

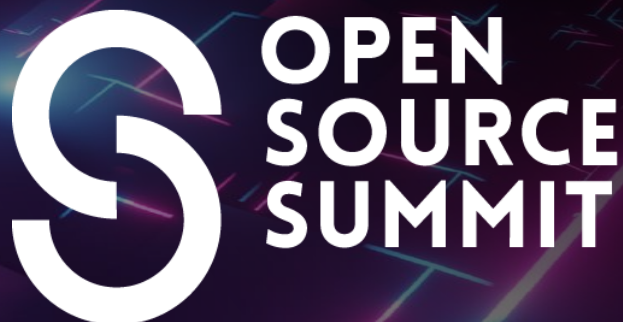


**KubeCon**



**CloudNativeCon**

THE LINUX FOUNDATION



**AI\_dev**  
Open Source GenAI & ML Summit

---

**China 2024**

---



KubeCon



CloudNativeCon



China 2024

# 深入探讨Windows CSI driver HostProcess容器

Andy Zhang, Principal Software Engineer @Microsoft

# Intro

## Andy Zhang (张夏)

- 存储项目负责人@AKS(Azure Kubernetes Service) 组
- 多个k8s开源项目维护者
  - [Azure Disk/File/Blob](#) CSI drivers
  - [SMB, NFS](#) under [github.com/kubernetes-csi](https://github.com/kubernetes-csi)
  - [Local volume provisioner](#)
  - [Azure cloud provider](#)
  - [csi-proxy](#)



Andy Zhang  
andyzhangx

- Github: [andyzhangx](#)
- Email: [xiazhang@microsoft.com](mailto:xiazhang@microsoft.com)
- Slack: andyzhang [@kubernetes](#)

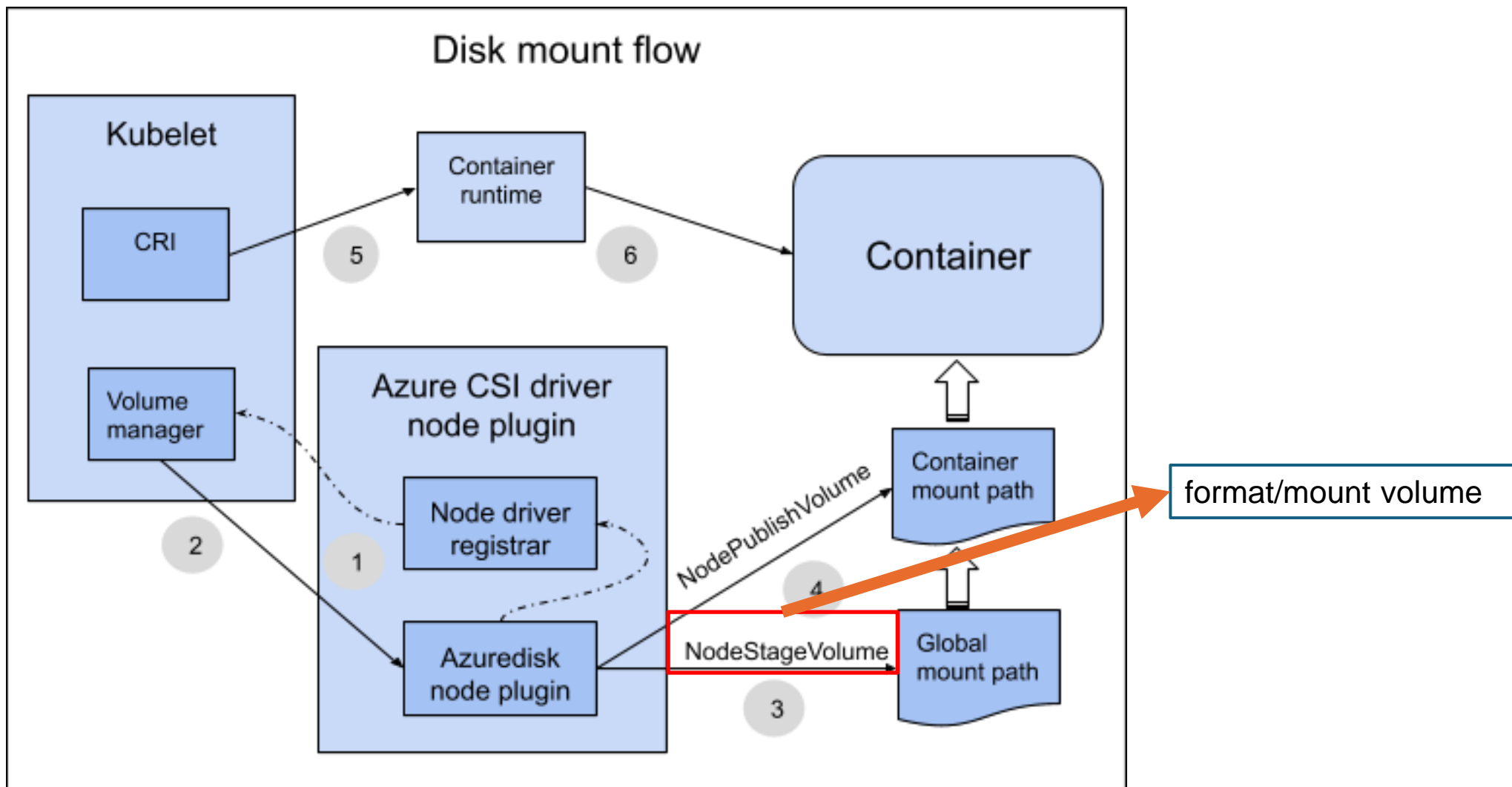
- 介绍
- Windows csi-proxy 项目背景
- HostProcess 容器的介绍
- 将 Windows 上的 CSI 驱动从 csi-proxy 依赖迁移到 HostProcess 容器部署
- 关键的学习和注意事项
- 云服务提供商的迁移进展
- 结论和未来工作

- 大多数 Windows 节点上的 CSI 驱动依赖于 csi-proxy 进程
  - csi-proxy 代表 CSI 驱动容器执行特权操作
- Windows HostProcess 容器
  - 在 k8s 1.23 中引入，在 k8s 1.26 中稳定
  - 作为具有特权的常规进程直接在 Windows 主机上运行
  - 切换到 HostProcess 容器部署将使 Windows CSI 驱动的开发和部署更加容易

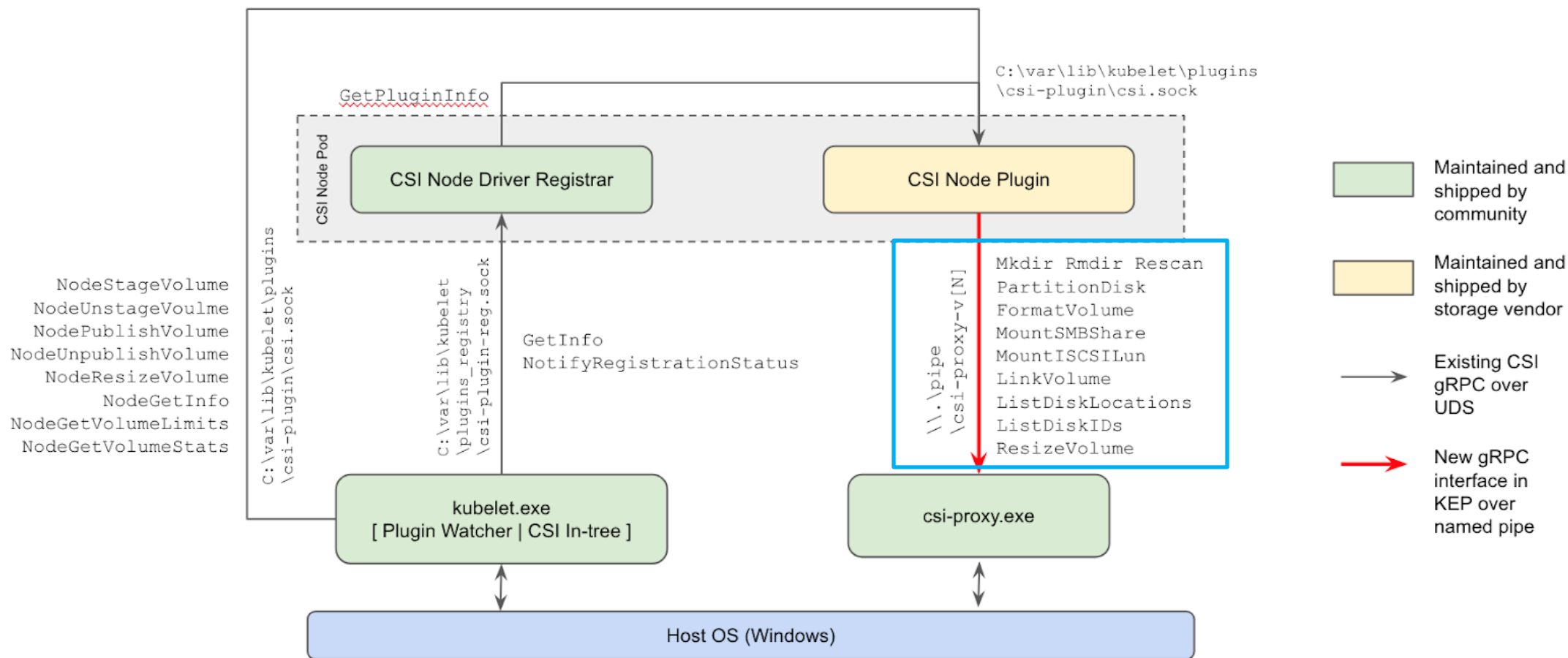


- 内置存储驱动程序
  - Kubelet 在 Linux 和 Windows 节点上处理内置存储驱动程序的特权操作
    - 例如，磁盘格式化、挂载/卸载等操作
- CSI 驱动程序
  - 在 k8s 1.13 中引入
  - 内置驱动程序的移除和迁移到 CSI 驱动程序
  - 节点插件容器需要特权权限才能执行存储操作
  - 在 Kubernetes v1.23 之前，特权 pod 仅支持 Linux 节点，而不支持 Windows 节点

# CSI 驱动程序在Linux节点上与k8s的交互方式



# CSI驱动程序在Windows节点上使用csi-proxy与k8s的交互方式



参考: <https://raw.githubusercontent.com/kubernetes/enhancements/master/keps/sig-windows/csi-proxy3.png>





KubeCon



CloudNativeCon



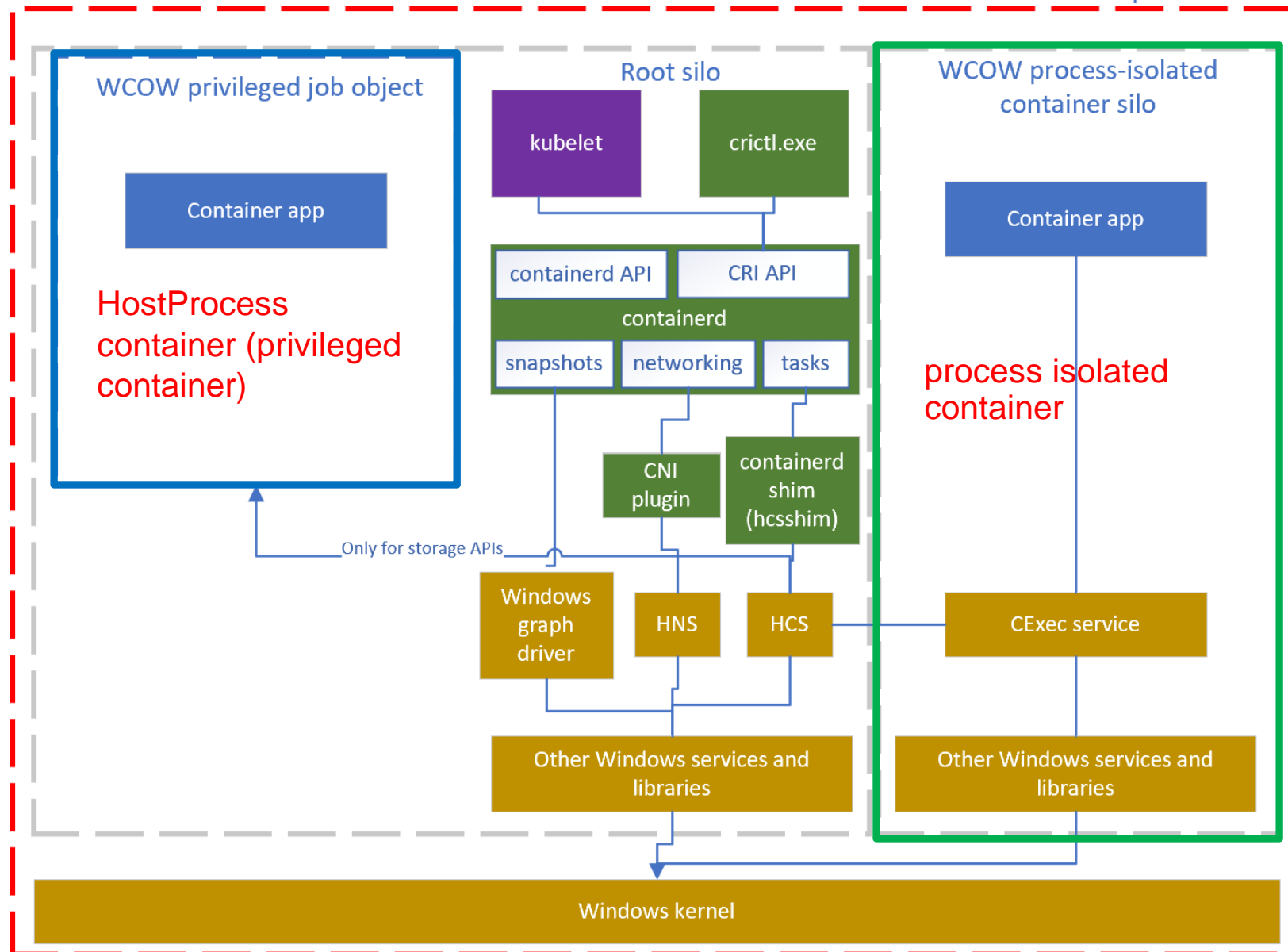
China 2024



- HostProcess 容器
  - 在 k8s v1.23 中引入，在 v1.26 中正式发布
  - 作为具有特权的常规进程直接在主机上运行
- 相对于Windows 节点上的csi-proxy，HostProcess 容器的优势
  - 简化开发和部署流程
  - 简化故障排除过程
  - 镜像大小更小

## Window Process isolated Containers vs Windows Privileged Containers

Root partition



# HostProcess 相对于csi-proxy的优势



China 2024

- 使用 csi-proxy 的开发和部署流程
  1. 在 csi-proxy 服务器和客户端中添加新的 API
  2. 发布新的 csi-proxy 版本
  3. 在 CSI 驱动程序中采用新的 csi-proxy 版本
  4. 将新的 csi-proxy 版本部署到 Windows 节点上
  5. 在所有 Windows 节点上升级新的 CSI 驱动程序版本

痛点:

- 在某些 Kubernetes 集群中，不支持就地替换，需要升级所有节点的整个镜像
- 新的 CSI 驱动程序版本必须与所有最近的 csi-proxy 版本兼容

- 使用 csi-proxy 进行故障排除的痛点
  - 用户需要登录到节点，确保 csi-proxy 服务正常运行，然后在故障排除过程中额外获取 csi-proxy 日志

## [csi-proxy](#) troubleshooting

- ssh to the windows node(e.g. 512k8s010), and then run Powershell command

```
azureuser@2512k8s010 C:\Users\azureuser>powershell

cd c:\k
ls csi-proxy.exe

PS C:\k> Get-Service csi-proxy

Status   Name           DisplayName
-----
Running  csi-proxy      csi-proxy

PS C:\k> cat C:\k\csi-proxy.err.log
I0905 13:22:22.341048    3880 main.go:55] Starting CSI-Proxy Server ...
I0905 13:22:22.398066    3880 main.go:56] Version: v0.2.2-0-gffb169f
```

# HostProcess 相对于csi-proxy的优势

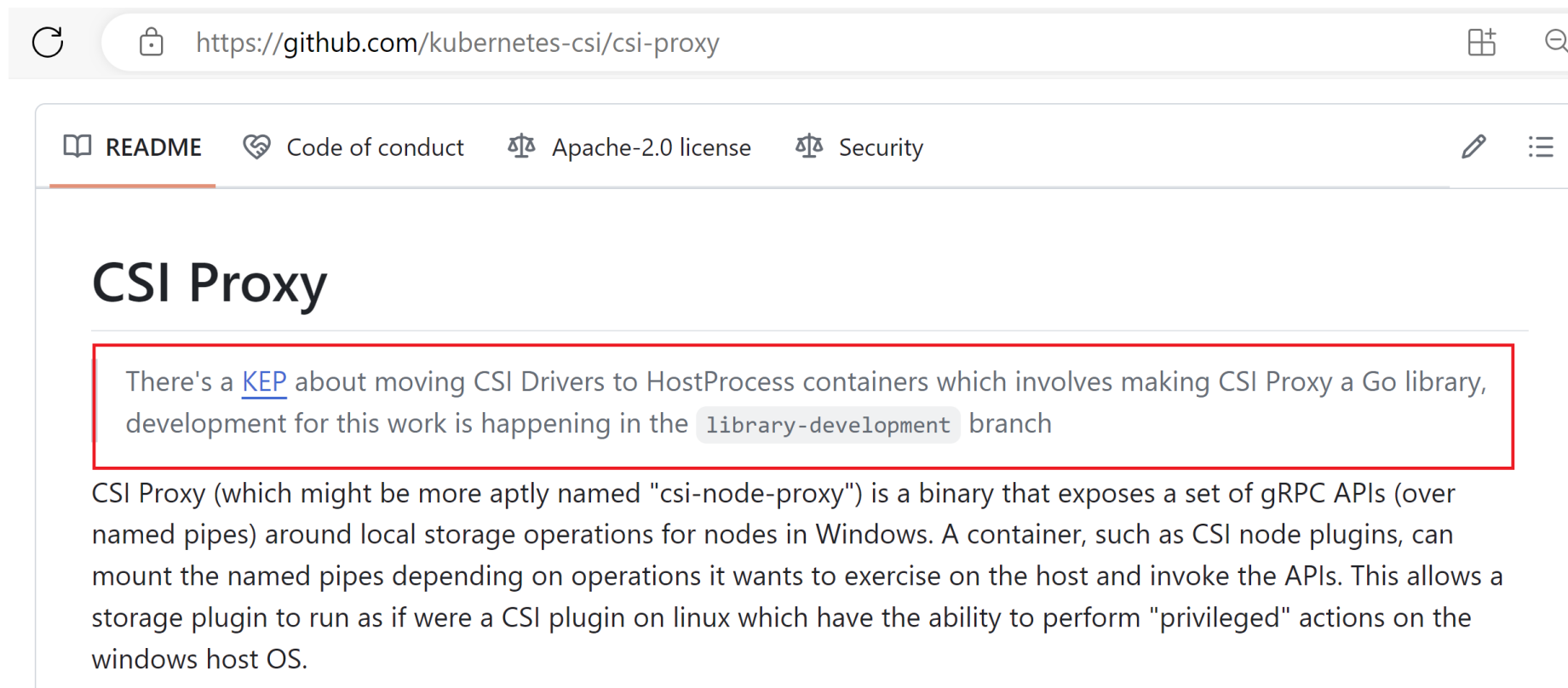


China 2024

- 使用 HostProcess CSI 驱动程序的开发和部署流程
  1. 在 CSI 驱动程序中添加新功能（无需新的 API）
  2. 发布新的 CSI 驱动程序版本
  3. 在所有 Windows 节点上部署新的 CSI 驱动程序版本



- <https://github.com/kubernetes-csi/csi-proxy>



# HostProcess 容器限制



KubeCon



CloudNativeCon



China 2024



- 需要 Kubernetes 1.23+
- 需要 containerd 1.6+
- 默认情况下，Host Process容器只提供有限的主机用户帐户
  - CSI驱动程序使用本地系统帐户
- HostProcess容器不支持文件系统和Hyper-V隔离
- HostProcess pod只能包含HostProcess容器

# 迁移到 HostProcess 容器部署的步骤



China 2024

- HostProcess 容器接口更改
- HostProcess 容器部署更改
- 为 HostProcess 容器镜像构建 Windows 镜像

- CSI驱动程序接口中的代码更改
  - NodeStageVolume
  - NodeUnstageVolume
  - NodePublishVolume
  - NodeUnpublishVolume
  - NodeExpandVolume
  - **NodeGetVolumeStats**

- 用容器内的直接调用替换csi-proxy调用

## AzureDisk change:

FormatAndMount  
ExistsPath  
Rmdir  
Rescan  
FindDiskByLun  
GetDeviceNameFromMount  
GetVolumeSizeInBytes  
ResizeVolume  
**GetVolumeStats**

## AzureFile(SMB mount) change:

SMBMount  
ExistsPath  
MakeDir  
Rmdir  
IsMountPointMatch  
EvalHostSymlinks  
**GetVolumeStats**



# 迁移 - HostProcess容器部署



China 2024

<https://kubernetes.io/docs/tasks/configure-pod-container/create-hostprocess-pod/>

kind: DaemonSet

apiVersion: apps/v1

metadata:

name: csi-azurefile-node-win

spec:

...

securityContext:

windowsOptions:

**hostProcess: true**

**runAsUserName: "NT AUTHORITY\\SYSTEM"**

hostNetwork: true

- <https://github.com/microsoft/windows-host-process-containers-base-image>

- HostProcess容器基础镜像

*mcr.microsoft.com/oss/kubernetes/windows-host-process-containers-base-image:v1.0.0*

- 与nanoserver镜像相比的优势

- 大小

- 几KB与几百MB

- 操作系统兼容

- 单个HostProcess容器镜像标签与所有流行版本的Windows操作系统兼容，例如 Windows Server 2019、2022等

- Windows HostProcess镜像的Dockerfile示例

```
FROM mcr.microsoft.com/oss/kubernetes/windows-host-process-containers-base-image:v1.0.0
LABEL description="CSI Azure disk plugin"

ARG ARCH=amd64
ARG binary=./_output/${ARCH}/azurediskplugin.exe
COPY ${binary} /azurediskplugin.exe
ENV PATH="C:\Windows\system32;C:\Windows;C:\WINDOWS\System32\WindowsPowerShell\v1.0\;"
USER ContainerAdministrator
ENTRYPOINT ["/azurediskplugin.exe"]
```

# 关键学习和注意事项



China 2024

```
PS C:\Windows\System32> kubectl top po --containers -n kube-system | Select-String -Pattern "disk|file"
```

csi-azure←[7mdisk←[0m-node-win-m2vrh	azuredisk	841m	707Mi
csi-azure←[7mdisk←[0m-node-win-m2vrh	node-driver-registrar	0m	5Mi
csi-azure←[7mfile←[0m-node-win-xwp4r	azurefile	0m	10Mi
csi-azure←[7mfile←[0m-node-win-xwp4r	node-driver-registrar	0m	5Mi

有关详细信息，请参阅 <https://github.com/kubernetes-sigs/azuredisk-csi-driver/issues/2235>

- 解决HostProcess容器性能问题的提示

- 使用wpr（Windows性能记录器）在Windows节点上捕获资源使用情况

有关详细信息，请参阅 <https://github.com/kubernetes-sigs/azuredisk-csi-driver/issues/2235>

```
kubectl exec -it csi-azurefile-node-win-xxx -n kube-system -c azurefile -- cmd  
wpr -start cpu  
# sleep 30s  
wpr -stop 'c:\cpu.etl'  
kubectl cp kube-system/csi-azurefile-node-win-xxx:/cpu.etl -c azurefile /tmp/cpu.etl
```



# 关键学习和注意事项

- 使用Windows性能分析器按进程基础分析CPU使用情况
  - <https://learn.microsoft.com/en-us/windows-hardware/test/wpt/windows-performance-analyzer>

Line #	Process Name	Process	% Weight	Sum S
1	▼ powershell.exe		60.08	
2		▷ powershell.exe (15096)	44.15	
3		▷ powershell.exe (6424)	11.66	
4		▷ powershell.exe (9488)	2.55	
5		▷ powershell.exe (13660)	1.57	
6		▷ powershell.exe (1828)	0.13	
7	▷ svchost.exe		19.01	

GetVolumeStats

(Get-Item -Path \$Env:mount).Target

Line #	Process	Sessio...	Duration (s)	Sum	Command Line	Start Time (s)	Min	End Time (s)	Max
1	powershell.exe (16260)	11	137.082586900		powershell.exe -File C:\entrypoint\entrypoint.ps1	0.000000000		137.082586900	
2	powershell.exe (8516)	15	137.082586900		powershell.exe -File C:\entrypoint\entrypoint.ps1	0.000000000		137.082586900	
3	powershell.exe (6424)	10	137.082586900		powershell.exe -File C:\entrypoint\entrypoint.ps1	0.000000000		137.082586900	
4	powershell.exe (6228)	8	137.082586900		powershell C:\opt\scripts\main.ps1	0.000000000		137.082586900	
5	powershell.exe (4296)	16	137.082586900		C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe -noexit -Command "Invoke-SCo	0.000000000		137.082586900	
6	powershell.exe (4192)	0	137.082586900		"C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe" c:\kubeproxy\start.ps1	0.000000000		137.082586900	
7	powershell.exe (1828)	0	137.082586900		"C:\Windows\System32\WindowsPowerShell\v1.0\powershell.exe" c:\k\kubelet\start.ps1	0.000000000		137.082586900	
8	powershell.exe (9488)	0	0.683055300		powershell -Mta -NoProfile -Command "(Get-Item -Path \$Env:mount).Target"	97.892886300		98.575941600	
9	powershell.exe (15096)	0	0.600230200		powershell -Mta -NoProfile -Command "(Get-Item -Path \$Env:mount).Target"	98.602900200		99.203130400	
10	powershell.exe (13660)	0	0.700041500		powershell -Mta -NoProfile -Command "(Get-Item -Path \$Env:mount).Target"	99.240264500		99.940306000	

# 关键学习和注意事项



China 2024

- 如何减少Windows节点上GetVolumeStats的CPU和内存使用示例
  - 提示: PowerShell命令调用开销较大, 请尽可能减少使用
  - 在GetVolumeStats中添加10分钟过期缓存
    - <https://github.com/kubernetes-sigs/azuredisk-csi-driver/pull/2273>
    - (Get-Item -Path \$Env:mount).Target

# 关键学习和注意事项



China 2024

- 尽可能多地使用 Golang API 调用来替换 PowerShell 命令调用

Powershell 命令	Golang API
Get-Volume -Uniqueid \$volumeID   Select SizeRemaining,Size	windows.GetDiskFreeSpaceEx
New-Item -ItemType SymbolicLink	os.Symlink



KubeCon



CloudNativeCon



China 2024

- HostProcess容器部署的好处
  - 简化开发和部署过程
  - 简化故障排除
  - 镜像大小更小

# 云服务提供商的迁移进展



China 2024

- **Azure disk and file CSI drivers**
  - 已从Kubernetes v1.27迁移并在AKS上推出
- **AWS ebs driver**
  - Alpha阶段
- **GCP disk driver**
  - 正在进行中



- 未来工作
  - 将csi-proxy作为基于主机进程的CSI驱动程序的Golang库
    - <https://github.com/kubernetes-csi/csi-proxy>
    - 将性能优化更改上游到csi-proxy项目
  - 在上游csi-release-tools库中支持host process镜像构建
    - <https://github.com/kubernetes-csi/csi-release-tools>
  - 将SMB CSI驱动程序迁移到HostProcess容器部署
    - <https://github.com/kubernetes-csi/csi-driver-smb>
  - 在CSI驱动程序中使用Golang API替换更多的PowerShell cmdlet
    - (Get-Item -Path \$Env:mount).Target
    - Test-Path \$Env:remotepath
    - <https://github.com/kubernetes-sigs/azuredisk-csi-driver/issues/2436>

- <https://github.com/kubernetes-csi/csi-proxy>
- <https://kubernetes.io/docs/tasks/configure-pod-container/create-hostprocess-pod/>
- <https://github.com/microsoft/windows-host-process-containers-base-image>
- 基于HostProcess容器的CSI驱动程序示例
  - [https://github.com/kubernetes-sigs/azuredisk-csi-driver/blob/master/pkg/mounter/safe\\_mounter\\_host\\_process\\_windows.go](https://github.com/kubernetes-sigs/azuredisk-csi-driver/blob/master/pkg/mounter/safe_mounter_host_process_windows.go)
  - <https://github.com/kubernetes-sigs/azuredisk-csi-driver/blob/master/deploy/csi-azuredisk-node-windows-hostprocess.yaml>



A high-altitude mountain landscape featuring a vibrant turquoise lake nestled in a valley. In the foreground, a hiker wearing an orange jacket stands on a rocky outcrop, holding a trekking pole. The surrounding slopes are covered in low-lying green vegetation and scattered rocks. In the background, majestic snow-capped mountain peaks rise against a clear blue sky with some light clouds. The overall scene conveys a sense of adventure and natural beauty.

Thanks!  
Q & A