# Riser 1OU test

Author : Van Liefde Thomas, Electrical engineer

## Process :

We will first test one Ava Board with 4 NVMe on a 2 OU riser from Wywinn or Quanta and check the results we have after a burn

We will then test a GPU Nvidia Tesla T4 with a 2 OU riser from Quanta or Wywinn and check results after a burn.

Only the x16 PCIe slot of the 2 OU riser will be used during those test

Those results will be the baseline for our test with the 1 OU riser for Leopard

## Test with a 2 OU riser

Une image contenant texte, équipement électronique, circuit

Description générée automatiquement

### Ava Board result

Une image contenant texte, équipement électronique

Description générée automatiquement

* Check that every drive is present :

ubuntu@ubuntu:~$ lsblk

NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT

sda 8:0 0 111.8G 0 disk

|-sda1 8:1 0 512M 0 part /boot/efi

|-sda2 8:2 0 110.3G 0 part /

`-sda3 8:3 0 976M 0 part [SWAP]

nvme0n1 259:1 0 1.9T 0 disk

nvme1n1 259:3 0 1.9T 0 disk

nvme2n1 259:2 0 1.9T 0 disk

nvme3n1 259:0 0 1.9T 0 disk

* We run different performance test on the drives and check the results:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BW randrw (MB/s) | BW write (MB/s) | IOPS randread | IOPS randrw  (read/write) | Latency (usec) |
| Nvme0n1 | 2711.7 | 1369.7 | 342324 | 80293/26736 | 157.63 |
| Nvme1n1 | 2666.7 | 1324.5 | 345215 | 81278/27068 | 132.68 |
| Nvme2n1 | 2731.3 | 1338.1 | 345726 | 80156/26696 | 132.81 |
| Nvme3n1 | 2694.5 | 1353.1 | 367096 | 86178/28700 | 128.11 |

### T4 GPU result

Fan speed has been set up to 50% of their max speed

* **Device before burn:**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 2.5GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 38C P8 9W / 70W | 63MiB / 15109MiB | 0% Default |

* **Device during burn :**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 8GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 74C P0 69W / 70W | 13628MiB / 15109MiB | 100% Default |

* **Device after burn :**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 2.5GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 56C P0 10W / 70W | 13628MiB / 15109MiB | 100% Default |

* **Burn result :**

ubuntu@ubuntu:~/gpu-burn$ ./gpu\_burn 300

GPU 0: Tesla T4 (UUID: GPU-a60cd8a3-1e2c-75ef-9996-20ab27d8f938)

Initialized device 0 with 15109 MB of memory (14776 MB available, using 13299 MB of it), using FLOATS

10.3% proc'd: 6632 (4156 Gflop/s) errors: 0 temps: 53 C

21.0% proc'd: 14093 (4058 Gflop/s) errors: 0 temps: 64 C

31.7% proc'd: 21554 (3851 Gflop/s) errors: 0 temps: 73 C

42.0% proc'd: 28186 (3591 Gflop/s) errors: 0 temps: 82 C

52.3% proc'd: 32331 (1242 Gflop/s) errors: 0 temps: 84 C

63.3% proc'd: 33989 (921 Gflop/s) errors: 0 temps: 84 C

73.3% proc'd: 34818 (778 Gflop/s) errors: 0 temps: 84 C

85.0% proc'd: 36476 (741 Gflop/s) errors: 0 temps: 85 C

96.7% proc'd: 38134 (741 Gflop/s) errors: 0 temps: 85 C

100.0% proc'd: 38963 (741 Gflop/s) errors: 0 temps: 86 C

Killing processes.. Freed memory for dev 0

Uninitted cublas

done

Tested 1 GPUs:

GPU 0: OK

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Before burn** | **During burn** | **After burn** |
| **LnkCap** | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 |
| **LnkSta** | Speed 2.5GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 2.5GT/s, Width x16 |

## Test with 1OU riser

Une image contenant texte, équipement électronique, circuit

Description générée automatiquement

### Ava Board result

* Check that every drive is present :

ubuntu@ubuntu:~$ lsblk

NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT

sda 8:0 0 111.8G 0 disk

|-sda1 8:1 0 512M 0 part /boot/efi

|-sda2 8:2 0 110.3G 0 part /

`-sda3 8:3 0 976M 0 part [SWAP]

nvme0n1 259:1 0 1.9T 0 disk

nvme1n1 259:3 0 1.9T 0 disk

nvme2n1 259:2 0 1.9T 0 disk

nvme3n1 259:0 0 1.9T 0 disk

* We run different performance test on the drives and check the results:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BW read (MB/s) | BW write (MB/s) | IOPS randread | IOPS randwrite | Latency (usec) |
| Nvme0n1 | 2815.4 | 1309.3 | 335537 | 112630/37524 | 147.81 |
| Nvme1n1 | 2783.4 | 1287.5 | 348800 | 113241/37729 | 147.38 |
| Nvme2n1 | 2974.1 | 1217.5 | 347284 | 113297/37738 | 152.38 |
| Nvme3n1 | 2783.2 | 1449.8 | 349198 | 85866/28596 | 148.49 |

### T4 GPU result

Fan speed has been set up to 50% of their max speed

* **Device before burn:**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 2.5GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 42C P8 9W / 70W | 63MiB / 15109MiB | 0% Default |

* **Device during burn :**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 8GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 74C P0 70W / 70W | 13628MiB / 15109MiB | 100% Default |

* **Device after burn :**

ubuntu@ubuntu:~$ sudo lspci -vv -s 01:00

01:00.0 3D controller: NVIDIA Corporation Device 1eb8 (rev a1)

LnkCap: Port #0, Speed 8GT/s, Width x16

LnkSta: Speed 2.5GT/s, Width x16

ubuntu@ubuntu:~$ nvidia-smi

| NVIDIA-SMI 460.27.04 Driver Version: 460.27.04 CUDA Version: 11.2 |

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

| 0 Tesla T4 Off | 00000000:01:00.0 Off | 0 |

| N/A 58C P0 10W / 70W | 13628MiB / 15109MiB | 100% Default |

* **Burn result :**

ubuntu@ubuntu:~/gpu-burn$ ./gpu\_burn 300

GPU 0: Tesla T4 (UUID: GPU-a60cd8a3-1e2c-75ef-9996-20ab27d8f938)

Initialized device 0 with 15109 MB of memory (14766 MB available, using 13290 MB of it), using FLOATS

10.3% proc'd: 6624 (4118 Gflop/s) errors: 0 temps: 59 C

21.0% proc'd: 14076 (3949 Gflop/s) errors: 0 temps: 70 C

31.7% proc'd: 20700 (3674 Gflop/s) errors: 0 temps: 79 C

42.0% proc'd: 25668 (1146 Gflop/s) errors: 0 temps: 84 C

53.3% proc'd: 26496 (912 Gflop/s) errors: 0 temps: 84 C

65.0% proc'd: 28152 (741 Gflop/s) errors: 0 temps: 86 C

76.7% proc'd: 29808 (741 Gflop/s) errors: 0 temps: 87 C

88.3% proc'd: 31464 (741 Gflop/s) errors: 0 temps: 89 C

100.0% proc'd: 33120 (741 Gflop/s) errors: 0 temps: 91 C

100.0% proc'd: 33120 (741 Gflop/s) errors: 0 temps: 91 C

Killing processes.. Freed memory for dev 0

Uninitted cublas

done

Tested 1 GPUs:

GPU 0: OK

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Before burn** | **During burn** | **After burn** |
| **LnkCap** | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 |
| **LnkSta** | Speed 2.5GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 2.5GT/s, Width x16 |

## Conclusion

* **For Ava Board:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Drives** | **BW randrw (MB/s)** | **BW write (MB/s)** | **IOPS randread** | **IOPS randrw**  **(read/write)** | **Latency (usec)** |
| **Riser 2OU** | Nvme0n1 | 2711.7 | 1369.7 | 342324 | 80293/26736 | 157.63 |
| Nvme1n1 | 2666.7 | 1324.5 | 345215 | 81278/27068 | 132.68 |
| Nvme2n1 | 2731.3 | 1338.1 | 345726 | 80156/26696 | 132.81 |
| Nvme3n1 | 2694.5 | 1353.1 | 367096 | 86178/28700 | 128.11 |
| **Average** | **2700** | **1346** | **350090** | **81976/27300** | **137.80** |
| **Riser 1OU** | Nvme0n1 | 2815.4 | 1309.3 | 335537 | 112630/37524 | 147.81 |
| Nvme1n1 | 2783.4 | 1287.5 | 348800 | 113241/37729 | 147.38 |
| Nvme2n1 | 2974.1 | 1217.5 | 347284 | 113297/37738 | 152.38 |
| Nvme3n1 | 2783.2 | 1449.8 | 349198 | 85866/28596 | 148.49 |
| **Average** | **2838** | **1315** | **345204** | **106258/35396** | **149** |

Average performances between 2OU riser and 1OU riser are similar (less than 2% difference when performance is less for 1OU riser).

* **For T4 GPU:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **Before burn** | **During burn** | **After burn** |
| **Riser 2OU** | **LnkCap** | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 |
| **LnkSta** | Speed 2.5GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 2.5GT/s, Width x16 |
| **Riser 1OU** | **LnkCap** | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 8GT/s, Width x16 |
| **LnkSta** | Speed 2.5GT/s, Width x16 | Speed 8GT/s, Width x16 | Speed 2.5GT/s, Width x16 |

Performance for GPU are the same, no matter we use riser 2OU or riser 1OU so the riser 1OU is good in term of performance.

There is no downgrade in term of performance when using a 1OU riser instead of a 2OU riser so we can validate the electrical design of the 1OU riser.