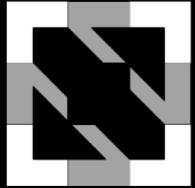




KubeCon



CloudNativeCon

China 2018

深入了解 Kubernetes 上的 Windows 容器

Pengfei Ni

Microsoft Azure

Kubernetes Maintainer



KubeCon



CloudNativeCon

China 2018

1. Windows 容器简介

什么是容器



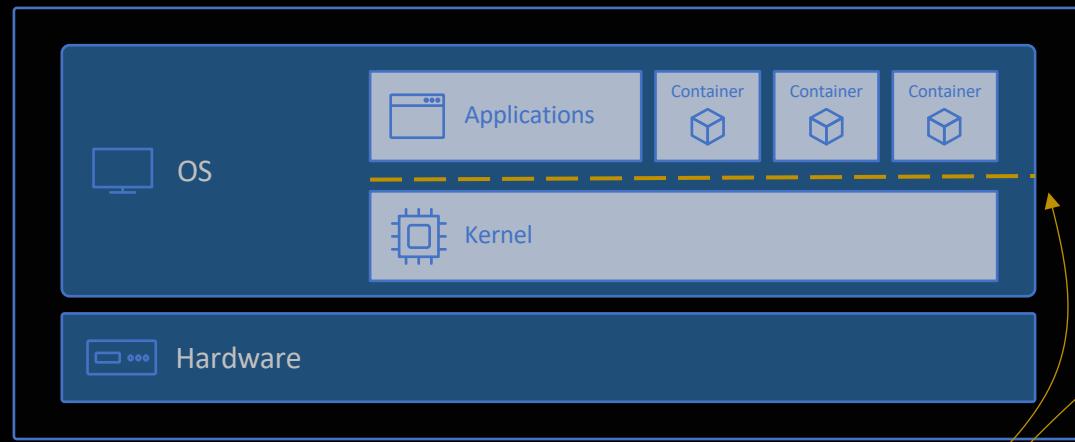
KubeCon



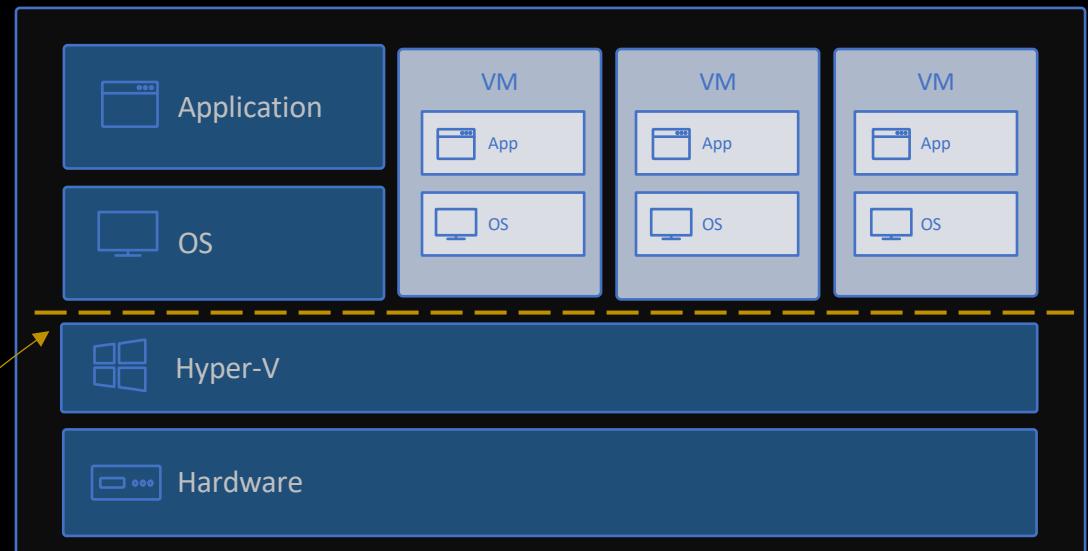
CloudNativeCon

China 2018

容器 = 操作系统虚拟化



传统虚拟机 = 硬件虚拟化

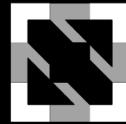


虚拟化边界

为什么要用容器

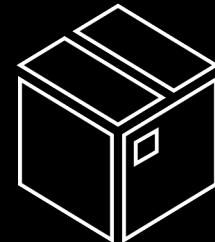


KubeCon



CloudNativeCon

China 2018



快速迭代



敏捷部署



不可变性



节省成本



高效部署



弹性突发

开发者

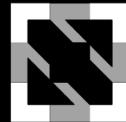
IT

容器的构成

```
PS> docker build -t myimage .
PS> docker run -itd myimage
PS> docker ps
```

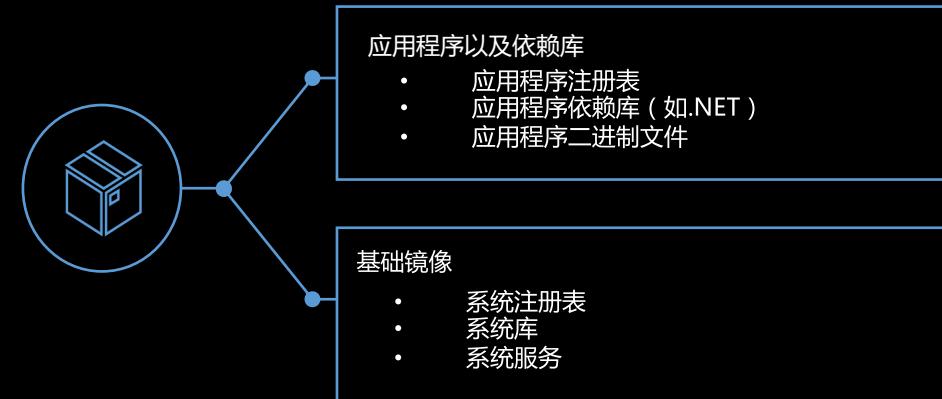


KubeCon



CloudNativeCon

China 2018



#基础镜像

```
FROM microsoft/windowsservercore:1803
```

#安装 IIS

```
RUN powershell -Command Add-WindowsFeature Web-Server
```

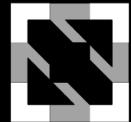
#定义容器启动时的默认命令

```
CMD [ "ping", "localhost", "-t" ]
```

Windows 基础镜像



KubeCon



CloudNativeCon

China 2018

nanoserver

为云应用而生

支持 .NET core

93MB

windowsservercore

兼容已有应用

支持 Full .NET
framework

1.4GB

Windows

自动化负载

包含 Windows OS 的大部
分组件

3.4GB

windowsservercore-insider
nanoserver-insider
windows-insider

仅用于开发和测试

[microsoft/windowsservercore:1803](https://mcr.microsoft.com/windows/nanoserver:1803) ⇒ mcr.microsoft.com/windows/nanoserver:1803

```
PS C:\> ls "C:\ProgramData\docker\windowsfilter\646c94c76da839f088e2e4a5209100496d7f4ae5b01802fa6a9209f530661293\"
```

```
Directory: C:\ProgramData\docker\windowsfilter\646c94c76da839f088e2e4a5209100496d7f4ae5b01802fa6a9209f530661293
```

Mode	LastWriteTime	Length	Name
----	-----	-----	-----
d----	10/9/2018 8:41 PM		Files
d----	10/9/2018 8:41 PM		Hives
-a---	10/16/2018 12:18 PM	429	layerchain.json

```
PS C:\> ls "C:\ProgramData\docker\windowsfilter\04f084cfcd981103e0efc7da2b8c341913d63f36ebcaa537e9cae2899912bdad\"
```

```
Directory: C:\ProgramData\docker\windowsfilter\04f084cfcd981103e0efc7da2b8c341913d63f36ebcaa537e9cae2899912bdad
```

Mode	LastWriteTime	Length	Name
----	-----	-----	-----
d----	4/12/2018 10:29 AM		Files
d----	9/13/2018 4:58 PM		Hives
d----	9/13/2018 4:58 PM		UtilityVM
-a---	9/13/2018 4:58 PM	16384	bcd.bak
-a---	9/13/2018 4:58 PM	16384	bcd.log.bak
-a---	9/13/2018 4:58 PM	0	bcd.log1.bak
-a---	9/13/2018 4:58 PM	0	bcd.log2.bak
-a---	9/13/2018 4:58 PM	41943040	blank-base.vhdx
-a---	9/13/2018 4:58 PM	4194304	blank.vhdx
-a---	9/13/2018 4:58 PM	4	layerchain.json
-a---	9/13/2018 4:58 PM	15	layout

容器的运行



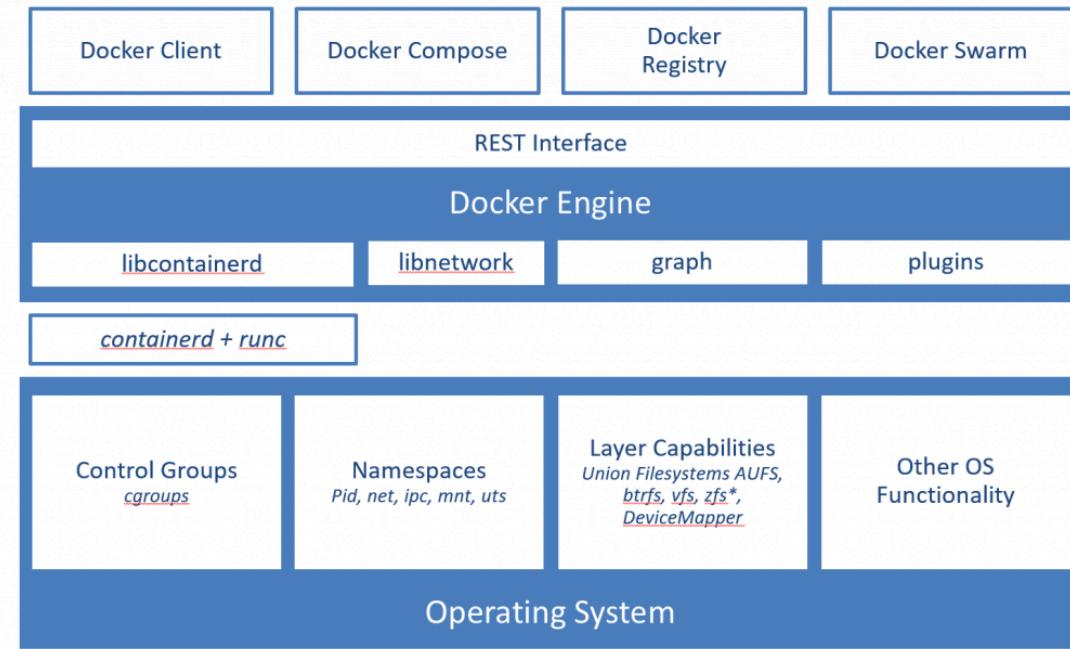
KubeCon



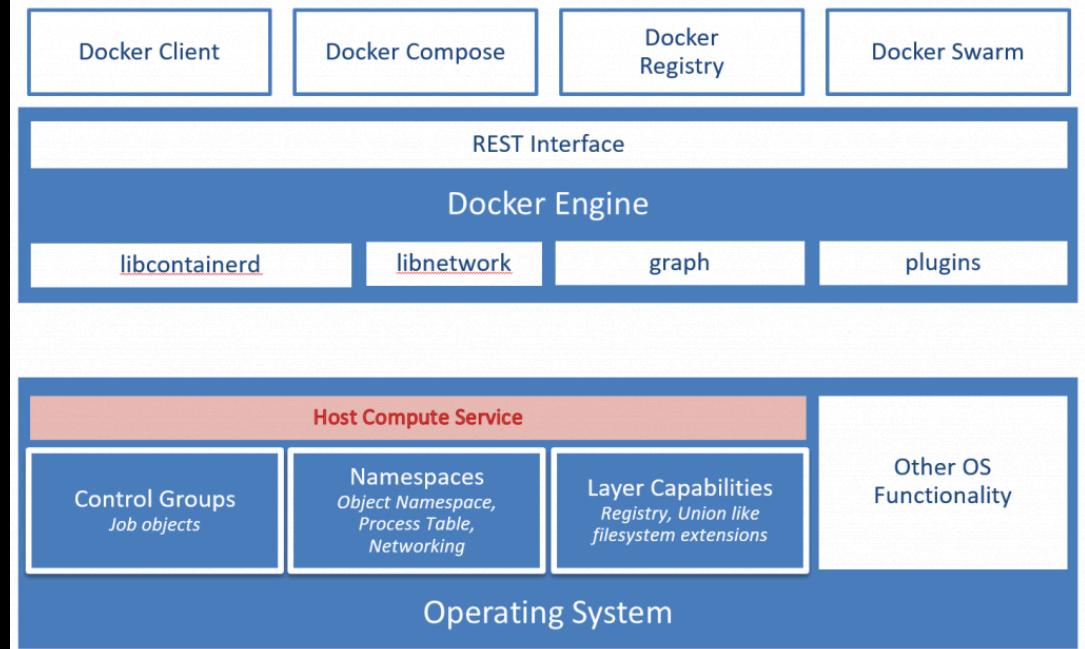
CloudNativeCon

China 2018

Architecture In Linux



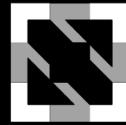
Architecture In Windows



隔离模型



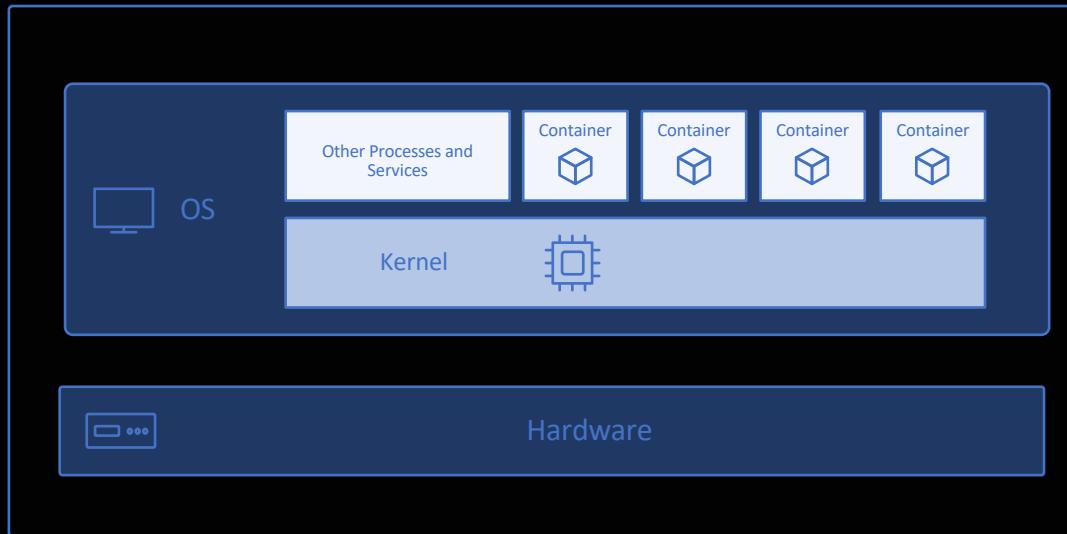
KubeCon



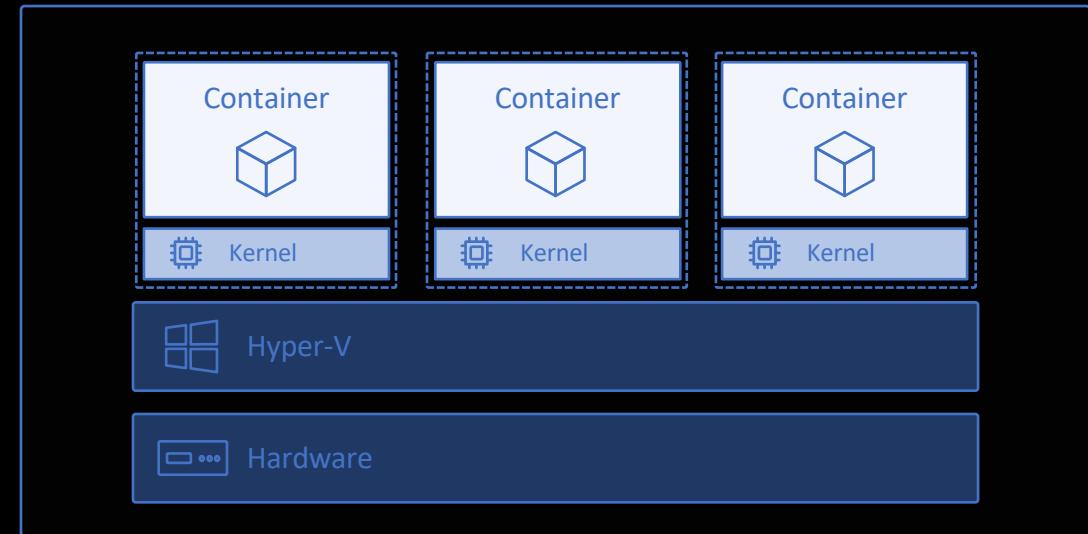
CloudNativeCon

China 2018

进程隔离: 共享内核



Hyper-V 隔离: 独立内核



Window Server 版本



KubeCon



CloudNativeCon

China 2018

Windows Server 2016

Initial launch of containers

Process and Hyper-V isolation

Docker EE Basic Included at no additional cost

Windows Server 1709

Optimized container images for Nano Server and Server Core

Platform level support for Linux containers

Windows Subsystem for Linux

Networking enhancements for overlays and SDN

Windows Server 1803

Optimized Server Core image

App compat improvements

Native command line tools – curl.exe, tar.exe and SSH

Enhancements to the Windows Subsystem for Linux

Networking enhancements for greater density and quicker endpoint creation

Improved network security with Calico Open source storage plugins for Kubernetes

Platform functionality required for Kubernetes conformance

Windows Server 2019

Optimized Server Core image

App compat improvements

Enhanced Group Managed Service Account support

Platform functionality for Kubernetes and Microsoft Service Fabric

Performance and density improvements

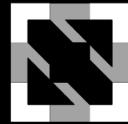
Platform and open source work on CNI networking plugins such as Calico and Flannel

Enhancements to the Windows Subsystem for Linux

版本兼容性



KubeCon



CloudNativeCon

China 2018

Container Version	Host OS Version			
	Windows Server 2016 Builds: 14393	Windows Server 1709 Builds 16299	Windows Server 1803 Builds 17134	Windows Server 2019 Builds 17763
Windows Server 2016	Both Hyper-V and Process Isolation	Hyper-V Isolation only	Hyper-V Isolation only	Hyper-V Isolation only
Windows Server 1709	Not supported	Both Hyper-V and Process Isolation	Hyper-V Isolation only	Hyper-V Isolation only
Windows Server 1803	Not supported	Not supported	Both Hyper-V and Process Isolation	Hyper-V Isolation only
Windows Server 1809	Not supported	Not supported	Not supported	Both Hyper-V and Process Isolation



KubeCon



CloudNativeCon

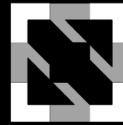
China 2018

2. Windows Server 上的 Kubernetes

容器的挑战



KubeCon

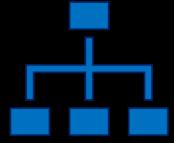


CloudNativeCon

China 2018



调度编排



服务发现



负载均衡



监控告警

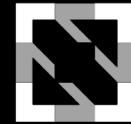


故障恢复

为什么要用Kubernetes



KubeCon



CloudNativeCon

China 2018

The Production-Grade
Container Orchestration



Portable

Public, private, hybrid, multi-
cloud

Extensible

Modular, pluggable,
hookable, composable

Self-healing

Auto-placement, auto-restart, auto-
replication, auto-scaling

Kubernetes 101

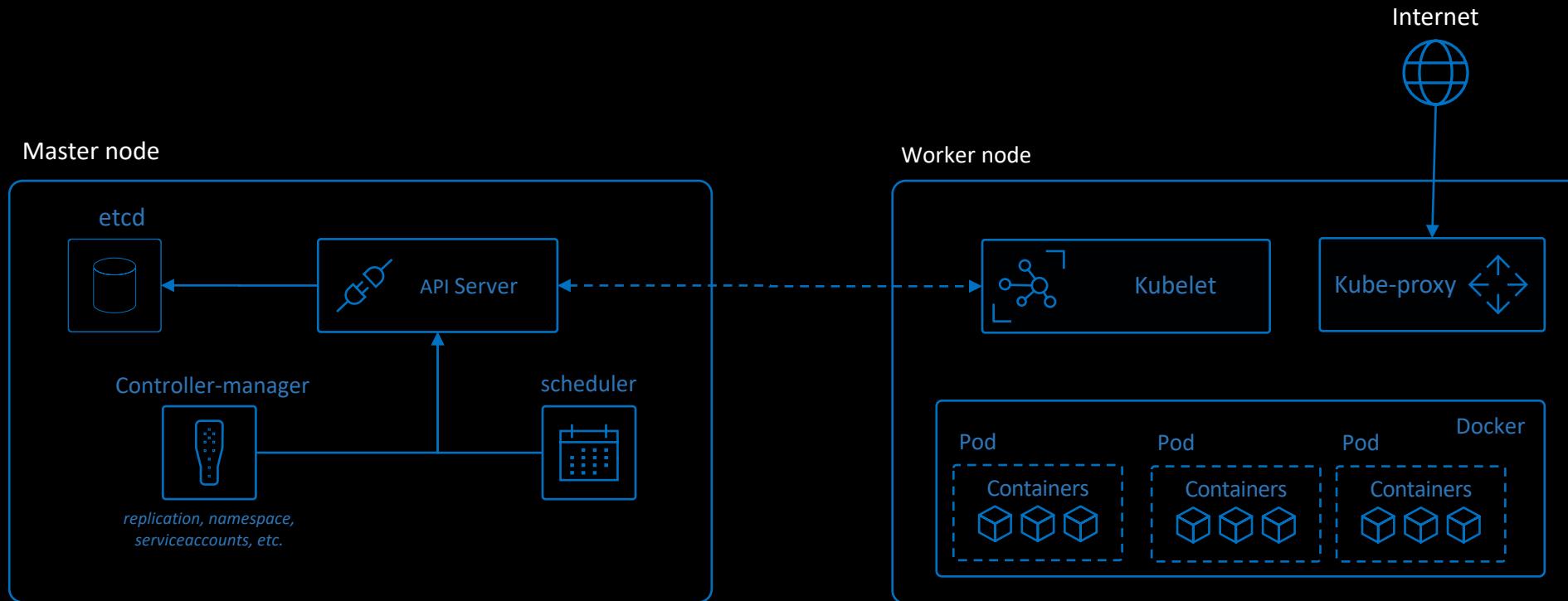


KubeCon



CloudNativeCon

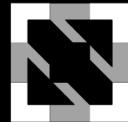
China 2018



Kubernetes+Windows 发展历程

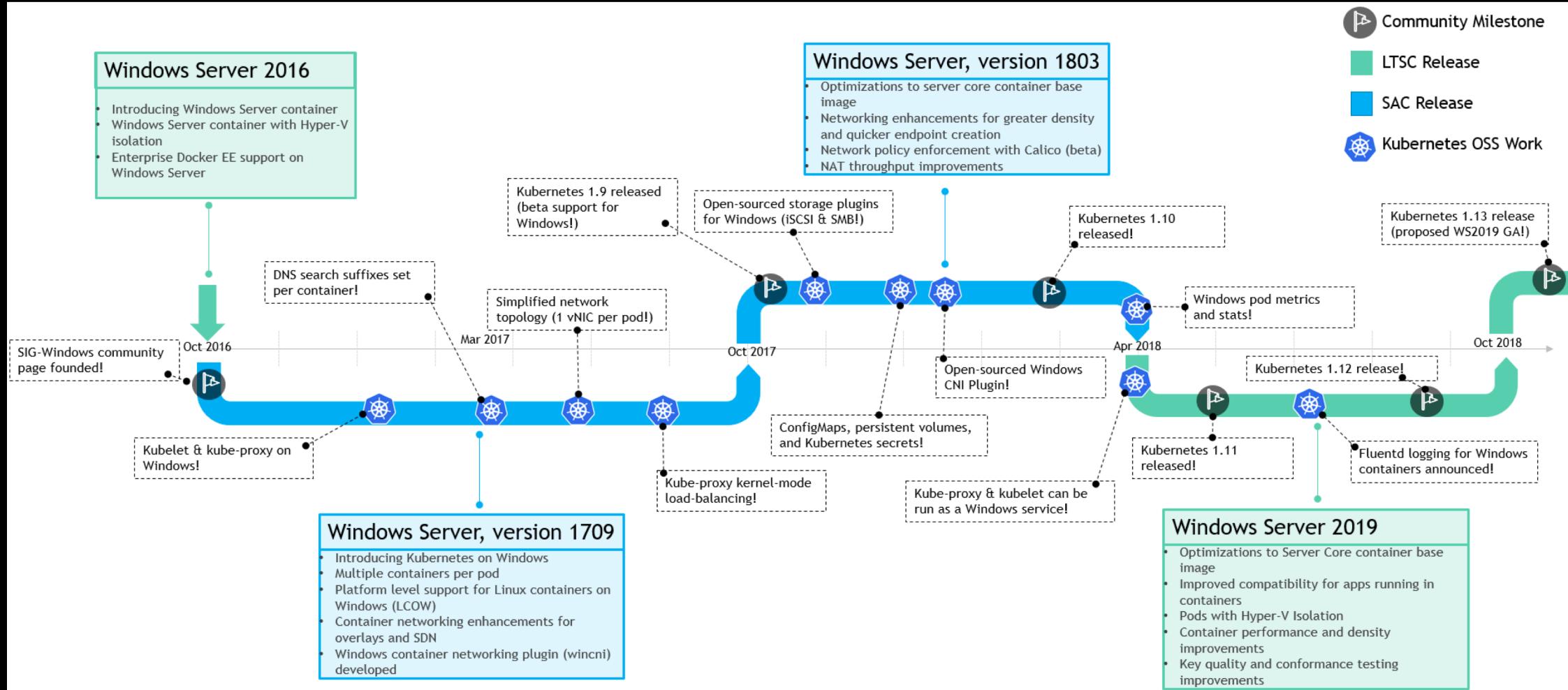


KubeCon



CloudNativeCon

China 2018



核心特性



Alpha: v1.5
Beta: v1.9
GA: ETA v1.14



支持共享相同命名空间的
多容器 Pod



基于 Virtual Filtering
Platform (VFP) 的负载均衡



容器运行时接口 (CRI)
容器网络插件 (CNI)



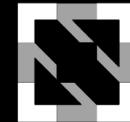
可使用 kubeadm 添加
Windows Server 节点



支持进程隔离和 Hyper-V
隔离



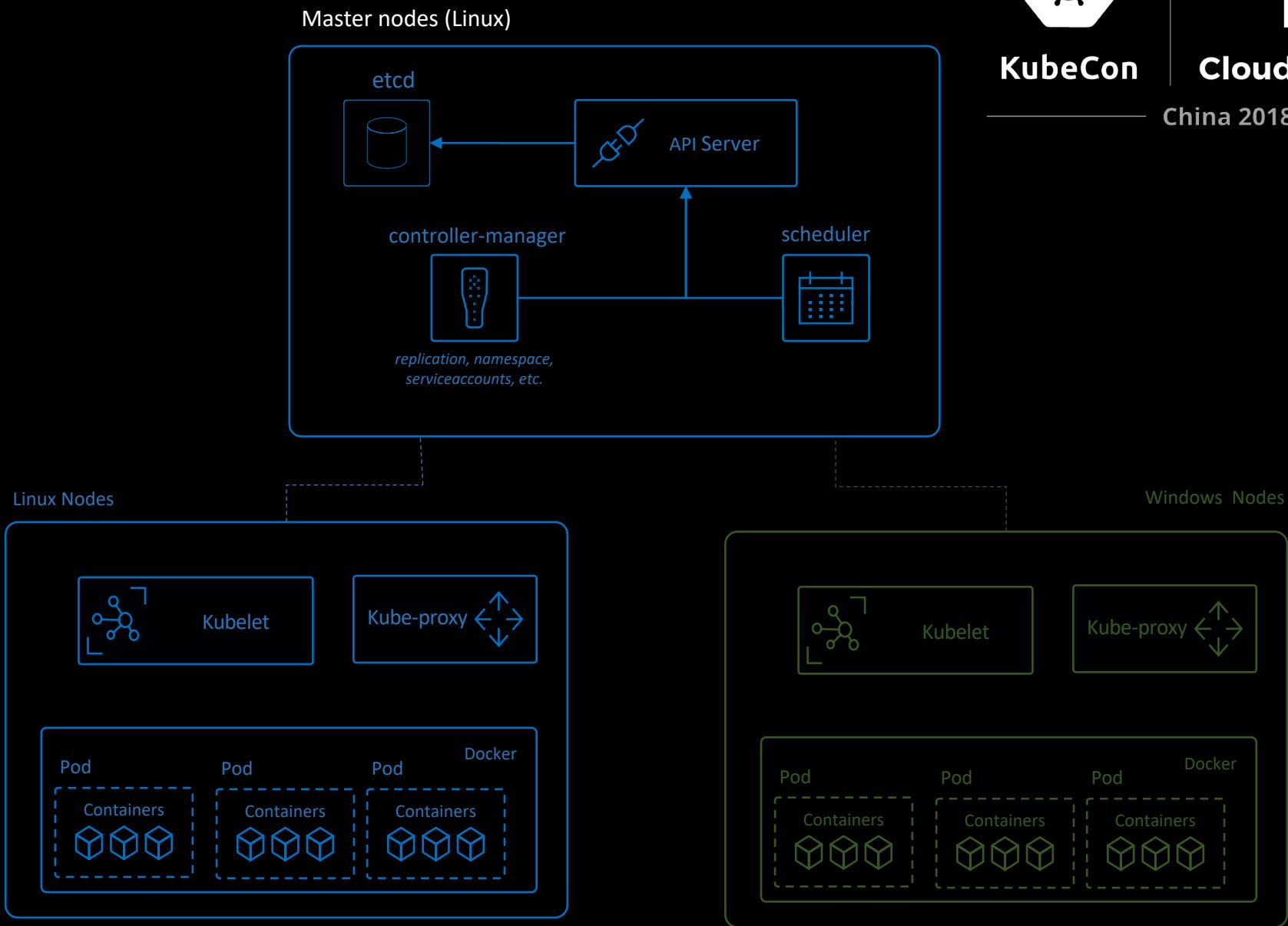
KubeCon



CloudNativeCon

China 2018

集群



KubeCon



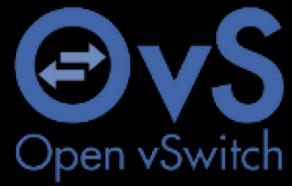
CloudNativeCon

China 2018

网络插件



Azure CNI



ovn-kubernetes

wincni

win-bridge
win-overlay

flannel

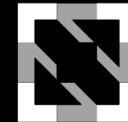
overlay
l2bridge



Calico
(Policy Only)

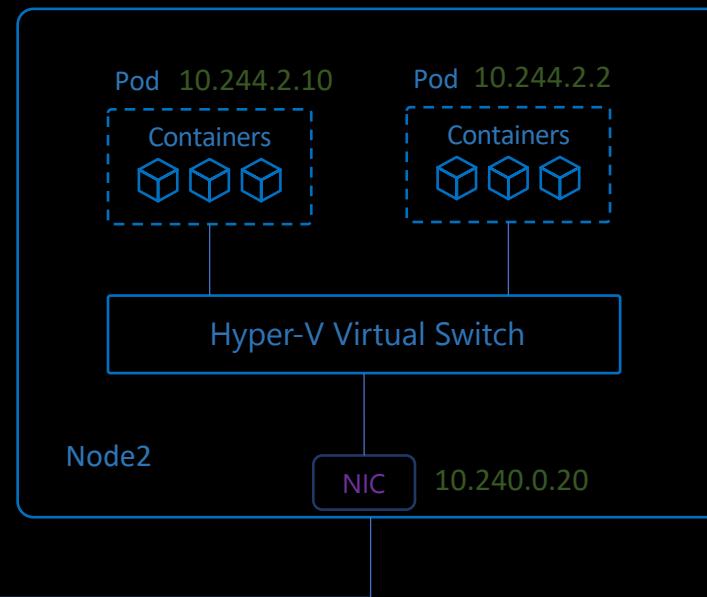
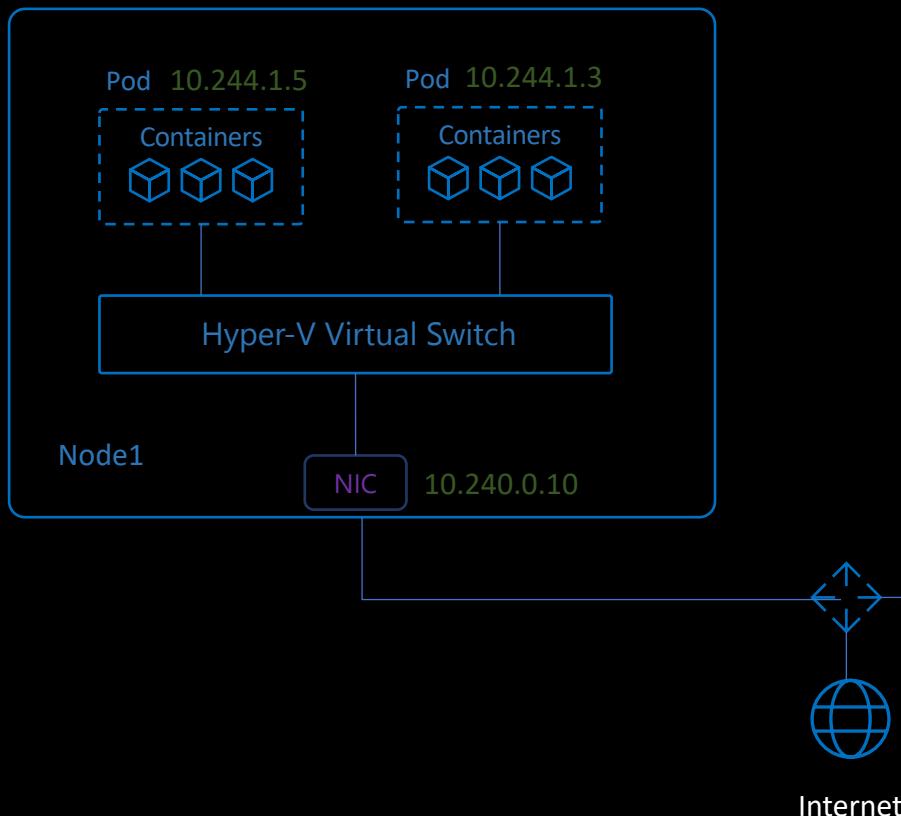


KubeCon



CloudNativeCon

China 2018

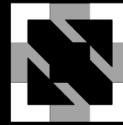


Prefix	Next Hop
10.244.1.0/24	10.240.0.10
10.244.2.0/24	10.240.0.20

存储卷



KubeCon



CloudNativeCon

China 2018



In-tree volume

AzureDisk, AzureFile

HostPath, ...



Flex volume

KeyVault, blobfuse, SMB



CSI

Working In Process

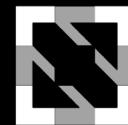
Pod示例

```
apiVersion: v1
kind: Pod
metadata:
  name: iis
  labels:
    name: iis
spec:
  containers:
    - name: iis
      image: microsoft/iis:windowsservercore-1709
      ports:
        - containerPort: 80
nodeSelector:
  "beta.kubernetes.io/os": windows
```

```
apiVersion: apps/v1
kind: DaemonSet
metadata:
  name: my-daemonset
  labels:
    app: foo
spec:
  selector:
    matchLabels:
      app: foo
  template:
    metadata:
      labels:
        app: foo
    spec:
      containers:
        - name: foo
          image: microsoft/windowsservercore:1709
nodeSelector:
  beta.kubernetes.io/os: windows
```



KubeCon



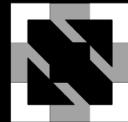
CloudNativeCon

China 2018

已知问题



KubeCon



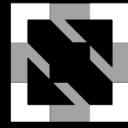
CloudNativeCon

China 2018

- 仅在 Windows Server 1709 或者以后版本的进程隔离模式才支持多容器 Pod
- Windows 容器的操作系统版本需要跟 Host 的操作系统版本匹配
- 通过网络插件为容器配置 DNS (而不是 resolv 文件)
- 仅支持有限的存储插件
- 不支持 Weave Net 和 Calico 网络插件
- 不支持本地网络策略 (Local traffic policy)



KubeCon



CloudNativeCon

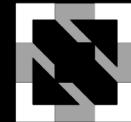
China 2018

3. 未来展望

预计 v1.14 GA



KubeCon



CloudNativeCon

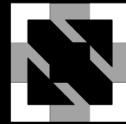
China 2018

- 支持 Windows Server 2019
- 更多的 Flex 存储卷: SMB, iSCSI
- 支持 Hyper-V 隔离的多容器 Pod
- 测试版的 Containerd 集成
- 集成测试完善，如 conformance tests 和 node e2e tests

未来工作



KubeCon



CloudNativeCon

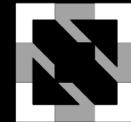
China 2018

- Windows Server 中的 Linux 容器
- Containerd 默认容器运行时
- 更多的存储卷: flexvolume, CSI
- 特权模式的 DaemonSet

参考文档



KubeCon



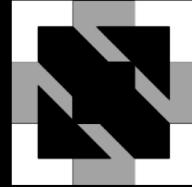
CloudNativeCon

China 2018

- <https://docs.microsoft.com/en-us/virtualization/windowscontainers/kubernetes/getting-started-kubernetes-windows>
- <https://kubernetes.io/docs/getting-started-guides/windows/>
- <https://blogs.technet.microsoft.com/networking/2018/09/19/ws2019-kubernetes/>
- <https://www.youtube.com/watch?v=j2B7cLdTXMw>
- <https://github.com/Microsoft/SDN/tree/master/Kubernetes>
- <https://trello.com/b/rjTqrwjI/windows-k8s-roadmap>



KubeCon



CloudNativeCon

China 2018

Thanks & QA