

Power コラム 「NVMe」市場にでて1年、その評価は？

2022年1月

株式会社イグアス/J B C C株式会社

NVMeが市場に投入されてから1年

JBCC（西日本地区）で調べたところ 約半数のPowerユーザーがSSDに代わりNVMeを選択されていました。

ここではNVMe選択の理由と実際にお客様に導入しての評価など次の内容をご紹介します。

- ・ NVMe概要・IBM i NVMeサポート
- ・ NVMeのメリット（選択の理由）
 - ・ 可用性
 - ・ 価格
 - ・ パフォーマンス（実際に導入しての評価含む）
- ・ よくある質問：書込み限界値について
- ・ Namespaceにつて



POWER9 G モデルでの NVMe サポートの拡張

高速、大容量を実現、G モデル IBM i 内蔵ディスクに最適

エンタープライズ NVMe ドライブ

- SAS SSD と比較して、書き込み、読み取り IOPS が大幅に向上

G モデルの NVMe サポート拡張

- フロントアクセス可能で保守が容易な PCIe4 U.2 ドライブをサポート
- PCIe NVMe アダプターの導入可能数が増加
- G モデルでサポートされる NVMe ドライブ数
 - S914 4コア：最大 4 ドライブ
 - S914 6,8コア：最大 11 ドライブ
 - S922/S924：最大 14 ドライブ

IBM i での NVMe ドライブ最大実効容量

- S914 4コア

➡ **3.2TB**

- S914 6, 8コア

➡ **35.2TB**

- S922, S924

➡ **44.8TB**

IBM i NVMe サポート

- IBM i ネイティブでサポートされる NVMe デバイス

	PCIe Gen4 アダプター	PCIe Gen4 U.2 デバイス
容量	1.6, 3.2, 6.4 TB	0.8, 1.6, 3.2, 6.4 TB
サポートモデル	9009-41G*, 42G, 22G 9080-M9S	9009-41G, 42G, 22G
OS サポート	IBM i 7.4 TR3 以降	IBM i 7.4 TR2 以降 (0.8TB : IBM i 7.4 TR4 以降)
構成位置	本体 (CEC) PCIe スロット	本体 (CEC) ディスク・バック プレーン
ディスク保護	OS ミラー	OS ミラー
書き込み限界	3 DWPD	3 DWPD
コンカレント・メ タデータ	可能	可能
書き込み限界での 交換	有償で交換	有償で交換

* 9009-41G タワー型では、PCIe Gen4 NVMe アダプターの構成は不可

出展：「IBM i NVMe サポート」
2021年4月
日本アイ・ビー・エム株式会社
テクノロジー事業本部
IBM Power Systems 事業部

IBM i NVMe サポート

- サポートされる NVMe デバイス数

	PCIe Gen3, Gen4 アダプター (1.6, 3.2, 6.4 TB)	PCIe Gen4 U.2 デバイス (0.8, 1.6, 3.2, 6.4 TB)
9009-41G 4 コア	4 (0.8, 1.6, 3.2 TB) *	
9009-41G 6, 8 コア	7	4
9009-42G 1 ソケット	7	4
9009-42G 2 ソケット	10	4
9009-22G 1 コア	4 (0.8, 1.6 TB) **	

* 9009-41G 4 コアは異なるタイプのデバイスの混在不可、最大容量 6.4 TB

**9009-22G 1 コアは異なるタイプのデバイスの混在不可、最大容量 3.2 TB

出展：「IBM i NVMe サポート」
2021年4月
日本アイ・ビー・エム株式会社
テクノロジー事業本部
IBM Power Systems 事業部

IBM i での NVMe デバイスの構成

- U.2, PCIe アダプター毎に、同じ容量のペアを構成する
 - デバイス・タイプが異なるペアの構成は不可
- ミラー保護 (OS ミラーリング) が必要
- LPAR はミラーペアのデバイス単位で構成
- ASP バランス
 - NVMe と SSD は同一階層として扱われる
 - NVMe と SSD 間では、オブジェクト配置の最適化は実行されない
- IBM i プロセッサ・グループ (9009-41G 4core, 9009-22G 1core) の NVMe デバイス構成
 - 異なるタイプのデバイスの混在は不可
 - 構成可能な最大容量
 - 9009-41G 4core : 6.4TB
 - 9009-22G 1core : 3.2TB

NVMeのメリット：可用性が高い

- そもそもNVMe接続SSDはHDDに比べ、駆動機構(ヘッド、モーター、サスペンションなど)がなく、フラッシュメモリーのため、障害発生率は低いと言えます。

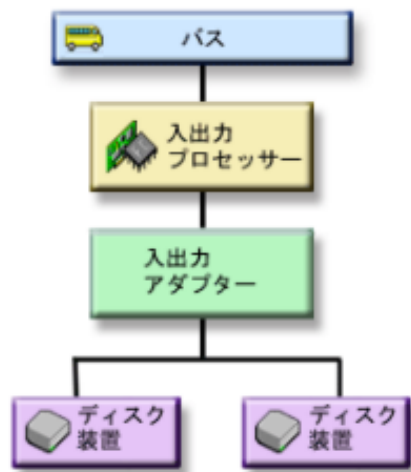
HDDやSSDは別途ディスクコントローラーが必要となります。

NVMeは、デバイス自身でディスク領域とディスクコントローラーが一体になってます。

例えば、Power S914 4coreモデルで比較しますと、HDDやSSDは本体バックプレーンのディスクコントローラーで制御のため、“ディスク装置レベルの保護(図1)”となります。

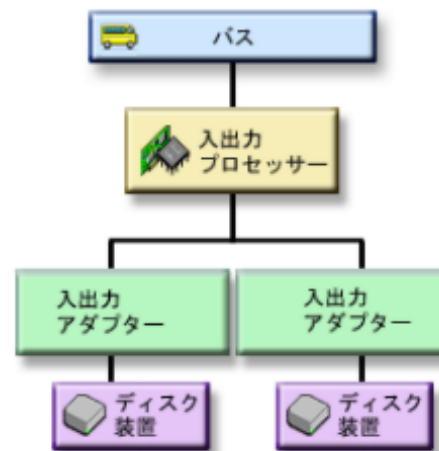
一方、NVMeはデバイス毎にディスクコントローラーがありますので、“IOAレベルの保護(図2)”となります。

図1. ディスク装置レベルの保護



<https://www.ibm.com/docs/ja/i/7.4?topic=want-disk-unit-level-protection>

図2. IOAレベルの保護



<https://www.ibm.com/docs/ja/i/7.4?topic=want-ioa-level-protection>

NVMeのメリット：価格

- SSDに比べ容量単価が安価になります。

SSDは保証満了後、別途保守料が発生するためランニングコストも高価になります。

また、構成にもよりますが、**HDDよりも容量単価が安価になります。**

以下、Power S914 4coreモデルでのディスクと制御機構のみの価格(定価)比較となります。

ディスクタイプ	製品名	本数	保護レベル	総容量	実行容量	単価	価格	備考
HDD	283GB 15krpm SAS HDD	10	Mirror	2,830	1,415	¥131,300	¥1,313,000	
	Backplane(EJ1M)	1	-	-	-	¥860,800	¥860,800	
							¥2,173,800	
SSD(Ent)	387GB Enterprise SSD	8	Mirror	3,096	1,548	¥409,300	¥3,274,400	保証期間後別途保守料(Base) @33,600/年
	Backplane(EJ1M)	1	-	-	-	¥860,800	¥860,800	
							¥4,135,200	HDD構成との差額：¥1,961,400
SSD(Main)	931GB Mainstream SSD	4	Mirror	3,724	1,862	¥251,800	¥1,007,200	保証期間後別途保守料(Base) @19,200/年
	Backplane(EJ1M)	1	-	-	-	¥860,800	¥860,800	
							¥1,868,000	HDD構成との差額：¥305,800
NVMe(PCI)	1.6TB NVMe Flash Adapter	2	Mirror	3,200	1,600	¥545,800	¥1,091,600	
							¥1,091,600	HDD構成との差額：¥-1,082,200
NVMe(U.2)	1.6TB SSD PCIe4 NVMe U.2 module	2	Mirror	3,200	1,600	¥587,800	¥1,175,600	
	Backplane(EJ1T)	1	-	-	-	¥125,800	¥125,800	
							¥1,301,400	HDD構成との差額：¥-872,400



NVMeのメリット：パフォーマンス

IBM提供：NVMeと SSD（SAS）接続とのパフォーマンス比較

IBM i : NVMe, SAS SSD パフォーマンス比較

Protection	Parameter	Units	RAID-5 vs. Mirrored			Mirrored vs. Mirrored		
			SAS	NVMe	Difference	SAS	NVMe	Difference
Application	RAID Type		0,1,5,6,10	RAID 5	OS Mirror	RAID 10	OS Mirror	
Speed	Random 255B Reads	Trans/s	273,000	390,000	42.0%	276,185	387,284	40.2%
	Random 255B Writes	Trans/s	2,898,000	5,824,000	101.0%	6,674,419	5,785,859	-13.3%
	Sequential 32KB Reads	GB/s	3.51	8.22	134.2%	3.33	8.22	146.8%
	Sequential 32KB Writes	GB/s	1.01	2.46	143.6%	1.79	2.48	38.5%
Storage Subsystem	Random 4KB Reads	io/s	218,000	368,000	68.8%	217,629	385,410	77.1%
	Random 4KB Writes	io/s	167,000	690,000	313.2%	385,088	666,441	73.1%
	Sequential 32KB Reads	GB/s	3.48	8.05	131.3%	3.22	7.588	135.4%
	Sequential 32KB Writes	GB/s	1.00	4.56	356.0%	1.80	4.969	175.4%
	Random 4KB Reads	us/io	123	99	-19.5%	133	99	-25.6%
	Random 4KB Writes	us/io	59	24	-59.3%	34	28	-17.6%

• Conclusions: Lots of green for NVMe indicates consistent better performance
 • Random 255B Writes see a slower NVMe performance because the Mirror code is offloaded to the CPUs instead of the SAS Controllers

Color Key	Meaning
Green	Advantage NVMe
Yellow	Within +/-10%
Red	Advantage SAS

← 10個のテスト項目において赤字部分以外はすべてNVMeが上回っていることを示しています。

出展：「IBM i NVMe サポート」
2021年4月
日本アイ・ビー・エム株式会社
テクノロジー事業本部
IBM Power Systems 事業部

- J B C C 導入実績からお客様の声
 - バッチ処理が速くなった
 - オンライン系での速度が改善された
 などの反響があります。

• J B C C 導入実績から

相対的な性能イメージ (Read&Writeを平準化した値)

HDD : SSD : NVMe = 1 : 6 : 15



よくある質問：書込み限界値

NVMe、SSDには書込み限界値（DWPD）があります。

DWPD（Drive Write Day）は、ドライブの寿命が計算される際の基となる使用率です。計算方法は、毎日NVMeの全容量を書込み(Write)して5年間で寿命となります。

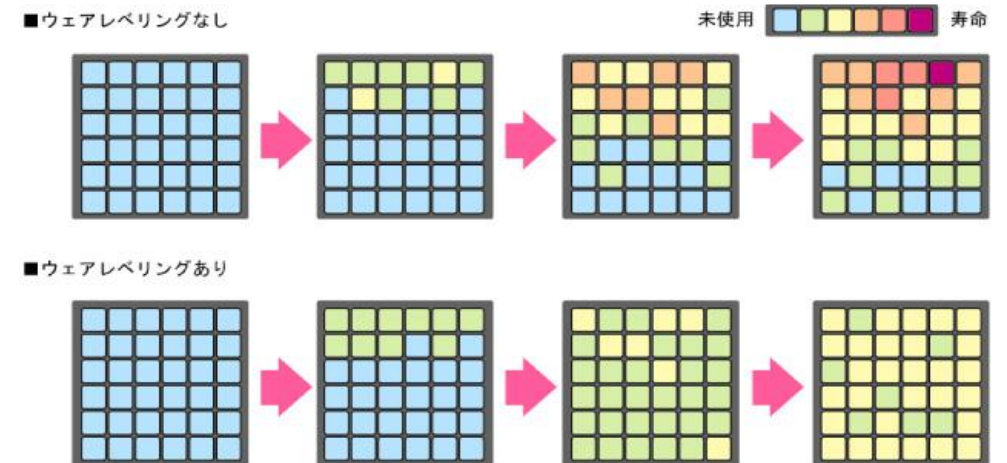
例えば、Power搭載の3.2TB U.2 NVMeを例とします。

U.2 NVMeは3DWPDの仕様となっており、毎日NVMeの全容量3.2TBを書込み(Write)して最低でも15年間の耐久性があることとなります。

また、ウェアレベリング（コンピュータの記憶媒体に用いられる技術の1つ）によって、寿命を延ばすテクノロジーが採用されています。

単純計算参考値	
1日に 1TB書き込み想定年数	47.3年
1日に 2TB書き込み想定年数	23.7年
1日に 3TB書き込み想定年数	15.8年

通常の利用方法で毎日全容量を書き込みすることはほぼないと言えますし、もしあったとしても最低でも15年はもつ計算となっています。NVMeでは書込み限界を心配する必要はないと言えます。（これは保証することなく計算上はそうなるということですよ）



各メモリセルの利用状況を記録・管理して、一部のセルに書き換えが集中しないようにセルを並べ替える技術が「ウェアレベリング」です。これによってセルの消去回数を均一化し、ストレージ全体の寿命を延ばします。

出展：ロジテック

NVMeの名前空間 (Namespace)

次ページ (8ページ) からは2021年4月にIBM社が提供している「**IBM i NVMe サポート**」から抜粋したNamespaceの設定の説明です。

NVMeの導入やNamespaceの設定など J B C C ではすでに実績がございますので NVMe導入検討のお客様はぜひご相談ください。

【お問合せ】

J B C C 株式会社 マーケティング
mail : mktg@jbcc.co.jp
URL : <https://www.jbcc.co.jp/>

IBM i NVMe サポート

2021年4月
日本アイ・ビー・エム株式会社
テクノロジー事業本部
IBM Power Systems 事業部

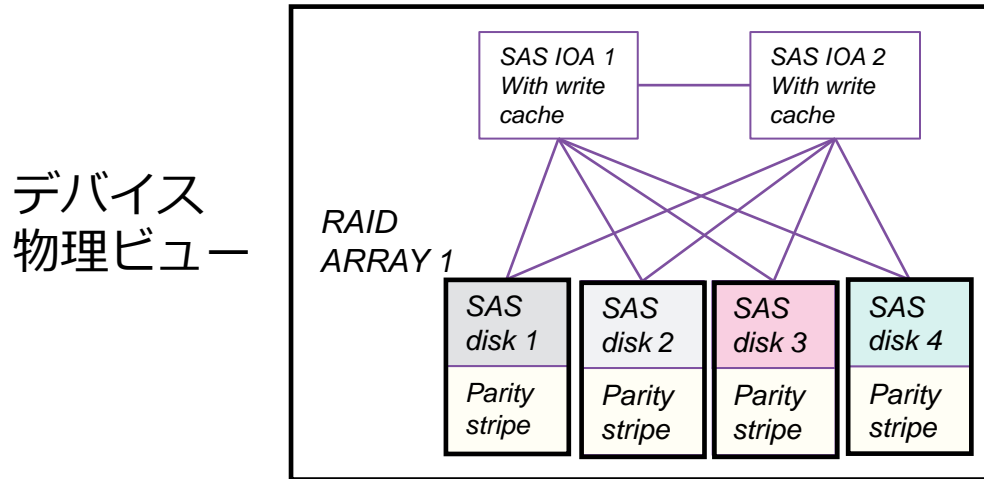
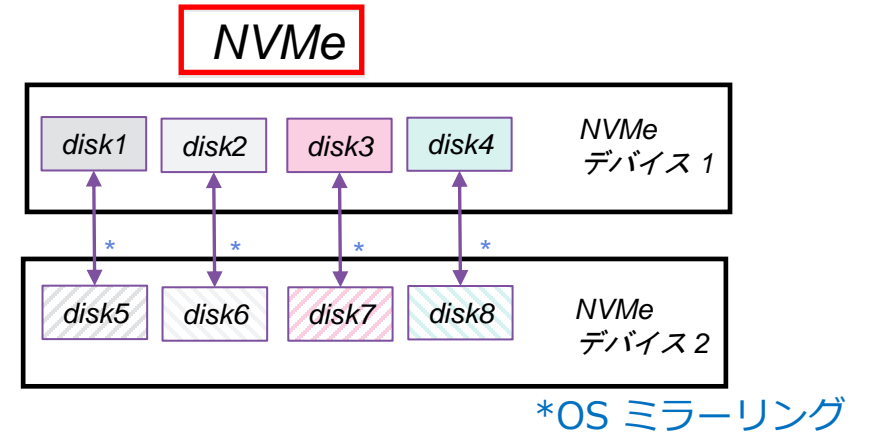
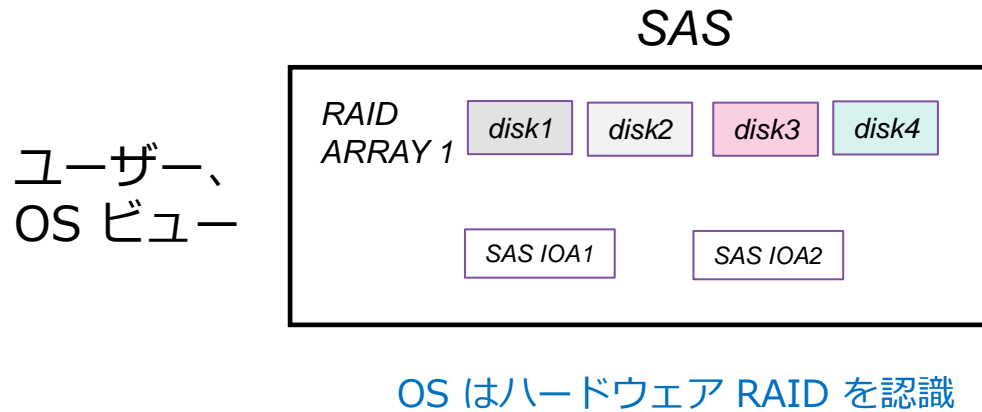


NVMe 名前空間 (Namespace)

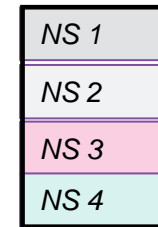
- NVMe 名前空間
 - NVMe デバイスの論理ブロック
 - 1つの NVMe デバイスに、最大 32 個の名前空間を作成可能
 - 名前空間のサイズは 64GB から 16TB
 - 現在提供されている NVMe デバイスの最大容量は 6.4TB のため、実際の名前空間最大サイズは 6.4TB
 - 既存の名前空間のサイズ変更は不可（削除して再作成する）
- IBM i は、名前空間を 1台のディスク・ドライブとして扱う
 - パフォーマンスを考慮して、複数の名前空間を構成することを推奨
- NVMe デバイスを IBM i ロードソースとしてプリロード構成とする場合
 - 工場出荷時にロードソース用の名前空間を作成するため、ロードソース用の名前空間サイズを指定
 - U.2 デバイス、PCIe Gen4 アダプターの場合
 - #ENSA : 200 GB IBM i NVMe Load Source Namespace size
 - #ENSB : 400 GB IBM i NVMe Load Source Namespace size
 - PCIe3 アダプターの場合
 - #ENS1 : 188 GB IBM i NVMe Load Source Namespace size
 - #ENS2 : 393 GB IBM i NVMe Load Source Namespace size

NVMe 名前空間

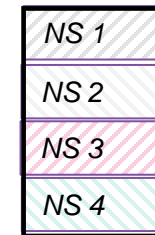
- NVMe 名前空間と SAS デバイスとの比較



NVMe デバイス 1

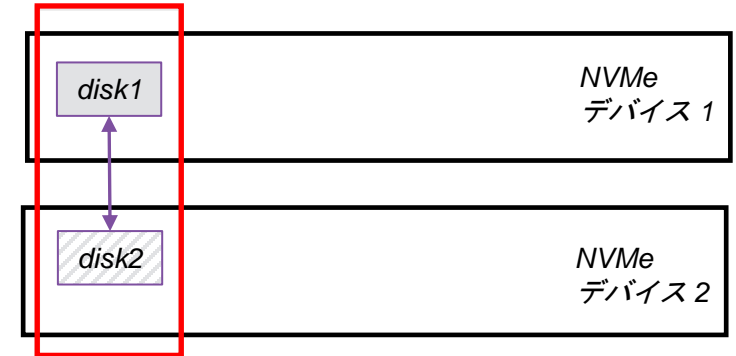


NVMe デバイス 2



工場出荷時の NVMe デバイスの名前空間

- NVMe デバイスをロードソースに指定して、IBM i をプリロードする場合
 - ロードソース・ディスク用の名前空間のみ作成される
 - ロードソース以外のディスク用の名前空間は、サービスツール (SST) で作成



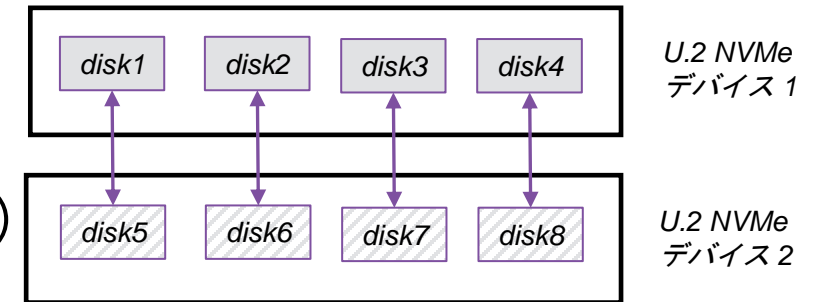
ロードソース・ディスク用
名前空間

- NVMe デバイスがロードソース指定されない場合
- IBM i プリロードが指定されない場合
 - 名前空間は作成されない
 - サービスツールで名前空間の作成が必要

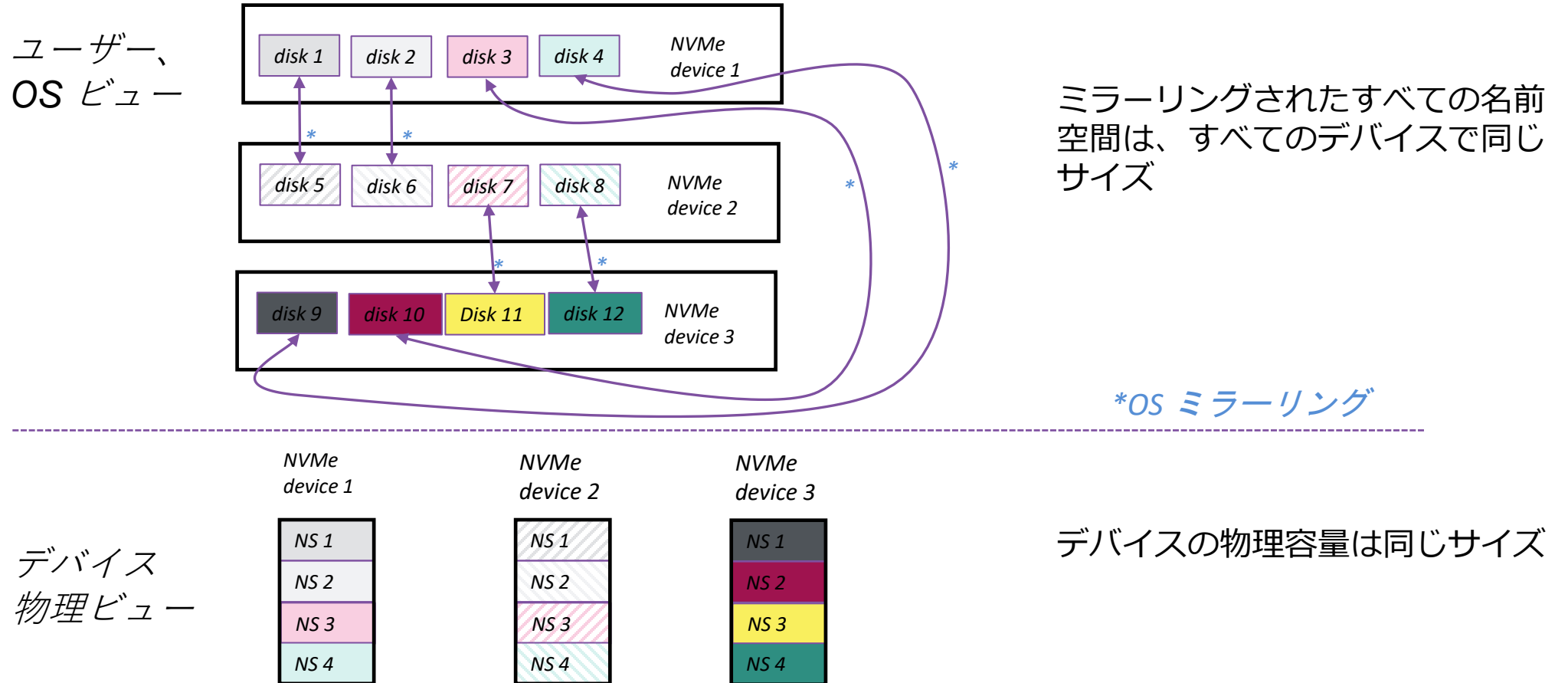


IBM i U.2 NVMe ドライブ Namespace (名前空間) 作成指定

- P05 マシンで IBM i ロードソースとして U.2 NVMe ドライブを構成する際に、名前空間が作成されて出荷されます
 - ハードウェア構成に #ENSM を含めます
- IBM i NVMe Load Source Namespace size の指定と同じサイズの名前空間を最大数作成します
 - #ENSA : 200GB, #ENSB : 400GB
- 名前空間を事前作成するための前提条件
 - 機械グループ P05 (9009-41G 4コア、9009-22G 1コア)
 - U.2 NVMe ドライブを IBM i ロードソースに指定
 - プライマリー OS - IBM i を指定 (#2145)
 - IBM i プリロードを指定
 - #5000 Software Preload
 - #0205 RISC to RISC data migration

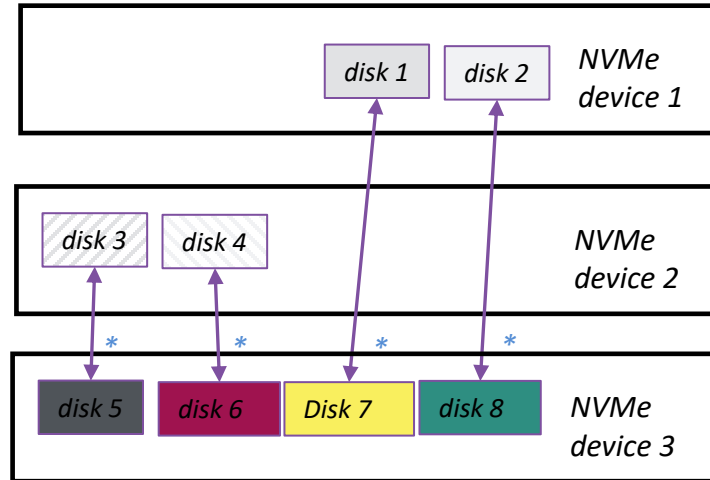


IBM i NVMe デバイスのミラーリング構成 (3 デバイス構成)



IBM i NVMe デバイスのミラーリング構成 (3 デバイス構成)

ユーザー、
OS ビュー

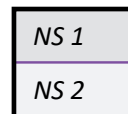


ミラーリングされたすべての名前空間は、すべてのデバイスで同じサイズ

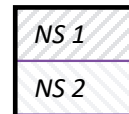
*OS ミラーリング

デバイス
物理ビュー

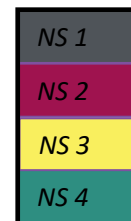
NVMe
device 1



NVMe
device 2



NVMe
device 3



デバイスの物理容量は同じサイズである必要はない

例 : device 1と device 2 は 1.6TB
device 3 は 3.2TB

IBM i : NVMe デバイス、名前空間の処理

- システム・サービス・ツール (SST) から名前空間の処理を行う
 - STRSST で SST を開始
 - Work with disk units を選択
 - Work with disk configuration を選択
 - Work with NVM Devices を選択

```
Work with Disk Configuration

Select one of the following:

1. Display disk configuration
2. Add units to ASPs
3. Work with ASP threshold
4. Add units to ASPs and balance data
5. Enable remote load source mirroring
6. Disable remote load source mirroring
7. Start compression on non-configured units
8. Work with device parity protection
9. Start hot spare
10. Stop hot spare
11. Work with encryption
12. Work with removing units from configuration
13. Work with NVM Devices

Selection
—
F3=Exit      F12=Cancel
```

IBM i : NVMe デバイス、名前空間の処理

- Work with NVM Devices 画面

```
Work with NVM Device

Select one of the following:

1. Display NVM namespaces
2. Display NVM devices
3. Create NVM namespaces
4. Create NVM namespaces to match another NVM device
5. Create NVM namespaces to pair with Active mirror protected NVM units
6. Delete existing NVM Namespaces
7. Sanitize/Erase NVM device
8. Format NVM device to prepare device for IBM i

Selection
-

F3=Exit          F12=Cancel
```

IBM i : NVMe デバイス、名前空間の処理

- Work with NVM Devices 画面
 1. Display NVM namespaces
 - IBM i 区画に存在するすべての NVMe デバイスと名前空間が表示されます
 2. Display NVM devices
 - IBM i 区画に存在するすべての NVMe デバイスが表示されます
 3. Create NVM namespaces
 - NVMe デバイスに名前空間を作成します
 4. Create NVM namespaces to match another NVM device
 - 別の NVMe デバイスの名前空間の容量と一致する名前空間を NVMe デバイスに作成します
 5. Create NVM namespaces to pair with Active mirror protected NVM units
 - NVMe デバイスに名前空間を作成し、アクティブなミラー保護された NVMe ユニットとペアリングします
 6. Delete existing NVM namespaces
 - NVMe 名前空間に関連付けられている IBM i ディスク装置を削除し、名前空間ストレージを NVMe デバイスの使用可能なプールに戻します
 7. Sanitize/Erase NVM device
 - NVMe メディアとキャッシュからすべてのデータを削除および破棄します
 8. Format NVM device to prepare device for IBM i
 - NVMe デバイスのブロックサイズを IBM i パーティションで使用できる形式に変更します

IBM i : NVMe デバイス、名前空間の作成

- Create NVM namespaces
 - 名前空間の数と容量を指定して作成

```
                Create NVM Namespaces

Device          Serial          Resource
Number          Number          Name          Type Logical Address
NVM             YOYACBYCB075    DC03          58FD U78D2.001.WZS0ET2-P1-C8

-----Capacity in GB-----    ---Namespaces---
NVM Configuration          Used Available          Total Used Avail Total
Current . . . . . :          0          3151          3151          0          32          32

Type choices, press Enter.
Quantity of namespaces to create . . . . . : 0 (1 - 32)
Capacity of each namespace . . . . . : 0 (64 - 3151) GB

F3=Exit          F5=Refresh          F12=Cancel

MA + A          12/050
```

IBM i : NVMe デバイス、名前空間の作成

- Create NVM namespaces to match another NVM device
 - 他の NVMe デバイスを指定して、そのデバイスの名前空間と同じ容量の名前空間を作成
- 作成手順
 - 名前空間を持つ ソース NVMe デバイスを選択します
 - 選択した NVMe デバイスの名前空間が表示されますので、Enter で続行します
 - 名前空間を作成するターゲット NVMe デバイスを選択します
 - 選択内容を確認する画面が表示されますので、F10 で名前空間を作成します

IBM i : NVMe デバイス、名前空間の作成

- Create NVM namespaces to pair with Active mirror protected NVM units
 - NVMe デバイスに名前空間を作成し、アクティブなミラー保護された NVMe ユニットとペアリングします
 - NVMe デバイスに障害が発生し、新しいデバイスに交換した後に、ミラー保護を復旧するために使用します
- 作成手順
 - 「Select NVM Device」画面で、交換された NVMe デバイスを選択します
 - 「Create NVM Namespaces - Show the Units Suspended from Mirror Protection」画面で、F7 を押してミラー保護が中断されたすべてのユニットを選択します
 - ミラー保護が中断されたディスク装置とペアになるリストされた名前空間の作成を確認します
 - F10 を押して操作を確認します。以下の手順が実行されます
 - 選択した NVMe デバイスに新しい名前空間を作成します
 - 新しい名前空間を、ミラー保護が中断された NVMe ディスクユニットとペアリングします
 - 新しく作成されたディスク装置でミラー保護の再開を開始します

この資料へのお問合せ

J B C C 株式会社 マーケティング

mail: mktg@jbcc.co.jp

URL : <https://www.jbcc.co.jp/>