

KMEM_DAX mode

출처:

pmem.io: Memkind support for KMEM DAX option
Posted January 20, 2020 " Previous post Next post " Linux kernel version 5.1 brings in support for the volatile-use of persistent memory as a hotplugged memory region (KMEM DAX). When this feature is enabled, persistent memory is seen as a separate memory-only NUMA node(s).
 <https://pmem.io/2020/01/20/memkind-dax-kmem.html>

kernel 에서 KMEM_DAX feature을 enable하게 되면 persistent memory가 seperate한 memory-only Numa node로 보여지게 됨.

Requirements


1. Kernel 5.1 이상에서 KMEM DAX driver을 enable해야함.
 - a. 처음에 KMEM DAX가 enable되어있는지 확인하는 방법을 몰라서 nconfig를 이용하여 알맞은 driver 설치를 위한 configure을 진행함.

```
sudo apt-get install flex
sudo apt-get install bison
make nconfig
```

- make nconfig를 하고 나면 KMEM DAX driver를 enable할 수 있는 configuration창이 나옴
- configuration창에서 "Device Drivers -->"라는 메뉴를 찾아서 enter한 다음에 "-*- DAX: direct access to differentiated memory -->"라는 메뉴로 들어가기
- <M> PMEM DAX: support the deprecated /sys/class/dax interface 라는 메뉴에서 스페이스바를 눌러 <> PMEM DAX: support the deprecated /sys/class/dax interface로만 들어준 후 저장하기(이때 kmemdax.config로 저장을 해줬음. 이름은 상관없을 것 같음)

- b. ndctl 와 daxctl을 버전 66 이상으로 업그레이드하기(ndctl --version으로 확인)

- ndctl을 업그레이드 하면 daxctl도 업그레이드 됨
- 참고 사이트:

GitHub - pmem/ndctl: Helper tools and libraries for managing the libnvdimm (non-volatile memory device) sub-system in the Linux kernel
Utility library for managing the libnvdimm (non-volatile memory device) sub-system in the Linux kernel ./autogen.sh ./configure CFLAGS="-g -O2" --prefix=/usr --sysconfdir=/etc --libdir=/usr/lib64 make make check sudo make install There are a number of packages required for the build steps that may not be installed by default.
 <https://github.com/pmem/ndctl>

pmem/ndctl
Helper tools and libraries for managing the libnvdimm (non-volatile memory device) sub-system in the Linux kernel

42 56 142 85


```
git clone -b v71 https://github.com/pmem/ndctl
cd ndctl
./autogen.sh
sudo apt install libkeyutils-dev
./configure CFLAGS="-g -O2" --prefix=/usr --sysconfdir=/etc --libdir=/usr/lib64
make
make check
sudo make install
ndctl --version
```

2. System configuration

- a. Device-Dax의 reconfiguration은 dax-bus device model에 의존함.
- b. 커널이 /sys/bus/dax 모델을 지원해야함
- c. /sys/class/dax에서 /sys/bus/dax 모델로 migrate하기 위해 다음 명령이 필요

```
daxctl migrate-device-model
```

- d. 참고사이트:

pmem.io: ndctl
daxctl-migrate-device-model - Opt-in to the /sys/bus/dax device-model, allow for alternative Device-DAX instance drivers. Arrange for modprobe to disable the dax_pmem_compat, if present, and instead deploy the dax_pmem module to convert to the /sys/bus/dax model. Kernel versions prior to v5.1 may not support /sys/bus/dax in which case the result of this command is a nop until the kernel is updated.
 <https://pmem.io/ndctl/daxctl-migrate-device-model.html>

3. Reconfigure DAX devices

- a. 다음 명령을 통해 numa node가 하나 더 생성된 것을 확인할 수 있음(pmem이 하나의 memory-only numa node로 잡힌 것)

```
numactl -H
```

- b. 기존 app direct mode에서 생성시켜준 namespace 지워주기

```
ndctl list -iN
umount /dev/pmem0
ndctl disable-namespace namespace0.0
ndctl destroy-namespace namespace0.0
ndctl list -R

#namespace 실행시켜주기
ndctl create-namespace --mode=devdax --map=mem
daxctl list
```

c. 시스템 재부팅

d. DAX device를 devdax mode에서 system-ram mode로 reconfigure하기

```
daxctl reconfigure-device dax0.0 --mode=system-ram
```

e. free -g를 통해 pmem과 dram 이 모두 메모리로 잡힌 것을 확인할 수 있음

4. 다시 devdax로 모드를 바꾸고 싶은 경우

<https://pmem.io/ndctl/daxctl-reconfigure-device.html>

```
sudo -i #root로 권한 바꾸기  
daxctl reconfigure-device --human --mode=devdax --force dax0.0
```

위의 명령을 실행한 후 memory is used 등의 오류가 나타날 수도 있으므로 그럴 때는 reboot를 진행하여 다시 위의 명령을 실행시켜주면 dax0.0 was already in devdax mode라는 메시지를 받을 수 있을 것

그 다음에 namespace를 disable하여 fsdax로 바꾸든지 아니면 아예 모드를 바꾸든지 하면 된다.

```
ndctl list  
ndctl disable-namespace namespace0.0  
ndctl destroy-namespace namespace0.0
```