Kubernetes
An open platform for container orchestration

Johannes M. Scheuermann
Karlsruhe, 30.08.2017
Johannes M. Scheuermann
Cloud Platform Engineer @ inovex

› Software-Defined Datacenters
› Infrastructure as Code
› Cloud technologies
› High Availability & Scalability
› @johscheuer
How to manage millions of containers/machines?
“Containers”

- cgroups and namespaces
- Application package management
- (nearly) Platform independent
- Resource separation
- Resource monitoring
“History”

- Unified container-management Borg
  - Before: Babysitter + Global Work Queue
- Autopilot
  - predicting resource requirements
- Omega
  - Splitting parts of the control plane
  - Multiple Schedulers
Architecture
Architecture

Master

API
Controller-Manager
Scheduler
etcd

Store

Node
Node
Node

Kubelet
Kubelet
Kubelet

Kube-proxy
Kube-proxy
Kube-proxy
Design Goals

- Portable
- General-Purpose
- Flexible
- Extensible
- Automatable
Principals
Pods

Container
- Process
- Resources
- Ports

Container
- Process
- Resources
- Ports

Container
- Process
- Resources
- Ports

Shared Resources
Deployments

ReplicaSet r1 → Pod → Pod → Pod

Deployment
Deployments

Deployment

ReplicaSet r1

ReplicaSet r2

Pod

Pod

Pod
Deployments

- Deployment
  - ReplicaSet r1
    - Pod
    - Pod
  - ReplicaSet r2
    - Pod
Deployments

- Deployment
- ReplicaSet r1
  - Pod
  - Pod
- ReplicaSet r2
  - Pod
  - Pod
Deployments

- Deployment
  - ReplicaSet r1
    - Pod
      - Pod
      - Pod
      - Pod
  - ReplicaSet r2
    - Pod
      - Pod
      - Pod
      - Pod
Deployments

Deployment

ReplicaSet r1

ReplicaSet r2

Pod

Pod

Pod
The magic glue - Labels

• Simple Key-Value pairs
• Can be attached to anything
• LabelSelectors used for selection
• Can be used for scheduling choices
• Extremely simple and mighty
Services

Pod x → Service
Student

Pod 1
app:Student,v:1

Pod 2
app:Student,v:1

Pod 3
app:Student,v:1

Pod 4
app:Student,v:2
Persistent Volumes

Admin

User

PVClaim

Pod

Persistent Volumes

1 2 3 4

StorageClass

User
Further principals

- Namespaces
- StatefulSet
- DaemonSet
- Batch/Scheduled Jobs
- Ingress
- ...
An open platform
API driven

- Complete communication over API’s
- Well defined API’s
- gRPC
- Extensible platform
  - Custom Resource Definitions
  - Custom Schedulers
  - API aggregation layer
Built on standards (plugins)
Core Infrastructure

• Flexible infrastructure as building block
• Cluster add-ons
  • Logging / DNS / Monitoring
• Spark on Kubernetes
• Everything deployed as containers
• Service Mesh on top
  • https://istio.io / https://linkerd.io
Cloud-vendor neutral
Platform independent

• No dependency to any cloud provider
  • Doesn’t care if running on bare-metal or in the cloud
• Cloud provider will be moved out-of-tree
• Kubernetes can make use of the underlying infrastructure
• Basically needs only an OS that runs containers
  • Linux and Windows support (arm/amd64)
Outlook
Kubernetes Federation

- Manage multiple Kubernetes clusters
- One federated control plane
  - On-premise or multi-cloud
- Running global services
- Some known issues
- High Availability
- Challenge -> Storage
Kubernetes 1.8+

- Focus on security and stability
- Allowing more customization
- Improve cluster federation
- Kubernetes and Big Data
Conclusions
Conclusions

• Open platform (OSS)
• Can be used to run simulations/tests
  • Is used in some academic fields
  • e.g. https://lasp-lang.readme.io
• Can be extended for special needs
• Growing eco-system
We are hiring!

Become an inovexpert

www.inovexperts.com
Further reading

https://research.google.com/pubs/pub44843.html
http://queue.acm.org/detail.cfm?id=2898444
https://research.google.com/pubs/pub41684.html
https://speakerdeck.com/thockin/kubernetes-understanding-pods-vs-containers
https://kubernetes.io/docs/concepts/api-extension/custom-resources
Further reading

https://www.opencontainers.org/about

https://containerd.io

https://github.com/kubernetes/community/blob/master/contributors/devel/container-runtime-interface.md

https://docs.microsoft.com/en-us/azure/container-service/kubernetes/container-service-kubernetes-windows-walkthrough
Further reading

https://research.google.com/pubs/pub45499.html
https://github.com/kubernetes/kubernetes/issues/38893
https://docs.google.com/spreadsheets/d/1AFksRDgAt6BGA3OjRNliO3IyKmA-GU7CXaxbihy48ns/edit#gid=0
https://github.com/apache-spark-on-k8s
https://github.com/kubernetes-incubator