

ExpEther ボード (40G)

ExpEther I/O 拡張ユニット (40G)

動作検証レポート

検証デバイス : Intel DC P3700 (NVMe SSD)

目次

1 はじめに	3
2 検証対象の I/O デバイス	3
3 検証環境（サーバ装置・OS・周辺機器）	3
4 接続構成	4
5 検証内容	5
6 検証結果	6
7 問い合わせ先	10

商標について

Microsoft Windows7、Windows Server、Visual Studio は、米国 Microsoft Corporation の米国およびその他の国における登録商標または商標です。

Intel、Pentium、Xeon は米国 Intel Corporation の登録商標です。

PCI EXPRESS は Peripheral Component Interconnect Special Interest Group の商標です。

Linux®は、Linus Torvalds 氏の日本およびその他の国における商標または登録商標です。

Red Hat®, Red Hat Enterprise Linux は、米国 Red Hat, Inc.の米国およびその他の国における商標または登録商標です。

その他記載の会社名および製品名は各社の商標もしくは登録商標です。

1 はじめに

日本電気製 ExpEther ボード (40G) / ExpEther I/O 拡張ユニット (40G) に IHV ベンダのデバイスを搭載して動作検証を行った。本レポートでは、検証構成・検証結果を報告するものである。

2 検証対象の I/O デバイス

ベンダ	インテル (Intel Corporation)
品名	Intel® SSD DC P3700 Series (400GB, 1/2 Height PCIe 3.0 x4, 20nm, MLC) MM# 933088
型番	SSDPEDMD400G401
デバイス種別	NVMe SSD
仕様	PCIe Gen3 / x4
製品ページ (URL)	https://www-ssl.intel.com/content/www/us/en/solid-state-drives/solid-state-drives-dc-p3700-series.html
ファームウェア・バージョン	8DV101F0

3 検証環境 (サーバ装置・OS・周辺機器)

本検証で使用したハードウェア・オペレーティングシステムを以下に示す。

装置	名称	型番
サーバ装置	Express5800/R120g-2M	
CPU	Intel® Xeon® Processor E5-2699 v4	
メモリ	128 GB (32GB DDR4 DIMM x 4)	
NVMe SSD	Intel® SSD DC P3700 Series (400GB)	4 ドライブ使用
ExpEther ボード	ExpEther ボード(40G)	N8104-165、2 枚
ExpEther I/O 拡張ユニット	ExpEther I/O 拡張ユニット(40G)	N8141-81、1 台
40Gb イーサネットスイッチ	Mellanox 社 MSX1012B-2BFS	12 ポートスイッチ
40Gb イーサネットケーブル	Mellanox 社 MC2210310-003	光, 3m ケーブル
オペレーティングシステム	Red Hat Enterprise Linux 7.2	
カーネル・バージョン	3.10.0-327-36.3.el7.x 86_64	
NVMe ドライバー・バージョン	1.0	

- ExpEther ボード（40G）は、サーバ装置の PCI スロット #1c / #2d に搭載。
- ExpEther ボード（40G）と ExpEther I/O 拡張ユニット(40G)は、40Gb イーサネットスイッチを介して 40Gb イーサネットケーブルで接続。

4 接続構成

以下の 3 通りの接続構成で性能検証を実施した。

構成 A : SSD 四枚をサーバの内蔵 PCIe スロットに搭載する。

構成 B : サーバに ExpEther ボード（40G）を 1 枚搭載し、40G スイッチ・ケーブルを介して ExpEther I/O 拡張ユニット（40G）に接続する。I/O 拡張ユニットに SSD を 4 枚搭載する。



図 1、検証接続構成図（構成 B）

構成 C : サーバに ExpEther ボード（40G）を 2 枚搭載し、40G スイッチ・ケーブルを介して ExpEther I/O 拡張ユニット（40G）に接続する。I/O 拡張ユニットに SSD 四枚を搭載し、SSD は各 2 枚ずつを ExpEther ボードへ接続する（Group ID #1/#2 の指定によりグループを分ける）。



図 2、検証接続構成図（構成 C）

構成A－Cを下表にまとめる。

表 1 検証装置の構成一覧

構成	サーバ	OS	kernel	CPU	メモリ	nvme driver	NVMe SSD 搭載位置	EE ボード 搭載スロット	SSD 搭載 スロット
A	Express5800 R120g-2M	RHEL7.2	3.10.0-327- 36.3.el7. x 86_64	E5-2699 v4	32G x4	1.0	サーバ内蔵	-	#1c,#1d, #2c,# 2d
B	同上	同上	同上	同上	同上	同上	ExpEther I/O 拡張ユ ニット	slot#1c	#1,#2,#3,#4
C	同上	同上	同上	同上	同上	同上	ExpEther I/O 拡張ユ ニット	1 枚目 : slot#1c 2 枚目 : slot#2d	#1,#2,#3,#4

5 検証内容

構成 A～C において、ディスク装置の性能測定ツールである “fio” を用いて性能測定を実施した。

使用ツール : fio <https://github.com/axboe/fio>

バージョン : 2.2.8

性能測定で用いたパラメータ :

IOPS 試験

パラメータ	値
Variation	randread, randwrite
Block Size	4k
num jobs	4
queue depth	1, 4, 16, 64

Bandwidth 試験

パラメータ	値
Variation	randread, randwrite
Block Size	4096k
num jobs	4
queue depth	1, 4, 16, 64

採取データ処理方法：

- ・各パラメータにつき fio を 5 回実施した。
- ・IOPS と Bandwidth について最大最小を除いた 3 回の実行結果の平均を取得した。
- ・SSD 複数枚に対して同時実行を行った場合、取得した平均値の和を結果として取得した。

6 検証結果

項番	確認項目	結果
1	ExpEther ボード(40G)－ExpEther I/O 拡張ユニットの接続	問題なし
2	I/O デバイスの OS からの認識	問題なし
3	I/O デバイスに対する試験 (fio)	問題なし

性能測定結果を次ページ以降に以下に示す。

範例：

local：

サーバ本体装置内に SSD を搭載した構成を示す（構成 A）。

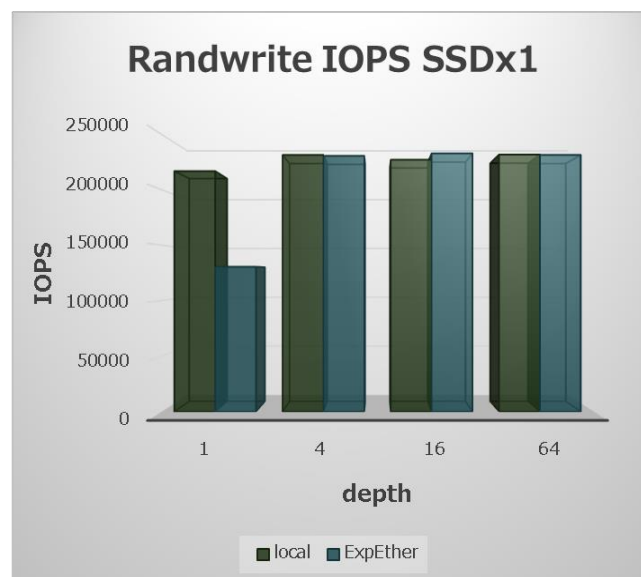
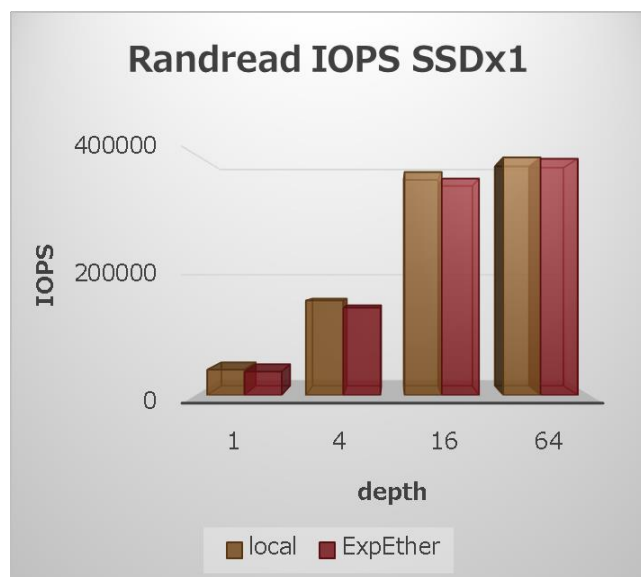
ExpEther、及び ExpEther (HBA1)：

サーバに ExpEther ボード（40G）を 1 枚搭載し、ExpEther I/O 拡張ユニット（40G）に SSD を搭載した構成を示す。（構成 B）

ExpEther (HBA2)：

サーバに ExpEther ボード（40G）を 2 枚搭載し、ExpEther I/O 拡張ユニット（40G）に SSD を搭載した構成を示す（構成 C）。

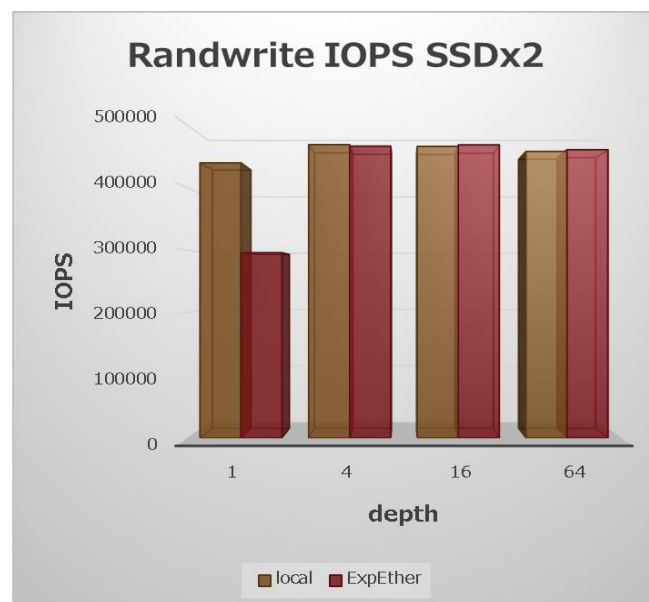
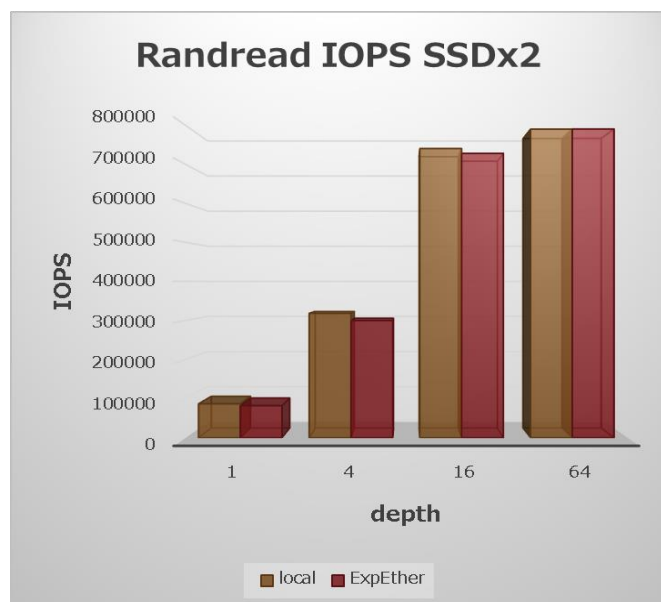
IOPS (SSD x 1 枚) :



name/depth	1	4	16	64
local	42467.67	157166.67	370562	394857.33
ExpEther	39113.67	145294.33	359693.67	392591
ExpEther/local (%)	92.10	92.45	97.07	99.43

name/depth	1	4	16	64
local	216616.67	231208.67	226872.67	231524
ExpEther	130586.67	230349.33	232602.67	231147.33
ExpEther/local (%)	60.28	99.63	102.53	99.84

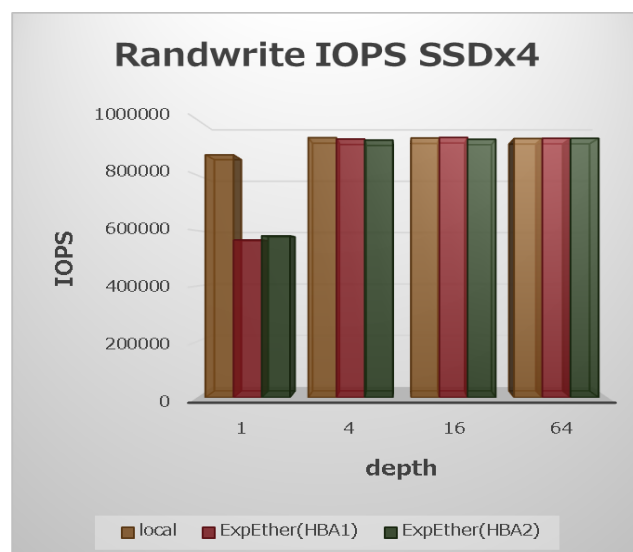
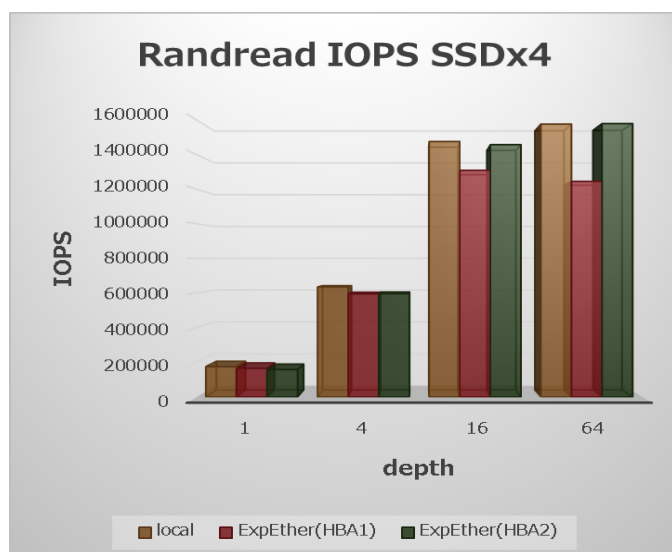
IOPS (SSD x 2 枚) :



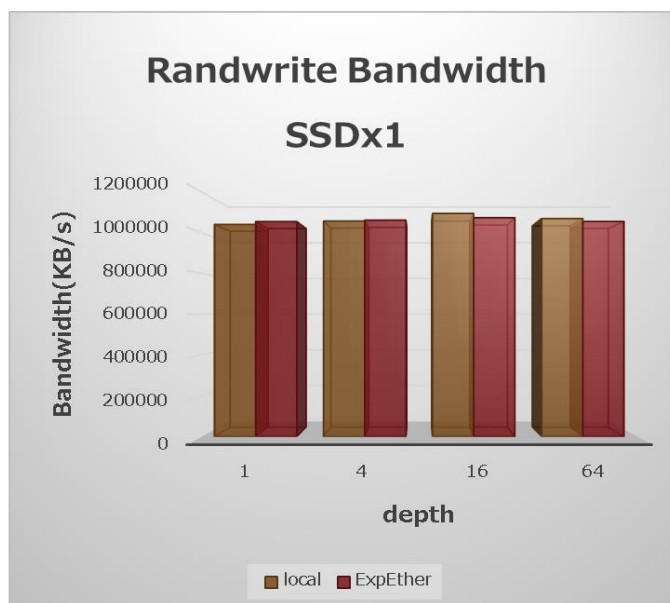
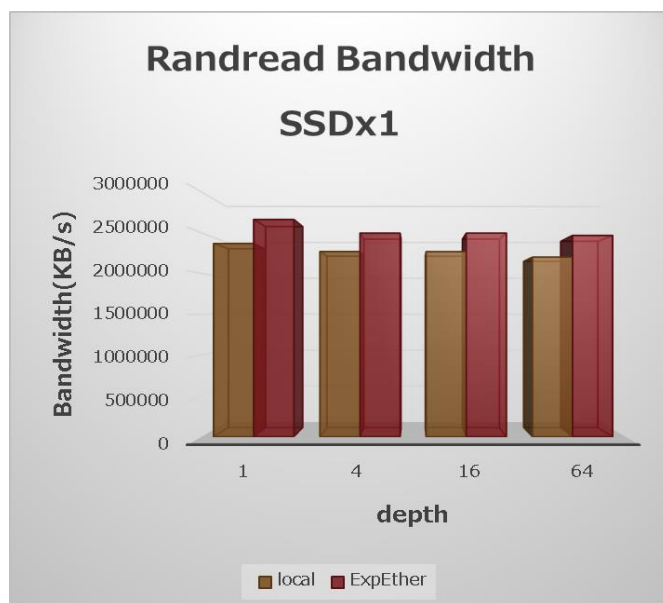
name/depth	1	4	16	64
local	87012.67	318238.66	739679	789092.67
ExpEther	82291	299715.67	727260.66	789575.33
ExpEther/local (%)	94.57	94.18	98.32	100.06

name/depth	1	4	16	64
local	439174.67	468321.33	465533	457413
ExpEther	295878.33	465887	467894.33	460132
ExpEther/local (%)	67.37	99.48	100.51	100.59

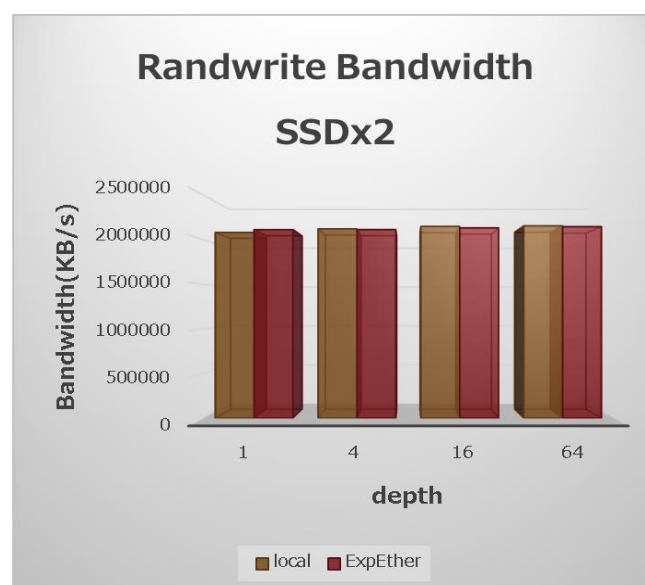
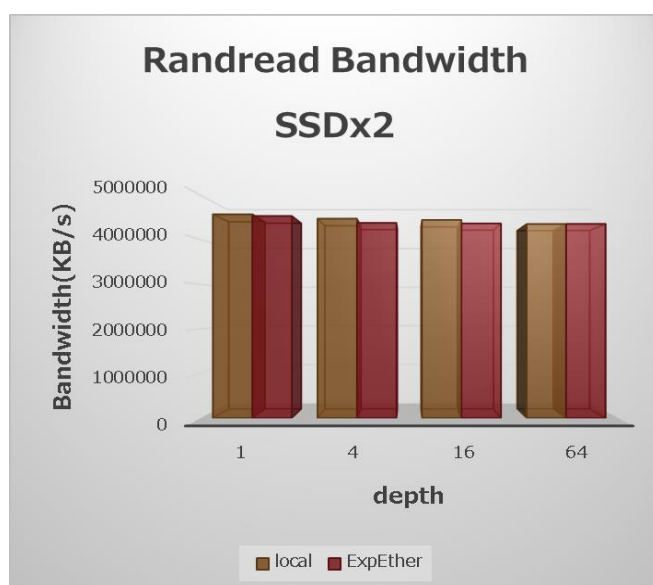
IOPS (SSD x 4 枚) :



Bandwidth (SSD x 1 枚) :



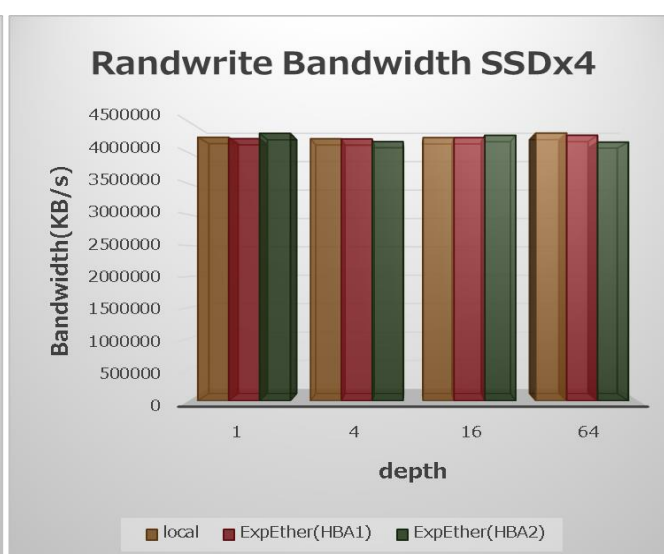
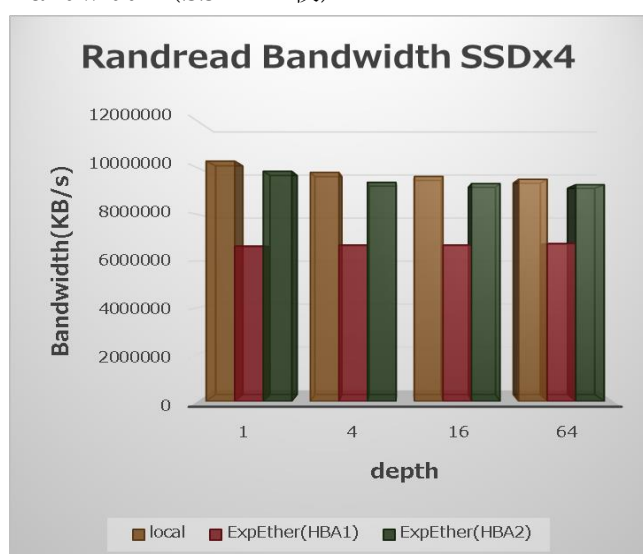
Bandwidth (SSD x 2 枚) :



name/depth	1	4	16	64
local	4566016	4468292.27	4440405.33	4348416
ExpEther	4528332.8	4373606.4	4361898.67	4351317.33
ExpEther/local (%)	99.17	97.88	98.23	100.07

name/depth	1	4	16	64
local	2078788.27	2119338.67	2146508.8	2152925.87
ExpEther	2110027.09	2110688.8	2129568.43	2143368.54
ExpEther/local (%)	101.50	99.59	99.21	99.56

Bandwidth (SSD x 4 枚) :



	1	4	16	64
local	10275089.06	9794184.52	9619694.93	9497088
ExpEther(HBA1)	6645691.73	6691191.46	6687778.14	6753689.59
ExpEther(HBA2)	9841595.73	9366391.47	9314201.59	9267063.46
EE(HBA1)/local (%)	64.68	68.32	69.52	71.11
EE(HBA2)/local (%)	95.78	95.63	96.82	97.58

	1	4	16	64
local	4242322.78	4219050.66	4236697.6	4309674.67
ExpEther(HBA1)	4221330.78	4216360.95	4236670.3	4274940.59
ExpEther(HBA2)	4306705.07	4173385.17	4274336.43	4165072.58
EE(HBA1)/local (%)	99.51	99.94	100.00	99.19
EE(HBA2)/local (%)	101.52	98.92	100.89	96.64

検証結果：

- ・ SSD 1 枚構成、及び 2 枚構成において、サーバ本体に内蔵した構成と、ExpEther I/O 拡張ユニット（40G）に搭載した構成を比較すると、ExpEther I/O 拡張ユニット（40G）の搭載において 95%以上の性能を発揮することが可能であることが検証された。
- ・ SSD 4 枚構成における Random Read において、ExpEther ボード(40G) 1 枚では内蔵構成に対する性能差が見られた。DC P3700 の性能はカタログ Spec から 3GB/s が上限と推定され、一方 ExpEther40G の Spec は Gen3/x8 (64Gb/s=8G/s) のため、ExpEther ボード(40G) 1 枚に DC P3700 を 2 枚まで接続することが適当と判断する。SSD が 2 枚を超える場合は、ExpEther ボード(40G)を 1 枚追加して、SSD を 2 枚に分散して使用することで SSD の性能を発揮できる。
- ・ ExpEther ボード（40G）の搭載枚数を追加することにより、より帯域を求められるシステム・アプリケーションにおいてもサーバ本体内蔵構成と同等の性能を発揮することが可能であると判断する。
- ・ SSD 1 枚構成の Randwrite 結果において、スペックシート上の値 17.5 万 IOPS に対して、20 万 IOPS を超える測定結果がでているが、書込キャッシュによるものが要因であると考える。

7 問い合わせ先

問い合わせ ML：ee-desk@ioth.jp.nec.com

改版履歴

版数	改版日	内容
1	2017 年 5 月	初版作成

[EOF]