AI400X2 24x13.9TiB nvme 8xHDR200

System description

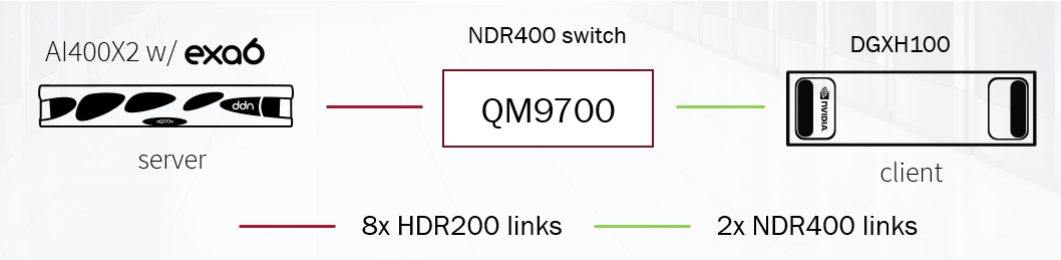
DDN

August 4, 2023

Contents

1 Hardware

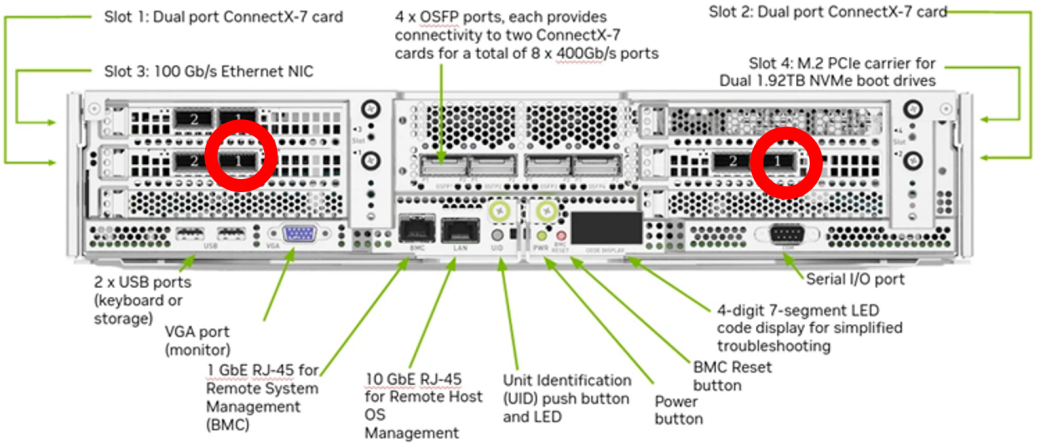
The hardware is composed of one client and one server connected to the same switch.



1.1 Client

The client is **a commercial NVIDIA DGX H100** node (compute). The configuration is the default one. The client communicate with the server through IB using the two links dedicated for the storage following the NVIDIA reference architecture.

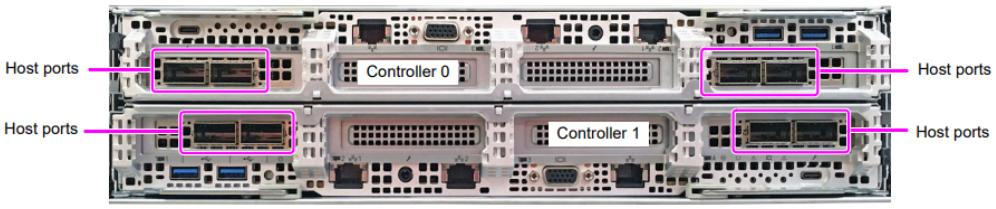
- 2x HDR400 links total from two different CX-7 network adapters (aggregated bandwidth of 800Gb/s)



1.2 Server

The server is **a commercial DDN AI400X2** appliance (storage). The configuration is the default one with 24x SAMSUMG NVME drives with 13.9TiB per drive. The server communicate with the client through IB:

- 4x dual ports HDR200 adapters for a total of 8x HDR200 links



1.3 Switch

The switch is **a commercial NVIDIA QM9700 (MQM9700-NS2R)**. The client and server are connected on it.

2 Software

2.1 Client

The NVIDIA DGX H100 uses DGX OS version 6.0.11. The distribution is Ubuntu 22.04.2 LTS (Jammy Jellyfish). The Mellanox OFED version is `MLNX_OFED_LINUX-5.9-0.5.6.0.113` (OFED-5.9-0.5.6.0.113)

2.2 Server

The AI400X2 uses EXAScaler SFA Rocky version 6.2.0-r9 on top of the SFAOS stack (SFAOS version 12.2.0) The EXAScaler driver uses on the server is `2.14.0_ddn85`. This is the GA version of the product with no modification applied. The OFED version is OFED-internal-5.8-2.0.3.

2.3 Switch

The firwmare of the NVIDIA QM9700 switch is 31.2010.2036.

3 Settings

3.1 Client

The standard EXAScaler 6.2.0 driver (GA) is installed on the client using the default installation procedure. This driver is used for the client to access the remote EXAScaler filesystem.

3.1.1 EXAScaler driver configuration

The EXAScaler configuration file has been generated automatically following the default installation procedure. No tuning is applied. Therefore checksums are enabled, which is known to decrease the performance. This configuration file is necessary to mount the filesystem and is generated automatically. It is not part of an optimization or customization. It is the default configuration file.

Content of /etc/modprobe.d/lustre.conf on the DGX H100:

```
# This file has been generated by exa-client-deploy
#
# Do not edit unless exa-client-deploy service is stopped & disabled
# e.g: 'systemctl status exa-client-deploy'
#

options lnet networks="o2ib(ibp41s0f0,ibp170s0f0)"
options lnet lnet_transaction_timeout=100
options lnet lnet_retry_count=2
options ko2iblnd peer_credits=32
options ko2iblnd peer_credits_hiw=16
options ko2iblnd concurrent_sends=64
options ksocklnd conns_per_peer=0
```

3.2 Server

No tuning applied on the DDN AI400X2 after the initial installation. The inital installation was automated using the latest version of our manufacturing script (mfg-e-dcr-20230627-1). The filesystem was mounted on the client after the end of this script.

3.3 Switch

No tuning applied.

4 Misc

4.1 Drop cache

The cache of both the client and the server was drop between two iterations using the following function for all tests. Since 'set -e' and 'clush -S' were used, both commands did work.

```
clear_caches()
{
    # Clear cache on client
    sync; echo 3 > /proc/sys/vm/drop_caches
    # Clear cache on EXAScaler VMs
    ssh root@${exafsip}  "clush -abS 'sync; echo 3 > /proc/sys/vm/drop_caches'"
}
```

4.2 No tunings

To ensure no tunings were applied, the EXAScaler filesystem was unmounted and the driver module was unloaded from the client at the start of the run.

```
mount_exa_fs()
{
    # If the filesystem is mounted: umount
    if mount -t lustre |grep "." ; then
  umount -t lustre -a
  lustre_rmmod
    fi
    # If the kernel module is loaded: unload
    if lsmod |grep lustre ; then
  lustre_rmmod
    fi

    # We make sure no optimization is applied
    # Regenerate the default lustre.conf
    rm /etc/modprobe.d/lustre.conf
    (cd /home/ddn/ldouriez/exa-client && ./exa_client_deploy.py -c \
                        --lnets "o2ib(ibp41s0f0,ibp170s0f0)" -y)

    # Mount the EXAScaler filesystem
    mount -t lustre (...)

}
```

4.3 Cgroups for single simulated GPU

cgroups memory limitation was used for single simulated GPU tests. The memory was reduced from 2TiB to 32GB to speed-up the testing. The command used is given below

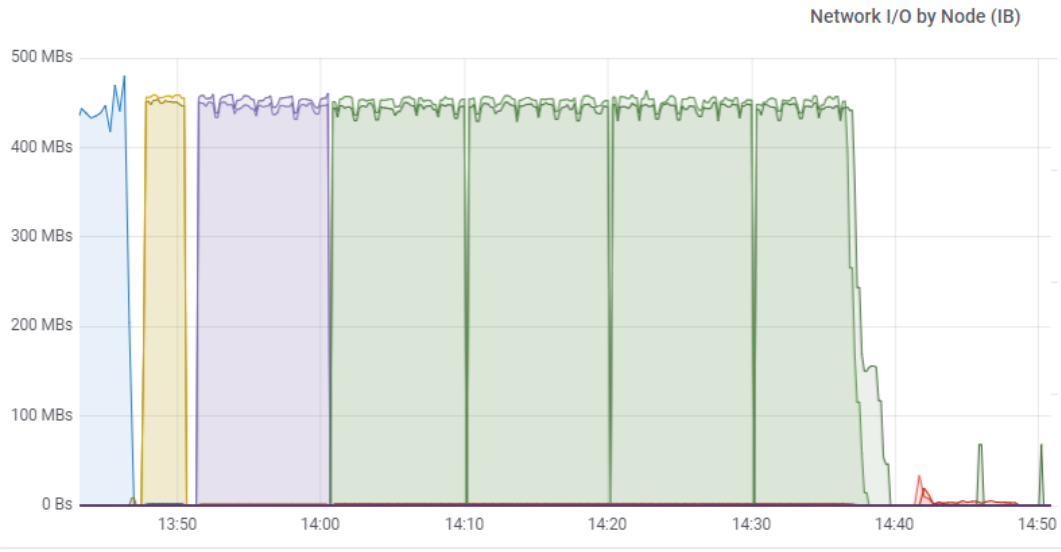
```
systemd-run --unit=mycgroup \
            --scope \
            -p MemoryMax=$((32*1024*1024*1024)) \
            ./single_host_closed_1_unet_PUB.sh
```

The free command was verified during the run along with the network traffic on the client to ensure that it worked correctly.

```
# The free command shows 32Gi of cached memory, which is correct
                total        used        free      shared  buff/cache   available
  Mem:            2.0Ti        13Gi       1.9Ti        21Mi        32Gi       1.9Ti
  Swap:             0B          0B          0B
```

The IB traffic on the client matches the output of the script

```
[METRIC] ============================================================
[METRIC] Training Accelerator Utilization [AU] (%): 99.7769 (0.0029)
[METRIC] Training Throughput (samples/second): 2.9096 (0.0015)
[METRIC] Training I/O Throughput (MB/second): 406.7823 (0.2112)
[METRIC] train_au_meet_expectation: success
[METRIC] ============================================================
```



4.4 Logs

All the scripts used have set -e and set -x at the beginning to ensure that all commands work and are logged. All scripts and logs are available in the code directory.