Pools	Single Cannot be changed	Multiple	Multiple	Multiple	Multiple
Worker Operating System	Ubuntu	Amazon Linux 2, Ubuntu(Partner)	Ubuntu	Container OS, Ubuntu(Partner)	CoreOS, SUSE JeOS Ubuntu, Flatcar (in process)
Worker Auto Repair	V	x	x	V	V
Monitoring	x	x	V	V	V
Logging	x	x	x	V	V
Regions	8 / 19	13 / 21	18 / 48	20 / 20	85 (all GA IaaS regions)
Hibernation	x	x	x	x	V

ADOPTERS.md

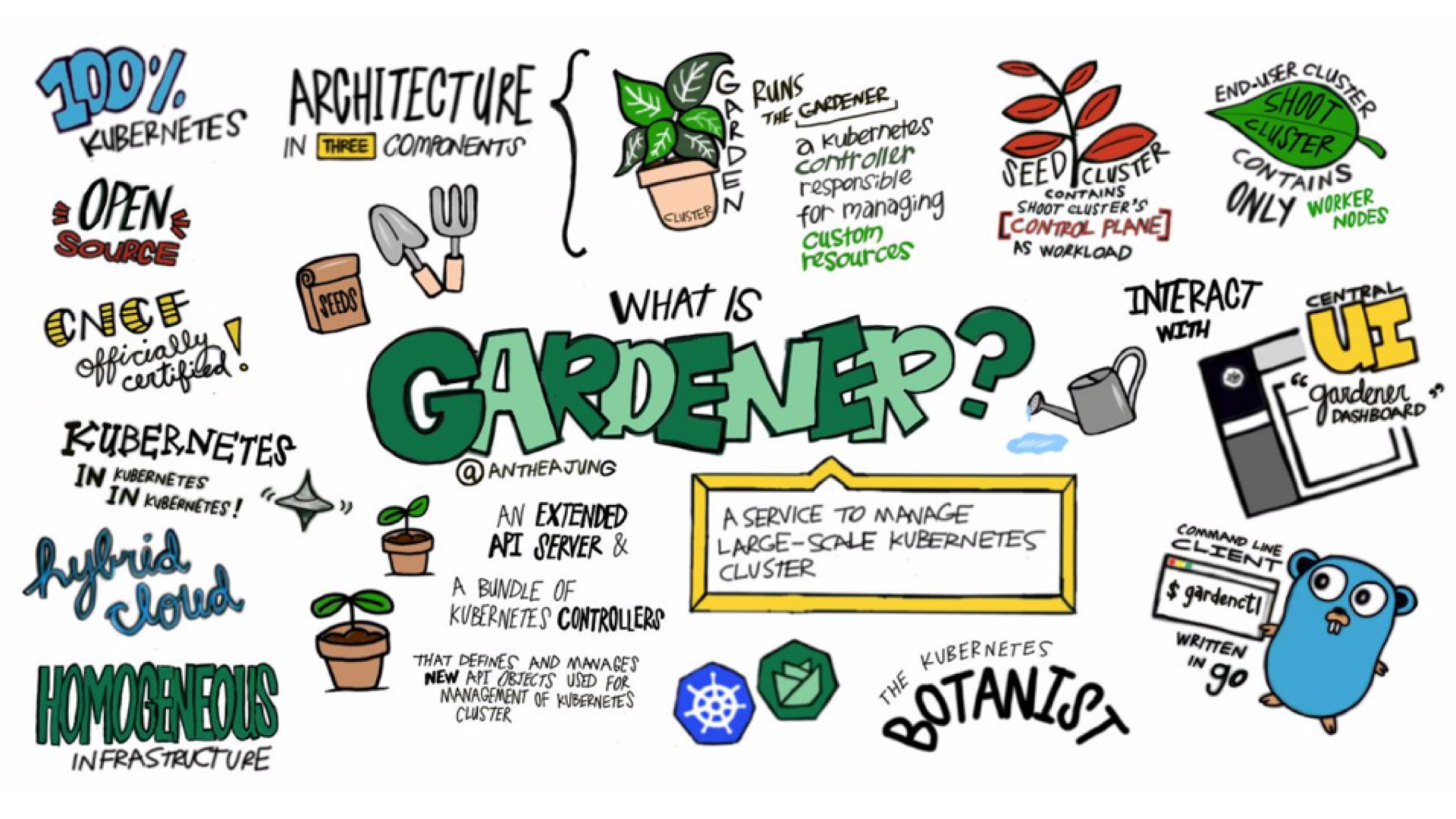
- SAP - thousands of clusters, **30K+** CPU, **150TB+** memory
- Finanz Informatik Technologie Services GmbH - uses Gardener on top of Metal as a Service
- PTC - uses Gardener to provide development environments and CI/CD systems internally
- b'nerd - uses Gardener as the core technology for its own managed Kubernetes as a Service offering

Ongoing improvements (GEPs)

- GEP-01: Extensibility (done)
  - vmware vsphere provider (in progress)
- GEP-06: HA setup for etcd (in progress)
- GEP-07: Shoot control plane migration (in progress)
- GEP-08: SNI pass-through proxy for kube-apiserver (in progress)
- (?): Shoot additional container runtimes (gvisor, kata-containers, …)

Thank you!

-  https://github.com/gardener/gardener
-  https://kubernetes.slack.com/messages/gardener
- Official website - https://gardener.cloud/
- Join our weekly community meetings


100% KUBERNETES
OPEN SOURCE
CNCF officially certified!
KUBERNETES IN KUBERNETES IN KUBERNETES!
hybrid cloud
HOMOGENEOUS INFRASTRUCTURE
ARCHITECTURE IN THREE COMPONENTS
SEEDS
GARDEN
CLUSTER
RUNS THE GARDENER
a kubernetes controller responsible for managing custom resources
SEED CLUSTER
CONTAINS SHOOT CLUSTER'S [CONTROL PLANE] AS WORKLOAD
END-USER CLUSTER
SHOOT CLUSTER
CONTAINS ONLY WORKER NODES
WHAT IS GARDENER?
@ANTHEAJUNG
AN EXTENDED API SERVER &
A BUNDLE OF KUBERNETES CONTROLLERS
THAT DEFINES AND MANAGES NEW API OBJECTS USED FOR MANAGEMENT OF KUBERNETES CLUSTER
A SERVICE TO MANAGE LARGE-SCALE KUBERNETES CLUSTER
THE KUBERNETES BOTANIST
INTERACT WITH
CENTRAL UI
"gardener DASHBOARD"
COMMAND LINE CLIENT
$ gardenctl
WRITTEN IN go