

# OpenShift Virtualizationへの RDMディスクの移行

レッドハット株式会社

2024/8/19

SA 田中司恩

shtanaka@redhat.com

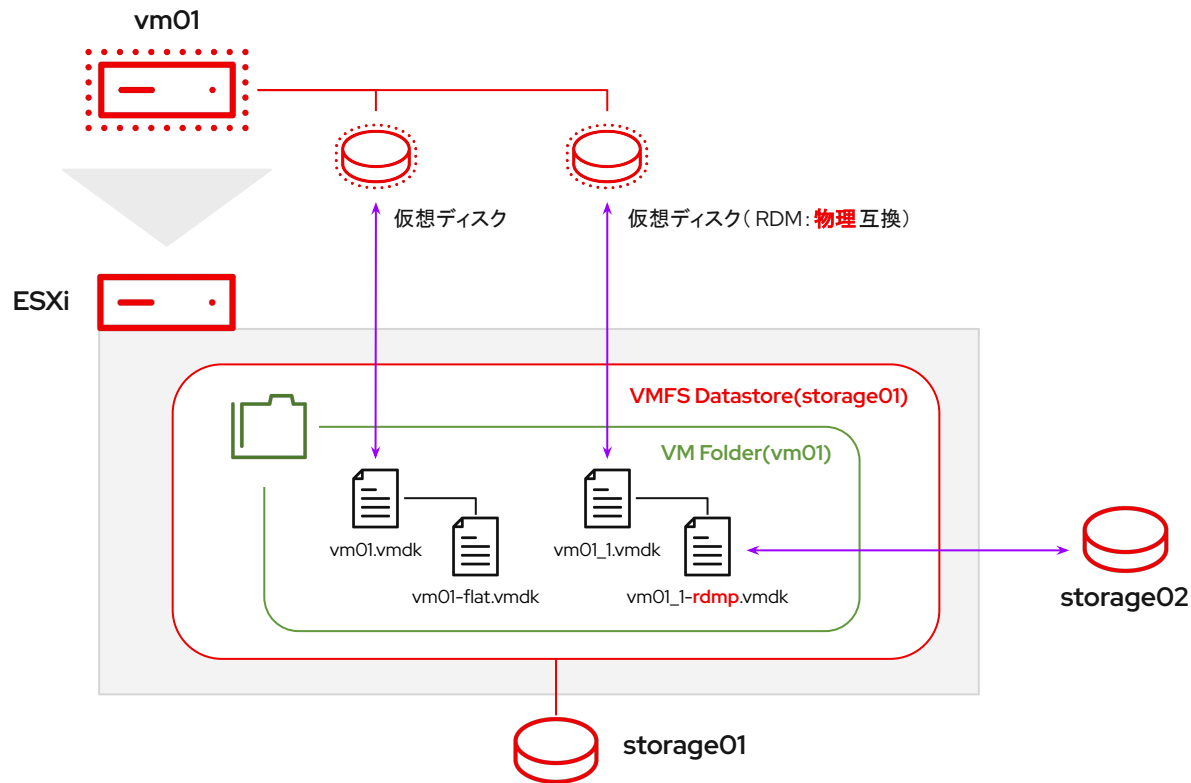
## はじめに

- ▶ 本資料に記載された内容は、公開情報および実際の挙動に基づいて作成されていますが、VMware (Broadcom) や Red Hat の公式情報ではありません。そのため、あくまで参考情報としてご利用ください。
- ▶ 本資料に基づくシステム設定や構成変更を行う際には、必ずご自身で十分な検証を行ってください。特に、本番環境におけるデータ移行作業を行う際は、事前にデータのバックアップを取得し、リスクを最小限に抑えるための対策を講じたうえで実施してください。
- ▶ 万が一、本資料の内容に従って実施した操作により生じた不具合や損害について、当方は一切の責任を負いかねますので、あらかじめご了承ください。

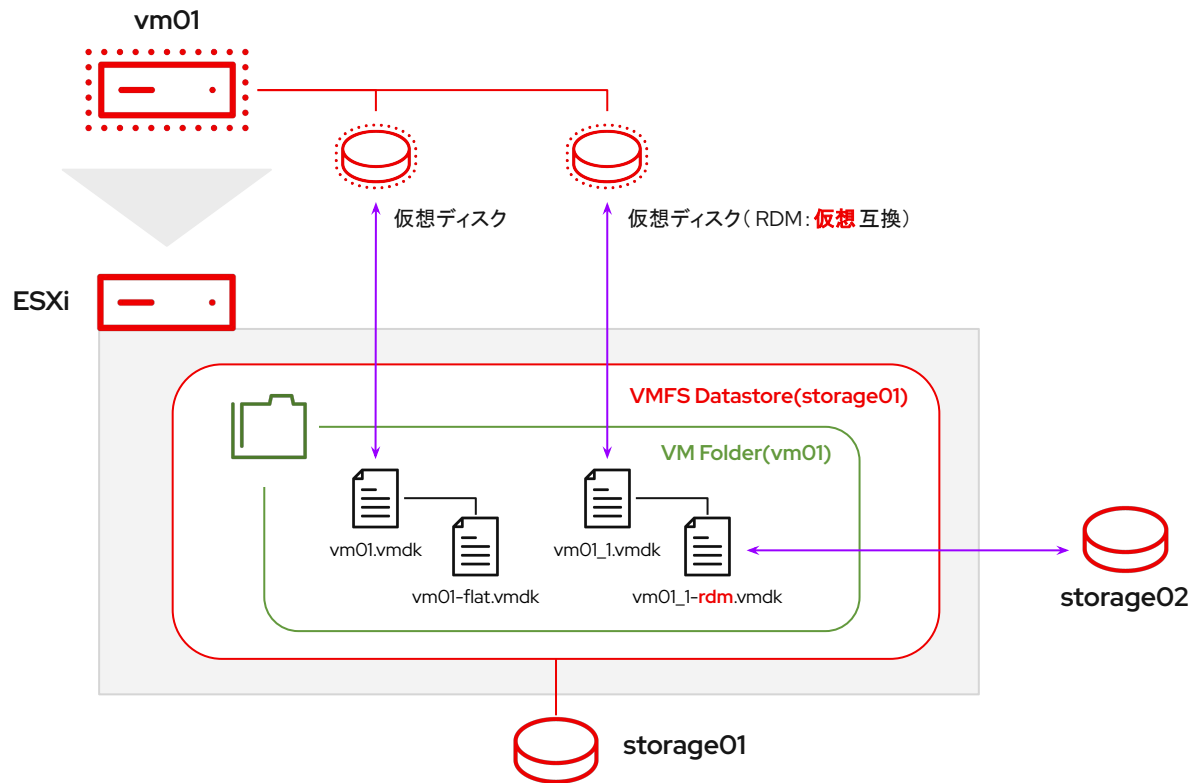
---

# RDMディスクの構成

## 物理互換モードのRDMディスクの構成



## 仮想互換モードのRDMディスクの構成



---

# RDMディスクの追加 (iSCSIデバイス)



# 仮想マシンの詳細設定

## 物理互換モードの RDM ディスク

Edit Settings | win2022-vmware

Virtual Hardware | VM Options | Advanced Parameters

ADD NEW DEVICE

> CPU	2	GB
> Memory	4	GB
> Hard disk 1	100	GB
> Hard disk 2	100	GB

VM storage policy: Datastore Default

Sharing: No sharing

Physical LUN: vml.020000000060014051662c5b5b5ec4f64a4e568a1d6469736b3120

Compatibility Mode: **Physical**

Disk File: [NVMe] win2022-vmware/win2022-vmware\_1.vmdk

Virtual Device Node: SCSI controller 0 | SCSI(0:2) Hard disk 2

> SCSI controller 0: VMware Paravirtual

> Network adapter 1: VM Network | Connected

> CD/DVD drive 1: Datastore ISO File | Connected

> USB xHCI controller: USB 3.2

> Video card: Auto-detect settings

CANCEL OK



# 物理互換モードのRDMディスクの詳細

## VMFSデータストア上のVMDKファイルの状態

### RDMディスク追加後のVMDKファイルの状態

```
[root@esxi:~] ls -ld /vmfs/volumes/NVMe/win2022-vmware/win2022-vmware_1*  
-rw----- 1 root    root      107374182400 Aug  9 09:56 /vmfs/volumes/NVMe/win2022-vmware/win2022-vmware_1-rdmp.vmdk  
-rw----- 1 root    root          484 Aug  9 09:58 /vmfs/volumes/NVMe/win2022-vmware/win2022-vmware_1.vmdk  
[root@esxi:~] vmkfstools -q /vmfs/volumes/NVMe/win2022-vmware/win2022-vmware_1.vmdk  
Disk /vmfs/volumes/NVMe/win2022-vmware/win2022-vmware_1.vmdk is a Passthrough Raw Device Mapping  
Maps to: vm1.020000000060014051662c5b5b5ec4f64a4e568a1d6469736b3120
```

# 物理互換モードのRDMディスクの詳細

## ESXi上のiSCSIデバイスの状態

### ESXi上のiSCSIデバイスの状態

```
[root@esxi:~] esxcli storage core device list -d naa.60014051662c5b5b5ec4f64a4e568a1d
naa.60014051662c5b5b5ec4f64a4e568a1d
  Display Name: LIO-ORG iSCSI Disk (naa.60014051662c5b5b5ec4f64a4e568a1d)
  Has Settable Display Name: true
  Size: 102400
  Device Type: Direct-Access
  Multipath Plugin: NMP
  Devfs Path: /vmfs/devices/disks/naa.60014051662c5b5b5ec4f64a4e568a1d
  Vendor: LIO-ORG
  Model: disk1
  Revision: 4.0
  SCSI Level: 6
  Is Pseudo: false
  Status: degraded
  Is RDM Capable: true
  Is Local: false
  Is Removable: false
  Is SSD: true
  Is VVOL PE: false
  Is Offline: false
  Is Perennially Reserved: false
  Queue Full Sample Size: 0
  Queue Full Threshold: 0
  Thin Provisioning Status: unknown
  Attached Filters:
  VAAI Status: supported
  Other UUIDs: vml.0200000000060014051662c5b5b5ec4f64a4e568a1d6469736b3120
  Is Shared Clusterwide: true
  Is SAS: false
  Is USB: false
  Is Boot Device: false
  Device Max Queue Depth: 128
```

```
Device Max Queue Depth: 128
No of outstanding I/Os with competing worlds: 32
Drive Type: unknown
RAID Level: unknown
Number of Physical Drives: unknown
Protection Enabled: false
PI Activated: false
PI Type: 0
PI Protection Mask: NO PROTECTION
Supported Guard Types: NO GUARD SUPPORT
DIX Enabled: false
DIX Guard Type: NO GUARD SUPPORT
Emulated DIX/DIF Enabled: false
```

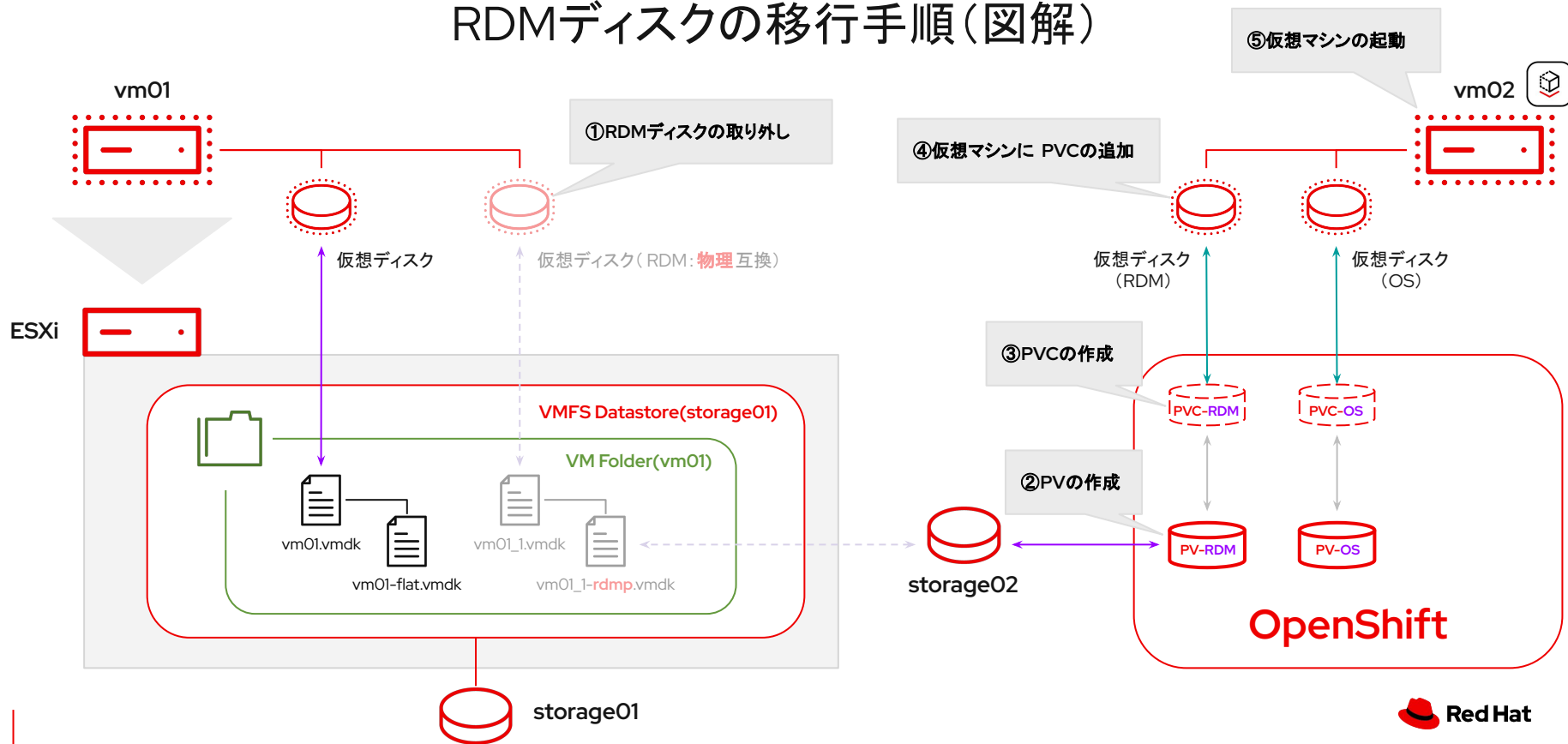
---

# RDMディスクの移行 (iSCSIデバイス)

## RDMディスクの移行手順

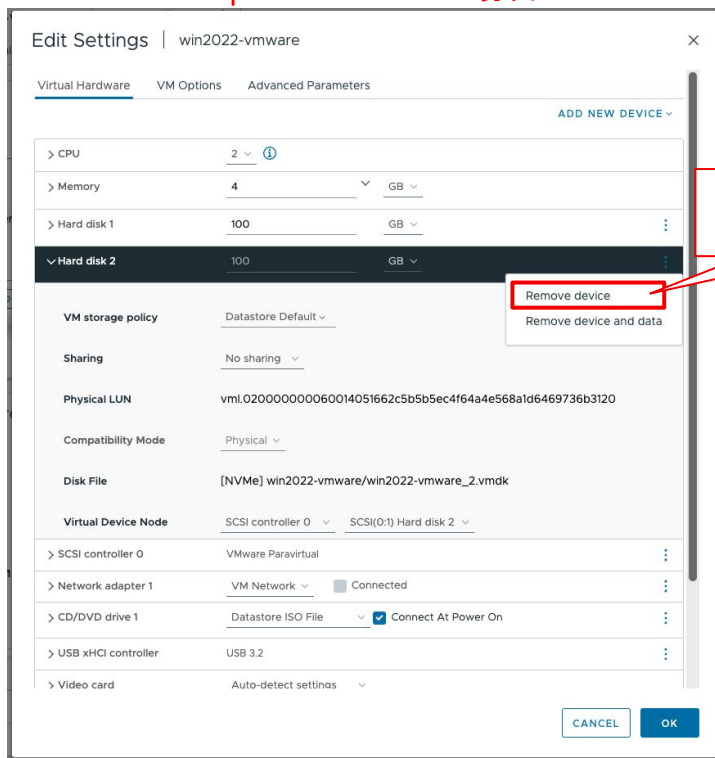
1. vSphere仮想マシンからRDMディスクを取り外す
2. OpenShift上にiSCSIディスクをPVとして手動で作成する
3. 手動作成したPVを使用するPVCを手動で作成する
4. OCP-Vの仮想マシンにPVCを追加する
5. 仮想マシンを起動する

# RDMディスクの移行手順(図解)



# 1. 仮想マシンからRDMディスクを取り外す

vSphere Clientの場合



## (参考)1. 仮想マシンからRDMディスクを取り外す

ESXi Host Clientの場合

Edit settings - win2022-vmware (ESXi 8.0 U2 virtual machine)

Virtual Hardware VM Options

Add hard disk Add network adapter Add other device

> CPU	2		
> Memory	4	GB	
> Hard disk 1	100	GB	
> Hard disk 2	100	GB	<input type="checkbox"/> Delete files from datastore
Device path vml.020000000060014051662c5b5b5ec4f64a4e568a1d6469736b3120			
Disk File [NVMe] win2022-vmware/win2022-vmware_1.vmdk			
Controller location		SCSI controller 0	SCSI (0:2)
Disk mode		Independent - persistent	
Disk compatibility		Physical	
> SCSI Controller 0	VMware Paravirtual		

CANCEL SAVE

データを保持するため  
チェックしない

## 2. PVを手動作成する

### PV (iscsi-pv.yaml)

```
apiVersion: v1
kind: PersistentVolume
metadata:
  name: iscsi-pv
spec:
  capacity:
    storage: 100Gi
  volumeMode: Block
  accessModes:
    - ReadWriteMany
  persistentVolumeReclaimPolicy: Retain
  storageClassName: iscsi-sc
  iscsi:
    targetPortal: 192.168.1.32:3260
    iqn: iqn.2024-08.lab.home:target1
    lun: 0
```

### PVの作成

```
$ oc create -f iscsi-pv.yaml
persistentvolume/iscsi-pv created
```

### (参考) iSCSI イニシエーターの確認

```
$ for node in node-01 node-02 node-03; do
  oc debug node/${node}.ocp.home.lab -- chroot /host \
    cat /etc/iscsi/initiatorname.iscsi
done
Temporary namespace openshift-debug-9ng7t is created for
debugging node...
Starting pod/node-01ocphomelab-debug-pslsd ...
To use host binaries, run `chroot /host`
InitiatorName=iqn.1994-05.com.redhat:bab67a50c1a6

Removing debug pod ...
Temporary namespace openshift-debug-9ng7t was removed.
Temporary namespace openshift-debug-9fxwr is created for
debugging node...
Starting pod/node-02ocphomelab-debug-gjjfm ...
To use host binaries, run `chroot /host`
InitiatorName=iqn.1994-05.com.redhat:edd53b17a1f

Removing debug pod ...
Temporary namespace openshift-debug-9fxwr was removed.
Temporary namespace openshift-debug-sjl4m is created for
debugging node...
Starting pod/node-03ocphomelab-debug-vx2pq ...
To use host binaries, run `chroot /host`
InitiatorName=iqn.1994-05.com.redhat:b3973d3ed24d

Removing debug pod ...
Temporary namespace openshift-debug-sjl4m was removed.
```



### 3. PVCを手動作成する

#### PVC (iscsi-pvc.yaml)

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: iscsi-pvc
spec:
  accessModes:
    - ReadWriteMany
  resources:
    requests:
      storage: 100Gi
  volumeMode: Block
  storageClassName: iscsi-sc
```

#### PVCの作成

```
$ oc create -f iscsi-pvc.yaml
persistentvolumeclaim/iscsi-pvc created
```

## 4. OCP-Vの仮想マシンにPVCを追加する

仮想マシン > Configuration > Storage > Disks > Add disk

The screenshot shows the OpenShift console interface. In the background, the 'Configuration' page for a virtual machine named 'win2k22-moccasin-elk-82' is visible, with the 'Disks' tab selected. A modal dialog titled 'Add disk' is open in the foreground. The dialog contains the following fields:

- Name**: rdm-iscsi
- Source**: PVC (highlighted with a red box)
- PersistentVolumeClaim project**: ocp-virt
- PersistentVolumeClaim name**: PVC iscsi-pvc (highlighted with a red box)
- Type**: Disk
- Interface**: VirtIO

Below the 'Interface' dropdown, a note states: 'Hot plug is enabled only for "SCSI" interface'. At the bottom of the dialog are 'Save' and 'Cancel' buttons. The background interface also shows a list of disks: 'rootdisk' (bootable) and 'windows-drivers-disk'.

## 4. OCP-Vの仮想マシンにPVCを追加する

### PVCの追加後

プロジェクト: ocp-virt ▾

VM win2k22-moccasin-elk-82 Stopped YAML ▶ Actions ▾

[Overview](#) [Metrics](#) [YAML](#) [Configuration](#) [Events](#) [Console](#) [Snapshots](#) [Diagnostics](#)

Q

Details

Storage Add disk

Network ▼ フィルター ▾  ☐ Mount Windows drivers disk

Name ↑	Source ↓	Size ↓	Drive ↓	Interface ↓	Storage class ↓	
rdm-iscsi	<span>PVC</span> iscsi-pvc	100.00 GiB	Disk	virtio	iscsi-sc	⋮
rootdisk <span>bootable</span>	<span>PVC</span> restore-9a5cf2f0-e707-4d5a-a9ec-3586835e3c7f-rootdisk	63.49 GiB	Disk	SATA	nfs-csi	⋮
windows-drivers-disk	Container (Ephemeral)	Dynamic	CD-ROM	SATA	-	⋮

元のディスクの内容のままマウントされている

元のディスクの内容のままマウントされている



## RDMディスクの移行に関するメモ

- ▶ 物理互換モードと仮想互換モードの違いは作成されるプロキシファイルの違い。物理互換モード→xxx-**rdmp**.vmdk、仮想互換モード→xxx-**rdm**.vmdk
- ▶ 同一のRawデバイスに対して、物理互換モードと仮想互換モードのプロキシファイルをそれぞれ作成することは可能
- ▶ ただし、同一のRawデバイスを物理互換モードと仮想互換モードの同時利用はできない(先にアタッチした方がロックする)
- ▶ vSphere上のプロキシファイルから切り離してしまえば純粋な Rawデバイスとしての取り扱いができる
  - OpenShiftのノードに Rawデバイスをアタッチしてデータ移行が可能 (物理互換モード、仮想互換モード問わず)
  - PV/PVCをRWXで作成すれば、そのままアタッチした Rawデバイスを使いつつ、ライブマイグレーションも使用可能

# Thank you



[linkedin.com/company/red-hat](https://linkedin.com/company/red-hat)



[youtube.com/user/RedHatVideos](https://youtube.com/user/RedHatVideos)



[facebook.com/redhatinc](https://facebook.com/redhatinc)



[twitter.com/RedHat](https://twitter.com/RedHat)

# Appendix

---

# RDMディスクの作成と追加 (ESXiのローカルディスク)



# RDMディスクの作成

## ESXiのローカルディスクをRDMディスクとして使用

### ESXiのローカルディスクの確認

```
[root@esxi:~] ls -l /vmfs/devices/disks/t10.ATA_____SAMSUNG_MMD0E28G5MPP2D0VA_____SY850A0123_____
-rw----- 1 root root 128035676160 Aug 8 17:32 /vmfs/devices/disks/t10.ATA_____SAMSUNG_MMD0E28G5MPP2D0VA_____SY850A0123_____
[root@esxi:~] esxcfg-mpath -L
vmhba32:C0:T0:L0 state:active mpx.vmhba32:C0:T0:L0 vmhba32 0 0 0 NMP active local usb.vmhba32 usb.0:0
vmhba0:C0:T0:L0 state:active t10.ATA_____SAMSUNG_MMD0E28G5MPP2D0VA_____SY850A0123_____ vmhba0 0 0 0 HPP active local sata.vmhba0 sata.0:0
vmhba1:C0:T0:L0 state:active t10.NVMe_____Samsung_SSD_970_EVO_Plus_1TB_____CEF2B49154382500 vmhba1 0 0 0 HPP active local pcie.200 pcie.0:0
```

### 物理互換モードのRDMディスクの作成

```
[root@esxi:~] vmkfstools -z /vmfs/devices/disks/t10.ATA_____SAMSUNG_MMD0E28G5MPP2D0VA_____SY850A0123_____ /vmfs/volumes/NVMe/rdm_p.vmdk
```

### 仮想互換モードのRDMディスクの作成

```
[root@esxi:~] vmkfstools -r /vmfs/devices/disks/t10.ATA_____SAMSUNG_MMD0E28G5MPP2D0VA_____SY850A0123_____ /vmfs/volumes/NVMe/rdm_v.vmdk
```

### RDM作成後のVMDKファイルの状態

```
[root@esxi:~] ls -ld /vmfs/volumes/NVMe/rdm_*
-rw----- 1 root root 128035676160 Aug 8 17:28 /vmfs/volumes/NVMe/rdm_p--rdmp.vmdk
-rw----- 1 root root 473 Aug 8 17:28 /vmfs/volumes/NVMe/rdm_p.vmdk
-rw----- 1 root root 128035676160 Aug 8 17:27 /vmfs/volumes/NVMe/rdm_v--rdm.vmdk
-rw----- 1 root root 461 Aug 8 17:27 /vmfs/volumes/NVMe/rdm_v.vmdk
```

※ RDMディスク用VMDKファイルはVMFSデータストア上にあればよく、仮想マシンフォルダ外に作成しても問題無い

## RDMディスクの追加（物理互換モード）

Edit settings - win2022-vmware (ESXI 8.0 U2 virtual machine)

Virtual Hardware VM Options

Add hard disk Add network adapter Add other device

> CPU 2

> Memory 4 GB

> Hard disk 1 100 GB

> Hard disk 2 119.24 GB

Device path vml.0100000005359383530413031323320202020202020202053414d53554e

Disk File [NVMe] rdm\_p.vmdk

Controller location SCSI controller 0 SCSI (0:1)

Disk mode Independent - persistent

Disk compatibility Physical

> SCSI Controller 0 VMware Paravirtual

CANCEL SAVE

## RDMディスクの追加(仮想互換モード)

Edit settings - win2022-vmware (ESXi 8.0 U2 virtual machine)

Virtual Hardware VM Options

Add hard disk Add network adapter Add other device

> CPU 2

> Memory 4 GB

> Hard disk 1 100 GB

> Hard disk 2 119.24 GB

Device path vml.01000000005359383530413031323320202020202020202053414d53554e

Disk File **[NVMe] rdm\_v.vmdk**

Controller location SCSI controller 0 SCSI (0:1)

Disk mode Dependent

Disk compatibility **Virtual**

> SCSI Controller 0 VMware Paravirtual

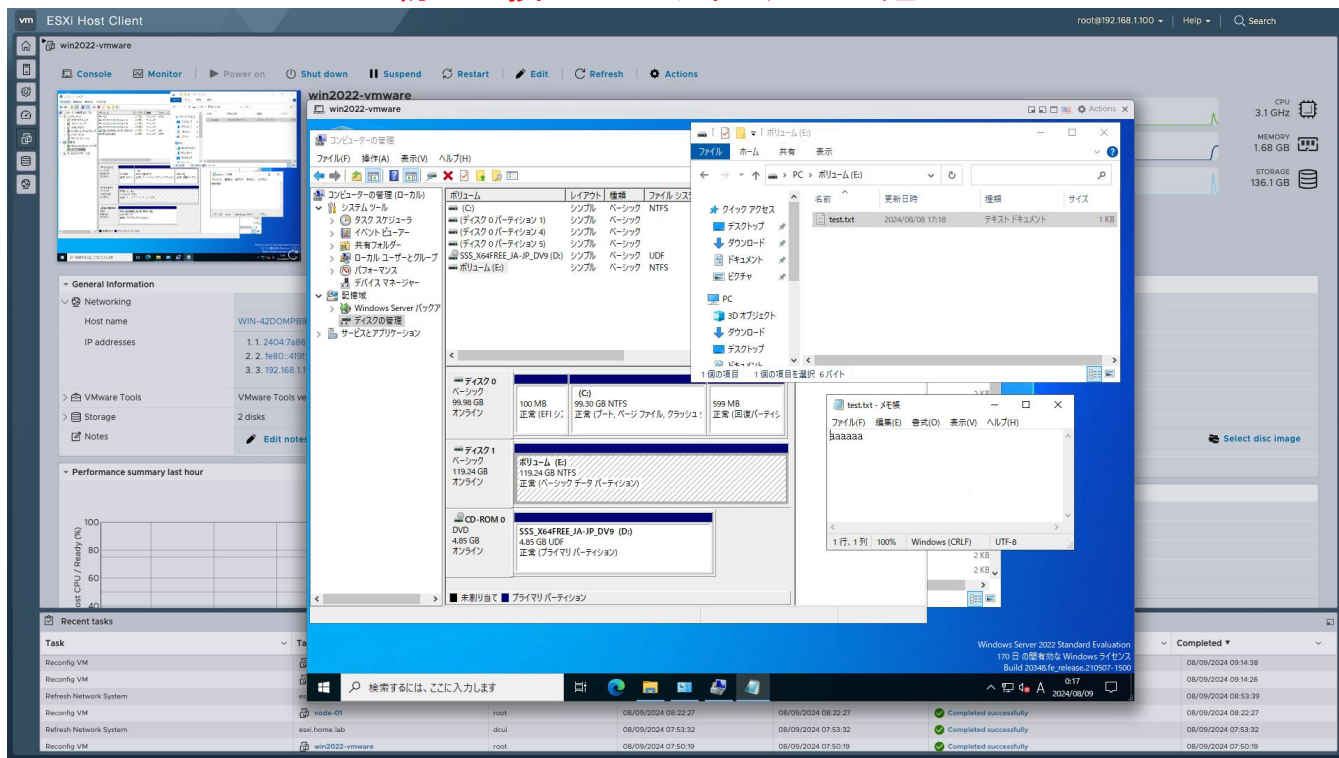
CANCEL SAVE

---

# RDMディスクの移行 (ESXiのローカルディスク)

# 移行元vSphere仮想マシン(Windows Server 2022)

物理互換のRDMディスクとして追加



# 移行先OCP-V仮想マシン(Windows Server 2022)

virtioディスクとして追加

Console

VNC console Send key Paste Disconnect

The screenshot shows a Windows Server 2022 virtual machine interface. On the left, the 'Computer Management' window is open, displaying the 'Disk Management' section. It shows two disks: Disk 0 (59.98 GB) and Disk 1 (119.24 GB). Disk 1 is currently unallocated. A table below lists the disks and their partitions:

ディスク	パーティション	容量	ファイルシステム	状態
ディスク 0	(C:)	100 MB	NTFS	正常
	(ディスク 0 パーティション 1)	59.30 GB	NTFS	正常
	(ディスク 0 パーティション 4)	599 MB	NTFS	正常
ディスク 1	virtio-win-1.9.40 (D:)	494 MB	CDFS	正常
	ボリューム (E:)	119.24 GB	NTFS	正常 (パーティションデータ パーティション)

On the right, a File Explorer window is open, showing the contents of the 'test' folder in the 'ボリューム (E:)' drive. The folder contains a single file named 'test' (Text Document, 1 KB). A small text box is overlaid on the File Explorer, displaying the text 'aaaaaa'.

Windows Server 2022 Standard Evaluation  
166 日の無償な Windows ライセンス  
Build 20348.1fe\_release.210907-1500  
019  
2024/08/09

# OCP-Vのディスク構成

事前に作成済みのPVCを指定

プロジェクト: ocp-virt ▾

VM win2k22-moccasin-elk-82 Running Not migratable YAML 🔍 🔄 ⏸ Actions ▾

[Overview](#) [Metrics](#) [YAML](#) [Configuration](#) [Events](#) [Console](#) [Snapshots](#) [Diagnostics](#)

🔍

**Disks**

Details

Storage Add disk

Network ▼ フィルター  / ☐ Mount Windows drivers disk

	Name ↑	Source ↑	Size ↑	Drive ↑	Interface ↑	Storage class ↑	
	rdm	<span>PVC</span> rdm-pvc	100.00 GiB	Disk	virtio	local-sc	⋮
SSH	rootdisk	<span>PVC</span> restore-9a5cf2f0-e707-4d5a-a9ec-3586835e3c7f-rootdisk	63.49 GiB	Disk	SATA	nfs-csi	⋮
Initial run	<span>bootable</span>						
Metadata	windows-drivers-disk	Container (Ephemeral)	Dynamic	CD-ROM	SATA	-	⋮

**Environment**

Include all values from existing config maps, secrets or service accounts (as disk) ⓘ

[+ Add Config Map, Secret or Service Account](#)

Save Reload

# PVとPVCのマニフェスト

LSOはパーティションがあるとPVが作成されないため、手動でPVとPVCを作成

## PV

```
apiVersion: v1
kind: PersistentVolume
metadata:
  name: rdm
spec:
  capacity:
    storage: 100Gi
  volumeMode: Block
  accessModes:
    - ReadWriteOnce
  persistentVolumeReclaimPolicy: Delete
  storageClassName: local-sc
  local:
    path: /dev/disk/by-id/wwn-0x50000f0056413031
  nodeAffinity:
    required:
      nodeSelectorTerms:
        - matchExpressions:
            - key: kubernetes.io/hostname
              operator: In
              values:
                - node-01.ocp.home.lab
```

## PVC

```
kind: PersistentVolumeClaim
apiVersion: v1
metadata:
  name: rdm-pvc
spec:
  accessModes:
    - ReadWriteOnce
  volumeMode: Block
  resources:
    requests:
      storage: 100Gi
  storageClassName: local-sc
```



# OpenShiftのノード上のディスク情報

## node-01.ocp.home.labにディスクをマウント

### lsblkの出力

```
sh-5.1# lsblk
NAME        MAJ:MIN RM   SIZE RO TYPE MOUNTPOINTS
sda          8:0    0 119.2G 0 disk
└─sda1       8:1    0 119.2G 0 part
sdb          8:16   0   100G 0 disk
├─sdb1       8:17   0     1M 0 part
├─sdb2       8:18   0   127M 0 part
├─sdb3       8:19   0   384M 0 part /boot
└─sdb4       8:20   0   99.5G 0 part /var

/sysroot/ostree/deploy/rhcos/var
                                /sysroot
                                /usr
                                /etc
                                /
```

### /dev/disk/by-id 配下のデバイス情報

```
sh-5.1# ls -l /dev/disk/by-id/
total 0
lrwxrwxrwx. 1 root root 10 Aug 8 15:39 ata-1ATA_SAMSUNG_MMD0E28G5MPP-0VA_SY850A0123-part1 → ../../sda1
lrwxrwxrwx. 1 root root 9 Aug 8 15:39 ata-SAMSUNG_MMD0E28G5MPP-0VA_SY850A0123 → ../../sda
lrwxrwxrwx. 1 root root 9 Aug 8 14:43 ata-VMware_Virtual_IDE_CDROM_Drive_00000000000000000001 → ../../sr0
lrwxrwxrwx. 1 root root 9 Aug 8 15:39 scsi-1ATA_SAMSUNG_MMD0E28G5MPP-0VA_SY850A0123 → ../../sda
lrwxrwxrwx. 1 root root 10 Aug 8 15:39 scsi-1ATA_SAMSUNG_MMD0E28G5MPP-0VA_SY850A0123-part1 → ../../sda1
lrwxrwxrwx. 1 root root 9 Aug 8 14:43 scsi-36000c29c4f9a862cc323a3702b7a9d3e → ../../sdb
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-36000c29c4f9a862cc323a3702b7a9d3e-part1 → ../../sdb1
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-36000c29c4f9a862cc323a3702b7a9d3e-part2 → ../../sdb2
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-36000c29c4f9a862cc323a3702b7a9d3e-part3 → ../../sdb3
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-36000c29c4f9a862cc323a3702b7a9d3e-part4 → ../../sdb4
lrwxrwxrwx. 1 root root 9 Aug 8 15:39 scsi-SATA_SAMSUNG_MMD0E28G_SY850A0123 → ../../sda
lrwxrwxrwx. 1 root root 10 Aug 8 15:39 scsi-SATA_SAMSUNG_MMD0E28G_SY850A0123-part1 → ../../sda1
lrwxrwxrwx. 1 root root 9 Aug 8 14:43 scsi-SVMware_Virtual_disk_6000c29c4f9a862cc323a3702b7a9d3e → ../../sdb
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-SVMware_Virtual_disk_6000c29c4f9a862cc323a3702b7a9d3e-part1 →
../../sdb1
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-SVMware_Virtual_disk_6000c29c4f9a862cc323a3702b7a9d3e-part2 →
../../sdb2
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-SVMware_Virtual_disk_6000c29c4f9a862cc323a3702b7a9d3e-part3 →
../../sdb3
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 scsi-SVMware_Virtual_disk_6000c29c4f9a862cc323a3702b7a9d3e-part4 →
../../sdb4
lrwxrwxrwx. 1 root root 9 Aug 8 15:39 wwn-0x50000f0056413031 → ../../sda
lrwxrwxrwx. 1 root root 10 Aug 8 15:39 wwn-0x50000f0056413031-part1 → ../../sda1
lrwxrwxrwx. 1 root root 9 Aug 8 14:43 wwn-0x6000c29c4f9a862cc323a3702b7a9d3e → ../../sdb
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 wwn-0x6000c29c4f9a862cc323a3702b7a9d3e-part1 → ../../sdb1
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 wwn-0x6000c29c4f9a862cc323a3702b7a9d3e-part2 → ../../sdb2
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 wwn-0x6000c29c4f9a862cc323a3702b7a9d3e-part3 → ../../sdb3
lrwxrwxrwx. 1 root root 10 Aug 8 14:43 wwn-0x6000c29c4f9a862cc323a3702b7a9d3e-part4 → ../../sdb4
```