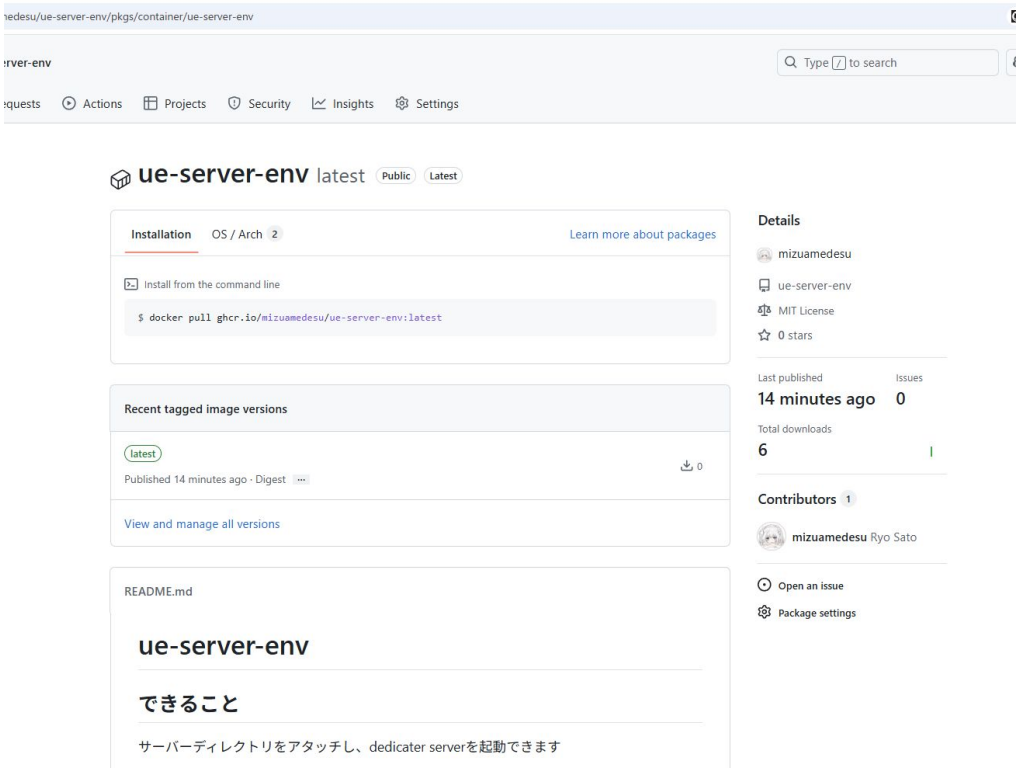
Unreal.k8s 中間報告

筑波大学情報学群情報科学類
2年 佐藤良 @2025/10/15

やったこと1

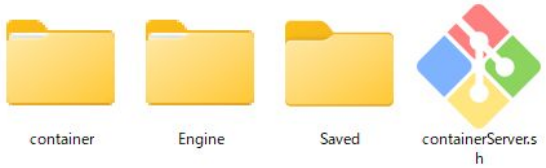


Linuxに書き出したUnrealEngineのdedicated serverを、dockerでー々ビルドせず、該当ディレクトリをマウントするだけで起動できるコンテナを公開しました

https://github.com/mizuamedesu/ue-server-env

やったこと1

```
#!/bin/bash
set -e
echo "Starting UE server setup..."

if [ ! -d "/game/Engine" ] || [ ! -d "/game/${PROJECT_DIR}" ]; then
  echo "ERROR: Required game files are not mounted. Please mount the game directory to /game."
  exit 1
fi

if [ ! -f "/game/${PROJECT_DIR}/${SERVER_SCRIPT}" ]; then
  echo "ERROR: Server start script (${SERVER_SCRIPT}) not found in /game/${PROJECT_DIR}/"
  exit 1
fi

# スクリプトに実行権限を付与
if [ -w "/game/${PROJECT_DIR}/${SERVER_SCRIPT}" ]; then
  chmod +x "/game/${PROJECT_DIR}/${SERVER_SCRIPT}" || true
fi

if [ ! -d "$SAVED_PATH" ]; then
  echo "Creating saves directory at $SAVED_PATH"
  mkdir -p "$SAVED_PATH"
fi

# 環境変数を使ってゲーム設定を修正（必要に応じて）
CONFIG_FILE="/game/${PROJECT_DIR}/Config/DefaultGame.ini"
if [ -f "$CONFIG_FILE" ]; then
  if grep -q "ServerName=" "$CONFIG_FILE"; then
    sed -i "s/ServerName=.*/ServerName=${SERVER_NAME}/" "$CONFIG_FILE"
  fi

  if grep -q "MaxPlayers=" "$CONFIG_FILE"; then
```



UnrealEngineのサーバー書き出しをすると、決まった形で書き出されるためこれを良しなにやるシェルスクリプトと、環境を提供するコンテナイメージです(MITで公開)

やったこと2

#Kata Containers の SandboxChanged を解決する

kataコンテナやりました！！！！！

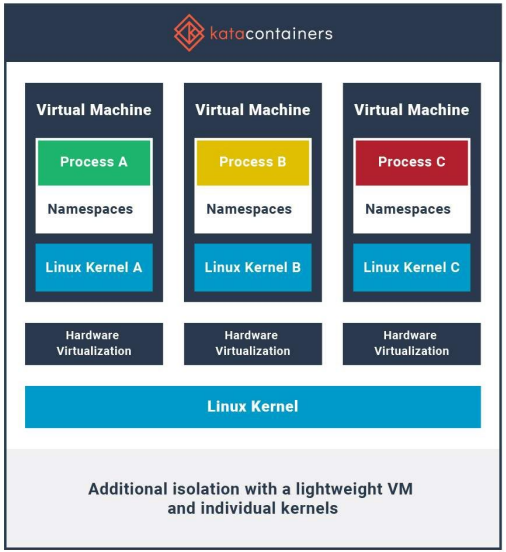
記事も書きました。
https://mizuame.works/blog/2025-10-12/

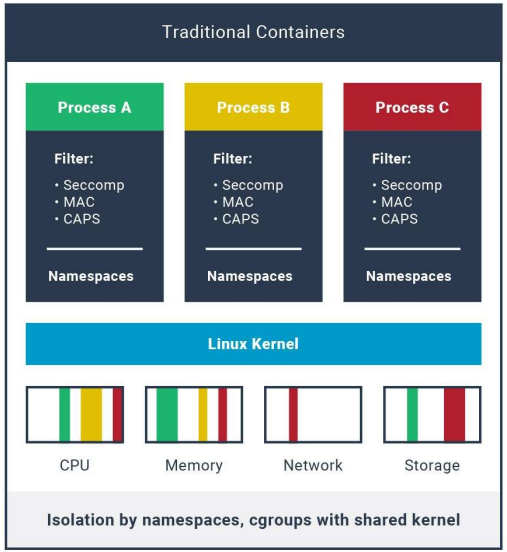
Kata Containersとは

The speed of containers, the security of VMs

↑ ok　　https://katacontainers.io/learn/

Kata Containersとは





OCI コンテナ形式のランタイム。
内部では基本的にQEMUを立ち上げている。
VM並みの隔離をコンテナ環境で実行できる。

https://katacontainers.io/static/6e497f9d3752ca1e354d0d2949abc020/8fef6/katacontainers_traditionalvskata_diagram.jpg

Kata Containersとは

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: ue5-gameserver
  namespace: game
spec:
  replicas: 2
  selector:
    matchLabels:
      app: ue5-server
  template:
    metadata:
      labels:
        app: ue5-server
    spec:
      runtimeClassName: kata
      nodeSelector:
        hardware: game-runner

      affinity:
        podAntiAffinity:
          preferredDuringSchedulingIgnoredDuringExecution:
          - weight: 100
            podAffinityTerm:
              labelSelector:
                matchExpressions:
                - key: app
                  operator: In
```

kata > ! runtimeclass.yaml

```yaml
apiVersion: node.k8s.io/v1
kind: RuntimeClass
metadata:
  name: kata
handler: kata
```

Alt runc
ランタイムクラスでkataを指定するだけでok

Kata Containersをworkerに導入

```
mizuame@k8s-gcp-test:~$ kata-runtime kata-check
No newer release available
ERRO[0000] CPU property not found                        arch=amd64 description="Virtualization support" name=vmx pid=92
12 source=runtime type=flag
ERRO[0000] Module is not loaded and it can not be inserted. Please consider running with sudo or as root  arch=amd64 mod
ule=kvm_intel name=kata-runtime pid=9212 source=runtime
ERRO[0000] kernel property kvm_intel not found           arch=amd64 description="Intel KVM" name=kvm_intel pid=9212 sour
ce=runtime type=module
ERRO[0000] Module is not loaded and it can not be inserted. Please consider running with sudo or as root  arch=amd64 mod
ule=kvm name=kata-runtime pid=9212 source=runtime
ERRO[0000] kernel property kvm not found                 arch=amd64 description="Kernel-based Virtual Machine" name=kvm
pid=9212 source=runtime type=module
ERRO[0000] ERROR: System is not capable of running Kata Containers  arch=amd64 name=kata-runtime pid=9212 source=runtime
ERROR: System is not capable of running Kata Containers
```

QEMUを立てている

→CPU が 仮想化をサポートしている必要がある
→kvm周りに対応している必要がある

Kata Containersをworkerに導入

いわゆる、
Nested Virtualization
VM in VM

Kata Containersをworkerに導入

- awsではベアメタルインスタンスでしかNestedVirtualization には対応してない
- 月1000ドル~
  ↑無理
- GCPのn1/n2-standardは月60ドル~でNestedVirtualizationに対応

Kata Containersをworkerに導入

1日前 14 views

GCP上でKata Containersを動かすための手順書

```
sudo modprobe vhost
sudo modprobe vhost_net
sudo modprobe vhost_vsock
```

Kata Containersのインストール

```
bash -c "$(curl -fsSL https://raw.githubusercontent.com/kata-containers/kata-containers/ma
```

インストール先の確認と移動

```
# ホームディレクトリにインストールされた場合は移動
if [ -d ~/opt/kata ]; then
  sudo mv ~/opt/kata /opt/
fi

echo 'export PATH=$PATH:/opt/kata/bin' | sudo tee /etc/profile.d/kata.sh
source /etc/profile.d/kata.sh
```

動作確認

```
kata-runtime --version
sudo kata-runtime kata-check
```

golang runtimeへのシンボリックリンク作成

```
sudo ln -sf /opt/kata/bin/containerd-shim-kata-v2 /usr/local/bin/containerd-shim-kata-v2

ls -l /usr/local/bin/containerd-shim-kata-v2
```

Kata Containers設定

導入に当たってかなり罠が多い

runtime-rsではなくgolang runtimeを使用する
- `/opt/kata/runtime-rs/bin/` ではなく `/opt/kata/bin/` を使用

enable_annotationsに use_vsock を追加
- デフォルトでは有効になっていない

ネストされた仮想化の確認
- `/dev/kvm` の存在確認

https://md.mizuame.app/s/aLluWmRdH

SandboxChanged問題

TL;DR

* k8s/k3sで起動したコンテナが 1~2分ほどで勝手に再起動 (死んで立ち直す)をする
* その際```SandboxChanged: Pod sandbox changed, it will be killed and re-created```と出る

結論
systemdとcontainer runtimeのcgroup管理の競合が発生しているため、ポッドが突然死する

ホスト側でcgroup v1を指定する必要がある。
```
GRUB_CMDLINE_LINUX="systemd.unified_cgroup_hierarchy=0"

sudo update-grub
sudo reboot
```

SandboxChanged問題

```
NAME                                READY   STATUS    RESTARTS   AGE
kata-test-nginx                     1/1     Running   0          5m7s
kata-test-stress                    1/1     Running   2          5m7s    # ← 2回再起動
ue5-gameserver-966c8c687-7m85p      1/1     Running   5          13m     # ← 5回再起動
ue5-gameserver-966c8c687-twhxr      1/1     Running   3          13m     # ← 3回再起動
```

軽量な方は落ちていなかったが、高負荷のものは落ちた。
UnrealEngineのdedicatedサーバーを最初動かしていたのだが、UEの問題なのかk3s側の問題なのか切り分けがしたかった。結果は k3s側の問題と分かった。

SandboxChanged問題

ログ

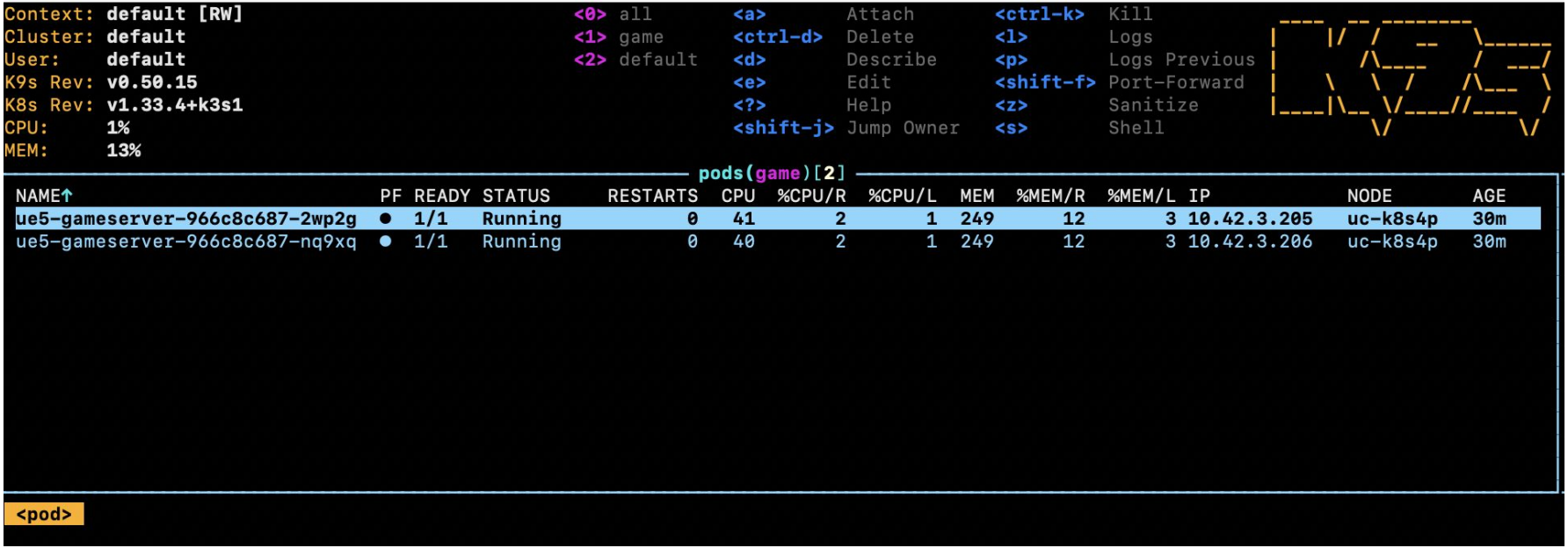
```
$ kubectl describe pod ue5-gameserver-966c8c687-qz4sn -n game | grep -A 10 "Events:"
Events:
  Type    Reason          Age                From               Message
  ----    ------          ----               ----               -------
  Normal  Scheduled       2m3s               default-scheduler  Successfully assigned
game/ue5-gameserver-966c8c687-qz4sn to uc-k8s4p
  Normal  Killing         49s                kubelet            Stopping container gameserver
  Normal  SandboxChanged  48s                kubelet            Pod sandbox changed, it will be killed
and re-created.
  Normal  Pulled          47s (x2 over 2m1s)  kubelet           Container image "ue-server-env:latest"
already present on machine
  Normal  Created         47s (x2 over 2m1s)  kubelet           Created container: gameserver
  Normal  Started         46s (x2 over 2m1s)  kubelet           Started container gameserver
```

SandboxChanged問題

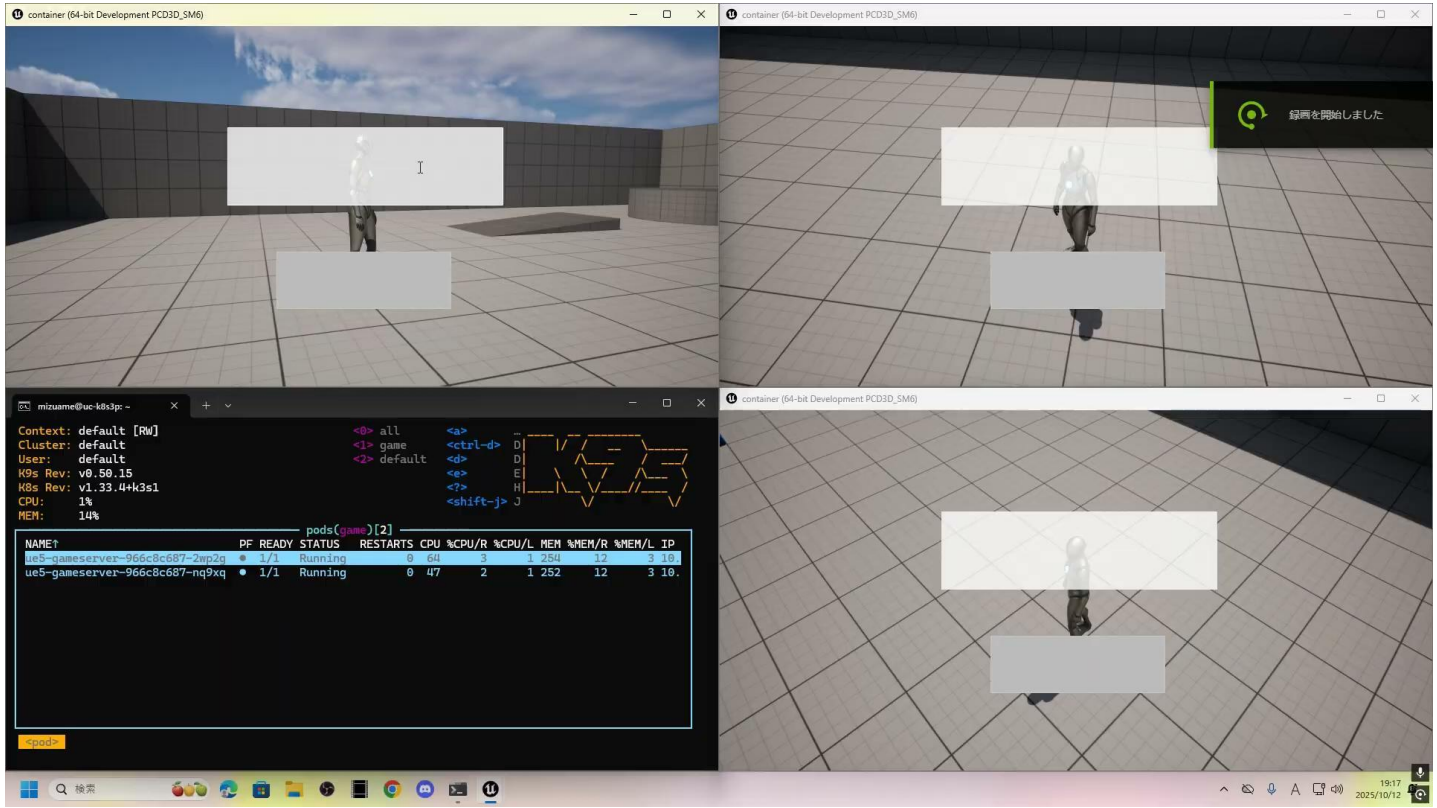
Kata ログ確認

```
$ sudo journalctl -u k3s-agent --since "10 minutes ago" | grep "9f04f8b941e445c5" | tail -50
Oct 12 04:19:50 uc-k8s4p kata[556809]: time="2025-10-12T04:19:50.989892702Z" level=warning
msg="Could not add /dev/mshv to the devices cgroup" name=containerd-shim-v2 pid=556809
sandbox=9f04f8b941e445c5...
Oct 12 04:21:06 uc-k8s4p k3s[556210]: I1012 04:21:06.672345  556210
pod_container_deletor.go:80] "Container not found in pod's containers"
containerID="9f04f8b941e445c5..."
```

SandboxChanged問題



k3sでとりあえずうごいた！



k8sに移行

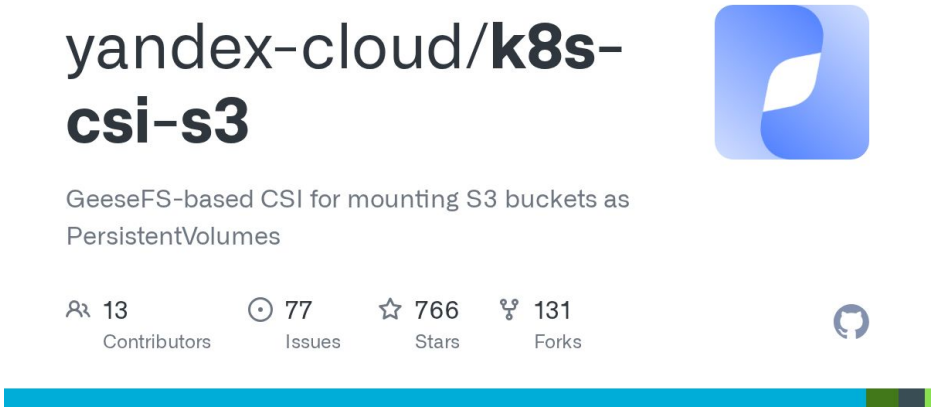
nfs > ! nfs-test.yaml

```yaml
apiVersion: v1
kind: Namespace
metadata:
  name: game
---
apiVersion: v1
kind: PersistentVolume
metadata:
  name: ue5-game-files-nfs
spec:
  capacity:
    storage: 100Gi
  accessModes:
    - ReadOnlyMany
  persistentVolumeReclaimPolicy: Retain
  storageClassName: ""
  nfs:
    server: 10.240.0.40
    path: /opt/ue5-games
    readOnly: true
---
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: ue5-game-files-pvc
  namespace: game
spec:
  accessModes:
    - ReadOnlyMany
  storageClassName: ""
  resources:
    requests:
      storage: 100Gi
```

NFSとPVどうにかしよう

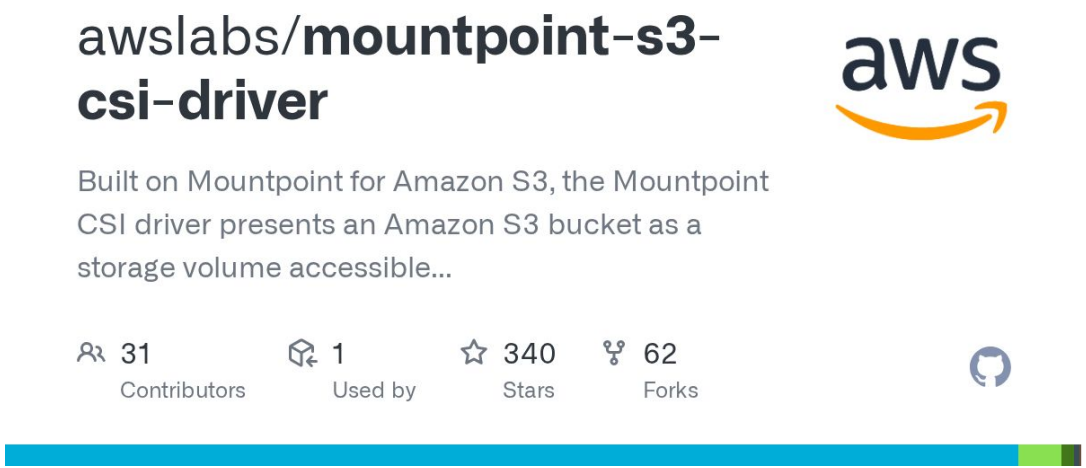
k3sの段階ではworkerノードにNFSを立てて、これをPVにしていた
流石に真面目にk8sやる状態でこのままなのは宜しくない
→オブジェクトストレージに移す

Container Storage Interface をやる



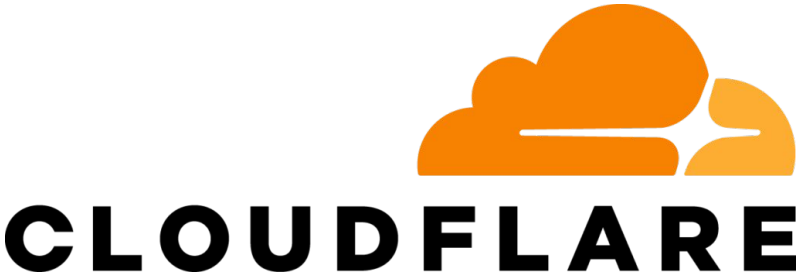
k8s-csi-s3という、
S3(オブジェクトストレージ)を
ディスクストレージのようにマウントしてくれるやつ

Container Storage Interface をやる



闇の技術っぽいが、
大本のAWSも同じようなものを公式提供しているので、割と最近はスタンダードぽい

Container Storage Interface をやる



k8s-csi-s3ではs3互換のオブジェクトストレージもサポートされている。
その為、エグレス料金0円で提供しているcloudflare R2オブジェクトストレージをPVに

Container Storage Interface をやる

pv > ! csi-s3-storageclass.yaml

```yaml
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: csi-s3
provisioner: ru.yandex.s3.csi
parameters:
  mounter: geesefs
  options: "--memory-limit 1000 --dir-mode 0777 --file-mode 0666"
  bucket: k8s
  csi.storage.k8s.io/provisioner-secret-name: csi-s3-secret
  csi.storage.k8s.io/provisioner-secret-namespace: kube-system
  csi.storage.k8s.io/controller-publish-secret-name: csi-s3-secret
  csi.storage.k8s.io/controller-publish-secret-namespace: kube-system
  csi.storage.k8s.io/node-stage-secret-name: csi-s3-secret
  csi.storage.k8s.io/node-stage-secret-namespace: kube-system
  csi.storage.k8s.io/node-publish-secret-name: csi-s3-secret
  csi.storage.k8s.io/node-publish-secret-namespace: kube-system
reclaimPolicy: Retain
volumeBindingMode: Immediate
```

```yaml
apiVersion: v1
kind: PersistentVolume
metadata:
  name: ue5-game-files-s3-pv
spec:
  storageClassName: csi-s3
  capacity:
    storage: 10Gi
  accessModes:
    - ReadOnlyMany
  claimRef:
    namespace: game
    name: ue5-game-files-s3-pvc
  csi:
    driver: ru.yandex.s3.csi
    controllerPublishSecretRef:
      name: csi-s3-secret
      namespace: kube-system
    nodePublishSecretRef:
      name: csi-s3-secret
      namespace: kube-system
    nodeStageSecretRef:
      name: csi-s3-secret
      namespace: kube-system
    volumeAttributes:
      capacity: 10Gi
      mounter: geesefs
      options: --memory-limit 1000 --dir-mode 0777 --file-mode 0777
    volumeHandle: k8s
```

k8sクラスタ構築

オンプレミス
├─ uc-k8s1p: (control-plane) i3-7100 4thread 4GB 128GB
├─ uc-k8s2p: (control-plane) i3-7100 4thread 4GB 128GB
├─ uc-k8s3p: (control-plane) i3-7100 4thread 4GB 128GB
└─ uc-k8s4p: (worker) xeon-2699v4 44thread 32GB 512GB
(game-runner)

GCP (asia-northeast1-b):
└─ k8s-gcp-test: (worker) 2thread 8GB 50GB
(game-runner)

Cloudflare R2(PV)

k8sクラスタ構築

BGP: Disabled (Calico in VXLAN mode)
→GCPとオンプレはL2がちがうので、無理ぽそう

Encapsulation: VXLAN
Network Policy: calico-apiserver のみ

Service Network:
  CIDR: 10.96.0.0/12
  DNS: 10.96.0.10

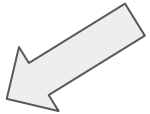
ロードバランサー(MetalLB L2モード→BGP無理なので)

```
IPAddressPool:
  name: onprem-pool
  addresses: 163.220.236.56-163.220.236.56
  autoAssign: true

L2Advertisement:
  name: onprem-l2
  ipAddressPools: [onprem-pool]
  nodeSelectors:
    - matchLabels:
        hardware: game-runner
```

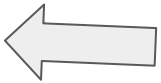
→マスター1~3にはgame-runnerを振っていないので、uc-k8s4pが単一障害点
(加えてそもそもL2ロードバランスだとECMPできていない)

ロードバランサー(MetalLB L2モード→BGP無理なので)

本当はここはk8s4pが死んでも1~3が代替できるが、
今はgame-runner指定しているので単一障害点

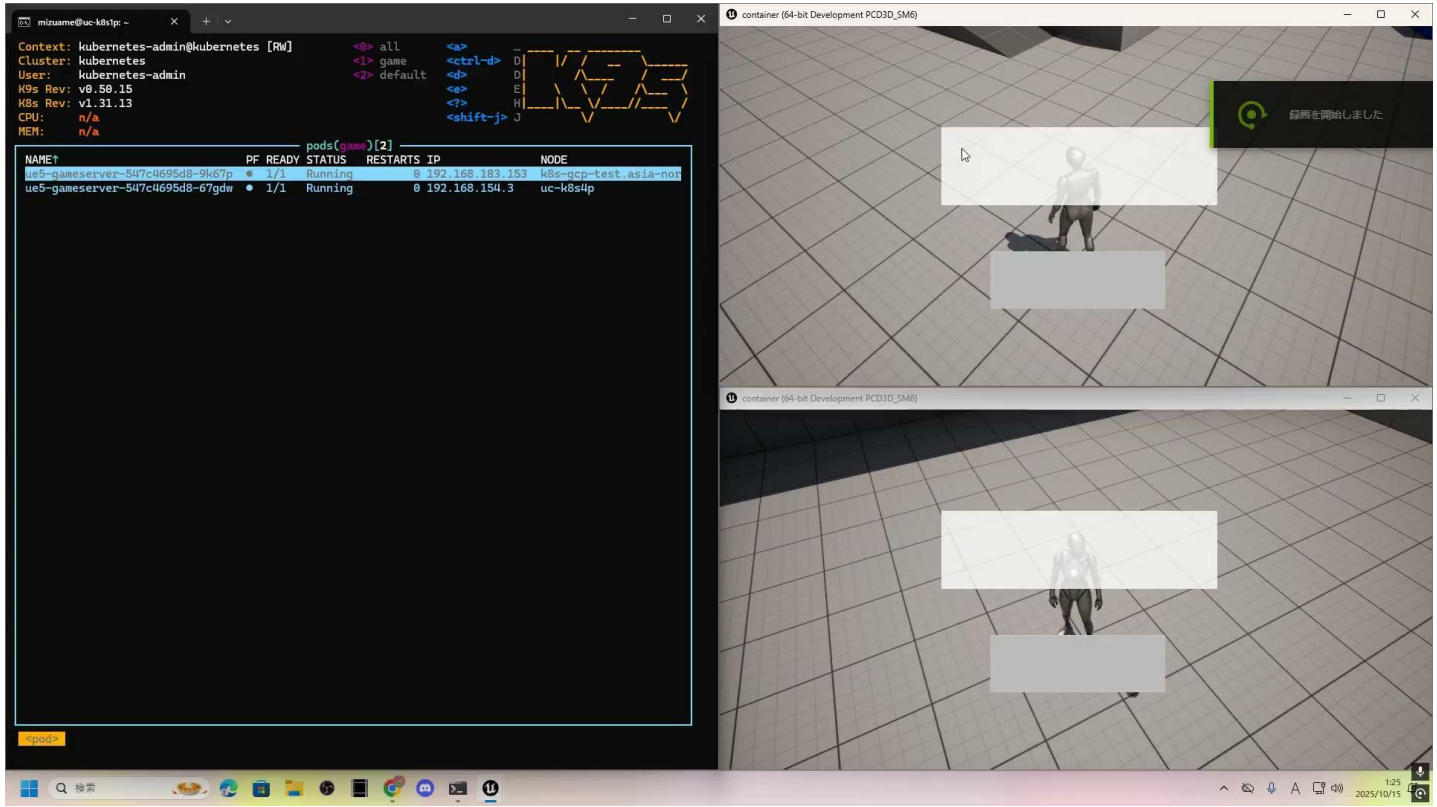


[Client]
↓
163.220.236.56:7777 (MetalLB VIP)
↓ ARP応答
[uc-k8s4p node] ← MetalLBがここまで運ぶ（負荷分散なし）
↓
iptables DNAT（kube-proxy）
├─ → 192.168.154.2:7777 (Pod 1) (uc-k8s4p)
└─ → 192.168.154.63:7777 (Pod 2) (GCP)

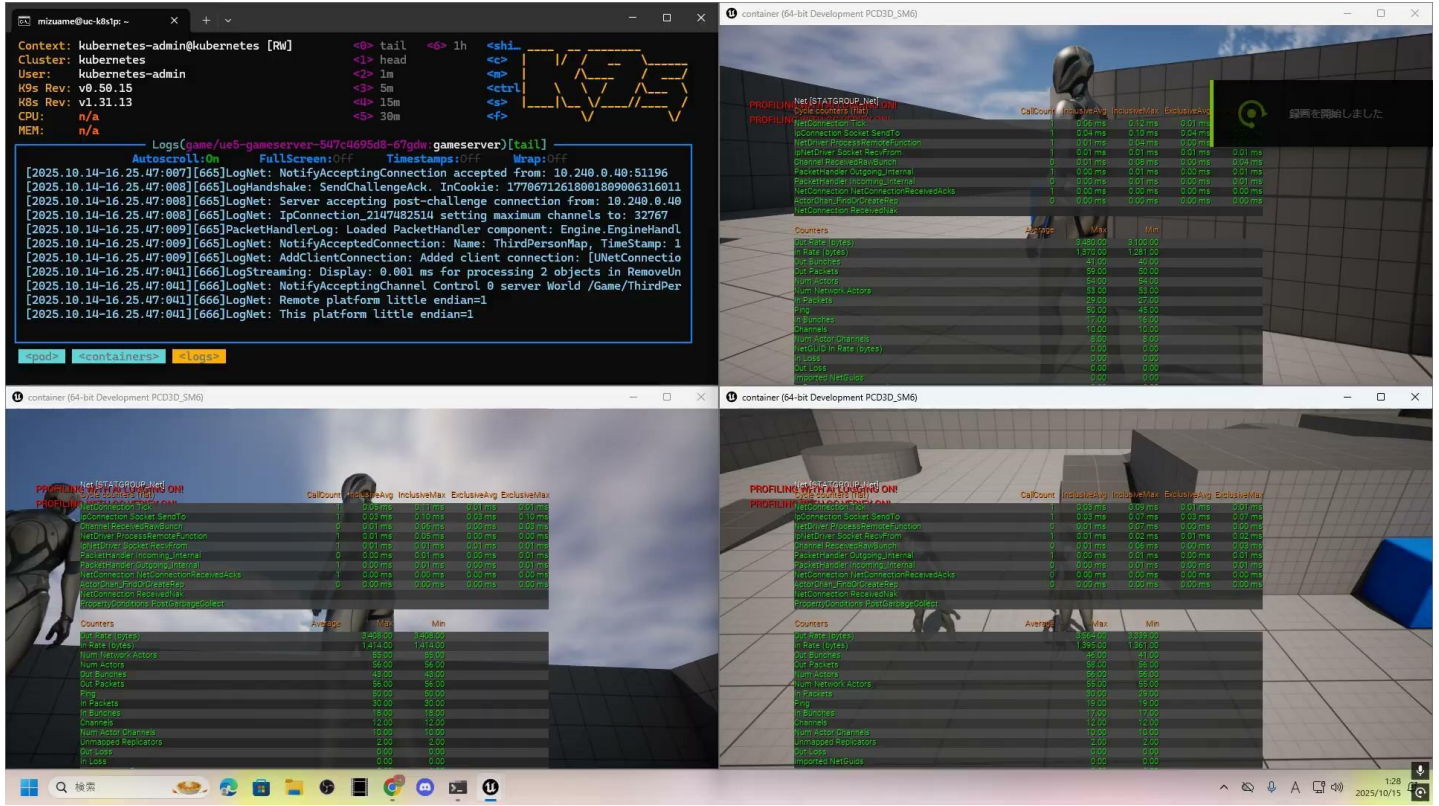


kube-proxy下の
podレベルでは分散

動画


mizuame@uc-k8s1p: ~
Context: kubernetes-admin@kubernetes [RW]
Cluster: kubernetes
User: kubernetes-admin
K9s Rev: v0.50.15
K8s Rev: v1.31.13
CPU: n/a
MEM: n/a
<0> all
<1> game
<2> default
<a>
<ctrl-d>
<d>
<e>
<?>
<shift-j>
pods(game)[2]
NAME↑ PF READY STATUS RESTARTS IP NODE
ue5-gameserver-547c4695d8-9k67p 1/1 Running 0 192.168.183.153 k8s-gcp-test.asia-nor
ue5-gameserver-547c4695d8-67gdw 1/1 Running 0 192.168.154.3 uc-k8s4p
<pod>
container (64-bit Development PCD3D_SM6)
録画を開始しました
container (64-bit Development PCD3D_SM6)

動画


mizuame@uc-k8s1p: ~
Context: kubernetes-admin@kubernetes [RW]
Cluster: kubernetes
User: kubernetes-admin
K9s Rev: v0.50.15
K8s Rev: v1.31.13
CPU: n/a
MEM: n/a
Logs(game/ue5-gameserver-547c4695d8-67gdw:gameserver)[tail]
[2025.10.14-16.25.47:007][665]LogNet: NotifyAcceptingConnection accepted from: 10.240.0.40:51196
[2025.10.14-16.25.47:008][665]LogHandshake: SendChallengeAck. InCookie: 17706712618001809006316011
[2025.10.14-16.25.47:008][665]LogNet: Server accepting post-challenge connection from: 10.240.0.40
[2025.10.14-16.25.47:008][665]LogNet: IpConnection_2147482514 setting maximum channels to: 32767
[2025.10.14-16.25.47:009][665]PacketHandlerLog: Loaded PacketHandler component: Engine.EngineHandl
[2025.10.14-16.25.47:009][665]LogNet: NotifyAcceptedConnection: Name: ThirdPersonMap, TimeStamp: 1
[2025.10.14-16.25.47:009][665]LogNet: AddClientConnection: Added client connection: [UNetConnectio
[2025.10.14-16.25.47:041][666]LogStreaming: Display: 0.001 ms for processing 2 objects in RemoveUn
[2025.10.14-16.25.47:041][666]LogNet: NotifyAcceptingChannel Control 0 server World /Game/ThirdPer
[2025.10.14-16.25.47:041][666]LogNet: Remote platform little endian=1
[2025.10.14-16.25.47:041][666]LogNet: This platform little endian=1
<pod> <containers> <logs>
container (64-bit Development PCD3D_SM6)
録画を開始しました

謝辞

間瀬bb
https://x.com/bb_mase

Ultra-Coins
https://ultra.coins.tsukuba.ac.jp/

登 大遊(https://x.com/dnobori)さんをはじめとする、グローバルIPの貸し出しや空間設備提供をして下さっているIPAの方々など