

# NTB 环境配置、相关资料整理

## 一、NTB 环境配置

---

1. 在 Bios 中将两台 B2B 相连的服务器设置为： ntb  
pcie port 为 NTB to NTB ； Enable NTB bars 为  
Enabled ； BAR size 设置为 29；
2. 系统后 执行

```
modprobe ntb_transport
```

之后 `lsmod | grep ntb` 应该有 3 个 mod

```
root@NPG_DPDK_VIRTIO_NTB_jiayuhu_2:/home/f-stack/dpdk# lsmod | grep ntb
ntb_transport      36864  0
ntb_hw_intel       53248  0
ntb                16384  2 ntb_transport,ntb_hw_intel
```

3. 执行

```
./dpdk/usertools/dpdk-setup.sh
export RTE_SDK=/home/ntb-server1/dpdk
export RTE_TARGET=x86_64-native-linuxapp-gcc
```

对 DPDK 进行编译，之后 Insert IGB UIO 模块，并  
Setup hugepage mappings

```

Step 1: Select the DPDK environment to build
-----
[1] arm64-armv8a-linuxapp-clang
[2] arm64-armv8a-linuxapp-gcc
[3] arm64-dpaa2-linuxapp-gcc
[4] arm64-dpaa-linuxapp-gcc
[5] arm64-stingray-linuxapp-gcc
[6] arm64-thunderx-linuxapp-gcc
[7] arm64-xgene1-linuxapp-gcc
[8] arm-armv7a-linuxapp-gcc
[9] i686-native-linuxapp-gcc
[10] i686-native-linuxapp-icc
[11] ppc_64-power8-linuxapp-gcc
[12] x86_64-native-bsdapp-clang
[13] x86_64-native-bsdapp-gcc
[14] x86_64-native-linuxapp-clang
[15] x86_64-native-linuxapp-gcc
[16] x86_64-native-linuxapp-icc
[17] x86_x32-native-linuxapp-gcc
-----

Step 2: Setup linuxapp environment
-----
[18] Insert IGB UIO module
[19] Insert VFIO module
[20] Insert KNI module
[21] Setup hugepage mappings for non-NUMA systems
[22] Setup hugepage mappings for NUMA systems
[23] Display current Ethernet/Crypto device settings
[24] Bind Ethernet/Crypto device to IGB UIO module
[25] Bind Ethernet/Crypto device to VFIO module
[26] Setup VFIO permissions
-----

Step 3: Run test application for linuxapp environment
-----
[27] Run test application ($RTE_TARGET/app/test)
[28] Run testpmd application in interactive mode ($RTE_TARGET/app/testpmd)
-----

Step 4: Other tools
-----
[29] List hugepage info from /proc/meminfo
-----

```

#### 4. 设置 1GB 大页，由于 Intel 的机器有两个 NUMA 节点，因此需要为两个节点各分配一页大页。

```

mount -t hugetlbfs -o pagesize=1GB nodev /mnt/huge_1GB/
echo 1 > /sys/devices/system/node/node0/hugepages/hugepages-
1048576kB/nr_hugepages
echo 1 > /sys/devices/system/node/node1/hugepages/hugepages-
1048576kB/nr_hugepages

```

分配后可使用

```
cat /sys/kernel/mm/hugepages/hugepages-1048576kB/free_hugepages
```

来查看 1GB 大页内存的数量，以及

```
mount | grep huge
```

查看挂载大页内存的目录

## 5. 将 NTB 设备绑定到 DPDK Driver，首先执行

```
./dpdk/usertools/dpdk-devbind.py -s
```

查看 NTB 设备的编号

```
=====
0000:3d:00.1 'Ethernet Connection X722 for 10GBASE-T 37d2' drv=igb_uio unused=i40e

Network devices using kernel driver
=====
0000:3d:00.0 'Ethernet Connection X722 for 10GBASE-T 37d2' if=enol drv=i40e unused=igb_uio *Active*
0000:af:00.0 'MT28800 Family [ConnectX-5 Ex] 1019' if=enpl175s0f0 drv=mlx5_core unused=igb_uio
0000:af:00.1 'MT28800 Family [ConnectX-5 Ex] 1019' if=enpl175s0f1 drv=mlx5_core unused=igb_uio

No 'Baseband' devices detected
=====

No 'Crypto' devices detected
=====

No 'Eventdev' devices detected
=====

No 'Mempool' devices detected
=====

No 'Compress' devices detected
=====

Misc (rawdev) devices using DPDK-compatible driver
=====
0000:17:00.0 'Sky Lake-E Non-Transparent Bridge Registers 201c' drv=igb_uio unused=ntb_hw_intel
=====
Misc (rawdev) devices using kernel driver
=====
0000:00:04.0 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.1 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.2 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.3 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.4 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.5 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.6 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:00:04.7 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.0 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.1 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.2 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.3 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.4 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.5 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.6 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
0000:80:04.7 'Sky Lake-E CBDMA Registers 2021' if= drv=ioatdma unused=igb_uio
```

之后执行

```
./dpdk/usertools/dpdk-devbind.py --bind=igb_uio 17:00.0
```

将 NTB 绑定到 DPDK Driver。如果想将 NTB 从 DPDK Driver 绑回 Kernel Driver，使用

```
./usertools/dpdk-devbind.py -b ntb_hw_intel 17:00.0
```

6. 为了使 NTB Remote Write 获得最佳性能，首先执行任意 NTB 可执行程序，获得其打印的 NTB 所属系统地址空间的地址，之后执行

```
echo "disable=1" >> /proc/mtrr  
echo "base=0x387fe8000000 size=0x8000000 type=write-combining" >>  
/proc/mtrr  
cat /proc/mtrr
```

第二条指令中的 Size 为 NTB 拥有的地址空间的大小

```
root@NPG_DPDK_VIRTIO NTB_jiayuhu 2:/home/ntb-server2/dpdk# cat /proc/mtrr  
reg00: base=0x000000000 ( 0MB), size= 2048MB, count=1: write-back  
reg01: base=0x387fe8000000 (59244160MB), size= 128MB, count=1: write-combining  
reg02: base=0x100000000 ( 4096MB), size= 4096MB, count=1: write-back  
reg03: base=0x200000000 ( 8192MB), size= 8192MB, count=1: write-back  
reg04: base=0x400000000 (16384MB), size=16384MB, count=1: write-back  
reg05: base=0x800000000 (32768MB), size=16384MB, count=1: write-back  
reg06: base=0xc00000000 (49152MB), size= 1024MB, count=1: write-back
```

## 二、相关网页链接

1. DPDK Programmer's Guide（包括 Ring、Mempool、Mbuf 相关介绍）

[https://doc.dpdk.org/guides/prog\\_guide/](https://doc.dpdk.org/guides/prog_guide/)

2. DPDK Multi-process Support（参考其多进程通讯的实现）

[http://doc.dpdk.org/guides/prog\\_guide/multi\\_proc\\_support.html](http://doc.dpdk.org/guides/prog_guide/multi_proc_support.html)

### 3. Intel NTB Startup Guide

<https://github.com/davejiang/linux/wiki/Intel-NTB-Startup-Guide>

### 4. Linux kernel.org NTB Drivers Documentation

<https://www.kernel.org/doc/html/latest/driver-api/ntb.html>

### 5. Arm 公司开源的 NTB Drivers

<https://github.com/jonmason/ntb>

### 6. Dolphin （一家 NTB 商用公司）

<https://www.dolphinics.com/index.html>