

# Configuración de DCPMM en VMware ESXi para el modo AppDirect

## Contenido

[Introducción](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Antecedentes](#)

[Configurar](#)

[Configuración del perfil de servicio](#)

[Verificar ESXi](#)

[Configurar NVDIMM de máquina virtual](#)

[Configurar espacios de nombres en la máquina virtual](#)

[Troubleshoot](#)

[Información Relacionada](#)

## Introducción

Este documento describe el proceso para configurar ESXi en servidores de la serie B de Unified Computing System (UCS) utilizando la memoria persistente Intel® Optane™ (PMEM) en modo administrado por host.

## Prerequisites

### Requirements

Cisco recomienda que tenga conocimiento sobre estos temas:

- UCS serie B
- Conceptos del módulo de memoria persistente (DCPMM) del Data Center Intel® Optane™
- Administración de VMware ESXi y vCenter Server

Asegúrese de cumplir estos requisitos antes de intentar esta configuración:

- Consulte las directrices de PMEM en la [guía de especificación](#) B200/B480 M5.
- Asegúrese de que la CPU sea la segunda generación de procesadores Intel® Xeon® escalables.
- La proporción PMEM/Memoria dinámica de acceso aleatorio (DRAM) cumple los requisitos según los [KB 67645](#).
- ESXi es 6,7 U2 + Express Patch 10 (ESXi670-201906002) o posterior. Las versiones anteriores 6.7 no son compatibles.
- UCS Manager y Server se encuentran en una versión 4.0(4) o posterior. Para consultar la

última versión recomendada, visite [www.software.cisco.com/](http://www.software.cisco.com/).

## Componentes Utilizados

La información que contiene este documento se basa en las siguientes versiones de software y hardware.

- UCS B480 M5
- UCS Manager 4.1(2b)

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

## Antecedentes

En los servidores UCS configurados para el modo App Direct, las máquinas virtuales VMware ESXi acceden a los módulos de memoria en línea duales (NVDIMM) Optane DCPMM Persistentes y no volátiles.

Intel Optane DCPMM se puede configurar a través de la utilidad de administración de IPMCTL a través del shell Unified Extensible Firmware Interface (UEFI) o a través de las utilidades del sistema operativo. Esta herramienta está diseñada para realizar algunas de las siguientes acciones:

- Detectar y administrar módulos
- Actualización y configuración del firmware del módulo
- Supervisión del estado
- Aprovisionar y configurar los espacios de nombres, región y objetivo
- Depurar y resolver problemas de PMEM

UCS se puede configurar mediante una política de memoria persistente asociada al perfil de servicio para facilitar su uso.

La utilidad de código abierto Non-Volatile Device Control (NDCTL) se utiliza para administrar el subsistema Linux Kernel de LIBNVDIMM. La utilidad NDCTL permite a un sistema aprovisionar y realizar configuraciones como regiones y espacios de nombres para uso del sistema operativo.

La memoria persistente agregada a un host ESXi es detectada por el host, formateada y montada como un almacén de datos local de PMem. Para utilizar el PMEM, ESXi utiliza el formato del sistema de archivos del sistema de volado de máquina virtual (VMFS)-L y sólo se admite un almacén de datos PMEM local por host.

A diferencia de otros almacenes de datos, el almacén de datos PMEM no admite tareas como almacenes de datos tradicionales. El directorio de inicio de VM con los archivos vmx y vmware.log no se puede colocar en el almacén de datos de PMEM.

PMEM se puede presentar a una VM en dos modos diferentes: Modo de acceso directo y modo de disco virtual.

- Modo de acceso directo

Las VM se pueden configurar para este modo presentando la región PMEM en forma de NVDIMM. El sistema operativo VM debe tener en cuenta PMem para utilizar este modo. Los datos almacenados en los módulos NVDIMM pueden persistir a lo largo de los ciclos de alimentación, ya que NVDIMM actúa como memoria direccionable por bytes. Los NVDIMM se almacenan automáticamente en el almacén de datos PMem creado por el ESXi cuando se crea el PMEM.

- Modo de disco virtual

Destinado a sistemas operativos tradicionales y antiguos que residen en VM para admitir cualquier versión de hardware. No es necesario que el SO de VM reconozca PMEM. En este modo, el SO de VM puede crear y utilizar un disco virtual tradicional de interfaz de sistemas informáticos pequeños (SCSI).

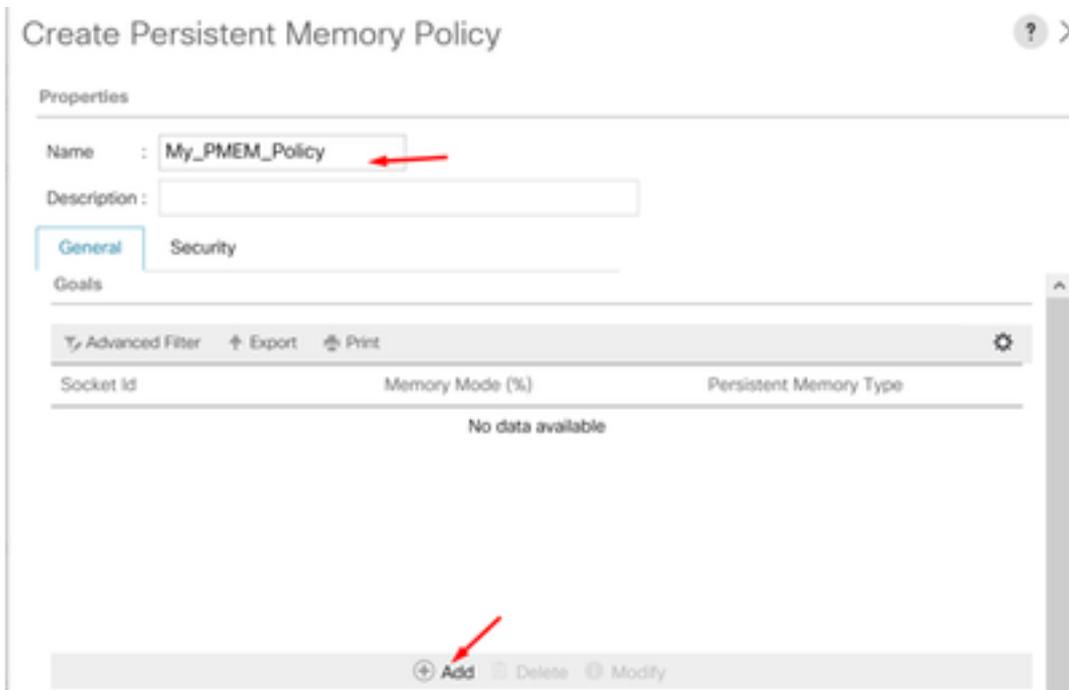
Este documento describe la configuración para utilizar una máquina virtual en modo de acceso directo.

## Configurar

Este procedimiento describe cómo configurar ESXi en servidores blade UCS con DCPMM Intel Optane.

### Configuración del perfil de servicio

1. En UCS Manager GUI, navegue hasta **Servers > Persistent Memory Policy** y haga clic en **Add** como se muestra en la imagen.



2. Crear **Objetivo**, asegúrese de que el **Modo de Memoria** sea del 0% como se muestra en la imagen.

## Create Goal



### Properties

Socket ID :  All Sockets

Memory Mode (%) :

Persistent Memory Type :  App Direct  App Direct Non Interleaved

OK

Cancel

3. Agregue la política PMEM al perfil de servicio deseado.

Navegue hasta **Perfil de servicio > Políticas > Política de memoria persistente** y adjunte la política creada.

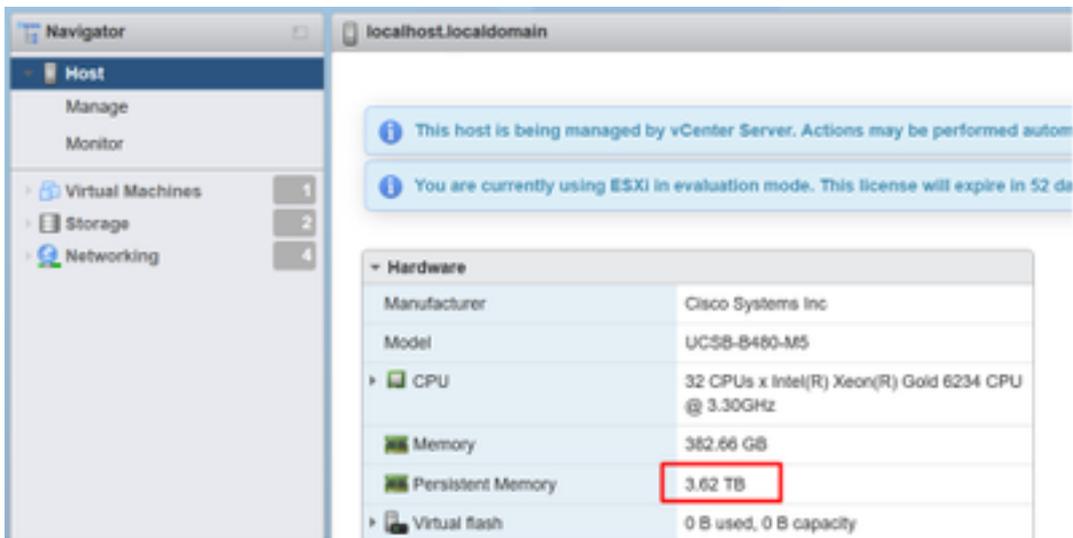
4. Verifique el estado de la región.

Navegue hasta el **Servidor** seleccionado > **Inventario > Memoria persistente > Regiones**. El tipo AppDirect está visible. Este método crea una región por socket de CPU.

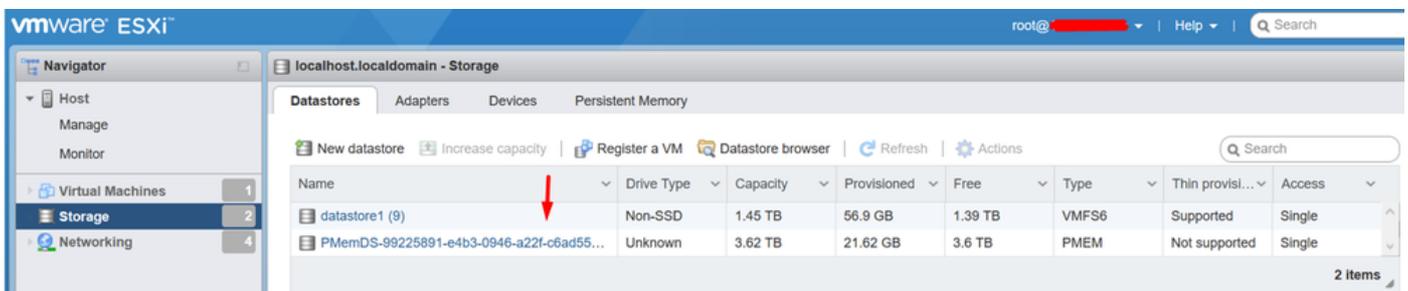
Id	Socket Id	Local DIMM Slot	DIMM Locator Id	Type	Total Capacity (L)	Free Capacity (L)	Health Status
1	Socket 1	Not Applicable	DIMM_A2.DIMM...	AppDirect	928	928	Healthy
2	Socket 2	Not Applicable	DIMM_G2.DIMM...	AppDirect	928	928	Healthy
3	Socket 3	Not Applicable	DIMM_N2.DIMM...	AppDirect	928	928	Healthy
4	Socket 4	Not Applicable	DIMM_U2.DIMM...	AppDirect	928	928	Healthy

## Verificar ESXi

1. En la consola Web, el host muestra el PMEM total disponible.



2. ESXi muestra un almacén de datos especial compuesto por la cantidad total de PMEM, como se muestra en la imagen.



## Configurar NVDIMM de máquina virtual

1. En ESXi, las máquinas virtuales acceden a Optane DCPMM PMEM como NVDIMM. Para asignar un NVMDIMM a una máquina virtual, acceda a la máquina virtual a través de vCenter y navegue hasta **Acciones > Editar configuración**, haga clic en **AGREGAR NUEVO DISPOSITIVO** y seleccione **NVDIMM** como se muestra en la imagen.



**Nota:** Cuando cree una máquina virtual, asegúrese de que la compatibilidad con el sistema operativo cumpla con la versión mínima requerida que admita la memoria persistente Intel® Optane™; de lo contrario, la opción **NVDIMM** no aparecerá en los elementos seleccionables.

2. Establezca el tamaño de NVDIMM como se muestra en la imagen.



## Configurar espacios de nombres en la máquina virtual

1. La utilidad **NDCTL** se utiliza para administrar y configurar PMEM o NVDIMM.

En el ejemplo, se utiliza Red Hat 8 para la configuración. Microsoft tiene cmdlets de PowerShell para la administración persistente del espacio de nombres de memoria.

Descargue la utilidad **NDCTL** usando la herramienta disponible según la distribución Linux

Por ejemplo:

```
# yum install ndctl # zypper install ndctl # apt-get install ndctl
```

2. Verifique la región NVDIMM y el espacio de nombres creados de forma predeterminada por ESXi, cuando el NVDIMM se asigna a la máquina virtual, verifique que el espacio coincida con la configuración. Asegúrese de que el modo del espacio de nombres esté establecido en **raw**, lo que significa que ESXi ha creado el espacio de nombres. Para verificar, utilice el comando:

```
# ndctl list -RuN
```

```
admin@localhost:/etc
[admin@localhost etc]$ ndctl list -RuN
{
  "regions": [
    {
      "dev": "region0",
      "size": "20.00 GiB (21.47 GB)",
      "available_size": 0,
      "max_available_extent": 0,
      "type": "pmem",
      "persistence_domain": "unknown",
      "namespaces": [
        {
          "dev": "namespace0.0",
          "mode": "raw",
          "size": "20.00 GiB (21.47 GB)",
          "blockdev": "pmem0"
        }
      ]
    }
  ]
}
```

3. (Opcional) Si el espacio de nombres no se ha creado ya, se puede crear un espacio de nombres con el comando:

```
# ndctl create-namespace
```

El comando **ndctl create-espacio de nombres** crea un nuevo espacio de nombres en el modo **fsdax** de forma predeterminada y crea un nuevo dispositivo **/dev/pmem([x].[y])**. Si ya se ha creado un espacio de nombres, se puede omitir este paso.

4. Seleccione el modo de acceso PMEM, los modos disponibles para la configuración son:

- Modo de sector:

Presenta el almacenamiento como un dispositivo de bloqueo rápido, lo que resulta útil para aplicaciones heredadas que aún no pueden utilizar la memoria persistente.

- Modo Fsdax:

Permite que los dispositivos de memoria persistentes admitan acceso directo al NVDIMM. El acceso directo al sistema de archivos requiere el uso del modo **fsdax**, para habilitar el uso del modelo de programación de acceso directo. Este modo permite la creación de un sistema de archivos sobre el NVDIMM.

- Modo Devdax:

Proporciona acceso sin formato a la memoria persistente mediante un dispositivo de caracteres DAX. Los sistemas de archivos no se pueden crear en los dispositivos que utilizan el modo **devdax**.

- Modo sin formato:

Este modo tiene varias limitaciones y no se recomienda para el uso de la Memoria Persistente.

Para cambiar el modo al modo **fsdax**, utilice el comando:

```
ndctl create-namespace -f -e
```

Si ya hay un **dev** creado, el espacio de nombres dev se utiliza para formatear y modificar el modo a **fsdax**.

```

admin@localhost:/etc
File Edit View Search Terminal Help
    "size": "20.00 GiB (21.47 GB)",
    "blockdev": "pmem0"
  }
}
}
}
}
[admin@localhost etc]$ ndctl create-namespace -f -e namespace0.0 --mode fsdax
failed to reconfigure namespace: Permission denied
[admin@localhost etc]$ sudo ndctl create-namespace -f -e namespace0.0 --mode fsdax
[sudo] password for admin:
{
  "dev": "namespace0.0",
  "mode": "fsdax",
  "map": "dev",
  "size": "19.69 GiB (21.14 GB)",
  "uuid": "09658ac7-16ea-4c3d-8fbc-e9dae854ddf0",
  "sector_size": 512,
  "blockdev": "pmem0",
  "numa_node": 0
}
[admin@localhost etc]$

```

**Nota:** Estos comandos requieren que la cuenta tenga privilegios de root, puede que se requiera comando **sudo**.

#### 5. Cree un directorio y un sistema de archivos.

Direct Access (Acceso directo) o DAX es un mecanismo que permite a las aplicaciones acceder directamente a los medios persistentes desde la CPU (a través de cargas y almacenes), evitando la pila de E/S tradicional. Los sistemas de archivos de memoria persistente habilitados para DAX incluyen ext4, XFS y Windows NTFS.

Ejemplo del sistema de archivos XFS creado y montado:

`sudo mkdir < directory route (e.g./mnt/pmem) > sudo mkfs.xfs < /dev/devicename (e.g. pmem0) >`

```

admin@localhost:/etc
File Edit View Search Terminal Help
}
[admin@localhost etc]$ mkdir /mnt/pmem
mkdir: cannot create directory '/mnt/pmem': Permission denied
[admin@localhost etc]$ sudo mkdir /mnt/pmem
[admin@localhost etc]$ sudo mkfs.xfs /dev/pmem0
meta-data=/dev/pmem0      isize=512    agcount=4, agsize=1290112 blks
=                sectsz=4096  attr=2, projid32bit=1
=                crc=1      finobt=1, sparse=1, rmapbt=0
=                reflink=1
data        =                bsize=4096  blocks=5160448, imaxpct=25
=                sunit=0    swidth=0 blks
naming     =version 2      bsize=4096  ascii-ci=0, ftype=1
log        =internal log  bsize=4096  blocks=2560, version=2
=                sectsz=4096  sunit=1 blks, lazy-count=1
realtime   =none        extsz=4096  blocks=0, rtextents=0
[admin@localhost etc]$

```

#### 6. Monte el sistema de archivos y verifique que se ha realizado correctamente.

`sudo mount`

```
admin@localhost:/etc
File Edit View Search Terminal Help
[admin@localhost etc]$ sudo mount /dev/pmem0 /mnt/pmem/
[admin@localhost etc]$ // verify the mount was successful
bash: //: Is a directory
[admin@localhost etc]$ df -h /mnt/pmem/
Filesystem      Size  Used Avail Use% Mounted on
/dev/pmem0      20G  173M   20G   1% /mnt/pmem
[admin@localhost etc]$
```

La máquina virtual está lista para utilizar PMEM.

## Troubleshoot

Se recomienda generalmente montar este sistema de archivos habilitado para DAX usando la opción de montaje **-o dax**, si se encuentra un error.

```
[admin@localhost etc]$ sudo mount -o dax /dev/pmem0 /mnt/pmem/
mount: /mnt/pmem: wrong fs type, bad option, bad superblock on /dev/pmem0, missing codepage or helper program, or other error.
```

La reparación del sistema de archivos se ejecuta para garantizar la integridad.

```
[admin@localhost etc]$ sudo xfs_repair /dev/pmem0
[sudo] password for admin:
Phase 1 - find and verify superblock...
Phase 2 - using internal log
- zero log...
- scan filesystem freespace and inode maps...
- found root inode chunk
Phase 3 - for each AG...
- scan and clear agi unlinked lists...
- process known inodes and perform inode discovery...
- agno = 0
- agno = 1
- agno = 2
- agno = 3
- process newly discovered inodes...
Phase 4 - check for duplicate blocks...
- setting up duplicate extent list...
- check for inodes claiming duplicate blocks...
- agno = 0
- agno = 1
- agno = 2
- agno = 3
Phase 5 - rebuild AG headers and trees...
- reset superblock..
Phase 6 - check inode connectivity...
- resetting contents of realtime bitmap and summary inodes
- traversing filesystem ...
- traversal finished ...
- moving disconnected inodes to lost+found ...
Phase 7 - verify and correct link counts...
done
[admin@localhost etc]$
```

Como solución alternativa, el montaje se puede montar sin la opción **-o dax**.

**Nota:** En la versión 5.1 de **xfsprogs**, el valor predeterminado es crear sistemas de archivos XFS con la opción **reflink** habilitada. Anteriormente estaba desactivado de forma predeterminada. Las opciones **reflink** y **dax** son mutuamente excluyentes, lo que hace que

falle el montaje.

"No se puede utilizar DAX y reflink juntos". el error se puede ver en **dmesg** cuando falla el comando mount:

```
admin@localhost:/etc
File Edit View Search Terminal Help
log      =internal log          bsize=4096   blocks=2560, version=2
         =                  sectsz=4096  sunit=1 blks, lazy-count=1
realtime =none              extsz=4096   blocks=0, rtextents=0
[admin@localhost etc]$ mount -o dax /dev/pmem0 /mnt/pmem
mount: only root can use "--options" option
[admin@localhost etc]$ sudo mount -o dax /dev/pmem0 /mnt/pmem/
mount: /mnt/pmem: wrong fs type, bad option, bad superblock on /dev/pmem0, missing
codepage or helper program, or other error.
[admin@localhost etc]$ dmesg -T | tail
[mar nov 10 00:12:18 2020] VFS: busy inodes on changed media or resized disk sr0
[mar nov 10 00:12:22 2020] ISO 9660 Extensions: Microsoft Joliet Level 3
[mar nov 10 00:12:22 2020] ISO 9660 Extensions: RRIP_1991A
[mar nov 10 01:47:35 2020] pmem0: detected capacity change from 0 to 21137195008
[mar nov 10 01:51:19 2020] XFS (pmem0): DAX enabled. Warning: EXPERIMENTAL, use
at your own risk
[mar nov 10 01:51:19 2020] XFS (pmem0): DAX and reflink cannot be used together!
[mar nov 10 01:53:06 2020] XFS (pmem0): DAX enabled. Warning: EXPERIMENTAL, use
at your own risk
[mar nov 10 01:53:06 2020] XFS (pmem0): DAX and reflink cannot be used together!
[mar nov 10 01:59:29 2020] XFS (pmem0): DAX enabled. Warning: EXPERIMENTAL, use
at your own risk
[mar nov 10 01:59:29 2020] XFS (pmem0): DAX and reflink cannot be used together!
[admin@localhost etc]$
```

Como solución alternativa, quite la opción **-o dax**.

```
admin@localhost:/etc
File Edit View Search Terminal Help
[admin@localhost etc]$ sudo mount /dev/pmem0 /mnt/pmem/
[admin@localhost etc]$ // verify the mount was successful
bash: //: Is a directory
[admin@localhost etc]$ df -h /mnt/pmem/
Filesystem      Size  Used Avail Use% Mounted on
/dev/pmem0      20G  173M   20G   1% /mnt/pmem
[admin@localhost etc]$
```

Montar con ext4 FS.

El sistema de archivos EXT4 se puede utilizar como alternativa porque no implementa la función de reflejo pero admite DAX.

```
[admin@localhost etc]$ sudo mkfs.ext4 /dev/pmem0
mke2fs 1.44.3 (10-July-2018)
/dev/pmem0 contains a xfs file system
Proceed anyway? (y,N) y
Creating filesystem with 5160448 4k blocks and 1291808 inodes
Filesystem UUID: 164c6d57-0462-45a0-9b94-703719272816
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

[admin@localhost etc]$ sudo mount /dev/pmem0 /mnt/pmem/
[admin@localhost etc]$ df -h /mnt/pmem/
Filesystem      Size  Used Avail Use% Mounted on
/dev/pmem0      20G   45M   19G   1% /mnt/pmem
[admin@localhost etc]$
```

## Información Relacionada

- [Guía de inicio rápido: Suministro de memoria persistente de CC Intel® Optane™](#)
- [Configuración de memoria persistente](#)
- [Utilidades de administración ipmctl y ndctl para memoria persistente Intel® Optane™](#)
- [Soporte Técnico y Documentación - Cisco Systems](#)