

Falcon 4010 GPU 拡張シャーシが PCIe 版 A100 の GPU 性能に与える影響

著者 ジーン・チョウ、ケビン・ツアイ

H3 Platform Inc. 2020 年 9 月

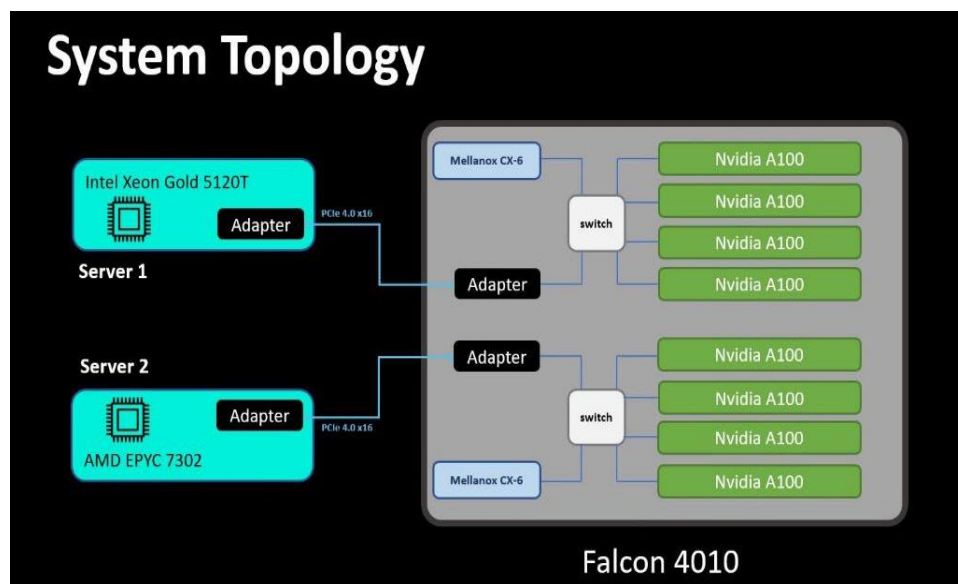
概要

PCIe 版 A100 GPU の登場に続き、異なるホストサーバ下での GPU ピアツーピア性能に対する当社の GPU ソリューションの効果をさらに理解するために、PCIe Gen.4 GPU を搭載した Falcon 4010 GPU 拡張シャーシをテストすることができました。この実験では、PCIe Gen.3 のインテル Xeon Gold 5120T と PCIe Gen.4 の AMD EPYC 7452 の 2 つの異なるサーバを使用し、PCIe 版 A100 GPU を Falcon 4010 プラットフォームを介してホストに接続し、両方の環境で CUDA サンプルコマンドを実行しました。この実験の結果、GPU は両方のサーバの下で PCIe Gen.4 のピアツーピア帯域幅を達成していますが、ホストとデバイスの通信はサーバ自体の PCIe 世代に制限されています。

テスト方法

Falcon 4010 が GPU ピアツーピアパフォーマンスにどのような影響を与えるかをテストするために、インテル Xeon Gold 5120T PCIe Gen.3 サーバーと AMD EPYC 7452 PCIe Gen.4 サーバーを同じ Falcon 4010 拡張シャーシに接続しました。各サーバには 4 つの PCIe 版 A100 GPU が割り当てられています。両方のホストサーバ上で CUDA サンプルコマンド

『p2pBandwidthLatencyTest』、『htod』、および『dtoh』を実行し、結果を比較します。



Copyright© 2020 H3 Platform Inc. All rights reserved.

The information is subject to change without notice

各サーバには PCIe 版 A100 GPU が割り当てられており、シャーシと各サーバ間の接続は PCIe 4.0 x4 ケーブル 4 本(PCIe Gen.4 x16)を介して行われます。ケーブルは miniSAS で、コネクタは SFF-8644。PCIe スイッチは Broadcom PEX88096、アダプタは Broadcom PEX88032 です。

手続き

1. AMD EPYC 7452 PCIe Gen.4 サーバーにログインします。
2. nvidia-smi コマンドを実行して、ホストが GPU を検出して認識していることを確認します。
3. p2pBandwidthLatencyTest コマンドを実行します。
4. htod コマンドを実行します。
5. dtod コマンドを実行します。
6. ホスト対デバイス、デバイス対ホスト、GPU ピア対ピア無効、GPU ピア対ピア有効の双方向帯域幅の結果を記録します。
7. インテル Xeon Gold 4120T PCIe Gen.3 サーバーにログインします。
8. 手順 2-6 を繰り返します。
9. 2 つのサーバーから得られたテスト結果を比較します。

テスト結果

AMD EPYC PCIe Gen.4 server

/p2pBandwidthLatencyTest

```

***NOTE: In case a device doesn't have P2P access to other one, it falls back to normal
memory procedure.
So you can see lesser Bandwidth (GB/s) and unstable Latency (us) in those cases.

P2P Connectivity Matrix
D\D 0 1 2 3
0 1 1 1 1
1 1 1 1 1
2 1 1 1 1
3 1 1 1 1

Unidirectional P2P=Disabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 1179.25 14.74 14.96 13.95
1 14.69 1166.04 14.62 14.16
2 15.52 17.10 1254.01 16.05
3 14.60 14.52 15.02 1287.07

Unidirectional P2P=Enabled Bandwidth Matrix (P2P Writes) Matrix (GB/s)
D\D 0 1 2 3
0 1287.07 24.60 24.60 24.48
1 24.60 1302.08 24.60 24.60
2 24.60 24.61 1301.00 24.60
3 24.60 24.60 24.61 1302.08

Bidirectional P2P=Disabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 1302.08 16.55 16.15 16.09
1 16.67 1300.46 16.16 16.09
2 15.90 15.94 1302.08 16.99
3 15.87 15.88 16.90 1303.71

Bidirectional P2P=Enabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 1296.68 48.62 48.65 48.65
1 48.65 1302.08 48.63 48.63
2 48.62 48.65 1299.38 48.66
3 48.63 48.62 48.64 1298.84

P2P=Disabled Latency Matrix (us)
GPU 0 1 2 3
0 2.74 25.60 25.54 25.53
1 25.62 2.75 25.68 24.60
2 25.62 25.62 2.55 24.60
3 25.52 24.71 24.59 2.49

CPU 0 1 2 3
0 4.05 11.77 11.69 11.67
1 11.67 4.00 11.67 11.60
2 11.58 11.56 4.01 11.65
3 11.75 11.54 11.69 4.00

P2P=Enabled Latency (P2P Writes) Matrix (us)
GPU 0 1 2 3
0 2.74 2.45 2.44 2.44
1 2.38 2.75 2.37 2.37
2 2.20 2.13 2.55 2.12
3 2.13 2.12 2.13 2.47

CPU 0 1 2 3
0 4.04 3.34 3.43 3.44
1 3.50 4.10 3.47 3.47
2 3.50 3.45 4.10 3.47
3 3.52 3.47 3.46 4.15

```

htod

```

root@h3:/root/ian/p2p# ../bandwidth/bandwidth -i 2 -t 5m htod
Host --> G2
Time,G2 [A100-PCIe-40GB]
2020-08-28 02:37:12,25.1152
2020-08-28 02:37:13,25.1148
2020-08-28 02:37:14,25.1145
2020-08-28 02:37:15,25.1145
2020-08-28 02:37:16,25.1171
2020-08-28 02:37:17,25.1173
2020-08-28 02:37:18,25.1168
2020-08-28 02:37:19,25.1178
2020-08-28 02:37:20,25.1173
2020-08-28 02:37:21,25.1173
2020-08-28 02:37:22,25.1167
2020-08-28 02:37:23,25.1174
2020-08-28 02:37:24,25.1178
2020-08-28 02:37:25,25.1166
2020-08-28 02:37:26,25.1180
2020-08-28 02:37:27,25.1174
2020-08-28 02:37:28,25.1170
2020-08-28 02:37:29,25.1176
2020-08-28 02:37:30,25.1170
2020-08-28 02:37:31,25.1176
2020-08-28 02:37:32,25.1176
2020-08-28 02:37:33,25.1174
2020-08-28 02:37:34,25.1181
2020-08-28 02:37:35,25.1169
2020-08-28 02:37:36,25.1185
2020-08-28 02:37:37,25.1174
2020-08-28 02:37:38,25.1170
2020-08-28 02:37:39,25.1177
2020-08-28 02:37:40,25.1169
2020-08-28 02:37:41,25.1169
2020-08-28 02:37:42,25.1174
2020-08-28 02:37:43,25.1176
2020-08-28 02:37:44,25.1173
2020-08-28 02:37:45,25.1170
2020-08-28 02:37:46,25.1174
2020-08-28 02:37:47,25.1171
2020-08-28 02:37:48,25.1174
2020-08-28 02:37:49,25.1167
2020-08-28 02:37:50,25.1173
2020-08-28 02:37:51,25.1169
2020-08-28 02:37:52,25.1174
2020-08-28 02:37:53,25.1168
2020-08-28 02:37:54,25.1173
2020-08-28 02:37:55,25.1175
2020-08-28 02:37:56,25.1179
2020-08-28 02:37:57,25.1169

```

Dtoh

```
root@h3:/root/ian/p2p# ../bandwidth/bandwidth -i 2 -t 5m dtoh
Host <--- G2
Time,G2 [AI00-PCIe-40GB]
2020-08-28 02:38:07,24.6109
2020-08-28 02:38:08,24.6108
2020-08-28 02:38:09,24.6108
2020-08-28 02:38:10,24.6110
2020-08-28 02:38:11,24.6107
2020-08-28 02:38:12,24.6109
2020-08-28 02:38:13,24.6109
2020-08-28 02:38:14,24.6108
2020-08-28 02:38:15,24.6108
2020-08-28 02:38:16,24.6109
2020-08-28 02:38:17,24.6110
2020-08-28 02:38:18,24.6108
2020-08-28 02:38:19,24.6107
2020-08-28 02:38:20,24.6110
2020-08-28 02:38:21,24.6111
2020-08-28 02:38:22,24.6107
2020-08-28 02:38:23,24.6109
2020-08-28 02:38:24,24.6109
2020-08-28 02:38:25,24.6108
2020-08-28 02:38:26,24.6109
2020-08-28 02:38:27,24.6110
2020-08-28 02:38:28,24.6110
2020-08-28 02:38:29,24.6108
2020-08-28 02:38:30,24.6109
2020-08-28 02:38:31,24.6108
2020-08-28 02:38:32,24.6109
2020-08-28 02:38:33,24.6111
2020-08-28 02:38:34,24.6109
2020-08-28 02:38:35,24.6109
2020-08-28 02:38:36,24.6109
2020-08-28 02:38:37,24.6107
2020-08-28 02:38:38,24.6110
2020-08-28 02:38:39,24.6107
2020-08-28 02:38:40,24.6108
2020-08-28 02:38:41,24.6108
2020-08-28 02:38:42,24.6110
2020-08-28 02:38:43,24.6108
2020-08-28 02:38:44,24.6110
2020-08-28 02:38:45,24.6110
2020-08-28 02:38:46,24.6109
2020-08-28 02:38:47,24.6108
2020-08-28 02:38:48,24.6109
2020-08-28 02:38:49,24.6106
2020-08-28 02:38:50,24.6108
2020-08-28 02:38:51,24.6110
2020-08-28 02:38:52,24.6110
2020-08-28 02:38:53,24.6111
2020-08-28 02:38:54,24.6109
2020-08-28 02:38:55,24.6107
2020-08-28 02:38:56,24.6108
2020-08-28 02:38:57,24.6109
```

インテル Xeon PCIe Gen.3 server

/p2pBandwidthLatencyTest

```
3 1 1 1 1
Unidirectional P2P=Disabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 991.43 6.78 7.10 7.09
1 6.75 1017.25 7.07 7.07
2 7.02 7.01 999.04 7.09
3 6.97 6.96 7.03 1038.90
Unidirectional P2P=Enabled Bandwidth (P2P Writes) Matrix (GB/s)
D\D 0 1 2 3
0 1061.48 13.97 15.39 14.51
1 15.33 1027.96 15.38 15.60
2 15.33 13.97 1073.15 15.61
3 15.33 13.97 15.39 1082.06
Bidirectional P2P=Disabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 1124.10 9.64 9.63 9.65
1 9.42 1094.19 9.78 9.65
2 9.34 9.44 1133.89 9.64
3 9.41 9.49 9.47 1209.37
Bidirectional P2P=Enabled Bandwidth Matrix (GB/s)
D\D 0 1 2 3
0 1236.16 38.89 48.47 40.97
1 38.84 1194.57 38.83 38.75
2 48.48 38.88 1218.80 41.02
3 42.06 43.15 48.59 1226.45
P2P=Disabled Latency Matrix (us)
GPU 0 1 2 3
0 2.65 24.60 22.74 23.30
1 24.60 2.67 24.59 24.79
2 24.16 24.54 2.39 22.05
3 24.50 24.50 22.51 2.29
CPU 0 1 2 3
0 3.11 8.38 8.31 8.37
1 8.46 3.02 8.25 8.67
2 8.36 8.16 3.05 8.23
3 8.38 8.27 8.24 3.06
P2P=Enabled Latency (P2P Writes) Matrix (us)
GPU 0 1 2 3
0 2.65 2.45 2.39 2.40
1 2.64 2.69 2.71 2.62
2 2.17 2.16 2.40 2.16
3 2.17 2.20 2.19 2.29
CPU 0 1 2 3
0 3.14 2.45 2.55 2.51
1 2.64 3.23 2.71 2.57
2 2.64 2.66 3.28 2.57
3 2.84 2.58 2.56 3.19
```

htod

```
root@ubuntu100:~/bandwidth# ../bandwidth -i 0 htod
Host ---> G0
Time,G0 [Graphics Device]
2020-08-28 12:09:45,12.3620
2020-08-28 12:09:46,12.3592
2020-08-28 12:09:47,12.3613
2020-08-28 12:09:48,12.3611
2020-08-28 12:09:49,12.3596
2020-08-28 12:09:50,12.3615
2020-08-28 12:09:51,12.3621
2020-08-28 12:09:52,12.3605
2020-08-28 12:09:53,12.3615
2020-08-28 12:09:54,12.3611
2020-08-28 12:09:55,12.3604
2020-08-28 12:09:56,12.3614
2020-08-28 12:09:57,12.3607
2020-08-28 12:09:58,12.3608
2020-08-28 12:09:59,12.3621
2020-08-28 12:10:00,12.3602
2020-08-28 12:10:01,12.3587
2020-08-28 12:10:02,12.3592
2020-08-28 12:10:03,12.3614
2020-08-28 12:10:04,12.3608
2020-08-28 12:10:05,12.3588
2020-08-28 12:10:06,12.3599
2020-08-28 12:10:07,12.3618
2020-08-28 12:10:08,12.3613
2020-08-28 12:10:09,12.3605
2020-08-28 12:10:10,12.3607
^Croot@ubuntu100:~/bandwidth#
root@ubuntu100:~/bandwidth# ../bandwidth -i 0 dtoh
```

株式会社 エルザ ジャパン

〒105-0014 東京都港区芝 3-42-10 三田 UT ビル TEL: 03-5765-7391(代表) / FAX: 03-5765-7235

ELSA(エルザ)はテクノロジー・ジョイント株式会社の登録商標です。NVIDIA、Quadro、GeForce は NVIDIA Corporation の商標です。

その他商品名は各社の商標または登録商標です。仕様などは予告無しに変更されることがあります。(141020)

<https://www.elsa-jp.co.jp>

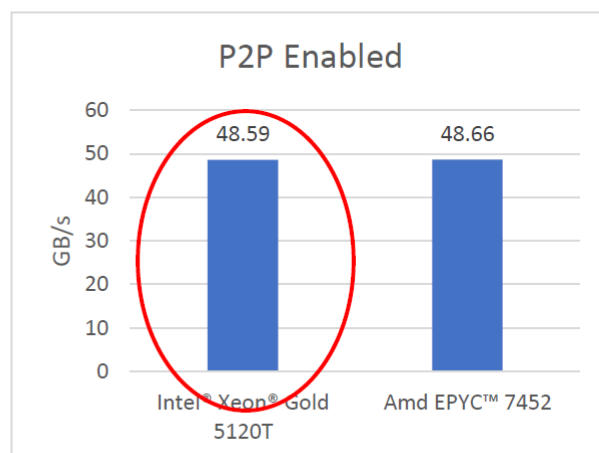
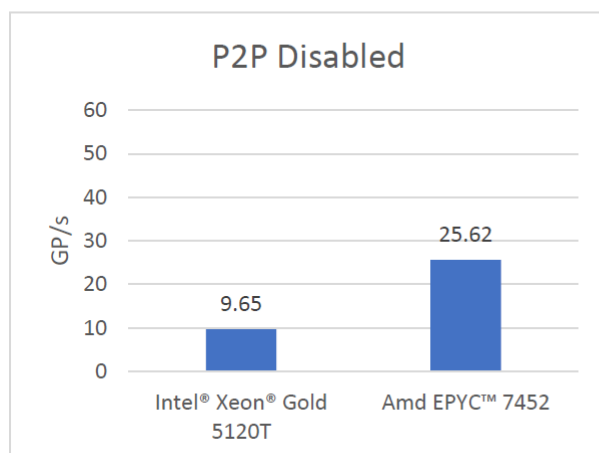
Dtoh

```
root@ubuntu100:~/bandwidth# ./bandwidth -i 0 dtch
Host <--- G0
Time,G0 [Graphics Device]
2020-08-28 12:10:30,12.5514
2020-08-28 12:10:31,12.5511
2020-08-28 12:10:32,12.5510
2020-08-28 12:10:33,12.5524
2020-08-28 12:10:34,12.5522
2020-08-28 12:10:35,12.5519
2020-08-28 12:10:36,12.5525
2020-08-28 12:10:37,12.5519
2020-08-28 12:10:38,12.5523
2020-08-28 12:10:39,12.5525
2020-08-28 12:10:40,12.5525
2020-08-28 12:10:41,12.5526
2020-08-28 12:10:42,12.5520
```

まとめ

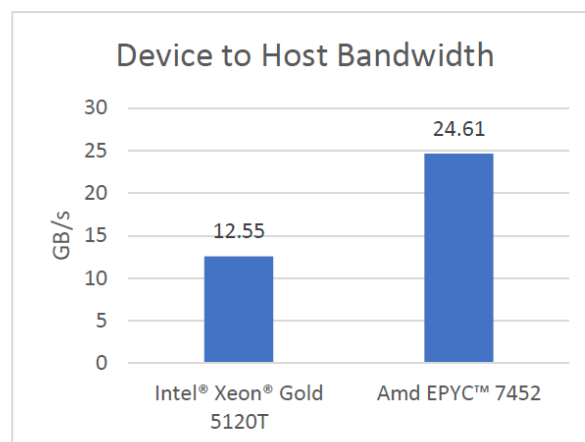
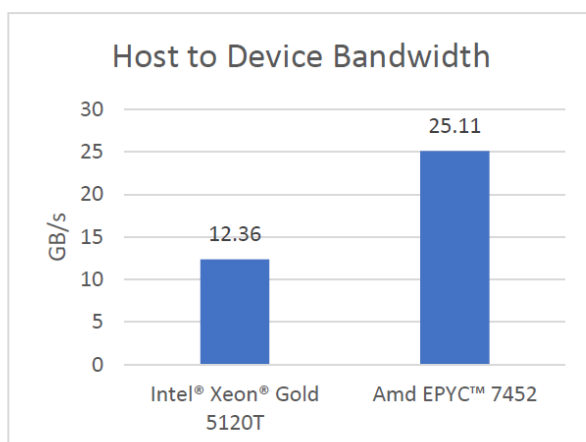
ホスト対デバイス、デバイス対ホストの最大帯域幅パフォーマンス

Maximum P2P bandwidth performance (bidirectional)



Copyright© 2020 H3 Platform Inc. All rights reserved. The information is subject to change without notice

Maximum host to device, device to host bandwidth performance



Copyright© 2020 H3 Platform Inc. All rights reserved. The information is subject to change without notice

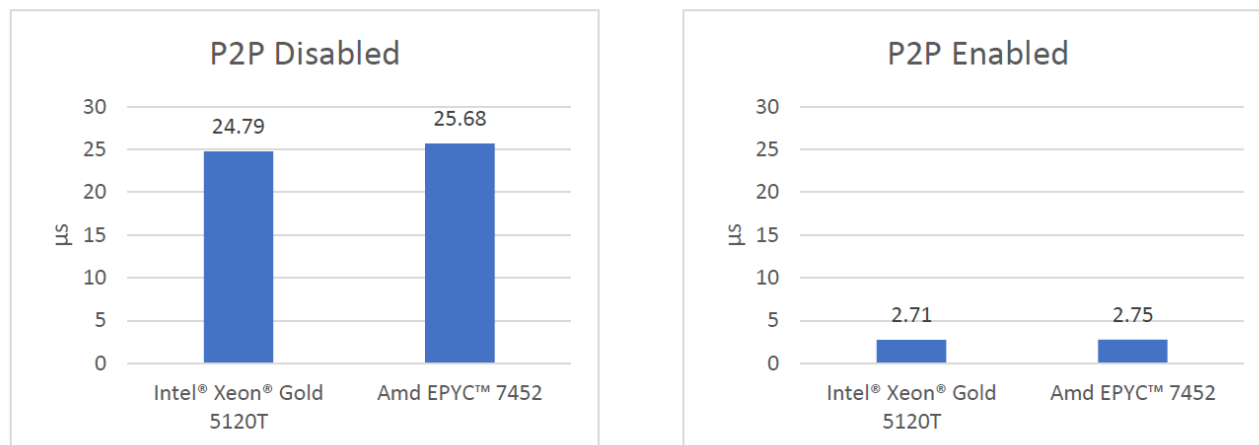
株式会社 エルザ ジャパン

〒105-0014 東京都港区芝 3-42-10 三田 UT ビル TEL: 03-5765-7391(代表) / FAX: 03-5765-7235

ELSA(エルザ)はテクノロジー・ジョイント株式会社の登録商標です。NVIDIA、Quadro、GeForce は NVIDIA Corporation の商標です。
 その他商品名は各社の商標または登録商標です。仕様などは予告無しに変更されることがあります。(141020)

<https://www.elsa-jp.co.jp>

Maximum P2P latency



Copyright© 2020 H3 Platform Inc. All rights reserved. The information is subject to change without notice

インテル Xeon Gold 5120T PCIe Gen.3 server.

ホストからデバイスへの帯域幅のピークはそれぞれ 12.36GB/s と 12.55GB/s で、双方向 p2p 無効時と有効時の帯域幅のピークはそれぞれ 9.65GB/s と 48.59GB/s でした。p2p を無効にした場合の GPU ピア間レイテンシは 24.79 μ s、有効にした場合の GPU ピア間レイテンシは 2.71 μ s とピークに達しています。

AMD EPYC 7452 PCIe Gen.4 server.

ホストからデバイスへの帯域幅とデバイスからホストへの帯域幅のピークはそれぞれ 25.11GB/s と 24.61GB/s で、双方向 p2p の無効化と有効化の帯域幅のピークはそれぞれ 25.62GB/s と 48.66GB/s でした。p2p を無効にした場合の GPU ピア間レイテンシは 25.68 μ s、有効にした場合の GPU ピア間レイテンシは 2.75 μ s とピークに達しています。

我々は、Falcon 4010 GPU による拡張がホストからデバイスへの帯域幅パフォーマンスや GPU ピアツーピアのレイテンシに影響を与えないことを発見しました。しかし、Falcon 4010 GPU プラットフォームは、ホストサーバ自体の PCIe バージョンにもかかわらず、PCIe Gen.4 標準の GPU ピアツーピアの帯域幅を可能にします。