

Université **IBM i**

7 novembre 2023

IBM Innovation Studio Paris

S15 – Comment implémenter le stockage interne NVMe dans l'environnement IBM i

14:45 / 15:45

Jean-Luc Bonhommet

IBM France

jeanluc_bonhommet@fr.ibm.com

 **infrasdufutur**

#ibmi

#uui2023

#infrastructuredufuturIBM23



Infrastructures du futur

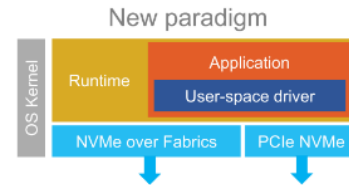
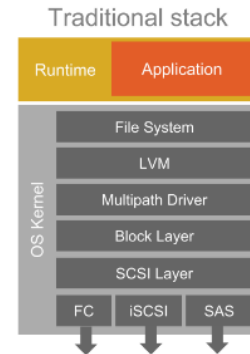


7 et 8 novembre 2023

Qu'est ce que NVMe? (Non-Volatile Memory Express)



- **NVMe est une interface ET un protocole de communication**
 - NVMe est conçu dès le départ pour offrir une bande passante élevée et un accès au stockage à faible latence
 - NVMe Express définit une interface efficace permettant au logiciel hôte de communiquer avec un sous-système de mémoire non volatile via PCI Express (NVMe sur PCIe)
 - Fonctionnellement analogue à SAS et SATA, mais conçu pour réduire les traitements supplémentaires des pilotes, du système d'exploitation et des applications
 - Exploite les complétions d'E / S basées sur l'interrogation par opposition aux complétions basées sur les interruptions
- **NVMe utilise la structure PCIe**
 - Plusieurs unités aujourd'hui sur le marché
 - Plusieurs formats, y compris des disques 2,5 pouces
- **NVMe a été conçu pour des performances élevées**
 - Augmentation des E/S, bande passante et latence plus faible
 - Exploite Flash et les mémoires non volatiles de nouvelle génération
 - Tire parti des environnements multi-cœurs, parallélisme d'E / S élevé



Composant NVMe



NVMe Controller

- PCIe Attached
- Parallel Architecture (Multi Q, Q pairs ...)
- Low Latency Design
- Fabrics Attach Friendly
- Self Encryption & Sanitize
- Virtualization (Multiple Namespace, SR-IOV)
- IO Determinism
- Zoned Namespace
- Management Interface Support (in & out of band)
- Computational Storage
- ... More Innovations



Media

- NAND TLC (most widely used)
- Optane 3DXP
- LL NAND
- NAND QLC
- DRAM – Flash backed
- MRAM
- .. More to come

Formats: M.2, U.2 (2.5” mince & épais), EDSFF (E1.S, E1.L, E3.S, E3.L mince & épais), Add in card

Infrastructure logicielle: investissement protégé par la réutilisation sur plusieurs appareils de fournisseurs

Caractéristiques NVMe et IBM i

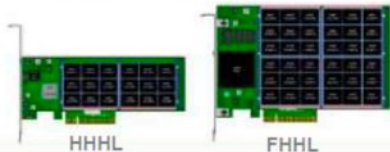


- NVMe est capable de fournir des performances plus élevées que les SSD. La technologie NVMe peut fournir beaucoup plus d'E/S en lecture ou en écriture et un débit nettement plus élevé (Go / s) par rapport aux SSD SAS / SATA. Les différences de performances réelles du système ou des applications varient en fonction du client et de la charge de travail.
- NVMe fournit des capacités de virtualisation supplémentaires puisque chaque périphérique est un point de terminaison PCIe qui peut être dédié à une partition / LPAR
- Au moins une paire d'adaptateurs NVMe identiques est requise; les paires d'adaptateurs NVMe suivantes peuvent être différentes de la première paire. Après la commande d'une paire identique, un adaptateur NVMe de capacité différente est autorisé.
- Les périphériques NVMe nécessitent la mise en miroir du système d'exploitation IBM i car il n'y a pas de prise en charge du RAID matériel. Les paires en miroir doivent se trouver sur des unités physiques différentes. NVMe ne peut être mis en miroir que sur NVMe et les disques SAS ne peuvent être mis en miroir que sur des disques SAS.
- Le disque de secours à chaud n'est pas pris en charge, mais un NVMe supplémentaire peut être sur le système en tant que disque de secours à froid pour accélérer le processus de réparation, et ce n'est qu'un disque de rechange dans le fait qu'un client n'a pas à le commander / le brancher. Le développement IBM i est conscient du désir de quelque chose de plus performant qu'une pièce de rechange froide.

Type de format NVMe – U.2 & AIC

Add-in Card (AIC)

- x4 / x8 PCIe Gen3 or Gen4
- High performance and power (25-50 Watts)
- Focus on high capacity - up to 6.4TB currently
- Best performance and latency



| Feature Code | PCIe3 Card NVMe Flash Adapter | | |
|-------------------|-------------------------------|---------------|---------------|
| | 1.6TB | 3.2TB | 6.4TB |
| AIX/Linux (LP/FH) | #EC5G / #EC5B | #EC5C / #EC5D | #EC5E / #EC5F |
| IBM i (LP/FH) | #EC6U / #EC6V | #EC6W / #EC6X | #EC6Y / #EC6Z |

Octobre 2019



| Feature Code | PCIe4 Card NVMe Flash Adapter | | |
|-------------------|-------------------------------|---------------|---------------|
| | 1.6TB | 3.2TB | 6.4TB |
| AIX/Linux (LP/FH) | #EC7A / #EC5B | #EC7C / #EC7D | #EC7E / #EC7F |
| IBM i (LP/FH) | #EC7J / #EC7K | #EC7L / #EC7M | #EC7N / #EC7P |

Octobre 2020

U.2

- x2 / x4 PCIe Gen3
- Higher power required to achieve max SSD performance
- Expect to be direct attached to CPU vs through a HBA



2.5" 7mm

E950/E980
800GB - #EC5J
1.6TB - #EC5K
3.2TB - #EC5L



Aout 2018



2.5" 15mm

S9xx G Models
800GB - #EC5X
1.6TB - #ES1F/ES1E
3.2TB - #ES1H/ES1G
6.4TB - #EC5W/EC5V



Juillet 2020

M.2

- x2 / x4 PCIe Gen3
- Different physical sizes
- Power less than 9 watts
- Limitations: Minimal Connector Plug Capability (<10), No hot swap
- Targeted as a replacement for SATA SSDs

■ Withdrawn from marketing 3/24/2020, announcement 920-065



Card #EC59 with 1-2 #ES14



S9xx
400GB #ES14



NVMe format U.2

Hardware Features

- NVMe Specs. 1.4 Compliant
- PCIe Gen 4 x 4 (1 x4 and 2 x2 modes)
- Multiple Name Space (64)
 - NS Granularity 8KB
- SFF U.2 (2.5") 15mm z-Height
- Power Modes (\leq 25W, 18W(default))
- Block Size 4096(AIX/Linux), 512, 4160 (IBMi)
- End-To-End Protection: T10 DIF & DIX
- Non Volatile Write Buffer
- Endurance 3 DWPD
- Boot: Option ROM BAR 128KB
- Warranty 5 years
- Hot Plug capable (Managed)
- ECC \geq 100 bits per 4KB
- RAIF: Tolerant of single flash die failures
- MTBF \geq 2 million hours
- End Of Life Data Retention \geq 3 months
- EEH Support
- Live Firmware update
- NVMe-MI Support (SMBUS)
- TCG OPAL SED

Supported Platforms

- P9 Scale Out Gen4 Refresh

| PCIe4 1.6TB NVMe Flash Adapter III x8 | | | | |
|---------------------------------------|-------|-------|-------|-------|
| | 800GB | 1.6TB | 3.2TB | 6.4TB |
| AIX/Linux | | ES3B | ES3D | ES3F |
| IBM i | ES3A | ES3C | ES3E | ES3G |

Power9 GA moved to July 29 Performance

| Workload | 800GB | 1.6 TB | 3.2 / 6.4 TB |
|----------------------|-------|--------|--------------|
| Read (IOPS) | 800K | 1.3M | 1.6M |
| Write (IOPS) | 140K | 260K | 280K |
| Mixed R/W (70/30) | 174K | 319K | 420K |
| Read Data Tp (GB/s) | 6.4 | 7.0 | |
| Write Data Tp (GB/s) | 1.3 | 2.5 | 4.1 |
| Read Latency (us) | 80 | | |
| Write Latency (us) | 15 | | |

Notes:

1. IOPs and Latency #'s on random 4K
2. Data throughput #'s are on sequential 256KB workload

Software Support

- Linux RHEL 8.1, SLES 15
- AIX (7.1c & 7.2Q), VIOS (2.2.6) or later
- IBM i 7.4 TR2 or later
- Load Source
- Software RAID 0, 1, 5 & 6 (Linux)
- OS Mirroring (AIX, IBMi)
- DIAG Support



Enterprise PCIe Gen4 Add In Card

Hardware Features

- NVMe Specs. 1.3 Compliant
- NVMe Over Fabrics 1.0 Capable
- **PCIe Gen 4 x 8**
- **Controller Memory Buffer (CMB) ≥ 4MB (Linux only)**
- Multiple Name Space (64)
 - NS Granularity 8KB
- Half Height Half Length (HH-HL)
- Power ≤ 25W
- Block Size 4096 (Default), 512, 4160
- End-To-End Protection: T10 DIF & DIX
- Non Volatile Write Buffer
- Endurance 3 DWPD
- PCIe Vendor VPD Support (IBM Provides content)
- Boot: Option ROM BAR 128KB (IBM Provides content)
- Warranty ≥ 5 years
- Hot Plug capable
- ECC ≥ 100 bits per 4KB
- RAIF: Tolerant of single flash die failures
- MTBF ≥ 2 million hours
- End Of Life Data Retention ≥ 3 months
- EEH Support
- Live Firmware update
- NVMe-MI (Optional)
- TCG OPAL SED

| | PCIe4 NVMe Flash Adapter III x8 | | |
|-------------------|---------------------------------|-------------|-------------|
| | 1.6TB | 3.2TB | 6.4TB |
| FC (LP/FH) | EC7A / EC5B | EC7C / EC7D | EC7E / EC7F |
| IBM i FCs (LP/FH) | EC7J / EC7K | EC7L / EC7M | EC7N / EC7P |

Performance

| Workload | Target (1.6 TB) | Target (3.2/6.4) |
|----------------------|-----------------|------------------|
| Read (IOPS) | 1M | 1.5M |
| Write (IOPS) | 250K | 250K |
| Mixed R/W (70/30) | 300K | 300K |
| Read Data Tp (GB/s) | 8.0 | 8.0 |
| Write Data Tp (GB/s) | 2.5 | 3.8 |
| Read Latency (us) | 80 | 80 |
| Write Latency (us) | 20 | 20 |

Notes:

1. IOPs and Latency #'s on random 4K
2. Data throughput #'s are on sequential 256KB workload

Software Support

- Linux
 - Power VM: RHEL 8.2LE, SLES 15 SP2 LE
 - Power NV: RHEL 8.2LE, Ubuntu 20.4
- AIX (7.2V), VIOS (2.2.6)
- IBM i 7.4TR3 or later
- Load Source
- Software RAID 0, 1, 5 & 6 (Linux)
- OS Mirroring (AIX, IBMi)
- DIAG Support



NVMe AIC retirée de commercialisation le 11 Aout 2023

Unité NVMe pour IBM i

| FC | Description | CAPACITY TB | Read BW GB/s | Write BW GB/s | Read IOPS (K) | Write IOPS (K) | Latence Lecture (us) | Latence écriture (us) | DWPD (5years) | PCIe Gen | Width (Lanes) | Form Factor | GA date |
|------------|-------------------------------------------------------------|----------------|-----------------|------------------|------------------|-------------------|----------------------------|-----------------------------|------------------|-------------|------------------|----------------|---------|
| EKF2 | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 800 | 6.8 | 1.4 | 880 | 100 | 90 | 15 | 3 | 4 | 4 | U.2 | Nov-22 |
| ES1K | 0.8TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i | 800 | 4.2 | 1.2 | 500 | 135 | 100 | 40 | 3 | 4 | 4 | U.2 | Apr-21 |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 for IBM i | 800 | 6.4 | 1.3 | 800 | 140 | 80 | 15 | 2.4 | 4 | 4 | U.2 | Jul-22 |
| EC6U, EC6V | PCIe3 x8 1.6 TB NVMe Flash adapter for IBM i | 1.6 | 4.7 | 1.9 | 700 | 100 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7J, EC7K | PCIe4 LP 1.6 TB NVMe Flash Adapter x8 for IBM i | 1.6 | 8.2 | 2.5 | 1000 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |
| EKF4 | Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i | 1.6 | 6.8 | 2.7 | 1300 | 215 | 90 | 15 | 3 | 4 | 4 | U.2 | Nov-22 |
| ES1F | 1.6 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i | 1.6 | 7.0 | 2.5 | 100 | 220 | 80 | 20 | 3 | 4 | 5 | U.2 | Jul-20 |
| ES3C | Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i | 1.6 | 7.0 | 2.5 | 1300 | 260 | 80 | 15 | 3 | 4 | 4 | U.2 | Jul-22 |
| EC6W, EC6X | PCIe3 x8 3.2 TB NVMe Flash adapter for IBM i | 3.2 | 6 | 3 | 910 | 170 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7L, EC7M | PCIe4 3.2 TB NVMe Flash Adapter for IBM i | 3.2 | 8.2 | 4.2 | 1500 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |
| EKF6 | Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i | 3.2 | 6.8 | 4.0 | 1400 | 350 | 90 | 15 | 3 | 4 | 4 | U.2 | Nov-22 |
| ES1H | 3.2 TB Enterprise 4K NVMe SFF U.2 15 mm SSD PCIe4 for IBM i | 3.2 | 7.4 | 3.4 | 1500 | 250 | 80 | 20 | 3 | 4 | 4 | U.2 | Jul-21 |
| ES3E | Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i | 3.2 | 7.0 | 4.1 | 1600 | 280 | 80 | 15 | 3 | 4 | 4 | U.2 | Jul-22 |
| EC5W | 6.4 TB Enterprise 4K NVMe SFF U.2 15 mm SSD PCIe4 for IBM i | 6.4 | 7.2 | 3.8 | 1500 | 250 | 80 | 20 | 3 | 4 | 4 | U.2 | Jul-20 |
| EC6Y, EC6Z | PCIe3 x8 6.4 TB NVMe Flash adapter for IBM i | 6.4 | 6 | 3 | 910 | 170 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7N, EC7P | PCIe4 6.4 TB NVMe Flash Adapter x8 for IBM i | 6.4 | 8.2 | 4.2 | 1500 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |
| EKF8 | Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i | 6.4 | 6.8 | 3.8 | 1400 | 325 | 90 | 15 | 3 | 4 | 4 | U.2 | Nov-22 |
| ES3G | Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i | 6.4 | 7.0 | 4.1 | 1600 | 280 | 80 | 15 | 3 | 4 | 4 | U.2 | Jul-22 |

Note : NVMe de capacité 1,6 TB ou supérieur sont plus performants que les NVMe de 800 Go

- Ex: EKF2 vs EKF4/EKF6/EKF8

■ **Cartes NVMe AIC retirées de commercialisation le 11 Aout 2023**

Simplification des unités NVMe

- **12 Mai 2023** – Les cartes PCIe NVMe AIC sont retirées de commercialisation
- **10 October 2023** - Simplification des unités NVMe U.2 pour IBM i and Power10 systems.
 - Les unités NVMe Enterprise SSD PCIe4 U.2 ES5B, ES5D, ES5F, ES5H, ES5A, ES5C, ES5E et ES5G remplaceront les unités NVMe existantes. NVMe AIC existe toujours. Effectif October 20, 2023
- **24 Octobre 24, 2023. Retrait des NVMe existants en format U.2 format, effectif October 24 !**

| NVMe Device Size | AIX / Linux Feature | IBM i Feature ¹ |
|------------------|---------------------|----------------------------|
| 800 GB | ES5A | ES5B |
| 1.6 TB | ES5C | ES5D |
| 3.2 TB | ES5E | ES5F |
| 6.4 TB | ES5G | ES5H |

IBM i supported¹

Feature codes for no-charge load source specify for IBM i features (#ELYA, #ELYC, #ELYE, and #ELYG) should be used when Enterprise PCIe4 SSD NVMe U.2 for IBM i is ordered.

Performances des derniers NVMe annoncées

| FC | Description | CAPACITY TB | Read BW GB/s | Write BW GB/s | Read IOPS (K) | Write IOPS (K) | Latence Lecture (us) | Latence écriture (us) | DWPD (5years) | PCIe Gen | Width (Lanes) |
|------|-------------------------------------------------------|----------------|-----------------|------------------|------------------|-------------------|----------------------------|-----------------------------|------------------|-------------|------------------|
| ES5B | Enterprise 800GB SSD PCIe4 NVMe U.2 for IBM i | 800 | 6.4 | 1.3 | 800 | 140 | 80 | 15 | 2.4 | 4 | 4 |
| ES5D | Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i | 1.6 | 7.0 | 2.5 | 1300 | 260 | 80 | 15 | 3 | 4 | 4 |
| ES5F | Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i | 3.2 | 7.0 | 4.1 | 1600 | 280 | 80 | 15 | 3 | 4 | 4 |
| ES5H | Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i | 6.4 | 7.0 | 4.1 | 1600 | 280 | 80 | 15 | 3 | 4 | 4 |

- ✓ NVMe U.2 format supporté par IBM i en natif
- ✓ Même performance NVMe que les dernières unités NVMe ES3A, ES3C, ES3E, ES3G

■ DWPD – Drive Writes per Day

- Définit le nombre de fois par jour que l'on peut écrire sur la totalité du NVMe pendant la période de garantie du NVMe (5 ans en général)
- Basé sur des écritures aléatoires. Si l'activité est plus sur des écritures séquentielles, la durée de vie du NVMe sera plus grande
- Les lectures séquentielles ou aléatoires n'ont aucune incidence sur la durée de vie du NVMe
- Exemple:
 - FC EKF2 - NVMe U.2 de 800Go. DWPD= 3
 - on peut écrire 3 x 800Go par jour pendant une durée de 5 ans
 - FC EC6UF4 – NVMe AIC de 1,6 To. DWPD=5
 - on peut écrire 5 x 1,6To par jour pendant 5 ans.

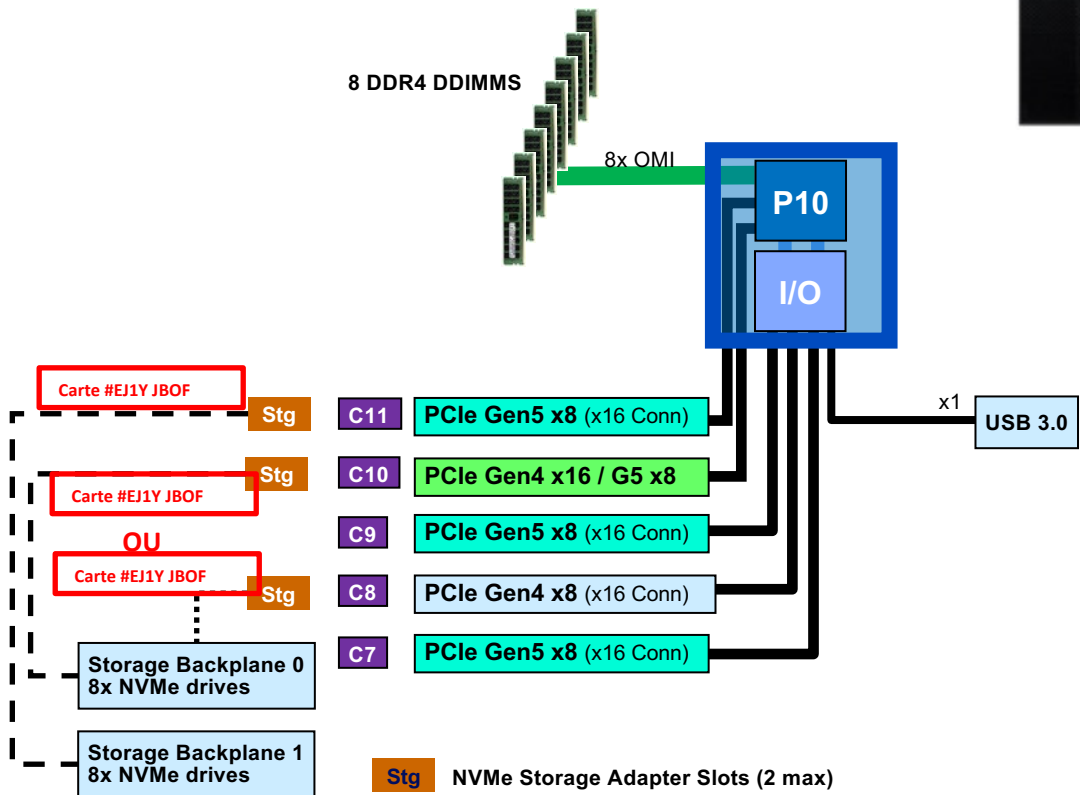
Versions IBM i et Technology Refresh supporté par NVMe et type de configuration

- iVirt = IBM i dans IBM i
- Native = dédié IBM i
- VIOS = disque virtuel via vSCSI
- All = tout type d'architecture

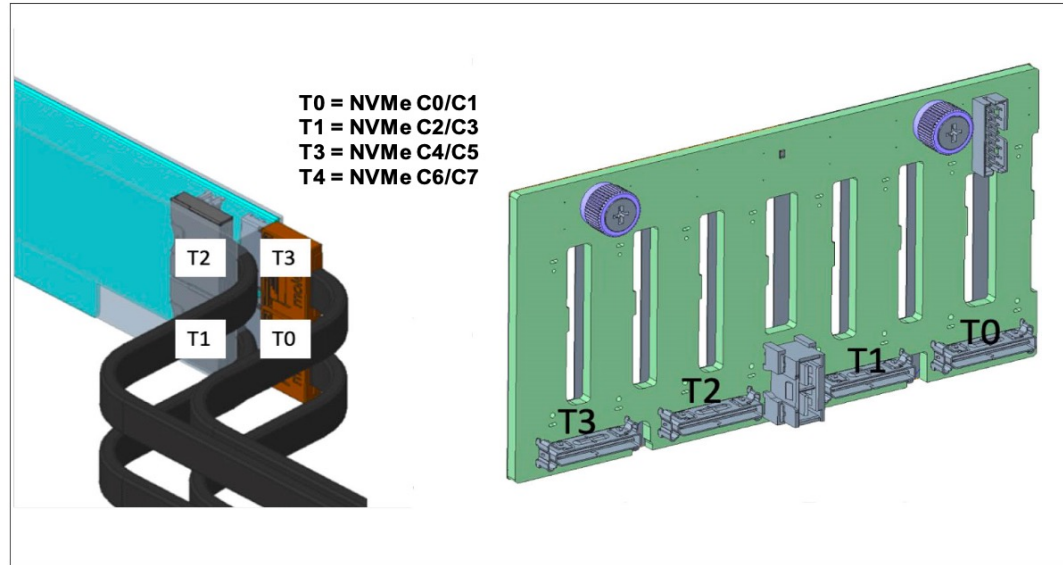
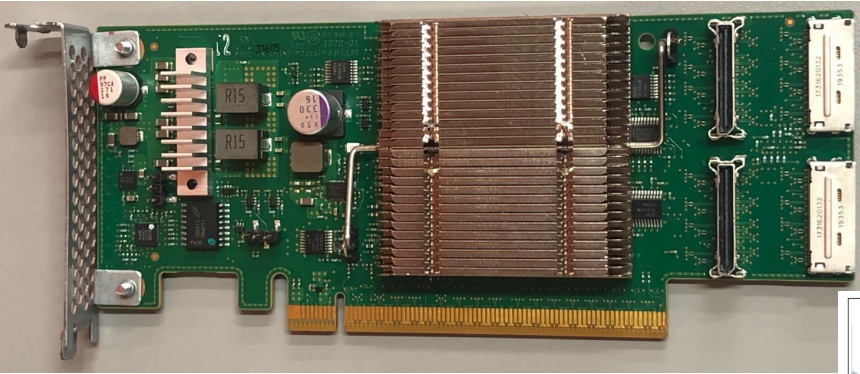
| IBM i I/O Support | Type of Configuration (Native, VIOS, iVirt, All) | IBM i 7.5 | IBM i 7.4 | IBM i 7.3 |
|-------------------------------------------------------------------|-----------------------------------------------------|----------------|----------------|-------------------------|
| Enhancements from November 2023 | | | | |
| #ENZO - PCIe Gen4 I/O Expansion Drawer | All | Tech Refresh 3 | Tech Refresh 9 | Base (iVirt, VIOS only) |
| #ENZF - PCIe Gen4 FanOut Module | All | Tech Refresh 3 | Tech Refresh 9 | Base (iVirt, VIOS only) |
| #EN24, #EN26 - PCIe 4-Port 25/10/1 GbE RoCE SFP28 Adapter | All | Tech Refresh 3 | Tech Refresh 9 | Base (iVirt, VIOS only) |
| #EC71, #EC72 - PCIe 2-Port 25/10/1 GbE RoCE SFP28 Adapter | All | Tech Refresh 3 | Tech Refresh 9 | Base (iVirt, VIOS only) |
| Enhancements from October 2023 | | | | |
| #ES5B - Enterprise 800 GB SSD PCIe4 NVMe U.2 module for IBM i | Native, iVirt | Base | Tech Refresh 6 | Tech Refresh 12 |
| #ES5D - Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i | Native, iVirt | Base | Tech Refresh 6 | Tech Refresh 12 |
| #ES5F - Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i | Native, iVirt | Base | Tech Refresh 6 | Tech Refresh 12 |
| #ES5H - Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i | Native, iVirt | Base | Tech Refresh 6 | Base (iVirt only) |
| #ES5A - Enterprise 800 GB SSD PCIe4 NVMe U.2 module for AIX/Linux | VIOS | Base | Base | Base |
| #ES5C - Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux | VIOS | Base | Base | Base |

<https://www.ibm.com/support/pages/node/1135378>

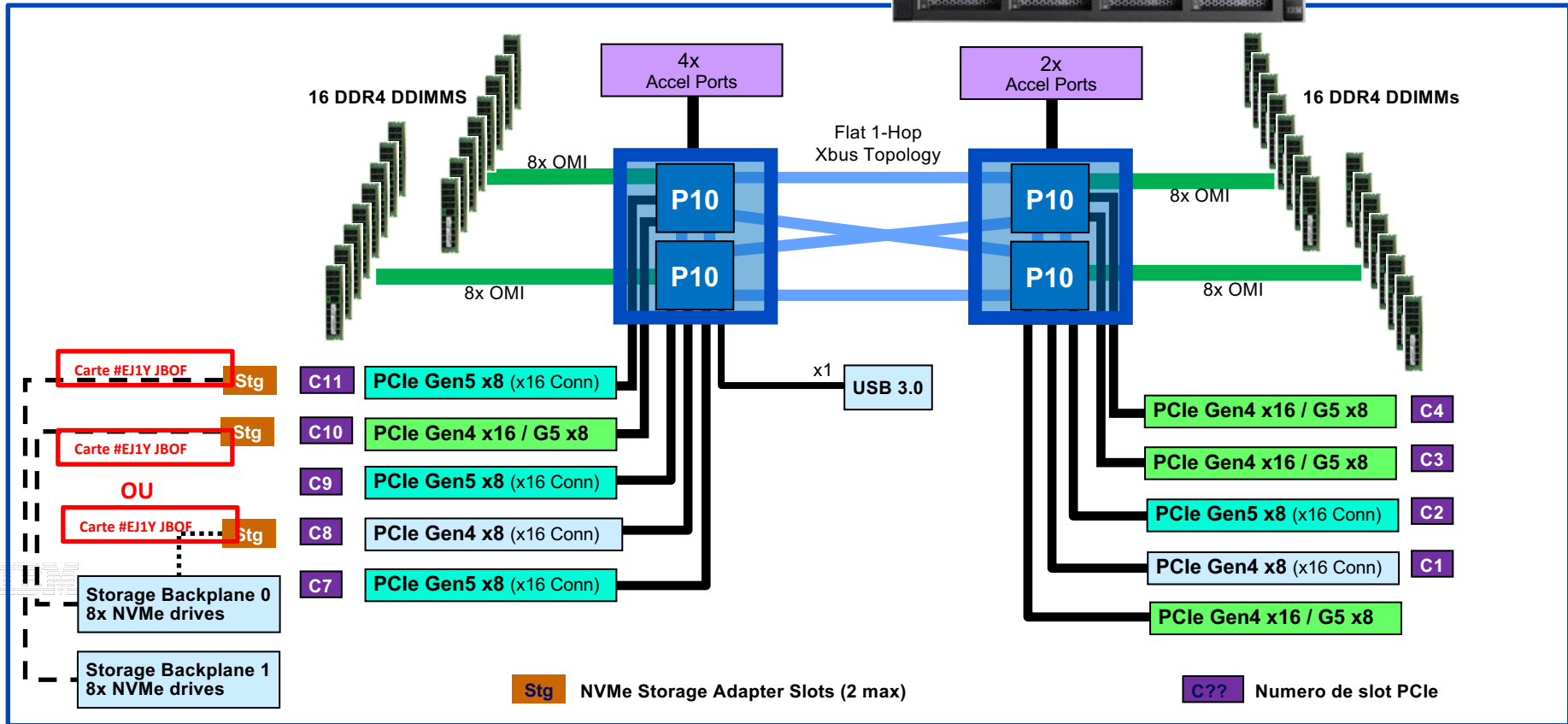
NVMe sur POWER10 S1014



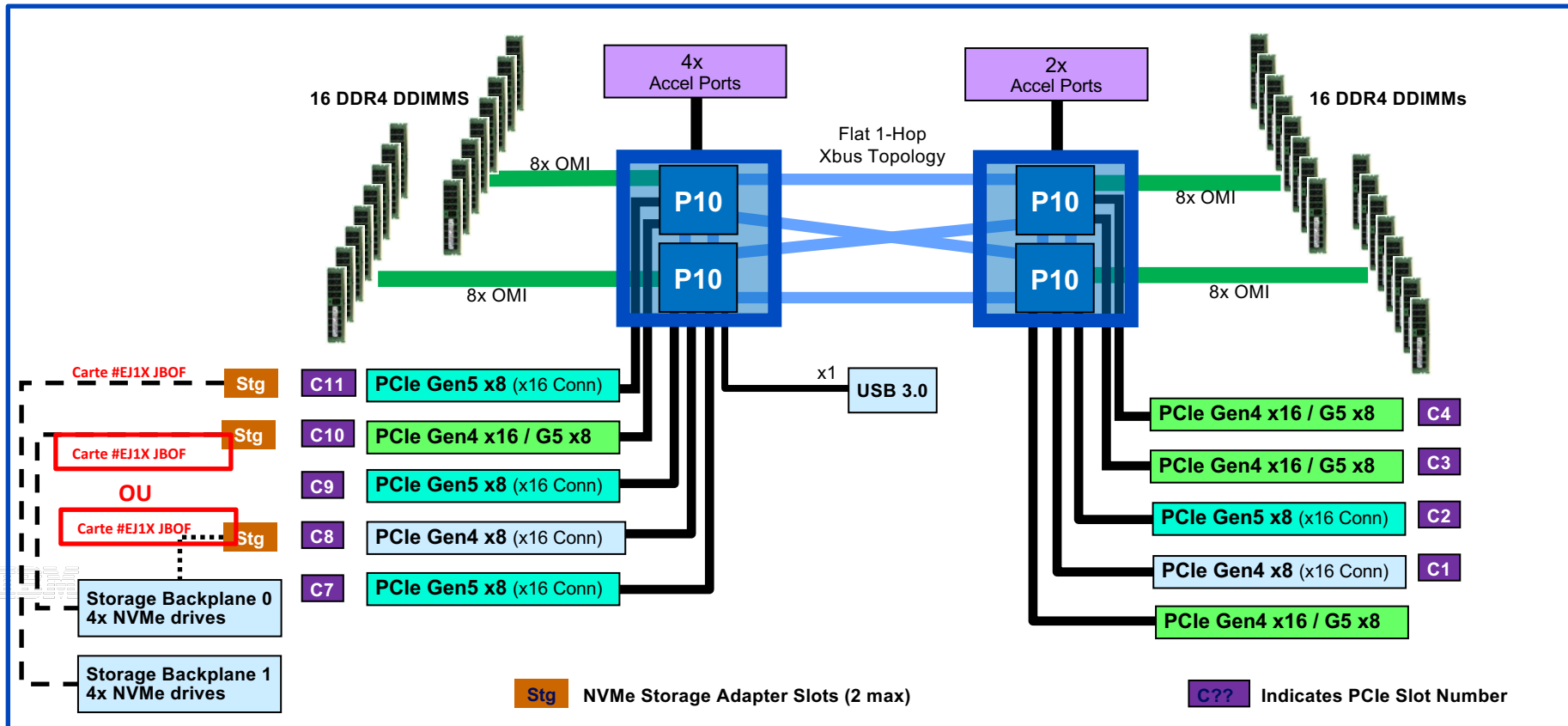
Carte PCIe4 4-port NVMe JBOF adapter #EJ1Y ou #EJ1X



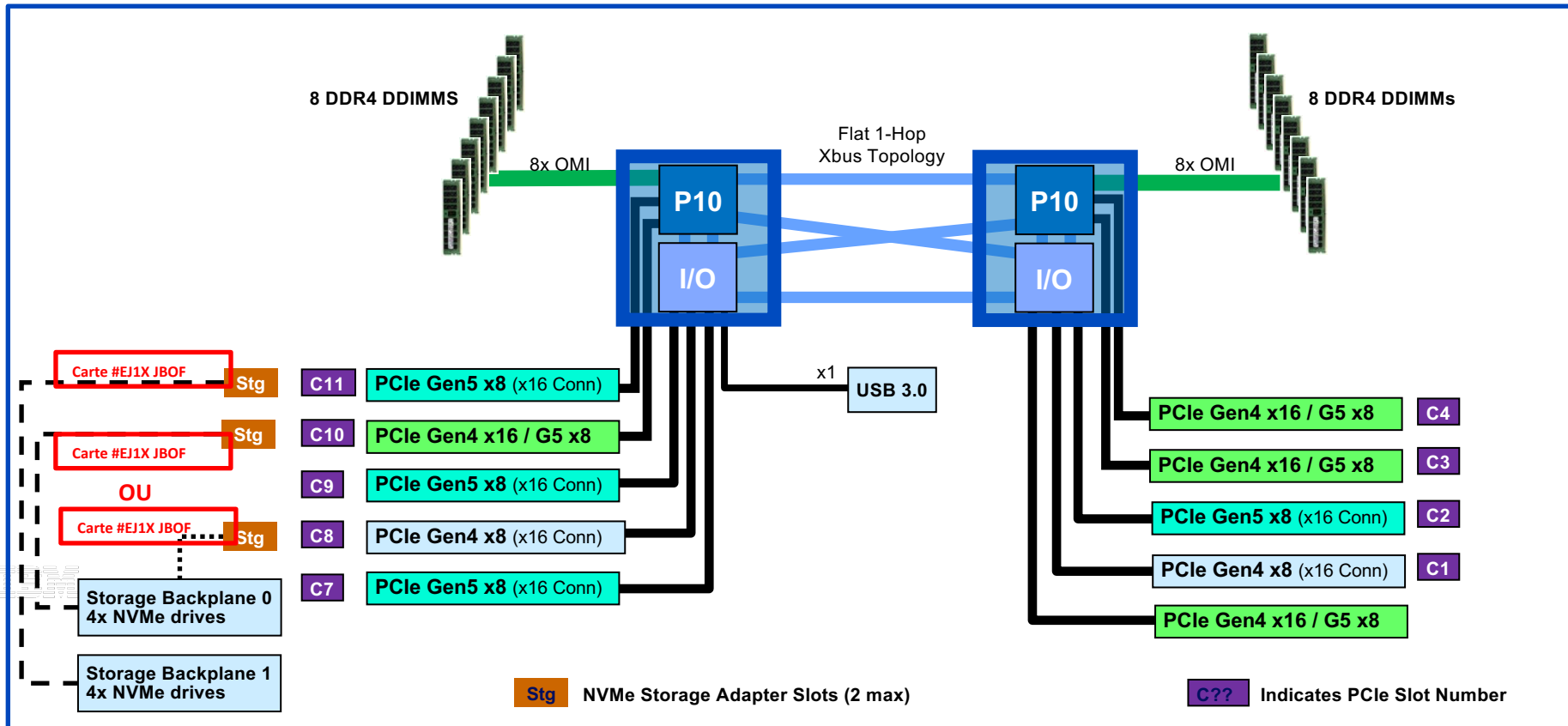
NVMe sur POWER10 S1024



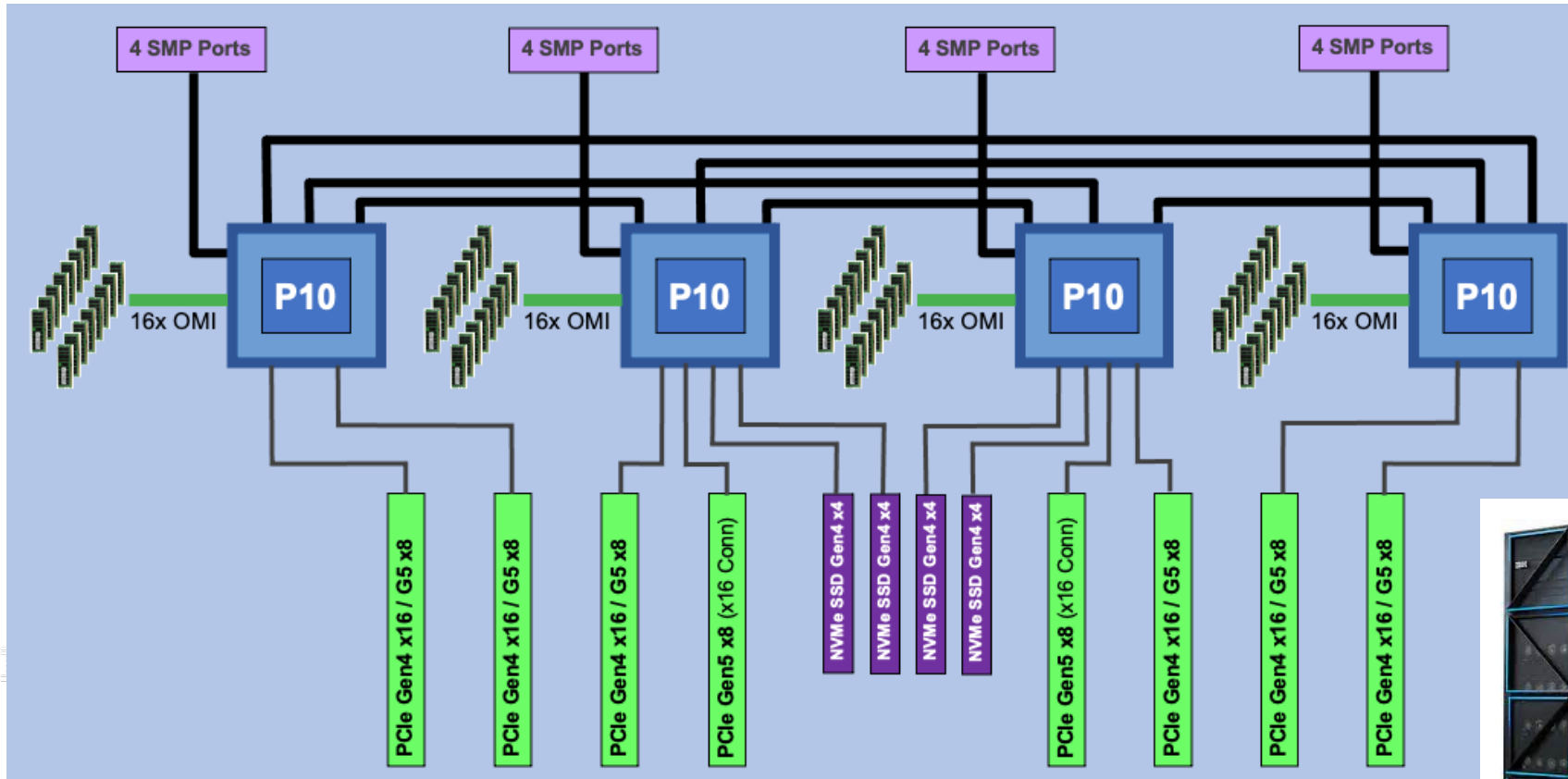
NVMe sur POWER10 S1022



NVMe sur POWER10 S1022s



NVMe sur POWER10 E1080



Note : Limité à 2 NVMe U.2 par noeud (cage 15mm). Maximum 8 NVMe U.2 sur E1080.

Namespaces sur unité NVMe et IBM i



- L'architecture NVMe utilise des **namespaces** qui sont une collection de blocs logiques dont les adresses de bloc logique vont de 0 à la taille du disque. Un ID de namespace(NSID) est un identificateur utilisé par un contrôleur pour fournir l'accès à un namespace.
- Avec NVMe, un namespace est un « bras » (unité logique), un disque pour IBM i. Un namespace est un morceau logique d'une unité NVMe physique et plusieurs namespaces sont autorisés sur un NVMe.
- La création et la gestion des namespaces est faite à travers l'IBM i.
- L'utilisation par IBM i de namespace multiples architecturés par NVMe fournit de nombreux «bras» sur un petit nombre de périphériques physiques NVMe haute capacité
- IBM i peut utiliser un périphérique NVMe (jusqu'à 16 To) avec un seul namespace pour l'ensemble du périphérique. Cependant, pour presque tous les clients, cela entraînera des performances sous-optimales car plus de (et plus petits) «bras» (unités logiques) valent mieux que moins et plus gros.
- Les NVMe **Mainstream** ne supporte qu'un seul namespace. Les NVMe **Enterprise** supportent multiples namespaces. Pas de NVMe Mainstream supporté en natif avec IBM i dû au bras disque ou namespace unique sur le NVMe.
- Notez que le mot «namespace» est utilisé dans l'industrie et par IBM i de différentes manières et dans différents contextes. Ainsi, par exemple, un namespace NVMe ne doit pas être confondu et n'a rien à voir avec les namespaces ASP

Namespace recommandé par IBM



- ✓ IBM recommande de créer des namespaces de taille de **200 Go** ou **400 Go**.
- ✓ Il est possible de créer des namespaces plus petit pour avoir plus de “bras disque” IBM i mais amélioration des performances non démontrée.
- ✓ Dans e-config on peut demander la création de namespace pour le “Load Source” de 200 Go ou 400 Go pour préchargement IBM i.
 - ✓ #ENSA 200Go IBM i NVMe LOAD SOURCE NAME SPACE SIZE
 - ✓ #ENSB 400 Go IBM i NVMe LOAD SOURCE NAME SPACE SIZE
- ✓ Les cartes AIC PCIe Gen3 NVMe ne peuvent être “mirroré” qu’avec une carte AIC PCIe Gen3 NVMe
- ✓ NVMe U.2 et nouvelles cartes AIC NVMe Gen4 peuvent être “mirroré” si utilisation même taille namespaces
- ✓ La recommandation des tailles de namespaces de 188 ou 393 Go n’est plus adequat avec les dernieres générations de NVMe (seulement avec la première génération)

| Unité NVMe U.2 et AIC | Max Namespace 200 Go | Max Namespace 400 Go |
|-----------------------|----------------------|----------------------|
| NVMe 800 Go | 4 | 2 |
| NVMe 1.6 To | 8 | 4 |
| NVMe 3.2 To | 16 | 8 |
| NVMe 6.4 To | 32 | 16 |

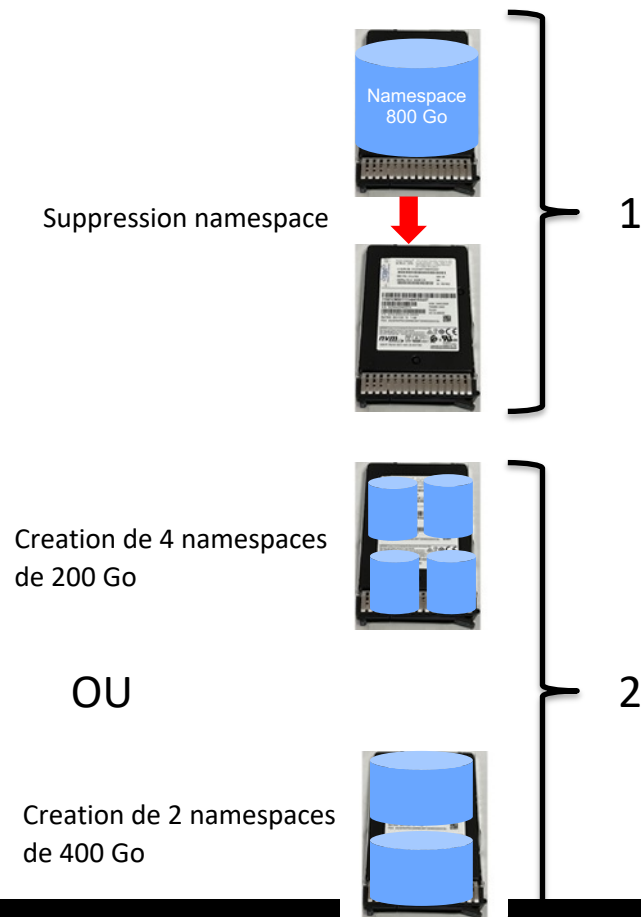
NVMe Namespace – exemple NVMe ES5B - 800GB



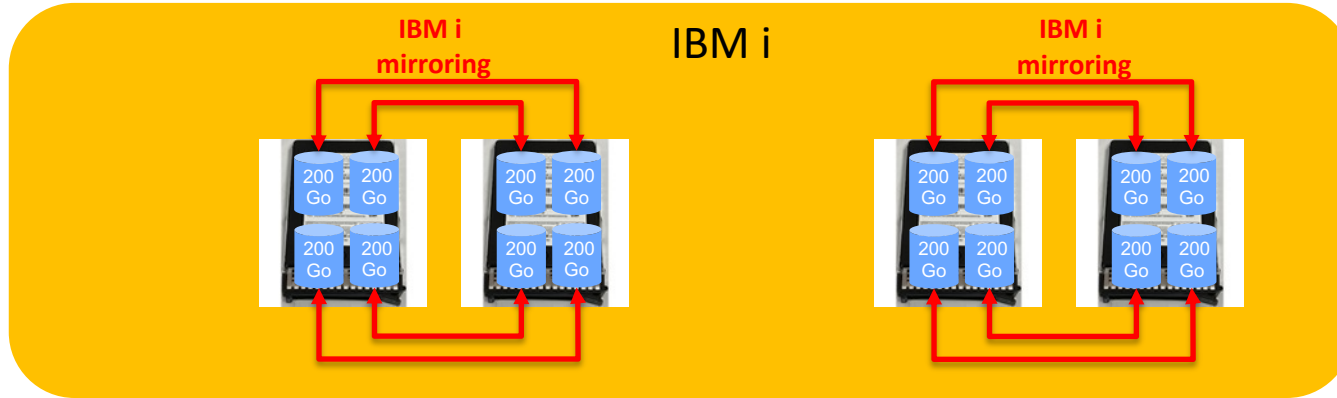
NVMe livrée avec 1 namespace configuré prenant tout l'espace de stockage par défaut

Configuration des disques NVMe pour IBM i

1. Suppression du namespace par défaut de l'unité
 2. Création des namespaces de 200 Go ou 400 Go sur l'unité NVMe
 - 1 namespace = 1 disque IBM i = 1 bras disque
- Possibilité de garder de l'espace non affecté à un namespace pour utilisation ultérieure
 - Même procédure pour les NVMe gérés par les VIOS



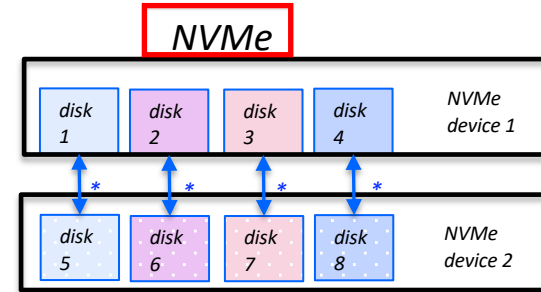
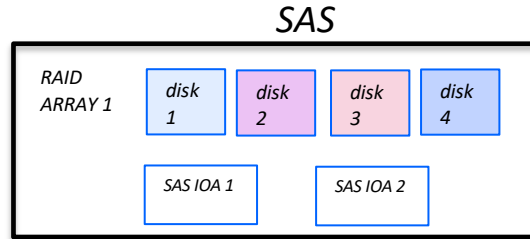
Namespace sur IBM i et protection (Mirroring)



- Exemple : 4 unités NVMe FC ES5B de 800 Go
 - création de 4 namespace de 200 Go par NVMe
 - IBM i voit donc 16 disques logiques
- Protection des namespaces par mirroring IBM i – seule protection disponible
 - Mirroring entre namespace d'unité NVMe différentes et de même taille
- Possibilité de créer des namespaces de tailles différentes sur les unités NVMe.
- Dans exemple : IBM i a 16 disques de 200 Go avec un espace de stockage de 1600 Go disponible

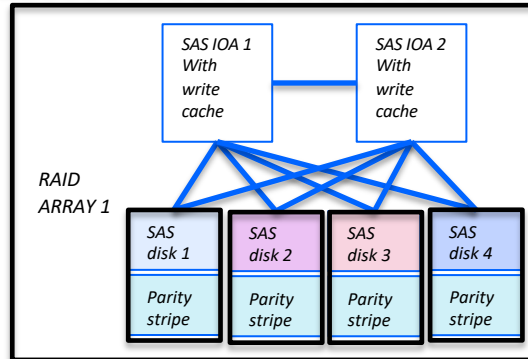
Stockage SAS versus NVMe avec IBM i

Vue système
d'exploitation

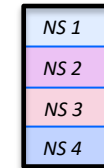


OS is aware of hardware RAID

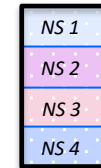
Vue unite
physique



NVMe device 1

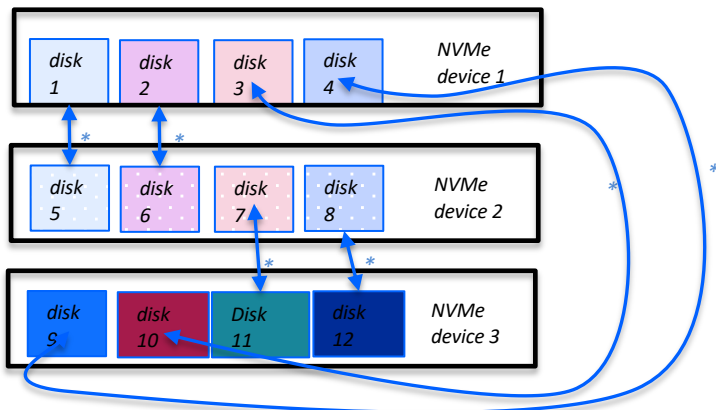


NVMe device 2



IBM i NVMe Mirroring (3 NVMe Option 1)

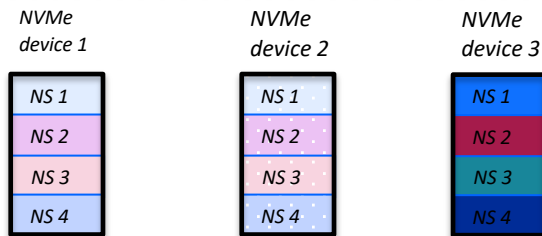
Vue Client et IBM i



Tous les namespaces en miroir ont la même taille sur toutes les unités.

**OS mirroring*

Vue Unité physique

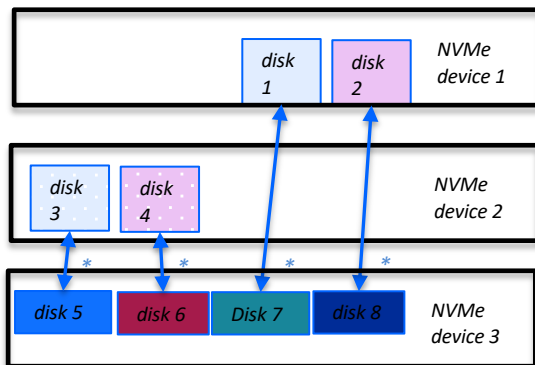


La capacité physique de l'unité ne doit pas nécessairement être de la même taille



IBM i NVMe Mirroring (3 NVMe Option 2)

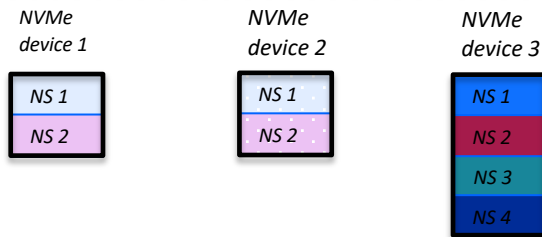
*Vue Client
et IBM i*



*Tous les namespaces
en miroir ont la
même taille sur
toutes les unités.*

**OS mirroring*

*Vue Unité
physique*

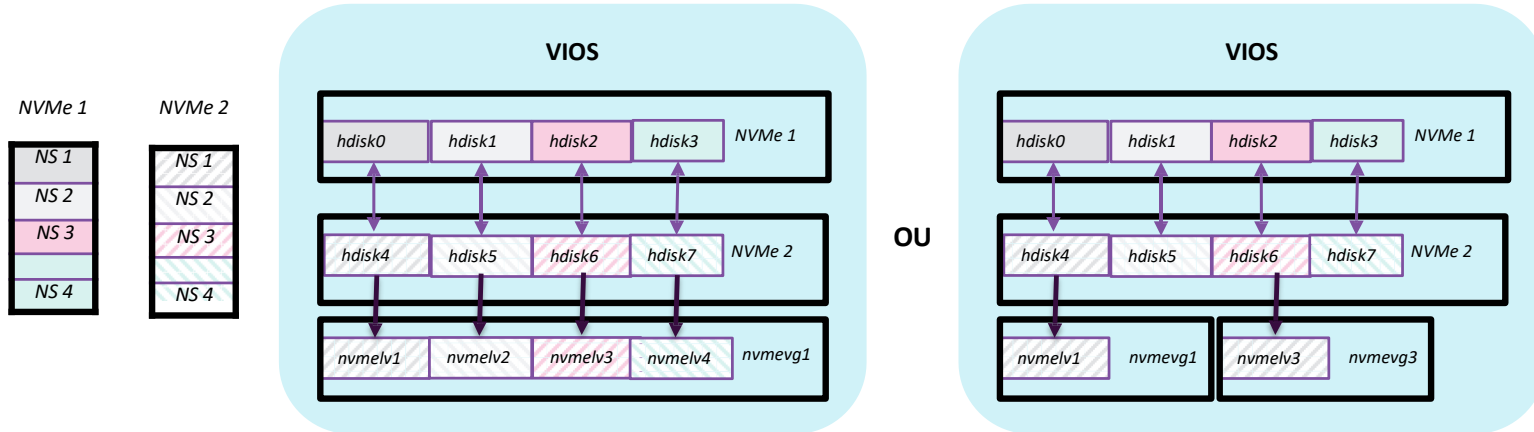


*La capacité physique
de l'unité ne doit pas
nécessairement être
de la même taille*



VIOS et NVMe - 1 VIOS et 2 NVMe

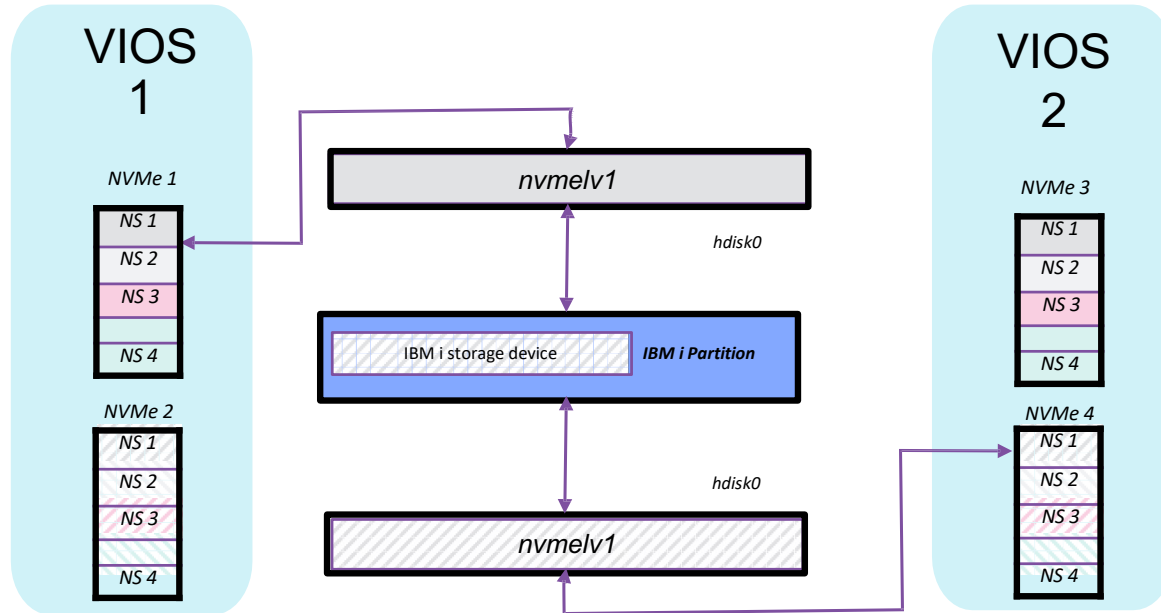
- Création Namespace **NS1** sur NVMe1 qui sera vue comme **hdisk0** par le VIOS
 - Création Namespace **NS1** sur NVMe2 qui sera vue comme **hdisk4** par le VIOS
 - Ajout **hdisk0** and **hdisk4** au volume group **nvmevg1** (configurez le mirroring sur VIOS – optionel)
 - Création volume logique **nvme1v1**
 - Création device vscsi **vnvme1** avec **nvme1v1** comme volume cible
 - Affectation **vnvme1** au vhost de la LPAR IBM i cliente.
- Vous decidez où faire le mirroring (VIOS ou IBM i) et les configurations des volumes groupes et logiques



ATTENTION : protocole de communication SCSI au lieu de NVMe sur le vSCSI, donc bande passante limité au SCSI !

VIOS et NVMe – VIOS redondant et 4 NVMe

- Chaque VIOS gère 2 unités NVMe
- Permet la redondance au niveau unité et instances VIOS
- Comme sur VIOS unique, créez namespace, volume logique et group sur 1 unité NVMe à partir de chaque instance VIOS et affectez les deux à la même LPAR IBM i
- Configurez le mirroring au niveau de l' IBM i

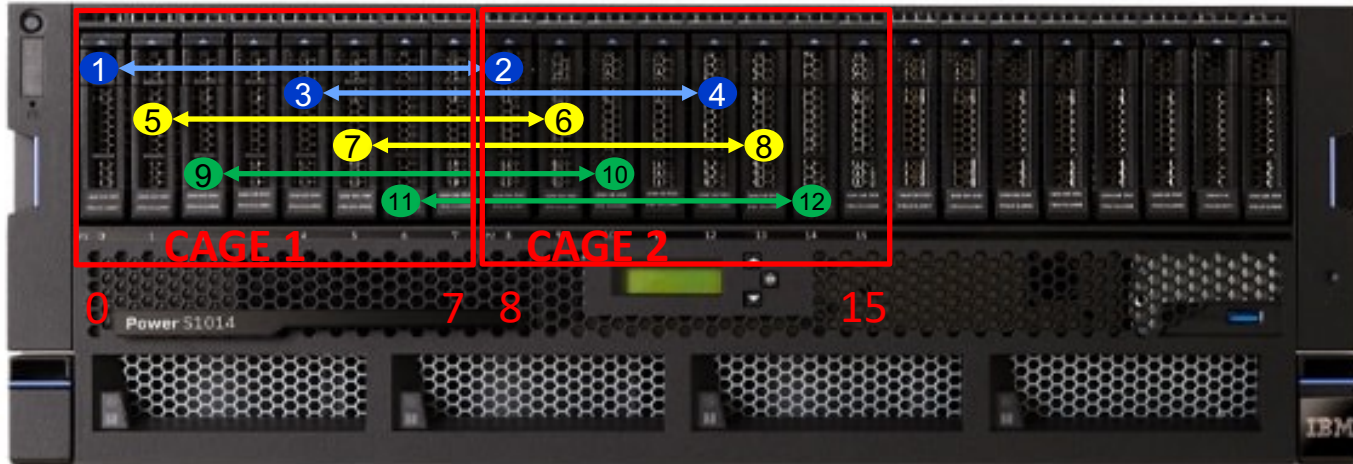


Règles de placement des NVMe pour un RAS maximum

- Avoir 2 cages de disques NVMe, donc 2 cartes JBOF #EJ1X ou EJ1Y
- Répartir les “miroirs” des unités NVMe dans une cage différente.
- Séparez NVMe dans une même cage pour une meilleure ventilation
- Utilisez le modèle suivant :

Positions dans les cages

0,8 4,12 1,9 5,13 2,10 6,14 3,11 7,15



Configuration NVMe sur 1024. Sur S1022, les cages contiennent 4 emplacements NVMe

IBM System Planning Tool – Configuration et placement NVMe

1. Placement de 1 ou 2 cartes JBOF EJ1Y
2. Ajout des unités NVMe U.2 et/ou AIC

IBM System Planning Tool
System plan: New system plan System: S1024 [IBM Power System S1024(9105-42A)]

System Partitions Hardware Networking Virtual Storage Installation Consoles Summary

9105_42A-0

Display hardware supported by: IBM I V7R5M0

Filter box: FC/CCIN

- External DiskCtrls
- LAN Adapters
- Removable MediaCtrls
- Fibre Adapters
- Backplane
- EJ1Y Storage Backplane with eight NVMe U.2 drive slots**
- USB Adapters
- Cryptographic Adapters
- I/O Expansion Adapters
- EJ2A PCIe3 x16 to CXF Converter Adapter (support AOC)
- NVMe Adapters
 - EC6V PCIe3 x8 1.6 TB NVMe Flash Adapter for IBM i
 - EC6X PCIe3 x8 3.2 TB NVMe Flash Adapter for IBM i
 - EC6Z PCIe3 x8 6.4 TB NVMe Flash Adapter for IBM i
 - EC7K PCIe4 LP 1.6TB NVMe Flash Adapter x8 for IBM i
 - EC7M PCIe4 3.2TB NVMe Flash Adapter x8 for IBM i
 - EC7P PCIe4 6.4TB NVMe Flash Adapter x8 for IBM i
- NVmemory module
 - EKF2 Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF4 Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF6 Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF8 Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES1K 0.8 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES1F 1.6 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES1H 3.2 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - EC5W 6.4 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES3A Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i
 - ES3C Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES3E Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES3G Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES5B Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i
 - ES5D Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES5F Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES5H Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i

| Action | Location | Bus Width | Bus Speed | Length | Priority | F/C | Partition / Port Connection | Order status | S/N | Description |
|--------|----------|-----------|-----------|--------|----------|------|-----------------------------|--------------|-----|---------------------------------------------------|
| | P0-C0 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #0 |
| | P0-C1 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe4 x16 Slot #0 |
| | P0-C2 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #1 |
| | P0-C3 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #2 |
| | P0-C4 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #3 |
| | P0-C6 | | | | | | | | | PCIe5 x16 Slot #5 |
| | P0-C7 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #6 |
| | P0-C8 | x16 | 2.5 GHz | Long | 8 | EJ1Y | Not Assignable | IBM | | Storage Backplane with eight NVMe U.2 drive slots |
| | P0-C9 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #8 |
| | P0-C10 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #9 |
| | P0-C11 | x16 | 2.5 GHz | Long | 8 | EJ1Y | Not Assignable | IBM | | Storage Backplane with eight NVMe U.2 drive slots |
| | P1-C0 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk |
| | P1-C1 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk |
| | P1-C2 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk |
| | P1-C3 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk |
| | P1-C4 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk |

Validation messages (10 total)

Outil gratuit – N’hésitez pas à l’utiliser pour vérifier configuration LPAR et positionnement des cartes et unités !

SPT – ajout NVMe

IBM System Planning Tool

System plan: S1024_NVMe

System: S1024 [IBM Power System S1024(9105-42A)]

System Partitions **Hardware** Networking Virtual Storage Installation Consoles Summary

9105_42A-0

Display hardware supported by: IBM i V7R5M0

Actions

Settings

Filter box: FC/CCIN

- External DiskCtrls
- LAN Adapters
- Removable MediaCtrls
- Fibre Adapters
- Backplane
 - EJ1Y Storage Backplane with eight NVMe U.2 drive slots
- USB Adapters
- Cryptographic Adapters
- I/O Expansion Adapters
 - EJ20 PCIe x16 to CXP Optical or CU converter Adapter for PCIe3 Expansion
 - EJ2A PCIe3 x16 to CXP Converter Adapter (support AOC)
- NVMe Adapters
 - EC6V PCIe3 x8 1.6 TB NVMe Flash Adapter for IBM i
 - EC6X PCIe3 x8 3.2 TB NVMe Flash Adapter for IBM i
 - EC6Z PCIe3 x8 6.4 TB NVMe Flash Adapter for IBM i
 - EC7K PCIe4 LP 1.6TB NVMe Flash Adapter x8 for IBM i
 - EC7M PCIe4 3.2TB NVMe Flash Adapter x8 for IBM i
 - EC7P PCIe4 6.4TB NVMe Flash Adapter x8 for IBM i
- NVmemory module
 - EKF2 Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF4 Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF6 Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i
 - EKF8 Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES1K 0.8 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES1F 1.6 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES1H 3.2 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES1W 6.4 TB Enterprise 4K NVMe SFF U.2 15mm SSD PCIe4 for IBM i
 - ES3A Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i**
 - ES3C Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES3E Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i
 - ES3G Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i

| Slot Characteristics | | | | | | | | | | | |
|----------------------|----------|-----------|-----------|--------|----------|------|-----------------------------|--------------|-----|---------------------------------------------------|--|
| Action | Location | Bus Width | Bus Speed | Length | Priority | F/C | Partition / Port Connection | Order status | S/N | Description | |
| ⬇ | P0-C0 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #0 | |
| ⬇ | P0-C1 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe4 x16 Slot #0 | |
| ⬇ | P0-C2 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #1 | |
| ⬇ | P0-C3 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #2 | |
| ⬇ | P0-C4 | x16 | 2.5 GHz | Long | 16 | | | | | PCIe5 x8 Slot #3 | |
| ⬇ | P0-C6 | | | | | | | | | PCIe5 x16 Slot #5 | |
| ⬇ | P0-C7 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #6 | |
| ⬇ | P0-C8 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe4 x16 Slot #7 | |
| ⬇ | P0-C9 | x16 | 2.5 GHz | Long | 8 | | | | | PCIe5 x16 Slot #8 | |
| ⬇ | P0-C10 | x16 | 2.5 GHz | Long | 16 | EJ1Y | Not Assignable | IBM | | Storage Backplane with eight NVMe U.2 drive slots | |
| ⬇ | P0-C11 | x16 | 2.5 GHz | Long | 8 | EJ1Y | Not Assignable | IBM | | Storage Backplane with eight NVMe U.2 drive slots | |
| ⬆ | P1-C0 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C1 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C2 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C3 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C4 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C5 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C6 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |
| ⬆ | P1-C7 | x16 | 2.5 GHz | Long | 16 | | | | | NVmemory Disk | |



IBM i



IBM e-config cloud

Pré-chargement IBM i avec option #ENSM

- **#ENSM - Specify Code Configure all IBM i Namespaces**
- **Création des namespaces maximum sur le Load source et son miroir.**
- Ce code ordonne à l'usine de créer un maximum de namespaces de même taille, comme indiqué sur le code de taille de Load Source #ENSA (200Go) ou #ENSB (400Go) définit, sur le disque NVMe Load Source et son miroir NVMe. Le NVMe load source et son miroir doivent être de même code dispo (FC).
- Ces namespaces créés seront ajouter à l'ASP 1
- **S'il y a plus de 2 NVMe U.2, les autres NVMe ne seront pas configurés et seront fournit avec 1 namespace configuré par défaut prenant toute la capacité du NVMe**
- #ENSM possible que sur commande initiale système et sur NVMe U.2 seulement. Non dispo sur NVMe AIC.
- #ENSM disponible que pour le IBM i processor Group P05:
 - Power9
 - S922 9009-22G avec 1 coeur (#EP5Y)
 - S914 9009-41G avec 4 coeurs (#EP50)
 - Power10
 - S1014 9105-41B avec 4 coeurs (#EPG0)
- **Pré-requis**
 - #2145 Primary operating system indicator – IBM i
 - #5000 Software preload required



IBM i NVMe options de pré-chargement



- **IBM i pré-chargement et configuration de tous les namespaces NVMe sur Load Source et Miroir**

- Code #ENSM
- 9009-22G avec 1 coeur #EP5Y
- 9009-41G avec 4 coeurs #EP50
- 9105-41B avec 4 coeurs #EPG0
- Option la plus simple, système prêt à l'emploi

- **IBM i pré-chargement et seul le premier namespace NVMe est configuré**

- Tous les autres namespaces NVMe seront créés et protégés par miroir par l'utilisateur.
- IBM i chargé, des configurations à finir

- **Pas de pré-chargement IBM**

- Client créera et protégera en miroir tous les namespaces
- Seule option si plusieurs LPAR configurées lors de la commande initiale

Power10 nombre NVMe et AIC par systeme

| Systeme P10 | PCIe Gen3, Gen4 adapteurs AIC (1.6, 3.2, 6.4 TB) * | NVMe U.2 (0.8, 1.6, 3.2, 6.4 TB) | NED24 – NVM U.2 |
|------------------------|-----------------------------------------------------|----------------------------------|-----------------|
| S1014 - 4 coeurs IBM i | Jusqu'à 6.4To** | | - |
| S1014 – 1 socket | 4 | 16 | - |
| S10124 - 1 socket | 4 | 16 | - |
| S1024 - 2 sockets | 9 | 16 | 24 |
| S1022 - 1 socket | 4 | 8 | |
| S1022 - 2 sockets | 9 | 8 | 24 |
| S1022s - 1 socket | 4 | 8 | - |
| S1022s - 2 sockets | 9 | 8 | 24 |



Power9 nombre NVMe et AIC par systeme



| Systeme P9 | PCIe Gen3, Gen4 adapteurs AIC (1.6, 3.2, 6.4 TB) * | NVMe U.2 (0.8, 1.6, 3.2, 6.4 TB) |
|------------------------------|-----------------------------------------------------|----------------------------------|
| S914 9009-41G 4 coeurs IBM i | Jusqu'à 4 (0.8, 1.6, 3.2 To)* | |
| S914 9009-41G 6, 8 coeurs | 7 | 4 |
| S10124 9009-42G 1 socket | 7 | 4 |
| S1024 9009-42G 2 sockets | 10 | 4 |
| S1022 9009-22G 1 coeur | 4 (0.8, 1.6 To)** | |
| S914 9009-41A 4 coeurs | 2 (1.6 To) | - |
| S914 9009-41A 6, 8 coeurs | 3 | - |
| S924 9009-42A 1 socket | 3 | - |
| S924 9009-42A 2 sockets | 5 | - |

© IBM Corp.



IBM i



Gestion NVMe sur IBM i

Gestion des disques NVMe – SST (IBM i 7.4 & 7.5)

- Create NVMe Namespace
- Delete NVMe Namespaces
- Sanitize/Erase NVMe device
 - écraser les données ou modifier la clé cryptographique des namespaces sur le NVMe.
- Format NVMe device to prepare device for IBM i
- Display NVMe namespaces
- Display NVMe device
- Create NVMe namespaces to match another NVMe device
- Create NVMe namespaces to pair with Active Mirror protected NVMe units
 - Option pour remplacer NVMe en erreur sur une paire miroir. Voici les étapes:
 1. ASP est examiné pour localiser le NVMe dont la protection miroir est suspendue
 2. Nouveaux namespaces sont créés sur le nouveau NVMe installé à la place de l'ancien
 3. Les nouveaux namespaces sont mise en pair avec les namespaces de l'unité NVMe de l'étape 1.
 4. La protection miroir est reconstruite sur le nouvel NVMe et les données sont copiées sur le nouvel NVMe.
- Display NVMe health

Gestion NVMe avec IBM i 7.3

- Pas de menu gestion NVMe sous SST comme sur IBM i 7.4 et 7.5
- Utilisation de macros commandes lancé sur le microcode sous SST.
- <https://www.ibm.com/docs/en/i/7.3?topic=management-nvme>

| Description | macro commande |
|-------------------------------------|----------------|
| Affichage unités NVMe et namespaces | NVMEDISPLAYNS |
| Affichage unités NVMe | NVMEDISPLAY |
| Création namespaces | NVMECREATENS |
| Formatage NVMe | NVMEFORMAT |
| Suppression namespaces | NVMEDELETENS |
| Nettoyage/suppression NVMe | NVMESANITIZE |

Endurance NVMe et durée de vie

```
MAIN                               IBM i Main Menu                               System:  IBMI4
Select one of the following:

  1. User tasks
  2. Office tasks
  3. General system tasks
  4. Files, libraries, and folders
  5. Programming
  6. Communications
  7. Define or change the system
  8. Problem handling
  9. Display a menu
 10. Information Assistant options
 11. IBM i Access tasks

 90. Sign off

Selection or command
==> CALL PGM(QSMGSSTD) PARM('NVMEGAUGE' X'00000009' 'SSTD0100' X'00000000')
```

```
F3=Exit   F4=Prompt   F9=Retrieve   F12=Cancel   F13=Information Assistant
F23=Set initial menu
```

```
MA + A                                     20/007
```



```
Display Spooled File
File . . . . . : QPCSMVRT                Page/Line  1/35
Control . . . . : _____            Columns   1 - 78
Find . . . . . :
* . . . . 1 . . . . 2 . . . . 3 . . . . 4 . . . . 5 . . . . + . . . . 6 . . . . + . . . . 7 . . . . + . . . .
PERCENTAGE USED (100% MEANS END OF LIFE): 0%
DATA UNITS READ (NUM 512 BYTE DATA UNITS): 53267
DATA UNITS WRITTEN (NUM 512 BYTE DATA UNITS): 6115427
HOST READ COMMANDS: 918982
HOST WRITE COMMANDS: 17685965
CONTROLLER BUSY TIME (MINUTES): 20
POWER CYCLES: 7
POWER ON HOURS: 27
UNSAFE SHUTDOWNS: 4
MEDIA AND DATA INTEGRITY ERRORS: 0
NUMBER OF ERROR INFORMATION LOGS: 65
WARNING COMPOSITE TEMPERATURE TIME (MINUTES): 0
CRITICAL COMPOSITE TEMPERATURE TIME (MINUTES): 0
TEMPERATURE SENSOR 1: 38 C
TEMPERATURE SENSOR 2: 35 C
TEMPERATURE SENSOR 3: 35 C
More . . .
F3=Exit  F12=Cancel  F19=Left  F20=Right  F24=More keys
Mâ + A 03/022
```

```

_
      Display NVMe Detailed Health Information

Serial      Resource
Number      Name
YOYAF1YBY08S  DC05
Type Model  Physical Location
58FE       U78CD.001.FZHBU50-P2-C6

NVMe serial number . . . . . : S4ECNYOM200292
Model information . . . . . : PCIe3 6.4TB NVMe Flash Adapter III x8
Firmware level . . . . . : MA25MA25

Warning flags . . . . . : 00

Composite temperature . . . . . : 34 C
Warning temperature time . . . . . : 0 minutes
Critical temperature time . . . . . : 0 minutes

Available spare space . . . . . : 100 %
Available spare threshold . . . . . : 10 %
Percentage life remaining . . . . . : 100 %

More...

F3=Exit      F5=Refresh      F12=Cancel
```



Commandes SQL

- La vue QSYS2.SYSDISKSTAT contient des informations sur les disques et reconnaît les unités NVMe

```
select disk_model, disk_type,  
       case when unit_type = 1 then 'SSD' else  
         'Spinning' end as "Type of disk",  
       case when unit_nvme = 1 then 'NVMe' else 'Not NVMe'  
       end as "NVMe indicator", percent_used from sysdiskstat;
```

- Exemple de mix SSD et NVMe

| DISK_MODEL | DISK_TYPE | Type of disk | NVMe indicator | PERCENT_USED |
|------------|-----------|--------------|----------------|--------------|
| 0099 | 198E | Spinning | Not NVMe | 7.848 |
| 0050 | 198C | Spinning | Not NVMe | 0.006 |
| 0205 | 6B7D | SSD | NVMe | 0.006 |
| 0205 | 6B7D | SSD | NVMe | 0.006 |
| 0205 | 6B7D | SSD | NVMe | 0.006 |

- Exemple Tous les NVMe du système (avec ACS Run SQL Scripts)

```
1 select disk_model, disk_type,  
2 case when unit_type = 1 then 'SSD' else 'Spinning'  
3 end as "Type of disk",  
4 case when unit_nvme = 1 then 'NVMe' else 'Not NVMe'  
5 end as "NVMe indicator", percent_used from qsys2.sysdiskstat;
```

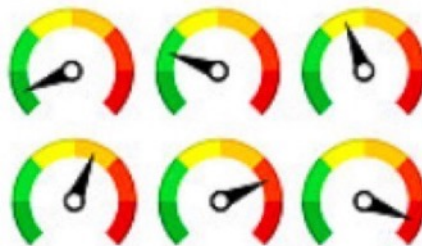
| DISK_MODEL | DISK_TYPE | Type of disk | NVMe indicator | PERCENT_USED |
|------------|-----------|--------------|----------------|--------------|
| 0205 | 6B7D | SSD | NVMe | 4.902 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.855 |
| 0205 | 6B7D | SSD | NVMe | 1.853 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |
| 0205 | 6B7D | SSD | NVMe | 1.852 |

Done: 16 rows retrieved.

NVME_INFO – View

Enhanced yet again to include more columns of insight

| | | |
|-----------------|----------------------|---------------|
| IOP_NAME | PAGES_MAXIMUM | FORMAT_CRYPTO |
| BUS_NUMBER | NAMESPACES_AVAILABLE | OVERWRITE |
| LOCATION_CODE | NAMESPACES_USER | SECURE_ERASE |
| PAGES_AVAILABLE | NAMESPACES_MAXIMUM | BLOCK_ERASE |
| PAGES_USED | NAMESPACES_CONFIGURE | CRYPTO_ERASE |



Navigator for i

IBM Navigator for i

Search RCHWACO.rch.stglabs.ibm.com timmr

NVMe Devices

Actions

| Resource Name | Hardware Model | Life Remaining (%) | Spare Capacity (%) | Spare Capacity Threshold (%) | Number Media Errors | Number Unsafe Shutdowns | Composite Temperature (Celsius) | Firmware Level |
|---------------|-------------------------|--------------------|--------------------|------------------------------|---------------------|-------------------------|---------------------------------|----------------|
| DC01 | 1.6TB NVMe Gen4 U.2 SSD | 100% | 100% | 10% | 0 | 1 | 35 | REV.SN49 |
| DC03 | 3.2TB NVMe Gen4 U.2 SSD | 100% | 100% | 10% | 0 | 1 | 33 | REV.SN49 |
| DC04 | 3.2TB NVMe Gen4 U.2 SSD | 100% | 100% | 10% | 0 | 1 | 35 | REV.SN49 |
| DC05 | 1.6TB NVMe Gen4 U.2 SSD | 100% | 100% | 10% | 0 | 1 | 33 | REV.SN49 |

Total Rows: 4

IBM Navigator for i

System

- System Status
- System Operator Messages
- History Log
- Disk Status
- NVMe Devices**
- SSD Status
- Watch Sessions
- Exit Programs
- SQL Services

Columns...

Available Columns

- Device Type
- Device Model
- Serial Number
- Hardware Serial
- Spare Capacity Threshold Met
- Backup Device Failed
- Persistent Memory Read-Only
- Device Read-Only
- Reliability Degraded
- Temperature Exceeded Threshold
- 512-Byte Data Units Read

Current Columns

- Resource Name
- Hardware Model
- Life Remaining (%)
- Spare Capacity (%)
- Spare Capacity Threshold (%)
- Number Media Errors
- Number Unsafe Shutdowns
- Composite Temperature (Celsius)
- Firmware Level

Reset Default OK

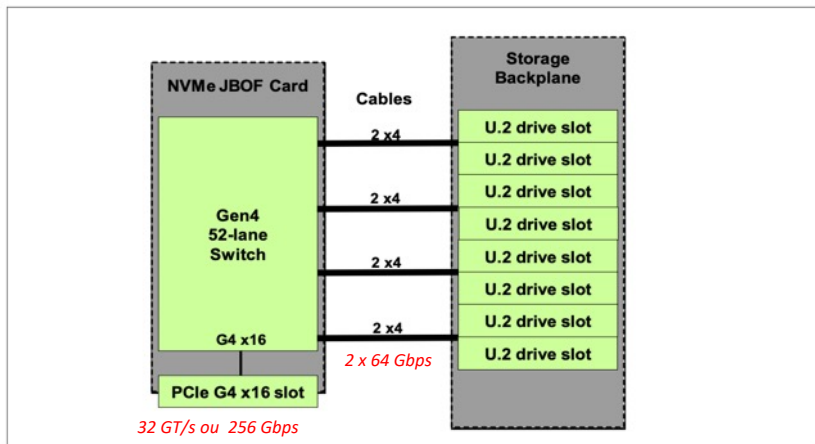
A decorative graphic on the left side of the slide. It features several blue, semi-transparent 3D cubes of varying sizes. One large cube is at the top left. Another medium cube is below it and to the right. A third medium cube is further down and to the left. A small cube is to the right of a large, light-blue, semi-transparent platform. The platform has a white top surface and a grey base. The text 'IBM i' is positioned to the right of the platform.

IBM i

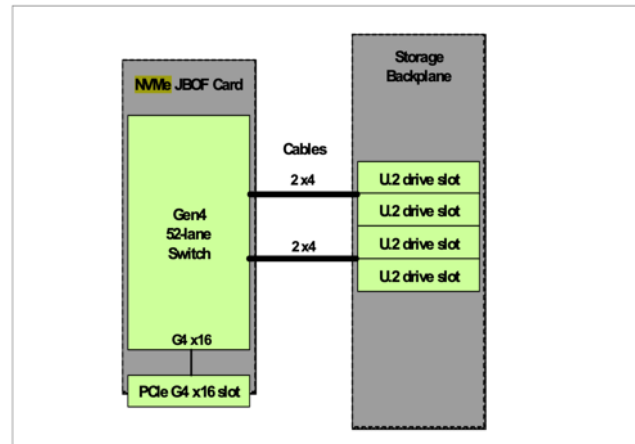
Performances NVMe

Architecture Cage de disque NVMe et carte JBOF*

S1024 / S1014



S1022 / S1022s



Cage de disque EJ1Y - Jusqu'à 8 NVMe

NVMe attaché au processeur par la carte JBOF

Chaque connecteur de la carte JBOF gère 2 NVMe

Carte sur slot Pcie Gen4 x16 = 16 lignes

Debit de 16 lignes x 16 Gbps/ ligne = **256 Gbps de bande passante**

Chaque NVMe a un debit theorique de 4 lignes * 16Gbps = **64 Gbps**

Le slot PCIe G4 x16 supportant la carte JBOF supporte une bande passante de 32 GT/s ou 256 Gbps

Cage de disque EJ1X

Jusqu'à 4 NVMe

NVMe attaché au processeur par la carte JBOF

Chaque connecteur de la carte JBOF gère 2 NVMe

* Just a Bunch Of = Juste un tas de

NVMe vs SATA / SAS

| | NVMe | SATA | SAS |
|-------------------------------------|-------------|----------|---------|
| Bande passante | 64 Gbps* | 6 Gps | 12 Gbps |
| Nombre E/S par seconde (IOPS) | > 1 000 000 | 100 000 | 400 000 |
| Nbre de commande par File d'attente | 64 000 | 32 | 256 |
| Nbre file d'attente commande | 64 000 | 1 | 1 |
| Latence lecture | 80 us | 1,8 ms** | 150 us |
| Latence écriture | 15 us | 3,6 ms** | 60 us |



* Bande passante pour 1 unité NVMe U.2 . Bande passant de 256 Gbps par carte JBOF

** Latence en ms dû aux contraintes physiques de déplacement bras disque et vitesse des plateaux disques. D'où la nécessité d'avoir des cartes controleurs disques avec cache pour masquer cela

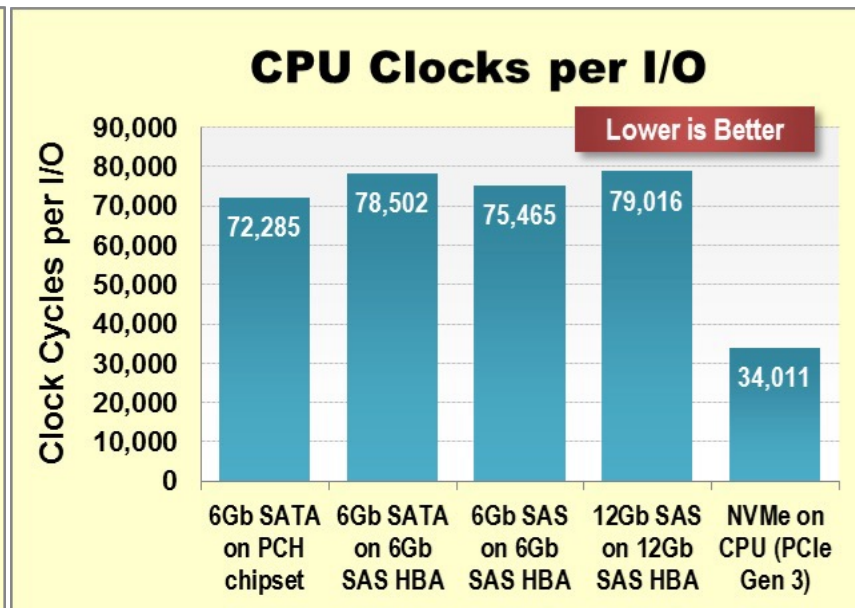
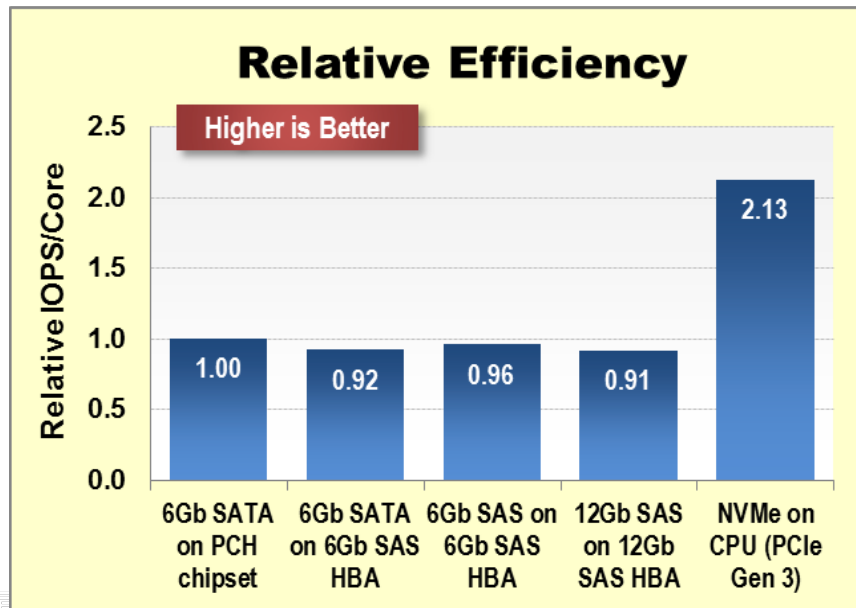
Comparatif SAS vs NVMe

- NVMe – chaque commande E/S et les réponses à ces commandes s'exécutent sur le même processeur et profite du traitement parallèle des processeurs.
- NVMe – chaque thread peuvent avoir leur propre file d'attente indépendante, ce qui permet de ne pas avoir de verouillage E/S.
- NVMe supporte MSI-X et des interruptions de direction prévenant les goulots d'étranglement au niveau processeur et permettre une évolutivité massive à mesure que les systèmes se développent
- NVMe dispose d'un jeu de commandes rationalisé et simple qui utilise **moins de la moitié du nombre d'instructions du processeur pour traiter une demande d'E/S effectuée par SAS ou SATA**, fournissant des E/S par seconde plus élevées par cycle d'instructions processeurs et une latence d'E/S plus faible dans la pile logicielle hôte.
- NVMe prend également en charge des fonctionnalités comme la gestion de l'alimentation,
- NVMe supporte différente technology de “non-volatile memory” comme “NAND flash” et “3D Xpoint” et plus...
- NVMe n'a pas de technology comme les bras de disques et les plateaux de disques tournant → accès plus rapide, moins de latence.
- Avec les performances NVMe plus besoin de cache en écriture pour pallier la latence des écritures dans les architectures SATA / SAS et surtout HDD.

Performances unités NVMe



| FC | CAPACITY | | Read BW | Write BW | Read | Write | DWPD | PCIe | Width | Form Factor | GA date | OS support |
|------------------------|----------|--|---------|----------|----------|----------|----------|------|---------|-------------|---------|---------------------|
| | TB | | GB/s | GB/s | IOPS (K) | IOPS (K) | (5years) | Gen | (Lanes) | | | |
| EC5J, EC5X | 800 | | 1.6 | 1.05 | 380 | 60 | 2.4 | 3 | 4 | U.2 | May-18 | AIX / Linux |
| EC7Q, EC7T | 800 | | 6.5 | 1.5 | 550 | 140 | 2.4 | 4 | 4 | U.2 7mm | Sep-21 | AIX / Linux |
| EKF2 | 800 | | 6.8 | 1.4 | 880 | 100 | 3 | 4 | 4 | U.2 | Nov-22 | IBM i |
| ES1K | 800 | | 4.2 | 1.2 | 500 | 135 | 3 | 4 | 4 | U.2 | Apr-21 | IBM i |
| ES3A / ES3H | 800 | | 6.4 | 1.3 | 800 | 140 | 2.4 | 4 | 4 | U.2 | Jul-22 | AIX / Linux / IBM i |
| EC5G, EC5B, EC6U, EC6V | 1.6 | | 4.7 | 1.9 | 700 | 100 | 5 | 3 | 8 | AIC | Nov-19 | AIX / Linux / IBM i |
| EC5K | 1.6 | | 2.6 | 1.9 | 500 | 80 | 2.4 | 3 | 4 | U.2 | May-18 | AIX / Linux |
| EC7A, EC7B, EC7J, EC7K | 1.6 | | 8.2 | 2.5 | 1000 | 250 | 3 | 4 | 8 | AIC | Nov-20 | AIX / Linux / IBM i |
| EKF3, EKF4 | 1.6 | | 6.8 | 2.7 | 1300 | 215 | 3 | 4 | 4 | U.2 | Nov-22 | AIX / Linux / IBM i |
| ES1E, ES1F | 1.6 | | 7.0 | 2.5 | 100 | 220 | 3 | 4 | 5 | U.2 | Jul-20 | AIX / Linux / IBM i |
| ES3B, ES3C | 1.6 | | 7.0 | 2.5 | 1300 | 260 | 3 | 4 | 4 | U.2 | Jul-22 | AIX / Linux / IBM i |
| EC5C, EC5D, EC6W, EC6X | 3.2 | | 6 | 3 | 910 | 170 | 5 | 3 | 8 | AIC | Nov-19 | AIX / Linux / IBM i |
| EC5L | 3.2 | | 3.2 | 1.9 | 500 | 80 | 2.4 | 3 | 4 | U.2 | May-18 | AIX / Linux |
| EC7C, EC7D, EC7L, EC7M | 3.2 | | 8.2 | 4.2 | 1500 | 250 | 3 | 4 | 8 | AIC | Nov-20 | AIX / Linux / IBM i |
| EKF5, EKF6 | 3.2 | | 6.8 | 4.0 | 1400 | 350 | 3 | 4 | 4 | U.2 | Nov-22 | AIX / Linux / IBM i |
| ES1G, ES1H | 3.2 | | 7.4 | 3.4 | 1500 | 250 | 3 | 4 | 4 | U.2 | Jul-21 | AIX / Linux / IBM i |
| ES3D, ES3E | 3.2 | | 7.0 | 4.1 | 1600 | 280 | 3 | 4 | 4 | U.2 | Jul-22 | AIX / Linux / IBM i |
| EC5E, EC5F, EC6Y, EC6Z | 6.4 | | 6 | 3 | 910 | 170 | 5 | 3 | 8 | AIC | Nov-19 | AIX / Linux / IBM i |
| EC5V, EC5W | 6.4 | | 7.2 | 3.8 | 1500 | 250 | 3 | 4 | 4 | U.2 | Jul-20 | AIX / Linux / IBM i |
| EC7E, EC7F, EC7N, EC7P | 6.4 | | 8.2 | 4.2 | 1500 | 250 | 3 | 4 | 8 | AIC | Nov-20 | AIX / Linux / IBM i |
| EKF7, EKF8 | 6.4 | | 6.8 | 3.8 | 1400 | 325 | 3 | 4 | 4 | U.2 | Nov-22 | AIX / Linux / IBM i |
| ES3F, ES3G | 6.4 | | 7.0 | 4.1 | 1600 | 280 | 3 | 4 | 4 | U.2 | Jul-22 | AIX / Linux / IBM i |



© IBM Corp. 2023

Comparatif performance NVMe / SAS / SATA

| Latency | Bus | Media | Read Lat. (us) | Write Lat. (us) | Read (IOPs) | Write (IOPs) | Read Tp (GB/s) | Write Tp (GB/s) | Approx. \$ Scale (2Q '19) | Cost |
|---------|-----------------|----------------|----------------|-----------------|--------------------------|--------------|----------------|-----------------|---------------------------|------|
| ↓ | Memory (in CEC) | DRAM | <1 | <1 | Not a Persistent Storage | | | | 20x | ↑ |
| | | 3DXP | <1 | <1 | Not a Persistent Storage | | | | 15x | |
| | | LL Flash (HMS) | TBD | TBD | Persistent Storage | | | | | |
| ↓ | PCIe (NVMe) | 3DXP | <10 | <10 | 550K | 500K | 2.4 | 2.0 | 10x | ↓ |
| | | LL Flash | <20 | <20 | 800K | 240K | 3.5 | 3.4 | 1.1x | |
| | | Flash | <90 | <25 | 1500K* | 250K | 6.4 | 3.8 | 1x | |
| | | QLC Flash | >150 | >60 | 80K | 25K | 2.0 | 0.6 | 0.5x | |
| ↓ | SAS | Flash | 150 | 60 | 420K | 50K | 2.2 | 1.6 | 0.7 to 1x | ↓ |
| | SATA | Flash | 1.8ms | 3.6ms | 93K | 25K | 0.5 | 0.5 | | |
| | NL-SAS / SATA | HDD | >ms | >ms | 200 | 200 | 0.15 | 0.15 | 0.1x | |
| ↓ | | TAPE | "secs" | "secs" | "slow" | "slow" | "slow" | "slow" | | ↓ |

SCM: 3DXP from Intel/Micron. Bytes addressable in DIMM (Apache Pass) and Block addressable(M.2/U.2/AIC..) in NVMe interface.

NVMe/SCM: Performance numbers are of Intel's Optane PCIe Gen 3 x4 Add in Card. Endurance 30 DWPD.

NVMe/LL Flash: Performance numbers are of Samsung's zSSD, Gen 3 x4. NVMe QLC Client Flash are based on Intel 660p M.2, Gen 3 x4.

NVMe/Flash*: PCIe Gen 4 x4 U.2 devices. Gen 4 x8 Add In Card will have additional bandwidth and IOPS performance.

SAS SSD: Assumes 12G dual port active/active. Performance of single port operation (typical) expected to be lower.

IOPs and Latencies: Normally measured on a random 4K ops. * <1us for 1K transfer utilizing Persistent Log Buffer feature

Data throughput: Normally measured on a large sequential 256KB ops

Generateur Workload pour test performance

Generateur utilisé : BatchWL

- Requêtes de base de données d'E/S DB2 ou natives
- La base de données utilise :
 - Fichiers avec clé
 - Petites bases de données utilisant des enregistrements de 256 octets
 - Grosses bases de données utilisant des enregistrements de 32 Ko
 - Les fichiers sont uniques pour chaque exécution de thread pour éliminer les verrouillages
 - Les fichiers sont en une taille en dessous de 3 Go pour réduire les effets de cache
- Chaque exécution de thread exécute une operation E/S sur son fichier unique
- Les opérations E/S séquentielles accèdent chaque enregistrement par clé incrémentale
- Les operations E/S aléatoires accident chaque enregistrement par une clé aléatoire
- Les données de performance sont collectées par thread et un total des threads est reporté comme résultat

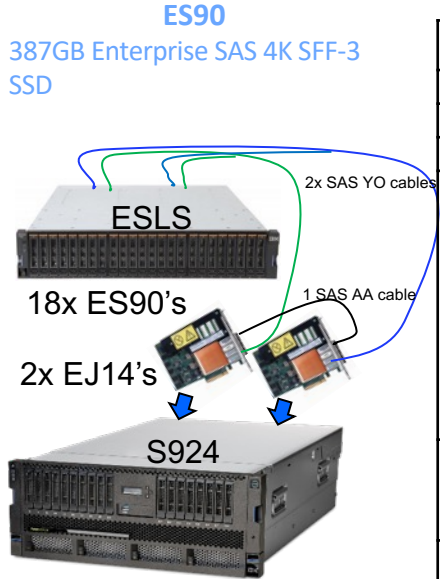
Workloads utilisés (6)

- Extreme “4 coins” pour stresser le software and le hardware,
 - via des petites operations aléatoires et des larges operations séquentielles avec 100% de lectures ou 100% d’écritures
 - + courtes mises à jour aléatoires (50/50 lecture/écriture) et grosses mises à jour séquentielles

Configurations Testés

RAID-5 vs. Mirrored

Mirrored vs. Mirrored

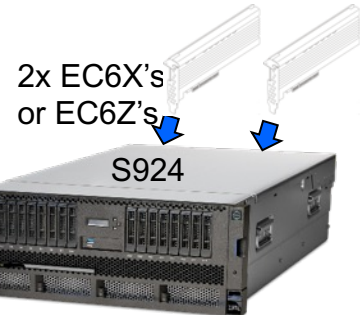


| SAS versus NVMe Comparison | | | | | | | | |
|----------------------------|---------------|------------|-----------|-----------|------------|-----------|-----------|------------|
| | Parameters | Units | SAS | NVMe | Difference | SAS | NVMe | Difference |
| Protection | RAID Type | 0,1,5,6,10 | RAID 5 | OS Mirror | | RAID 10 | OS Mirror | |
| Endurance | DWPD | | 10 | 5 | | 10 | 5 | |
| Bill Of Materials | Adapters | # | 2 | 2 | | 2 | 2 | |
| | | FC / CCIN | EJ14/57B1 | EC6Z/58FE | | EJ14/57B1 | EC6X/58FE | |
| | Signal Cables | # | 4 | 0 | | 4 | 0 | |
| | | FC | Various | N/A | | Various | N/A | |
| | Enclosure | # | 1 | 0 | | 1 | 0 | |
| | | FC | ESLS | N/A | | ESLS | N/A | |
| Drives | # | 18 | 2 | | 18 | 2 | | |
| | FC | ES90 | EC6Z | | ES90 | EC6X | | |
| Capacity | Per Drive | GB | 387 | 6400 | | 397 | 3200 | |
| | # of Arrays | # | 2 | 1 | | 2 | 1 | |
| | Total | TB | 6.192 | 6.4 | 3.4% | 3.483 | 3.2 | -8.1% |
| Real Estate | PCIe Slots | # | 2 | 2 | | 2 | 2 | |
| | Rack Space | Rack Us | 2 | 0 | | 2 | 0 | |
| | Total | Us | 2.2 | 0.2 | -90.9% | 2 | 0.2 | -90.9% |

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



EC6X – PCIe3
NVMe 3.2TB



List pricing as of 12/6/2019 assumes MTM = 9009-42A

All testing was done with native IBM i device driver.

| FC | Description | CAPACITY TB | Read BW GB/s | Write BW GB/s | Read IOPS (K) | Write IOPS (K) | Latence Lecture (us) | Latence écriture (us) | DWPD (5years) | PCIe Gen | Width (Lanes) |
|------------|----------------------------------------------|----------------|-----------------|------------------|------------------|-------------------|----------------------------|-----------------------------|------------------|----------|------------------|
| EC6W, EC6X | PCIe3 x8 3.2 TB NVMe Flash adapter for IBM i | 3.2 | 6 | 3 | 910 | 170 | 110 | 30 | 5 | 3 | 8 |
| EC7L, EC7M | PCIe4 3.2 TB NVMe Flash Adapter for IBM i | 3.2 | 8.2 | 4.2 | 1500 | 250 | 80 | 20 | 3 | 4 | 8 |

NVMe AIC plus récent – meilleures performances

Comparatif capacité et performance

RAID-5 vs. Mirroring

Mirroring vs. Mirroring

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



- Conclusion: Beaucoup de vert indique des performances consistantes bien meilleures du NVMe.
- Ecritures aléatoires 255B montre pour NVMe une performance moindre dû à la fonction miroir qui est exécuté par la CPU au lieu du contrôleur SAS

| Coleur | Description |
|--------|---------------|
| Verte | Avantage NVMe |
| Jaune | Entre +/-10% |
| Rouge | Avantage SAS |

NVMe performance P9 vs P10

| | | Max T/s | Max Go/s | Max T/s/coeur | Max T/s | Max Go/s | Max T/s/coeur | Max débit | Max efficacité CPU |
|------------|-------------------------------|------------|----------|---------------|-----------|----------|---------------|-----------|--------------------|
| Config | Workload | P10 | | | P9 | | | P9 → P10 | |
| NVMe natif | 32 Ko écritures séquentielles | 92 380 | 3,027 | 83 534 | 91173 | 2,987 | 71242 | 1% | 17% |
| | 4 Ko écritures aléatoires | 750 000 | 3,072 | 244 444 | 743 333 | 3,045 | 238 408 | 1% | 3% |
| | 32 Ko Lectures séquentielles | 446 667 | 14,635 | 81 858 | 366 667 | 12,014 | 61 013 | 22% | 34% |
| | 4 Ko Lectures aléatoires | 5 8000 000 | 23,757 | 406 505 | 1 370 000 | 5,612 | 355 844 | 323% | 14% |
| | 4 Ko Mises à jour aléatoires | 251 948 | 1,032 | 85 554 | 223 911 | 0,917 | 65 231 | 13% | 31% |

© 2023 IBM Corporation

Nbre de bras disques et taille

Tests réalisés avec un workload client. Ceci n'est pas un résultat officiel IBM

| | 32 namespaces 200GB 2 NVMe | 32 namespaces 200GB 8 NVMe | 64 namespaces 100GB 8 NVMe | 16 namespaces 400GB 2 NVMe | 1 namespace 3.2TB 2 NVMe | 12 namespaces 200GB 3 NVMe |
|---------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-----------------------------|-------------------------------|
| Temps de réponse Ecriture | 0,0357 ms | 0,0230 ms | 0,0232 ms | 0,0395 ms | 0,0275 ms | 0,0269 ms |
| Temps de réponse Lecture | 0.1629 ms | 0,1230 ms | 0,1221 ms | 0.1631 ms | 0.1530 ms | 0,1521 ms |
| Disque occupé % max Ecriture | 24,86 % | 20,51 % | 10,83 % | 42,74 % | 99,35 % | 41,17 % |
| Disque occupé % max Lecture | 18,69 % | 13,86 % | 7,034 % | 34,37 % | 100 % | 43,85 % |
| | 32 namespaces 200GB 2 NVMe | 32 namespaces 200GB 8 NVMe | 64 namespaces 100GB 8 NVMe | 16 namespaces 400GB 2 NVMe | 1 namespace 3,2TB 2 NVMe | 12 namespaces 200GB 3 NVMe |
| E/S max Ecriture | 504544 | 643437 | 655964 | 501399 | 566204 | 618092 |
| E/S max Lecture | 76021 | 77741 | 78235 | 77747 | 78232 | 76640 |

Tiroir d'extension NVMe - NED24

IBM i



Nouveau tiroir d'extension pour NVMe - Date de disponibilité 19 mai 2023



- Réduction de 62% du coût (\$/To) en connexion directe vs solutions SSD en mode SAS [1]
- Bande passante multipliée par 10 pour les E/S (Go/s) et opérations jusqu'à 3,7 (ES/s) vs SSD/SAS [2]
- Possibilité max de 1,2 Po de stockage sur un E1080 (153 To par tiroir et maximum 8 tiroirs) [3]
- Réduction de la consommation d'énergie et plus de 85 % de capacité supplémentaire par watt vs SSD/SAS
- Extension de la capacité des solutions Flash Cache requises pour les workloads critiques
- Capacité de démarrage natif pour un maximum de 24 partitions (sans protection miroir)

[1] Sur la base des tarifs et de la capacité actuels d'IBM pour les plus grands disques de stockage de classe entreprise dans l'actuel FC #ESLS (1,55 TB) par rapport au nouveau FC #ESR0 (6,4 TB).

[2] Basé sur les mesures internes d'IBM du nouveau FC #ESR0 attaché à un E1080 comparé au FC #ESLS actuel attaché à un E980 exécutant 100% des opérations de lecture (256KB pour GB/s et 4KB pour IO/s) dans un environnement IBM i.

[3] Basé sur la capacité maximale et la consommation d'énergie de 153,6 To et 1000 Watts pour FC #ESR0 utilisant des disques NVMe U.2 Enterprise par rapport à 37,2 To et 450 Watts pour FC #ESLS utilisant des disques SSD de classe entreprise.

Version IBM i supporté

- IBM i 7.5 TR2
- IBM i 7.4 TR8
- IBM i 7.3 Base – restrictions :
 - ✓ Supporte le mode iVirt (IBM i dans i)
 - ✓ Supporté à travers VIOS
 - ✓ Pas de protocole de communication NVMe – utilise le vSCSI (limité bande passante)
dans les 2 cas

NED24 vue arrière

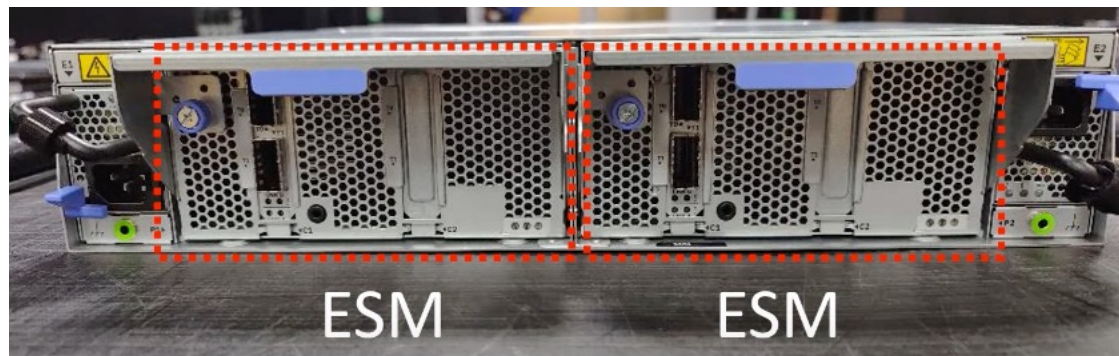
Rear

2 Power
Supplies

2 Cable Sockets
to Power10
server



Enclosure Services Manager
Minimum 2 par tiroir
Chaque ESM gère 12 NVMe



NVMe Expansion Drawer Devices

| Dispo | Description | Stockage Total (Go) |
|-------|-------------------------------------------------------|---------------------|
| ES5B | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 19,200 |
| ES5D | Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i | 38,400 |
| ES5F | Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i | 76,800 |
| ES5H | Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i | 153,600 |

Stockage maximum avec NED24 NVMe

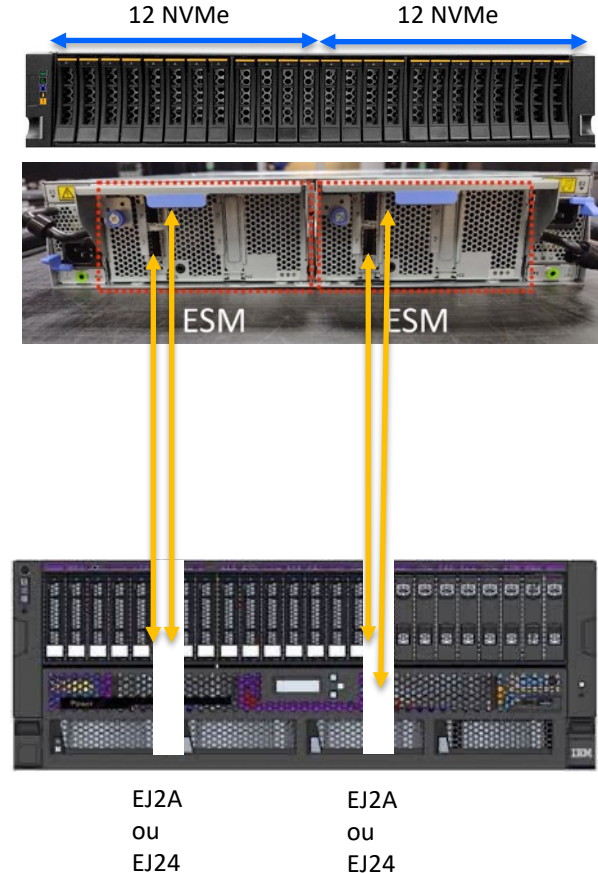
- E1080 - 2 NED24 par nœud, 8 max, 192 unités NVMe
- E1050 - 2 NED24 par système max, 48 unités NVMe*
- S1024 - 1 NED24 par système max, 24 unités NVMe
- S1022 – 1 NED24 par système max, 24 unités NVMe
- S1022s 2x4 coeurs – 1 NED24 par système max, 24 unités NVMe



* E1050 ne supporte pas IBM i

Architecture

- 2 ESM minimum dans chaque tiroir
- 2 cartes Hôte min. et max. sur le serveur (EJ2A ou EJ24)
- "Mode 1 Single path" supporté seulement :
 - chaque ESM est géré par une seule carte Hôte (EJ2A/EJ24)
 - chaque NVMe est géré par une seule carte Hôte
- "Mode dual path" support ultérieur peut-etre:
 - Chaque NVMe accède aux 2 cartes Hôtes eliminant la dépendance à la carte hôte

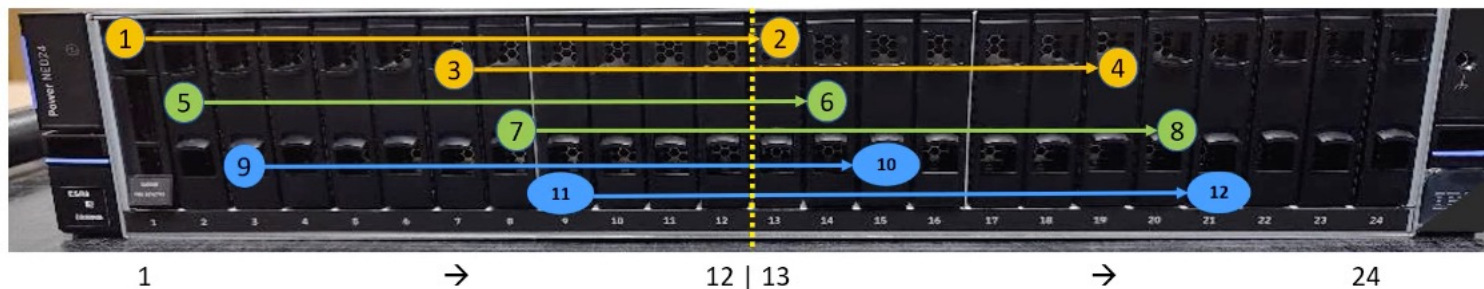


Règles de placement

1. Les 24 emplacements NVMe ont une performance indentique
2. Pour une fiabilité, disponibilité et sécurité (RAS) maximum :
 - Utilisez équitablement les 2 modules ESM1 pour une meilleure circulation de l'air. Ne pas remplir tous les slots du même côté
3. Mirroring ? Verifiez que les miroirs sont sur différents modules ESM1
4. Utilisez le modèle suivant:

Positions dans le tiroir

1,13, 7,19, 2,14 8,20 3,15 9,21
 4,16 10,22 5,17 11,23 6,18. 10,24



Questions fréquentes

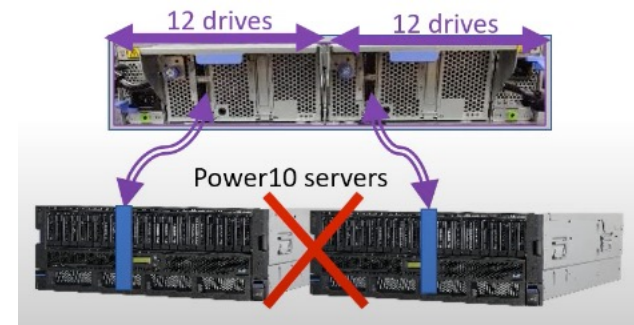
1. Est ce que on peut avoir 1 seul ESM dans le tiroir NED24
Pas actuellement

2. Cables Power10 <-> NED24
Achetés par paire (avec un seul Feature Code #FC)
Les 2 paires doivent être même Feature Code sur un NED24

3. Peut-on partager un NED24 entre 2 serveurs Power10 ?
Pas actuellement

4. Pourquoi des cables optiques et cuivres? Quel est le différence?
Les cables cuivres sont volumineux et on une longueur maximum de 3 mètres
Les cables optiques sont plus flexibles et ont une longueur de 3, 10 et 20 mètres

5. Quels serveurs supportent le tiroir NED24 ?
Serveurs Power10 seulement



Nombre maximum de NED24 par type de serveur P10?



1



1



2



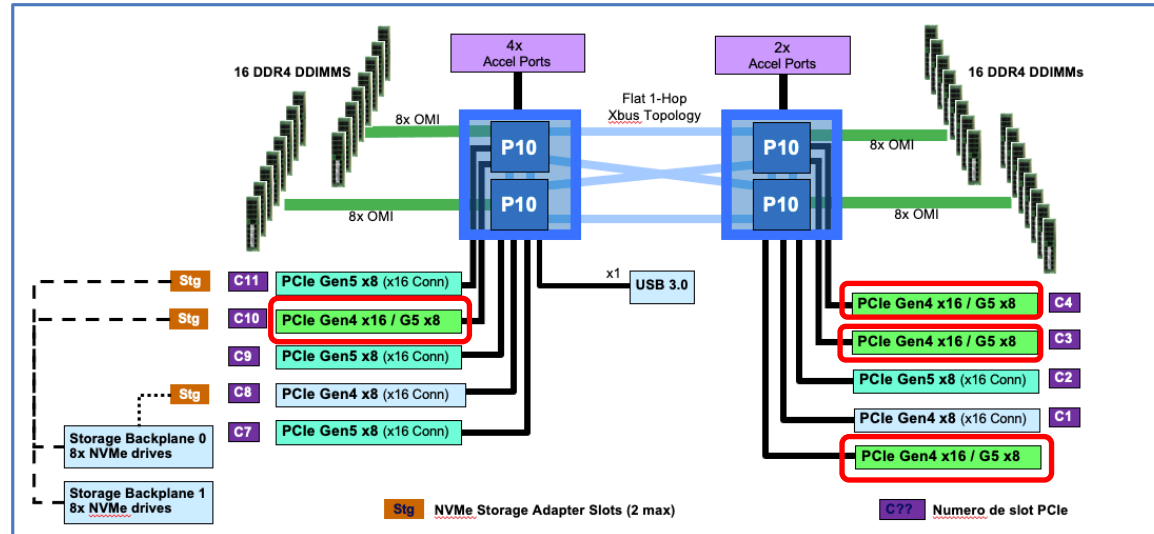
1



Seul les serveurs 2 sockets sont supportés : S1024, S1022 et S1022s 1 socket non supporté!

Restrictions emplacement carte connexion NED24

- Cartes de Connexion au tiroir NED24 se placent dans des slots spécifiques:
 - Slots PCIe5 x8 seulement
 - Slots C0, C3, C4, C10
- **Seul les systèmes 2 sockets sont supportés** car le 1 socket a seulement le slot C10 et le tiroir NED24 a besoin de 2 slots PCIe



Comparatif coûts pour même capacité de stockage

| DISPO | DESCRIPTION | Qté | Prix | Prix Total | Taille disque (GB) | Taille disque Totale (GB) | Coûts /GB |
|--------------------------------------|-------------------------------------------------------------------|-----|---------|----------------|--------------------|---------------------------|---------------------|
| CONFIGURATION DISK SAS 15KRPM | | | | | | | |
| ESLS | EXP24SX SAS Storage enclosure | 1 | €7 689 | €7 689 | | | |
| ESLN | 283GB 15K RPM SAS SFF-2 4k Block Cached Disk Drive IBM i | 24 | €871 | €20 904 | 283 | 6792 | |
| EJ14 | PCIe3 12GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8 | 2 | €11 472 | €22 944 | | | |
| ECE3 | 3.0M SAS AA12 Cable | 2 | €296 | €592 | | | |
| ECDU | 3.0M SAS YO12 Cable (Adapter to Enclosure) | 2 | €296 | €592 | | | |
| | TOTAL | | | €52 721 | | 6792 | 7,76 € / GB |
| CONFIGURATION DISK SSD | | | | | | | |
| ESLS | EXP24SX SAS Storage enclosure | 1 | €7 689 | €7 689 | | | |
| ETK9 | 387GB Enterprise SAS 4K SFF-2 SSD for IBM i | 18 | €2 719 | €48 942 | 387 | 6966 | |
| EJ14 | PCIe3 12GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8 | 2 | €11 472 | €22 944 | | | |
| ECE3 | 3.0M SAS AA12 Cable | 2 | €296 | €592 | | | |
| ECDU | 3.0M SAS YO12 Cable (Adapter to Enclosure) | 2 | €296 | €592 | | | |
| | TOTAL | | | €80 759 | | 6966 | 11,59 € / GB |
| CONFIGURATION EXTENSION NVMe | | | | | | | |
| ESR0 | EXP24U.2 NVMe SAS Storage Enclosure | 1 | €19 333 | €19 333 | | | |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 10 | €1 284 | €12 840 | 800 | 8000 | |
| ECLS | 3.0M CXP x16 Copper Cable Pair for PCIe4 Expansion | 2 | €1 437 | €2 874 | | | |
| EJ2A | PCIe4 x16 to CXP Converter Adapter (support AOC) | 2 | €2 773 | €5 546 | | | |
| ESM1 | ESM Module w/PCIe4 x16 2-port connection to CXP converter adapter | 2 | €6 495 | €12 990 | | | |
| | TOTAL | | | €53 583 | | 8000 | 6,70 € / GB |
| CONFIGURATION NVME INTERNE | | | | | | | |
| EJ1Y | Storage Backplane with eight NVMe U.2 drive slots | 2 | €3 168 | €6 336 | | | |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 10 | €1 284 | €12 840 | 800 | 8000 | |
| EJUU | Front IBM Bezel for 16 NVMe-bays Rack Mount | 1 | €317 | €317 | | | |
| | TOTAL | | | €19 493 | | 8000 | 2,44 € / GB |



Le futur du SAS?

- EXP24SX tiroir d'extension SAS (#ESLS) retiré de commercialisation
 - Code dispo Tiroir (30/6/23) et Load Source retiré (23/3/23)
- Pourquoi le retrait de la commercialisation?
 - NVMe est la direction stratégique d'IBM pour tout le stockage interne et externe attaché directement.
 - NVMe fournit un stockage plus rapide à un coût plus faible.





IBM i



Annexes

Protection par mot de passe - NVMe

IBM i 7.5 TR 1
IBM i 7.4 TR 7



- Les unités NVMe supportés par IBM i prennent en charge le chiffrement automatique (SED) sur Power10
- Cependant, les clés utilisées pour chiffrer et déchiffrer les données ne sont pas protégées, donc une fois qu'un "méchant" obtient l'appareil, il peut accéder à toutes vos données.
- IBM i peut désormais protéger par mot de passe les unités NVMe dédiés sur les serveurs Power10
- Utilise la spécification Opal Security Subsystem Class (SSC) du Trusted Computer Group (TCG) pour le stockage
- En créant une politique de mot de passe dans IBM i et en ajoutant les unités NVMe, les unités peuvent protéger la confidentialité des données contre des accès non autorisés quand les unités quittent le contrôle d'IBM i

<https://www.ibm.com/support/pages/how-set-nvme-locking-policy-password-ibm-i>

Protection par mot de passe NVMe - Verrouillage des appareils NVMe

- Les mots de passe sont stockés dans le PKS (Platform Keystore) des partitions LPAR gérées par PowerVM
- L'appareil NVMe se verrouille en cas de perte d'alimentation principale/réinitialisation à froid PCIe
- L'unité NVMe se verrouillera quand :
 - une opération de suppression DLPAR est effectuée sur une unité NVMe
 - une mise hors tension de maintenance simultanée de maintenance simultanée est effectuée sur l'appareil
 - La partition est démarré
 - L'unité NVMe est réinitialisé
- Tant que l'unité NVMe reste dans la partition, il se déverrouille automatiquement lorsque l'alimentation est rétablie sur l'unité à l'aide du mot de passe de stratégie stocké dans le PKS
- Les unités NVMe protégés par mot de passe sont verrouillés lorsqu'ils sont déplacés vers une autre partition LPAR
- La saisie du mot de passe à partir des outils de service système (SST) déverrouille l'appareil et le rend lisible et inscriptible, mais il est à nouveau verrouillé lorsque l'appareil est éteint et rallumé.



IBM i



IBM e-config cloud

Example S1024 48c – 16 NVMe ES1K 800Go

| Product | Description | Qty |
|----------|-------------------------------------------------------------------|-----|
| 9105-42A | S1024 48 cores:IBM Power S1024 | 1 |
| 0040 | Mirrored System Disk Level, Specify Code | 1 |
| 0267 | IBM i Operating System Partition Specify | 1 |
| 2145 | Primary OS - IBM i | 1 |
| 4650 | Rack Indicator- Not Factory Integrated | 1 |
| 4970 | One Processor of 5250 Enterprise Enablement | 1 |
| 4974 | Full 5250 Enterprise Enablement | 1 |
| 5000 | Software Preload Required | 1 |
| 5228 | PowerVM Enterprise Edition | 48 |
| 5550 | Sys Console On HMC | 1 |
| 6458 | Power Cord 4.3m (14-ft), Drawer to IBM PDU (250V/10A) | 4 |
| 9441 | New IBM i License Core Counter | 1 |
| 9448 | Other IBM i License Core Counter | 48 |
| 9703 | Language Group Specify - French | 1 |
| ALH0 | Expert Care Indicator | 1 |
| EB3S | AC Titanium Power Supply - 1600W for Server (200-240 VAC) | 4 |
| EB48 | 1GbE Base-T Transceiver RJ45 | 2 |
| EB74 | IBM i 7.4 Indicator | 1 |
| EC2U | PCIe3 2-Port 25/10Gb NIC&ROCE SR/Cu Adapter | 1 |
| EJ1Y | Storage Backplane with eight NVMe U.2 drive slots | 2 |
| EJUU | Front IBM Bezel for 16 NVMe-bays BackPlane Rack-Mount | 1 |
| ELSG | ES1K Load Source Specify (800 GB 4K NVMe U.2 SSD PCIe4 for IBM i) | 1 |
| EM6N | 32GB (2x16GB) DDIMMs, 3200 MHz, 8GBIT DDR4 Memory | 16 |
| ENSB | 400 GB IBM i NVMe Load Source Namespace size | 1 |
| EPFD | One CUoD Static Processor Core Activation for EPGD | 48 |
| EPGD | 24-core Typical 2.75 to 3.90 Ghz (max) Power10 Processor | 2 |
| ES1K | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 16 |
| ESC6 | S&H-b | 1 |
| EU0K | Operator Panel LCD Display | 1 |
| EU19 | Cable Ties & Labels | 1 |
| EXA3 | 3 YEAR, ADVANCED EXPERT CARE | 1 |

#0040 – Mirroring IBM i pour NVMe

#5000 - Pré-chargement IBM i

#EJ1Y - 2 cages NVMe U.2

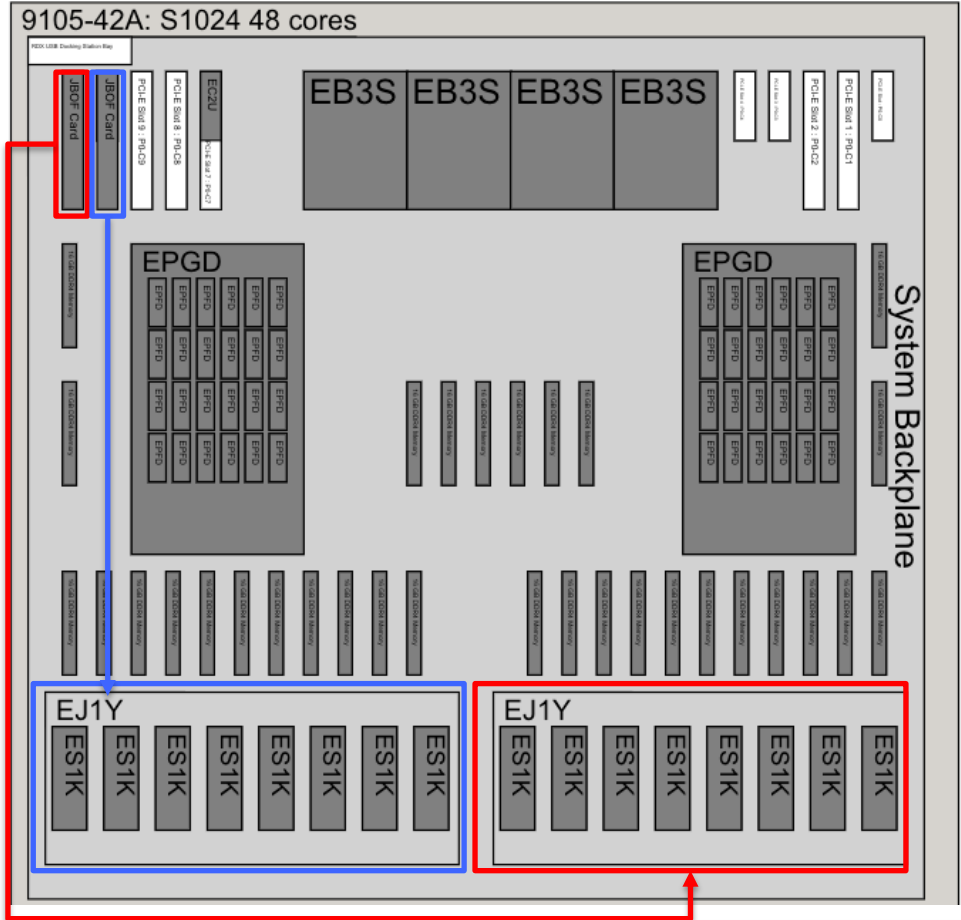
#ENSB – namespace 400 Go pour Load source. Un namespace de 400 Go est crée sur un autre NVMe pour le miroir du load source

Au total 2 namespaces de 400 Go seront créés sur NVMe load source et son NVMe miroir

#ELSG – Load Source sur NVMe ES1K

#ES1K – 16 NVMe 800Go

E-config vue placement



IBM
i
35
YEARS



E-config S1014 4 coeurs

#ENSM – Configuration des namespaces sur le load source et son miroir. Taille depend du dispo **#ENSB** (400Go) ou **#ENSA** (200Go). Seul les 2 premiers NVMe seront configurés avec des namespaces definit par **#ENSB** ou **#ENSA**. Namespaces ou disques logiques seront ajoutés à l'ASP 1

Dans l'exemple, on a des NVMe de 800Go : 2 namespaces de 400Go seront créés sur le NVMe load source et sont miroir (**#ENSB**).

Après démarrage du systeme et de la LPAR IBM i, on pourra créer d'autres namespaces sur le Load source, son miroir et les autres untiés NVMe.



| Product | Description | Qty |
|----------|-------------------------------------------------------------------|-----|
| 9105-41B | S1014 4 coeurs:IBM Power S1014 | 1 |
| 0040 | Mirrored System Disk Level, Specify Code | 1 |
| 0267 | IBM i Operating System Partition Specify | 1 |
| 2145 | Primary OS - IBM i | 1 |
| 4650 | Rack Indicator- Not Factory Integrated | 1 |
| 5000 | Software Preload Required | 1 |
| 5228 | PowerVM Enterprise Edition | 4 |
| 5550 | Sys Console On HMC | 1 |
| 5899 | PCIe2 4-port 1GbE Adapter | 1 |
| 6458 | Power Cord 4.3m (14-ft), Drawer to IBM PDU (250V/10A) | 2 |
| 9441 | New IBM i License Core Counter | 1 |
| 9448 | Other IBM i License Core Counter | 4 |
| 9703 | Language Group Specify - French | 1 |
| ALH0 | Expert Care Indicator | 1 |
| EB3S | AC Titanium Power Supply - 1600W for Server (200-240 VAC) | 2 |
| EB75 | IBM i 7.5 Indicator | 1 |
| EJ1Y | Storage Backplane with eight NVMe U.2 drive slots | 1 |
| EJXU | Front IBM Bezel for 16 NVMe-bays BackPlane Rack-Mount | 1 |
| ELSG | ES1K Load Source Specify (800 GB 4K NVMe U.2 SSD PCIe4 for IBM i) | 1 |
| EM6N | 32GB (2x16GB) DDIMMs, 3200 MHz, 8GBIT DDR4 Memory | 2 |
| ENSB | 400 GB IBM i NVMe Load Source Namespace size | 1 |
| ENSM | Specify Code Configure all IBM i Namespaces | 1 |
| EPFT | One Processor Core Activation for EPG0 | 4 |
| EPG0 | 4-core Typical 3.0 to 3.90 Ghz (max) Power10 Processor | 1 |
| ES1K | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 8 |
| ESC6 | S&H-b | 1 |
| EU0K | Operator Panel LCD Display | 1 |
| EU19 | Cable Ties & Labels | 1 |
| EXA3 | 3 YEAR, ADVANCED EXPERT CARE | 1 |

S1022s IBM i NVMe exemple (Proc #EPGR 2x4 coeurs)

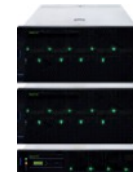


| Product | Description | Qty |
|----------|-------------------------------------------------------------------|-----|
| 9105-22B | S1022s IBM i natif:IBM Power S1022s | 1 |
| 0040 | Mirrored System Disk Level, Specify Code | 1 |
| 0267 | IBM i Operating System Partition Specify | 3 |
| 2145 | Primary OS - IBM i | 1 |
| 4650 | Rack Indicator- Not Factory Integrated | 1 |
| 5228 | PowerVM Enterprise Edition | 8 |
| 5550 | Sys Console On HMC | 1 |
| 6458 | Power Cord 4.3m (14-ft), Drawer to IBM PDU (250V/10A) | 2 |
| 9441 | New IBM i License Core Counter | 8 |
| 9448 | Other IBM i License Core Counter | 3 |
| 9703 | Language Group Specify - French | 1 |
| ALH0 | Expert Care Indicator | 1 |
| EB3N | AC Titanium Power Supply - 2000W for Server (200-240 VAC) | 2 |
| EB48 | 1GbE Base-T Transceiver RJ45 | 2 |
| EB74 | IBM i 7.4 Indicator | 1 |
| EB75 | IBM i 7.5 Indicator | 1 |
| EC2T | PCIe3 LP 2-Port 25/10Gb NIC&ROCE SR/Cu Adapter | 1 |
| EEPZ | IBM i Enablement for 9105-22B with two #EPGR Processors | 1 |
| EJ1X | Storage backplane with four NVMe U.2 drive slots | 2 |
| EJ05 | Front IBM Bezel for 8 NVMe-bays BackPlane Rack-Mount | 1 |
| ELYA | ES3A Load Source Specify (800 GB 4K NVMe U.2 SSD PCIe4 for IBM i) | 1 |
| EM6N | 32GB (2x16GB) DDIMMs, 3200 MHz, 8GBIT DDR4 Memory | 8 |
| ENSA | 200 GB IBM i NVMe Load Source Namespace size | 1 |
| EPFR | One Processor Core Activation for EPGR | 8 |
| EPGR | 4-core Typical 3.0 to 3.90 Ghz (max) Power10 Processor | 2 |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 8 |
| ESC5 | S&H-a | 1 |
| EU0K | Operator Panel LCD Display | 1 |
| EU19 | Cable Ties & Labels | 1 |
| EXA3 | 3 YEAR, ADVANCED EXPERT CARE | 1 |



Carte PCIe NVMe / IBM i / IBM Power Systems modèles

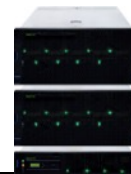
- NVMe est uniquement pris en charge dans l'unité centrale des serveurs POWER9. Il n'est pas pris en charge dans un tiroir d'E / S PCIe Gen3 en raison à la fois de problèmes de bande passante et de tests requis.
- Le S914 prend en charge jusqu'à 3 NVMe. Le S924 prend en charge jusqu'à 6 NVMe (3 avec un seul module processeur). **Pour serveur scale-out modèle A: 9009-41A (S914), 9009-42A (S924)**
- Le S914 prend en charge jusqu'à 7 NVMe. Le S924 prend en charge jusqu'à 7 NVMe (avec un seul module processeur). **Pour serveur scale-out modèle G: 9009-41G, 900942G**
- Le système S914 P05 à 4 cœurs est limité à 2 carte NVMe de 1,6 To ou 3,2To uniquement. Aucun mélange de disques NVMe et SAS n'est autorisé sur le 4-coeurs (10 disques SAS max ou 2 NVMe max).
- E980 prend en charge jusqu'à 8 NVMe par tiroir (6 dans premier tiroir, 8 dans chaque tiroir 2, 3 et 4 pour un total maximum de 30)
- Notez que l'utilisation de plusieurs cartes PCIe NVMe affectera les emplacements Pcie disponibles, ce qui peut affecter les configurations système avec plusieurs partitions et l'utilisation de tiroirs externes peut affecter la quantité maximale de PCIe NVMe.



Dernière annonce 2021 : plus de cartes NVMe supportées sur S924

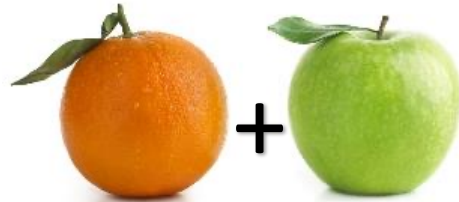


- S924 MTM 9009-42G mis à jour 12 Janvier 2021
- Total de 14 cartes NVMe supportées:
 - 10 NVMe AIC
 - 4 NVMe U.2
- Jusqu'à 89.6 TB de capacité de stockage NVMe Enterprise Class



Carte PCIe NVMe/IBM i/IBM Power Systems Hardware

- Les disques HDD / SSD internes sont autorisés dans l'unité centrale «avec» la carte PCIe NVMe (sauf le S914 4 cœurs).



- Un « storage backplane » n'est pas nécessaire lorsque des adaptateurs NVMe de carte PCIe sont utilisés comme source de charge pour IBM i, mais il peut être utile pour une future extension de stockage.

Protection par mot de passe NVMe - activée et désactivée

- Les unités NVMe configurées dans la partition sont protégés par mot de passe en créant une politique de protection par mot de passe NVMe sur la partition LPAR
- Création, modification ou suppression des stratégies de protection par mot de passe via System Service Tools (SST)
- Sauvegarde complète du système recommandée avant de créer une politique de protection par mot de passe
- Les mots de passe sont stockés sur le PKS et les NVMe dans les partitions LPAR

- Si l'unité NVMe est ajouté à une partition dotée d'une politique de protection par mot de passe, définissez un mot de passe sur le périphérique NVMe via SST

- Si un périphérique NVMe protégé par mot de passe est déplacé vers une autre partition LPAR, il sera verrouillé et l'espace de noms de ce périphérique NVMe sera illisible/inscriptible.
- La saisie du mot de passe à partir de System Service Tools (SST) déverrouille l'appareil et le rend lisible et inscriptible, mais il se verrouille à nouveau lorsque l'appareil est redémarré.
- La réinitialisation d'usine est le seul moyen de déverrouiller si vous ne connaissez pas le mot de passe NVMe
- La réinitialisation d'usine détruit toutes les données dans l'espace de noms

Protection par mot de passe NVMe - Services IBM i



- **CREATE_LOCKING_POLICY**
 - Créer une stratégie de mot de passe pour LPAR
- **DELETE_LOCKING_POLICY**
 - Supprimer la stratégie de mot de passe actuelle et supprimer la protection par mot de passe pour tous les périphériques NVMe sous la stratégie
- **ADD_DEVICE_LOCKING_POLICY**
 - Ajouter des périphériques NVMe à la stratégie de mot de passe actuelle
- **REMOVE_DEVICE_LOCKING_POLICY**
 - Supprimer la protection par mot de passe des appareils NVMe
- **CHANGE_DEVICE_LOCKING_POLICY**
 - Changer le mot de passe de la politique
- **UNLOCK_DEVICE**
 - Déverrouiller les appareils NVMe
- **LOCKING_POLICY_INFO**
 - Afficher une liste de tous les périphériques NVMe dans une partition logique, ainsi que leur capacité de verrouillage et leur état
- **FACTORY_RESET_DEVICE**
 - Effacez toutes les données des périphériques NVMe afin qu'elles puissent être utilisées dans les LPAR lorsque les mots de passe sont inconnus



IBM i



Performance NVMe

Carte NVMe AIC (Add In Card)



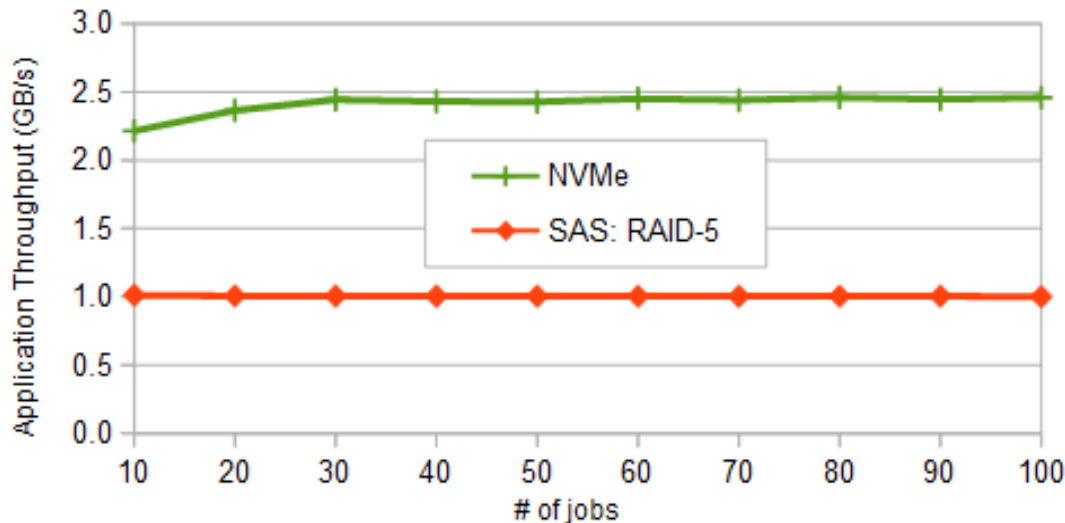
- Les cartes AIC NVMe existent en PCIe Gen3 ou Gen4
- PCIe Gen3 utilisent 8 lignes PCIe @ 8Gbps → **Debit max de 64 Gbps**
- PCIe Gen4 utilisent 8 lignes PCIe @16Gbps → **Débit max de 128 Gbps**
- Meilleures performances avec les cartes EC7J/EC7K, EC7L/EC7M, EC7N/EC7P**

| FC | Description | CAPACITY TB | Read BW GB/s | Write BW GB/s | Read IOPS (K) | Write IOPS (K) | Latence Lecture (us) | Latence écriture (us) | DWPD (5years) | PCIe Gen | Width (Lanes) | Form Factor | GA date |
|------------|-------------------------------------------------|----------------|-----------------|------------------|------------------|-------------------|----------------------------|-----------------------------|------------------|-------------|------------------|----------------|---------|
| EC6U, EC6V | PCIe3 x8 1.6 TB NVMe Flash adapter for IBM i | 1.6 | 4.7 | 1.9 | 700 | 100 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7J, EC7K | PCIe4 LP 1.6 TB NVMe Flash Adapter x8 for IBM i | 1.6 | 8.2 | 2.5 | 1000 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |
| EC6W, EC6X | PCIe3 x8 3.2 TB NVMe Flash adapter for IBM i | 3.2 | 6 | 3 | 910 | 170 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7L, EC7M | PCIe4 3.2 TB NVMe Flash Adapter for IBM i | 3.2 | 8.2 | 4.2 | 1500 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |
| EC6Y, EC6Z | PCIe3 x8 6.4 TB NVMe Flash adapter for IBM i | 6.4 | 6 | 3 | 910 | 170 | 110 | 30 | 5 | 3 | 8 | AIC | Nov-19 |
| EC7N, EC7P | PCIe4 6.4 TB NVMe Flash Adapter x8 for IBM i | 6.4 | 8.2 | 4.2 | 1500 | 250 | 80 | 20 | 3 | 4 | 8 | AIC | Nov-20 |

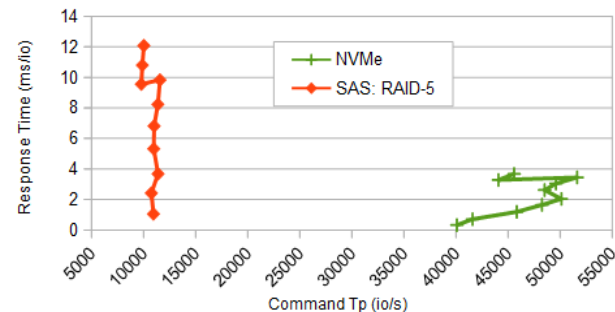
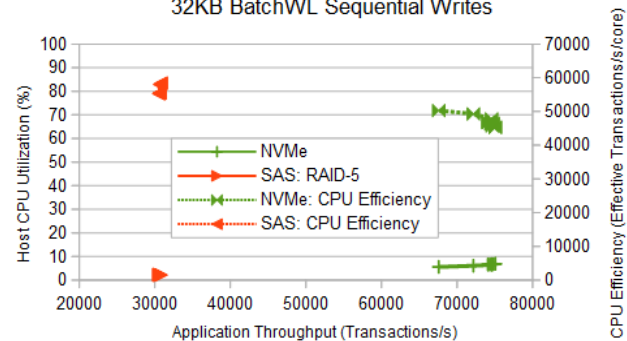
Recommandation – Privilégiez les cartes AIC PCIe Gen4 plutôt que Gen3

Large Sequential Writes: Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS
32KB BatchWL Sequential Writes



OS Mirror NVMe vs RAID-5 SAS
32KB BatchWL Sequential Writes

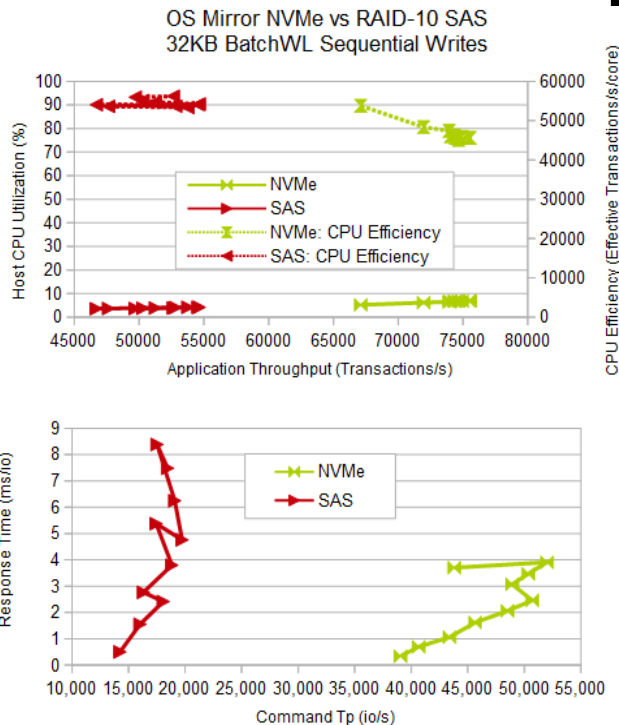
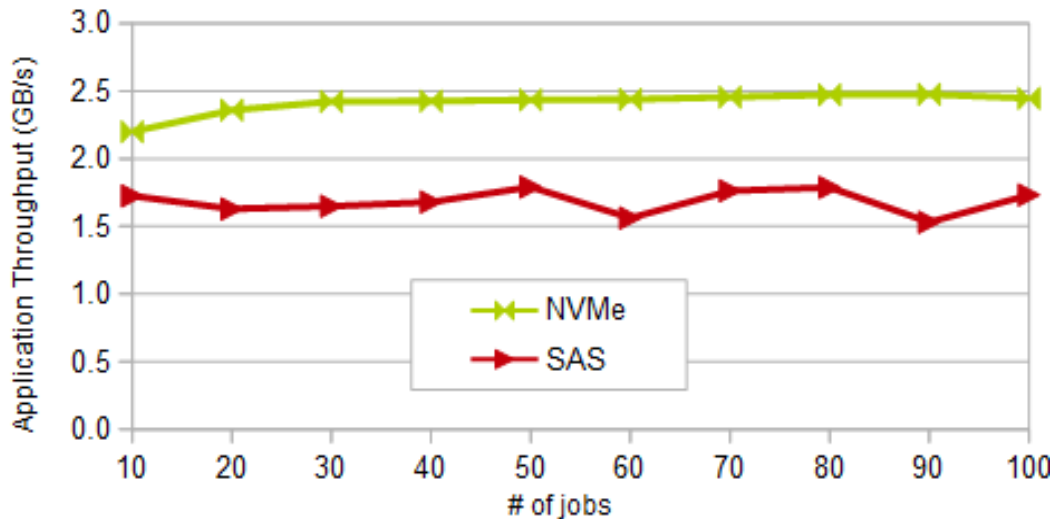


Take-aways:

1. Writing to a mirror is typically 2x faster than to RAID-5/6 ([not new news](#)).
2. There is a CPU utilization efficiency (amount of CPU per IO) cost for doing RAID (mirroring) function in the host CPUs.
3. NVMe write response times compare nicely with writes to the SAS IOA's DRAM cache.

Large Sequential Writes : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
32KB BatchWL Sequential Writes

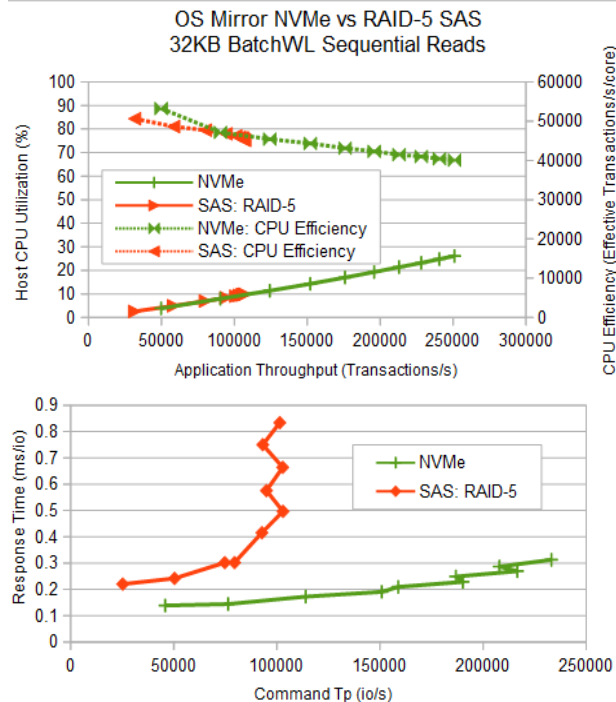
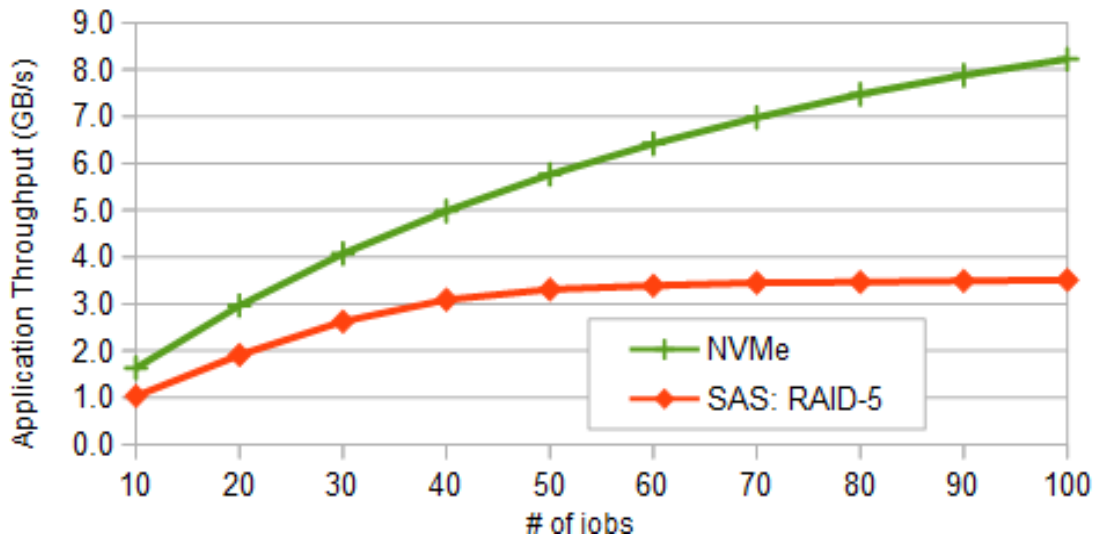


Take-aways:

1. One can make the SAS rig faster by throwing more \$ at it by adding another ESLS drawer (more SAS fabric bandwidth).
2. There is a CPU utilization efficiency cost for doing RAID (mirror) function in the host CPUs.
3. NVMe write response times compare nicely with writes to the SAS IOA's DRAM cache.

Large Sequential Reads : Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS
32KB BatchWL Sequential Reads

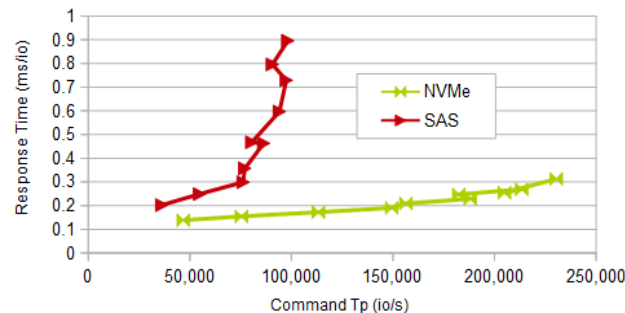
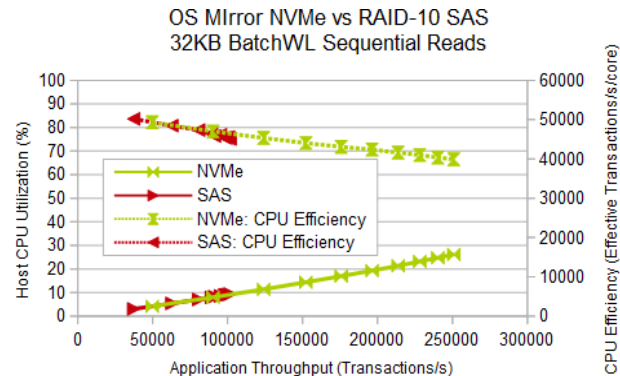
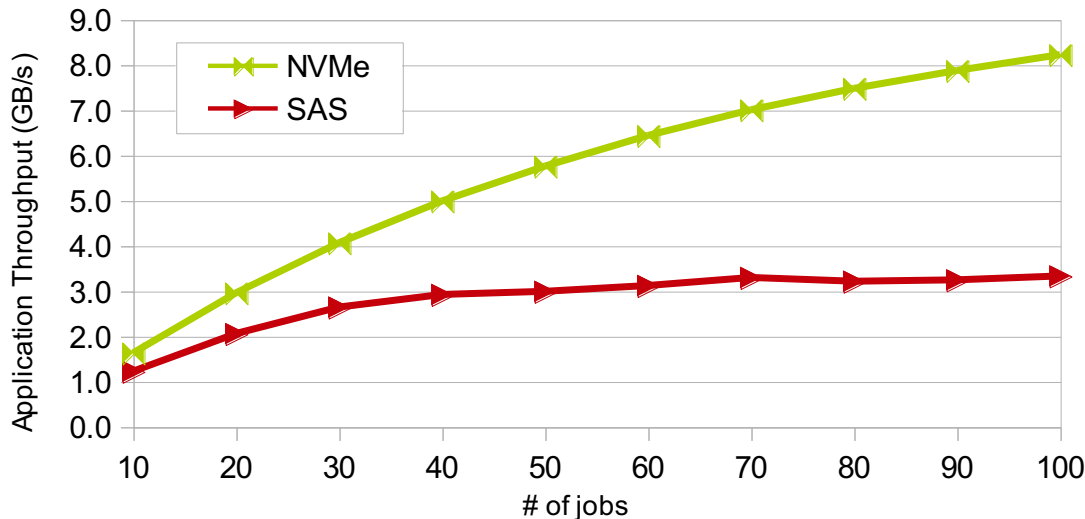


Take-aways:

1. The NVMe drive's PCIe bus is a clear winner over the SAS fabric.
 - Like the big writes, one could double the SAS throughput by adding a 2nd ESLS.
2. NVMe read response times are clear winners. (Thank you PCIe bus !)
3. There is still a slight CPU utilization efficiency cost for doing RAID function in the host CPUs.

Large Sequential Reads : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
32KB BatchWL Sequential Reads

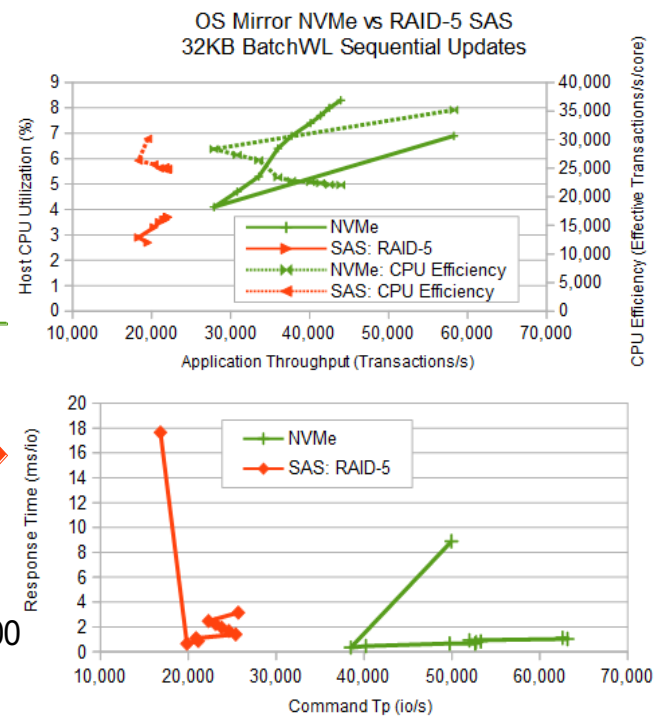
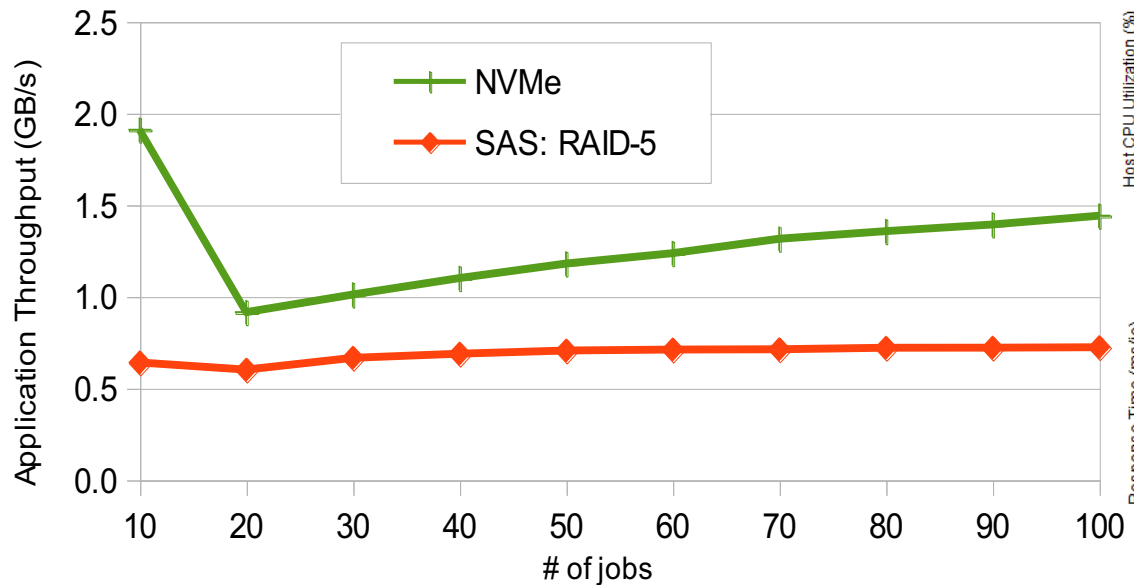


Take-aways:

1. The NVMe drive's PCIe bus is a clear winner over the SAS fabric.
 - Like the big writes, one could double the SAS throughput by adding a 2nd ESLS.
2. NVMe read response times are clear winners. (Thank you PCIe bus !)
3. There is still a slight CPU utilization efficiency cost for doing RAID function in the host CPUs.

Large Sequential Updates : Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS 32KB BatchWL Sequential Updates

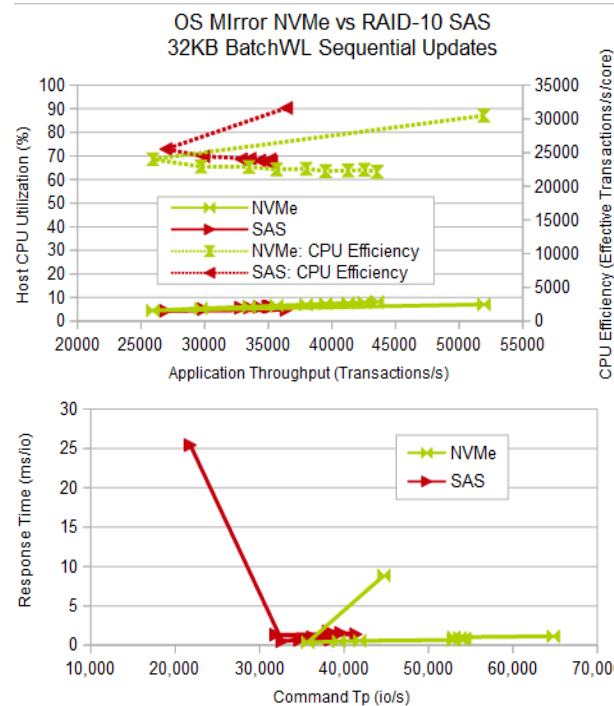
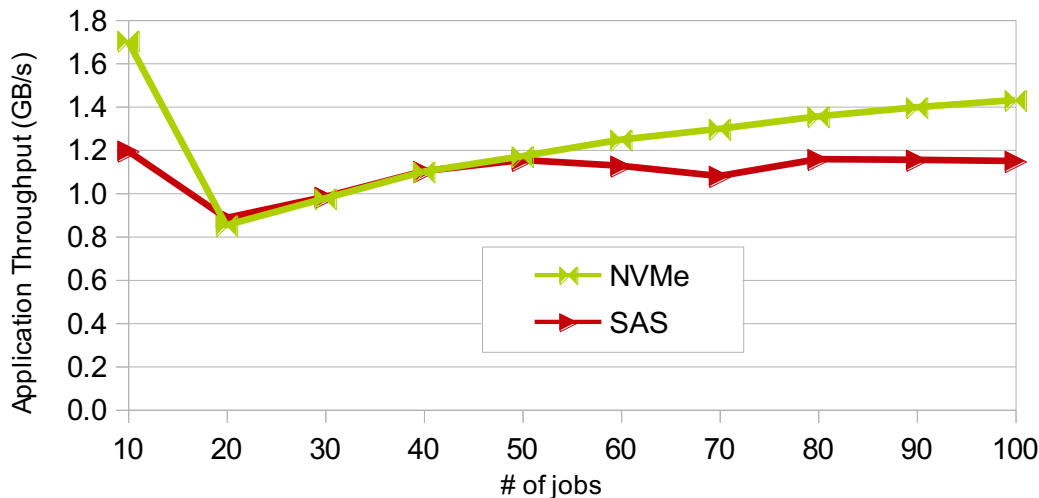


Take-aways:

- (1) The write portion of the update procedure benefits mirroring over RAID-5 (and -6).
- (2) Root cause for the low load (10 jobs) NVMe advantage has yet to be isolated.
 - We will study that closer when gen4 NVMe drives arrive on the scene.

Large Sequential Updates : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
32KB BatchWL Sequential Updates

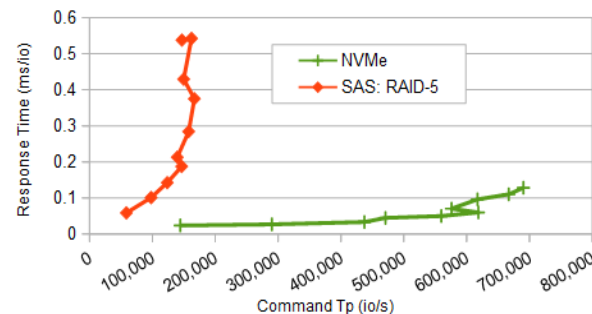
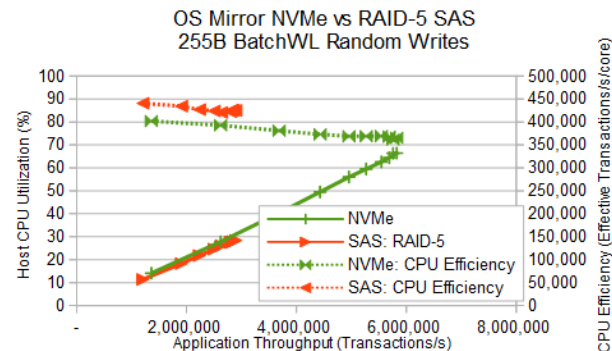
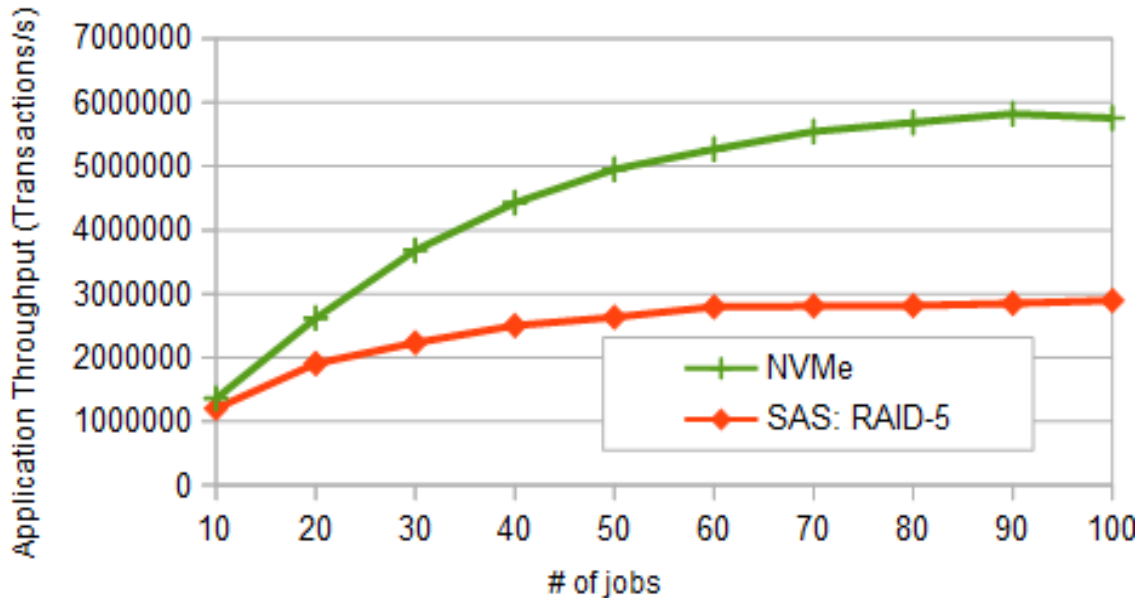


Take-aways:

- (1) A bit cloudier picture than for pure reads and pure writes, and versus RAID-5.
- (2) Yet the 50/50 read/write workload shows a clear advantage for NVMe over SAS in terms of throughput and response time.

Small Op Random Writes : Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS 255B BatchWL Random Writes

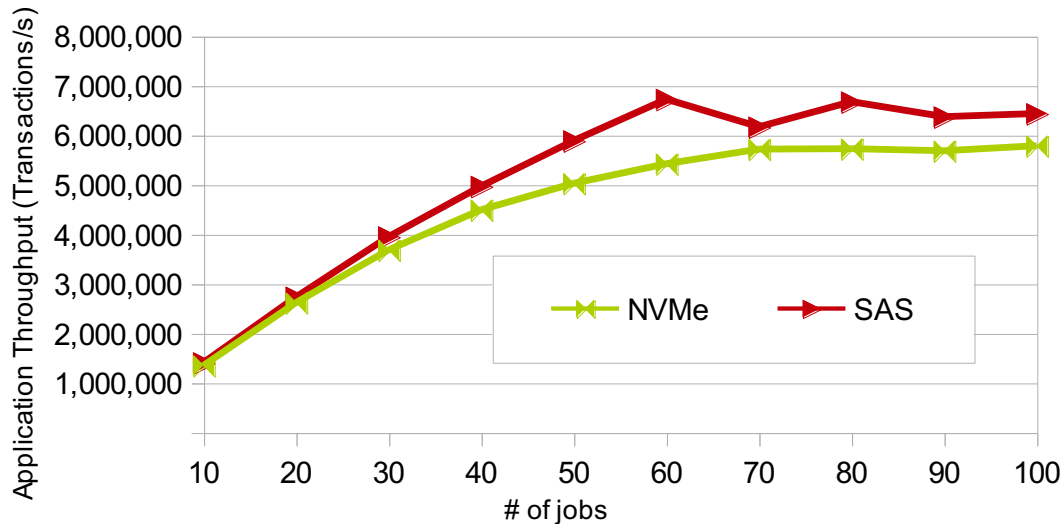


Take-aways:

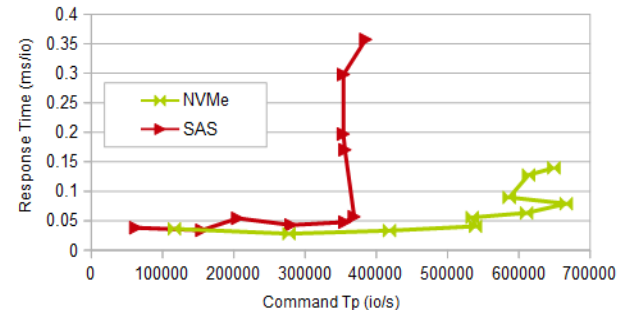
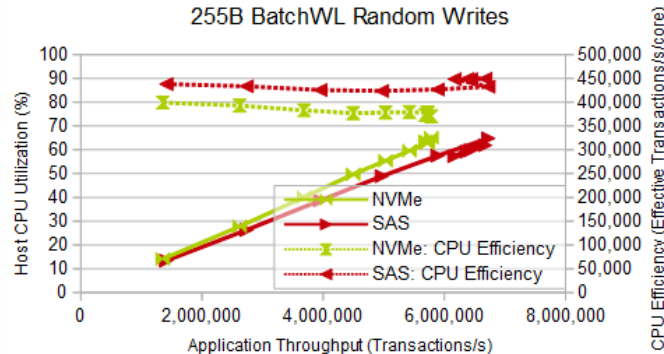
1. Writing to a mirror is **still** faster than to RAID-5/6 😊.
2. CPU utilization efficiency cost for doing RAID function in the host CPUs does allow SAS to use less CPU/io than NVMe.
3. Yet response times for NVMe writes still look A-OK. (i.e. lower than SAS.)

Small Op Random Writes : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
255B BatchWL Random Writes



OS Mirror NVMe vs RAID-10 SAS
255B BatchWL Random Writes

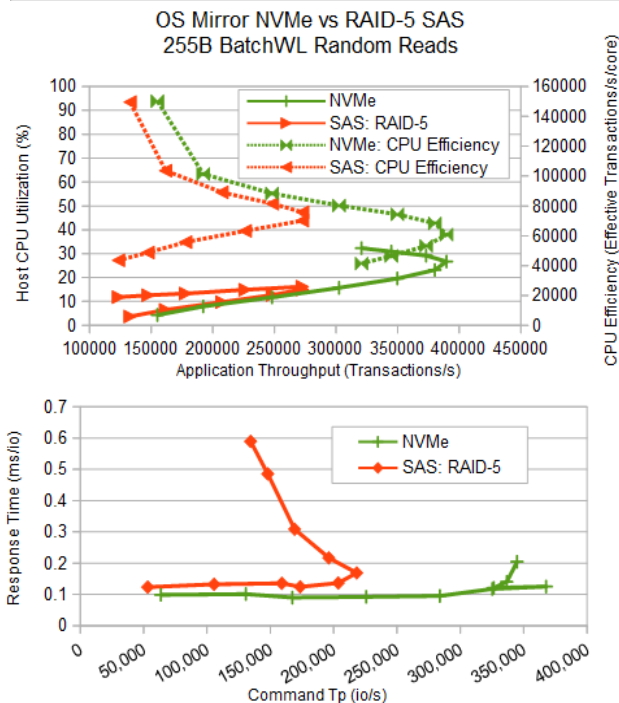
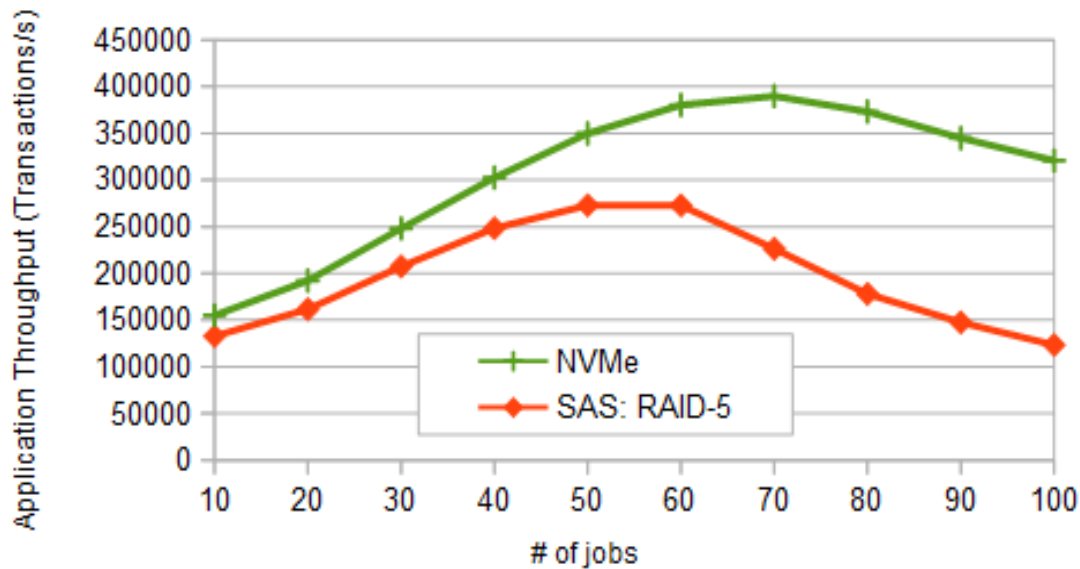


Take-aways:

1. Due to IBM i's Storage Management's excellent data placement (use of buffers), the high transaction rate stresses CPU.
2. CPU utilization efficiency cost for doing RAID function in the host CPUs does allow SAS to use less CPU/io than NVMe.
3. Yet response times for NVMe writes still look A-OK. (i.e. lower than SAS.)

Small Op Random Reads : Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS
255B BatchWL Random Reads

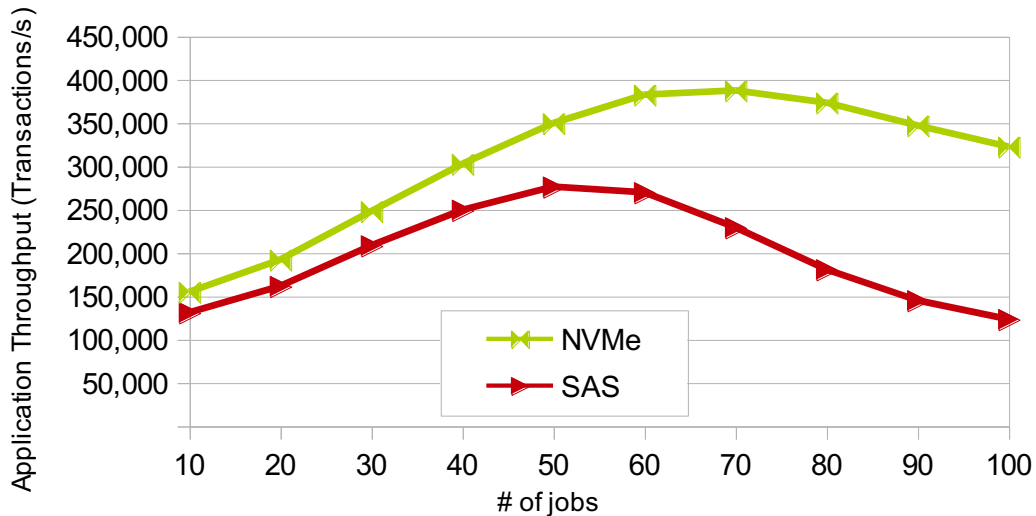


Take-aways:

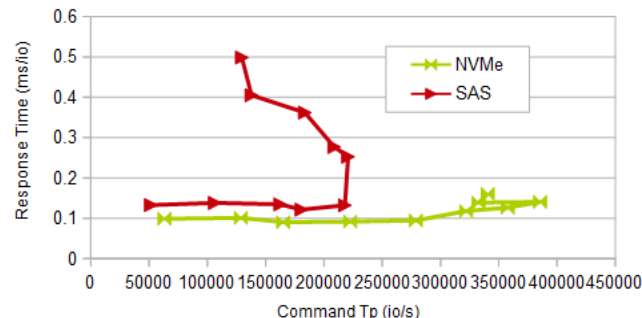
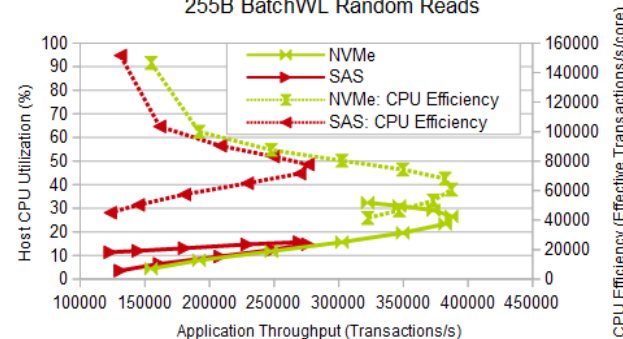
1. High correlation between application and subsystem command throughput confirms that caches do not help truly random reads. Thus the raw capabilities of the 2 NVMe drives versus the 18 SAS drives (and 2 IOAs) shines through.
2. Eliminating the SAS adapter chip and bus lowers response times (latency).
3. Nice to see NVMe host CPU mirroring code impact is actually lower !

Small Op Random Reads : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
255B BatchWL Random Reads



OS Mirror NVMe vs RAID-10 SAS
255B BatchWL Random Reads

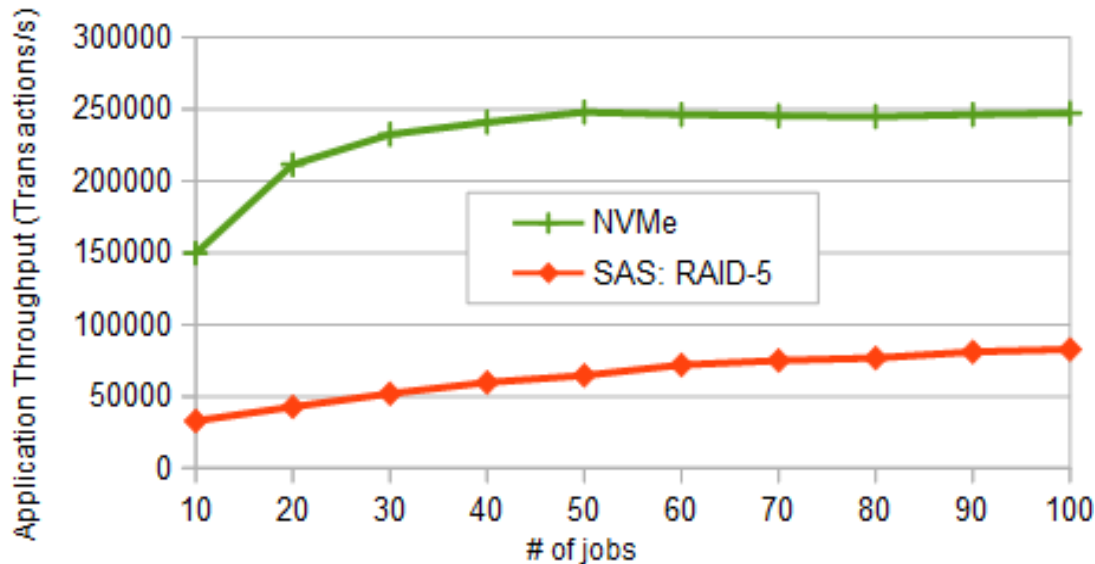


Take-aways:

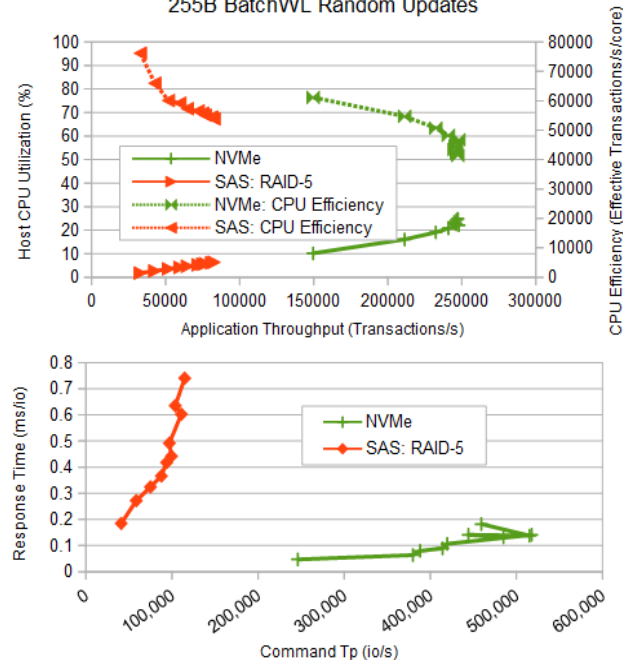
1. High correlation between application and subsystem command throughput confirms that caches do not help truly random reads. Thus the raw capabilities of the 2 NVMe drives versus the 18 SAS drives (and 2 IOAs) shines through.
2. Eliminating the SAS adapter chip and bus lowers response times (latency).
3. Nice to see NVMe host CPU mirroring code impact is actually lower !

Small Op Random Updates : Mirror vs RAID-5

OS Mirror NVMe vs RAID-5 SAS 255B BatchWL Random Updates



OS Mirror NVMe vs RAID-5 SAS 255B BatchWL Random Updates

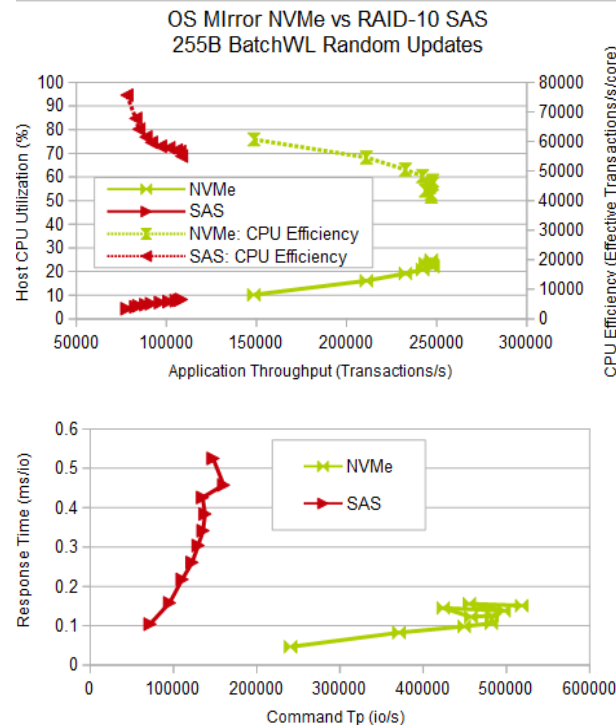
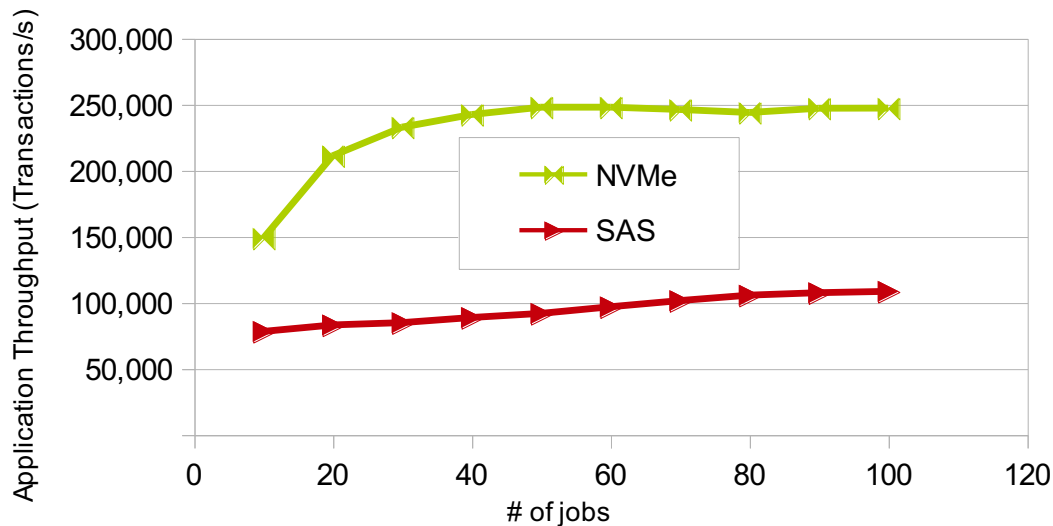


Take-aways:

1. Coupling the RAID-5 write cost and raw NVMe subsystem latency reductions, migrating from SAS RAID-5/6 to NVMe mirrored subsystem will provide LOTS of headroom for application throughput growth.
2. CPU Efficiency at a given throughput is fairly close despite the fact that NVMe uses CPU cycles to manage protection while neither adapter level protection does.

Small Op Random Updates : Mirror vs Mirror

OS Mirror NVMe vs RAID-10 SAS
255B BatchWL Random Updates



Take-aways:

1. Running SAS in a mirror does allow it to support more throughput than in RAID-5 mode, but still not nearly enough to keep up with the mirrored NVMe config.
2. CPU Efficiency at a given throughput is fairly close despite the fact that NVMe uses CPU cycles to manage protection while neither adapter level protection does.

Tiroir d'extension NVMe - NED24

IBM i



NED24 Tiroir d'extension NVMe

- IBM étend son leadership en matière de capacités d'E/S avec le 1er tiroir de stockage NVMe U.2 à connexion directe PCIe du marché. Le nouveau tiroir d'extension NED24 NVMe (#ESR0) est un boîtier d'extension de stockage doté de 24 baies U.2 NVMe de 15 mm.
 - Il prend en charge jusqu'à 24 unités NVME U.2 dans des supports Gen3 ou supérieur de 15 mm.
 - Les supports de 15 mm peuvent accueillir des périphériques NVMe de 7 mm ou de 15 mm.
- Le tiroir d'extension NVMe NED24 est pris en charge sur les serveurs IBM Power E1080, Power E1050, IBM Power S1024, IBM Power S1022, Power S1022s, IBM Power L1022 et IBM Power L1024 par IBM AIX®, IBM i, Linux® et VIOS.
- Chaque tiroir d'extension NVMe NED24 contient deux blocs d'alimentation CA redondants. Les blocs d'alimentation CA font partie de la base du boîtier.
- Le tiroir d'extension NVMe NED24 est connecté à un serveur Power via deux adaptateurs PCIe CXP Converter (#EJ24 ou #EJ2A). Les deux adaptateurs CXP Converter nécessitent l'une des caractéristiques de câble identiques suivantes :
 - #ECLR - 2.0 M Active Câble Optique x16 Paire for PCIe4 Expansion Drawer
 - #ECLS - 3.0 M CXP x16 Cable Pair for PCIe4 Expansion Drawer (paire de câbles en cuivre CXP x16 pour tiroir d'extension PCIe4)
 - #ECLX - Câble optique actif 3.0 M x16 paires pour tiroir d'extension PCIe4
 - #ECLY - Câble optique actif 10 M x16 paires pour tiroir d'extension PCIe4
 - #ECLZ - Câble optique actif 20 M x16 paires pour tiroir d'extension PCIe4
- Le tiroir d'extension NVMe NED24 prend en charge un chemin unique en mode 1 (#ECMS), qui consiste en une connexion x16 unique depuis le serveur hôte à travers chacun des deux ESM (#ESM1) vers l'ensemble des 24 périphériques NVMe.
 - Le commutateur de chaque ESM est configuré pour ne piloter logiquement que 12 des 24 disques NVMe.
 - Cela permet d'avoir un seul chemin d'accès à chacun des 24 périphériques NVMe depuis le serveur hôte.

Tiroir d'extension NVMe



Mécanique

- 2U, 19" Rack
- 24 SFF U.2, Accès frontal
- Installation par client, expédiable en rack

Design PCIe Gen4

- Cartes de connexion Hôte
 - PCIe Gen4 x16 I/F vers Hôte / extension à travers 2 câbles CXP
- ESM (Enclosure Service Module)
 - Jusqu'à 2 cartes PCIe G4 x16 I/Fs vers Hôte via Adapteur Plat
 - 48 lignes PCIe Gen 4
- Cablage CXP – Support PCIe G4 & SRIS
 - Câbles optiques et cuivres
 - U dépendant du hôte I/F adapter (Plate)

2 ESM

- Ajout / Retrait à chaud
- Tolérant aux pannes de courant

2 Bloc d'alimentation

- AC/DC Supply Option – Meet EU 2023 Efficiency Req
 - 180-264VAC 50/60Hz
- N+1 alimentation et refroidissement
- Changement à chaud

Unités de stockage

- 2U/24 SFF U.2 NVMe SSD, PCIe G4
- Périphériques double port
- Emplacement alimentation/Refroidissement >= 25W
- Maintenance concurrente supporté

Fond de panier

- Fond de panier à deux ports (ESM redondant)
- Contrôle de l'alimentation de chaque appareil

SE supporté

AIX 7.2, 7.3

IBM i 7.4, 7.5

Linux

- SLES 15, SLES 16
- RHEL 8, RHEL 9

VIOS 3.1.4.21

HMC 10.2.1040

Firmware **FW1040**

Logicile supporté

- PHYP détient tiroir et câble hôte
 - Chaque unite NVMe est partitionable
 - Supporte Boot

RAS

- Major FRUs
 - Cable Card, ESM, PSU & PDB, Câbles, Mid-Plane, Devices
- Maintenance concurrente
 - Unités, bloc d'alimentation
- R&V Maintenance
 - Cable Cartes, ESM, CXP Câbles
- Concurrent Code Download - Devices
- Redundant Cooling and Power
- Partitionable endpoint / Recovery Domain is the U.2 Device
- Host cable with single-cable failure recovery for each host IO I/F
- Redundant Sideband I/F to drawer
- Full Individual PERST to each Device PCIe I/F(s)

IBM Cable Cards

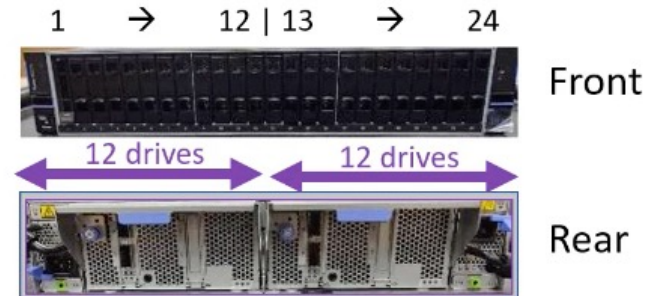
- EJ24 / EJ2A (HHHL/FHHL)
- PCIe Gen4 x16 ports required

Certifications

- FCC: Class A
- Acoustics: Data Center Category 1A
- Environment: ASHRAE A3, ASHRAE A2 with FCM 3 (30W)
- 5-40C, 8-85% RH, 3050m max
- RoHS
- IBM Anti-Smoke

Administration NED24 par HMC

- NED24 géré par HMC
 - Affichage tiroir NED24 et unités NVMe : HMC → Server → E/S (PCIe) ressources
 - Affectation unité NVMe à une LPAR : HMC → LPAR → Adapteurs E/S physiques → “Ajout adaptateur”
 - La même chose que pour les unités NVMe interne
- Chaque unité NVMe est affecté séparément à un différent VIOS, LPAR ou VM
- Unité NVMe peut être utiliser comme Load source ou boot pour tous les Systèmes d’exploitation
- JBOD seulement. (Just a bunch of drives)
- Pas d’options RAID car aucun controleur RAID
- Utilisation du “mirroring” au niveau Système d’exploitation
- Requis
 - HMC v10.2.1040
 - Power10 Firmware v1040



NED24 Systems Admin = HMC

Hardware Management Console YLB21C-GUI

Resources > All Systems > YLB210-9080-HEX-SN7893F38-J104 > Processor, Memory, I/O >

All Resources Search systems, partitions, help topics and more

Save Cancel

Processor, Memory, and I/O

View or change the memory, processor, and physical I/O resource settings for the managed system.
[Learn More](#) →

- Processors
- Memory
- Physical I/O Adapters**

The table displays the physical I/O resources for the managed system. Click the Adapter Info(i) to view information on the selected I/O device. Click I/O Pools to display all the I/O pools found in the managed system and the partitions that are participating in the pools.

I/O Pools Expand All

| Adapter Description | Details | Physical Location Code | Owner |
|----------------------------------------------|----------|----------------------------|--------------------|
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND0.FGD03FD-P0-C2-C0 | Unassigned |
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND0.FGD03FD-P0-C6-C0 | Unassigned |
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND3.FGD03EK-P0-C0-C0 | Unassigned |
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND3.FGD03EK-P0-C2-C0 | Unassigned |
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND3.FGD03EK-P0-C5-C0 | Unassigned |
| PCIe Gen4 PCIe Expansion Cable Card | <i>i</i> | U78D8.ND3.FGD03EK-P0-C6-C0 | Unassigned |
| PCIe3 2 PORT 25/10 Gb NIC&ROCE SFP28 ADAPTER | <i>i</i> | U78D8.ND3.FGD03EK-P0-C3-C0 | AIX 7.3 in IUM lab |
| PCIe3 2-Port 16Gb Fibre Channel Adapter | <i>i</i> | U78D8.ND3.FGD03EK-P0-C4-C0 | AIX 7.3 in IUM lab |
| PCIe3 2-Port 32Gb FC Adapter | <i>i</i> | U78D8.ND0.FGD03FD-P0-C3-C0 | AIX preloaded |
| PCIe3 4-port 10GbE Base-T Adapter | <i>i</i> | U78D8.ND3.FGD03EK-P0-C7-C0 | AIX 7.3 in IUM lab |

Power10 E1080 with multiple CECs
6 NED adapters for 3 NED24 drawers
First CEC PCIe slots C2 + C6
Second CEC PCIe slots C0 +C2 and C5 + C6

Adapters (indicated by a red dashed box around the first six rows of the table)

PCIe Slots (indicated by a blue dashed box around the physical location codes U78D8.ND3.FGD03EK-P0-C0-C0, U78D8.ND3.FGD03EK-P0-C2-C0, U78D8.ND3.FGD03EK-P0-C5-C0, and U78D8.ND3.FGD03EK-P0-C6-C0)

Resources

YLB210-9080-HEX

Console Management

Operating

Capacity

System Actions

Partitions

Properties

General Settings

Processor, Memory, I/O

Persistent Memory

Power VM

Virtual I/O servers

Virtual Networks

Virtual NICs

Virtual Storage

Hardware Virtualized I/O

Shared Processor Pool

Capacity On Demand

CoD Functions

Licensed Capabilities

Serviceability

Tasks Log

Serviceability

Reference Code Log...

NED24 Systems Admin = HMC

Hardware Management Console [gdliumhmc1.gdl.stglabs.ibm.com](#)

Resources > All Systems > YLB20Z-9043-MRX-78FF25X-J106 > Processor, Memory, I/O >

All Resources Search systems, partitions, help topics and more

Resources
Console Management
Users and Security
Serviceability

Operating

Capacity

System Actions

Partitions

Properties

General Settings

Processor, Memory, I/O

Processor, Memory, and I/O

View or change the memory, processor, and physical I/O resource settings for the managed system.
[Learn More](#)

Processors

Memory

Physical I/O Adapters

The table displays the physical I/O resources for the managed system.
Click the Adapter Info(i) to view information on the selected I/O device.
Click I/O Pools to display all the I/O pools found in the managed system and the partitions that are participating in the pools.

Save Cancel

Power10 E1050
Two NED24 drawers

U50EE.001.WZS0084-P3 NVMe slot C1 + C13 NED24
U780C.ND0.WZS015C-P3 → Internal NVMe slot
U50EE.001.WZS008J-P3 NVMe slot C1 + C13 NED24

IO Pools Expand All

| Adapter Description | Details | Physical Location Code | Owner |
|--------------------------------|---------|--------------------------|---------------------|
| 3.2TB NVMe Gen4 U.2 SSD III | | U50EE.001.WZS0084-P3-C1 | AIX/Linux preloaded |
| 3.2TB NVMe Gen4 U.2 SSD III | | U50EE.001.WZS0084-P3-C13 | AIX/Linux preloaded |
| PCIe4.4-Port 22Gb/s FC Adapter | i | U780C.ND0.WZS015C-P0-C10 | AIX/Linux preloaded |
| PCIe2.2-port USB 3.0 Adapter | i | U780C.ND0.WZS015C-P0-C7 | AIX/Linux preloaded |
| 6.4TB NVMe Gen4 U.2 SSD II | i | U780C.ND0.WZS015C-P1-C3 | AIX/Linux preloaded |
| 800GB NVMe Gen3 U.2 Slim SSD | i | U780C.ND0.WZS015C-P1-C4 | AIX/Linux preloaded |
| 800GB NVMe Gen3 U.2 Slim SSD | i | U780C.ND0.WZS015C-P1-C8 | AIX/Linux preloaded |
| 6.4TB NVMe Gen4 U.2 SSD | i | U780C.ND0.WZS015C-P1-C9 | AIX/Linux preloaded |
| 6.4TB NVMe Gen4 U.2 SSD III | i | U50EE.001.WZS008J-P3-C1 | Linux RH8.8 |
| 800GB NVMe Gen4 U.2 SSD III | i | U50EE.001.WZS008J-P3-C13 | Linux RH8.8 |

NVMe Drive Descriptions

NED24 Drive Slots

- ✓ L'amélioration matérielle suivante est disponible pour la famille de serveurs IBM® Power10 scale-up et Power10 scale-out utilisant le nouveau firmware Power® niveau **1040** :
 - Nouveau boîtier de stockage NVMe externe à connexion directe pour prendre en charge jusqu'à vingt-quatre périphériques NVMe d'entreprise U.2 de 15 mm.
- Le micrologiciel 1040 ne prend en charge que les fonctionnalités de la version de base du micrologiciel 1030 et non les fonctionnalités nouvellement annoncées qui nécessitent la version **FW1030.20**.
- **Remarque : FW1040 et FW1030.20** sont des versions distinctes du micrologiciel, leur prise en charge des fonctionnalités s'exclut mutuellement.
- ✓ Les améliorations matérielles suivantes sont disponibles pour la famille de serveurs Power10 scale-up et Power10 scale-out utilisant le niveau 1030.20 du micrologiciel Power. **Notez qu'elles ne sont pas prises en charge par le micrologiciel 1040** :
 - Nouvel adaptateur HBA PCIe3 SAS Tape pour prendre en charge les plates-formes IBM i (#EJ2B/EJ2C)
 - Nouvel adaptateur PCIe4 32 Gb 4 ports optiques Fibre Channel (FC) (#EN2L/EN2M)
 - Nouvel adaptateur FC optique PCIe4 64 Gb à 2 ports (#EN2N/EN2P)
 - Prise en charge des DDIMM mixtes pour le serveur IBM Power E1050 (#EMCM)
 - Prise en charge des alimentations 110 V pour le serveur IBM Power S1022s
- **Remarque** : le micrologiciel 1030.20 comprend la prise en charge des fonctionnalités prises en charge dans la version de base du micrologiciel 1030 et les fonctionnalités nouvellement annoncées, **à l'exception de #ESR0**.

Liste des codes du configurateur e-config pour IBM i

■ Tiroir

- #ESR0 EXP24U.2 NVMe SAS Storage enclosure Qté=1
- #ESM1 ESM module w/PCIe4 x16 2-port connection to CXP Converter Adapter. Qté=2 min et max

■ NVMe

- #ES3A Enterprise 800Go SSD PCIe4 NVMe U.2 module for IBM i 24 max
- #ES3C Enterprise 1.6To SSD PCIe4 NVMe U.2 module for IBM i 24 max
- #ES3E Enterprise 3.2To SSD PCIe4 NVMe U.2 module for IBM i 24 max
- #ES3G Enterprise 6.4To SSD PCIe4 NVMe U.2 module for IBM i 24 max

■ Cables & Adapteurs Power10

- #ECLS 3.0M CXP x16 Copper cable pair for PCIe4 Expansion drawer Qté=2
- #EJ2A PCIe4 x16 CXP Converter Adapter (support AOC) Qté=2

■ Administration

- #ECMS. Mode 1 Single Path Indicator for EXP24U.2 NVMe SAS Storage Enclosure Qté=1

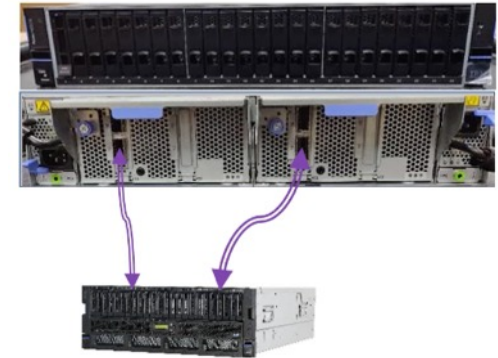
Fiabilité, disponibilité, sécurité et Environnement



- Maintenance concurrente (changement à chaud)
 - Unité NVMe – Inactif et déconfiguré au niveau de l'IBM i
 - Blocs d'alimentation sont N+1
 - Ventilateurs sont N+1
 - Note: 1 minute maximum avec le couvercle des ventilateurs ouvert
 - Nécessite l'accès au couvercle du dessus = donc le tiroir NED24 doit être sorti du rack
- Certifications
 - FCC: Class A
 - Acoustics : Data Center Category 1A
 - Environment : ASHRAE A3, ASHRAE A2 with FCM 3 (30W) = 5-40C, 8-85% RH, 3050m max
 - RoHS and IBM Anti-Smoke
- Dimension
 - Largeur = 19 inches (rack standard) x Hauteur=2U
- Support Boot
 - IBM i, AIX et Linux

Tiroir d'extension NED24

- ✓ Maximum : 153 To d'espace de stockage à travers 24 unités NVMe
- ✓ Simple à configurer et à utiliser
- ✓ Bande passante et nombre d'E/S par seconde élevés
- ✓ Coûts et consommation d'énergie bas par Go



NED24 vs solutions SAS

- Réduction par 62% du coût par To sur le stockage attaché directement
- Presque 10 fois plus de bande passante (Gops) et jusqu'à 3,7 fois plus d'E/S par seconde
- 153 To par tiroir d'extension NED24.
- Le Power10 E1080 supporte jusqu'à 8 tiroirs NED24 pour un maximum de 1,2 Po
- Délivre plus de 85% en plus de capacité/watt
- Étend la capacité du "Flash Cache" des systèmes nécessaire pour les applications avec grosse base de données critique
- Boot natif possible jusqu'à 24 partitions (1 par NVMe)

Comparatif coûts pour même nombre de disques

| DISPO | DESCRIPTION | Qté | Prix | Prix Total | Taille disque (GB) | Taille disque Totale (GB) | Coûts /GB |
|--------------------------------------|-------------------------------------------------------------------|-----|---------|----------------|--------------------|---------------------------|---------------------|
| CONFIGURATION DISK SAS 15KRPM | | | | | | | |
| ESLS | EXP24SX SAS Storage enclosure | 1 | €7 689 | €7 689 | | | |
| ESLN | 283GB 15K RPM SAS SFF-2 4k Block Cached Disk Drive IBM i | 24 | €871 | €20 904 | 283 | 6792 | |
| EJ14 | PCIe3 12GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8 | 2 | €11 472 | €22 944 | | | |
| ECE3 | 3.0M SAS AA12 Cable | 2 | €296 | €592 | | | |
| ECDU | 3.0M SAS YO12 Cable (Adapter to Enclosure) | 2 | €296 | €592 | | | |
| | TOTAL | | | €52 721 | | 6792 | 7,76 € / GB |
| CONFIGURATION DISK SSD | | | | | | | |
| ESLS | EXP24SX SAS Storage enclosure | 1 | €7 689 | €7 689 | | | |
| ETK9 | 387GB Enterprise SAS 4K SFF-2 SSD for IBM i | 24 | €2 719 | €65 256 | 387 | 9288 | |
| EJ14 | PCIe3 12GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8 | 2 | €11 472 | €22 944 | | | |
| ECE3 | 3.0M SAS AA12 Cable | 2 | €296 | €592 | | | |
| ECDU | 3.0M SAS YO12 Cable (Adapter to Enclosure) | 2 | €296 | €592 | | | |
| | TOTAL | | | €97 073 | | 9288 | 10,45 € / GB |
| CONFIGURATION EXTENSION NVMe | | | | | | | |
| ESR0 | EXP24U.2 NVMe SAS Storage Enclosure | 1 | €19 333 | €19 333 | | | |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 24 | €1 284 | €30 816 | 800 | 19200 | |
| ECLS | 3.0M CXP x16 Copper Cable Pair for PCIe4 Expansion | 2 | €1 437 | €2 874 | | | |
| EJ2A | PCIe4 x16 to CXP Converter Adapter (support AOC) | 2 | €2 773 | €5 546 | | | |
| ESM1 | ESM Module w/PCIe4 x16 2-port connection to CXP converter adapter | 2 | €6 495 | €12 990 | | | |
| | TOTAL | | | €71 559 | | 19200 | 3,73 € / GB |
| CONFIGURATION NVME INTERNE | | | | | | | |
| EJ1Y | Storage Backplane with eight NVMe U.2 drive slots | 2 | €3 168 | €6 336 | | | |
| ES3A | Enterprise 800GB SSD PCIe4 NVMe U.2 module for IBM i | 24 | €1 284 | €30 816 | 800 | 19200 | |
| EJUJ | Front IBM Bezel for 16 NVMe-bays Rack Mount | 1 | €317 | €317 | | | |
| | TOTAL | | | €37 469 | | 19200 | 1,95 € / GB |



IBM i



Questions / Réponses

■ Question

- Le nombre de bras de disque est très important sur IBM i. Une taille de namespaces plus petite, par exemple 80 Go au lieu de 200 Go et d'avantage de bras disque pour IBM i amélioreraient-ils les performances ? Avez-vous des recommandations concernant la taille de l'espace de noms pour IBM i ?

■ Réponse

- IBM i nécessite un nombre de bras de disque, mais nous n'avons rencontré aucun problème avec les clients configurant des tailles de namespaces de 200/400 Go avec PCIe Gen4 NVMe. Pour les petits clients avec des unités NVMe de 800 Go et 1,6 To, nous avons tendance à recommander la taille de 200 Go pour avoir plus de bras.

■ Question

- Je souhaite configurer un "remplacement à chaud" pour une unité NVMe en miroir, mais le groupe IBM i SW P05 a deux périphériques NVMe de 3,2 To configurés. Si un périphérique tombe en panne, est ce que les opérations de lecture/écriture sur l'autre namespace mirroré continuent ? Ou est-il en lecture seule jusqu'à ce qu'il soit corrigé ?

■ Règles de configuration P05 :

- Jusqu'à 6,4 To
- Ne peut pas dépasser 6,4 To pour 2 à 4 appareils
- L'utilisation de 2 disques de 3,2 To élimine la possibilité d'ajouter un autre disque pour le remplacement à chaud.

■ Réponse

- Le S1014 P05 dispose de plusieurs options de configuration tant que la capacité totale ne dépasse pas 6,4 To. Si un périphérique NVMe d'une paire miroir tombe en panne, il restera en mode lecture/écriture au lieu de passer en mode lecture seule.

■ Question

- En ce qui concerne le remplacement des modèles IBM i 4 cœurs, nous avons de nombreux petits clients avec 4 cœurs Power8 ou Power9 exécutant IBM i avec 6-10 disques durs 283G en CEC. Quelle est la bonne configuration de disque NVMe interne pour remplacer ces machines ?

■ Réponse

- Un S1014 (1S4U) à 4 cœurs a une capacité maximale de 6,4 To. Les configurations d'unités NVMe incluent jusqu'à 8 x 800 Go, 4 x 1,6 To ou 2 x 3,2 To. Voir si le client a besoin d'un disque de secours "à chaud", c'est-à-dire un NVMe qui n'est pas affecté à une partition. Il n'est pas possible dans e-config de commander plus de 6,4To de stockage NVMe dans une configuration initiale.

■ Question

- Prise en charge d'IBM i 7.3 pour NVMe - Existe-t-il des unités NVMe pris en charge de manière native par IBM i 7.3 ? Plus précisément, les unités ES3A et ES3C seront-elles prises en charge dans IBM i 7.3 ? Le manuel de vente indique simplement qu'il est pris en charge sur IBM i.

■ Réponse

- IBM i 7.3 supporte en natif les unités NVMe uniquement sur la plate-forme Power10. Pas de support natif sur la plate-forme Power9. Il est indépendant du facteur de forme de l'unité NVMe.

■ Question

- Quelle est la meilleure pratique de configuration pour S1024 avec 2 partitions VIOS ? Besoin d'un seul fond de panier FC EJ1Y et de 4 disques EC7T 800G NVMe ? Ou devrais-je avoir deux fonds de panier FC EJ1Y et configurer deux disques dans chacun ?

■ Réponse

- Tout d'abord, quel niveau de redondance VIOS avez-vous besoin ?
- Le niveau maximal possible est de 2 fonds de panier et de 4 unités NVMe (une unité utilisée par VIOS sur chaque fond de panier utilisant des emplacements avec différents câbles internes).
- Une configuration minimale est une simple configuration de mise en miroir de deux appareils NVMe (U.2 ou AIC).
- Lorsque vous utilisez VIOS et vscsi, la mise en miroir doit être effectuée au niveau du système d'exploitation invité, et non au sein de VIOS.

Questions & Réponses

■ Question

- IBM i prendra-t-il en charge FC-NVMe ?

■ Réponse

- Non, car FC-NVMe n'a pas d'avantage clair sur les performances d'efficacité par rapport à FCP (Fibre Channel Protocol). IBM i continue de se concentrer sur les améliorations FCP.

Questions & Réponses

■ Question

- Quelle est la prise en charge au niveau du système d'exploitation pour les disques NVMe ES1* et ES3* ?

■ Réponse

- Les disques ES1* sont les premiers périphériques U.2 à prendre en charge tous les systèmes d'exploitation Power10. Les disques ES3* sont les derniers appareils U.2 et sont pris en charge par les dernières versions de système d'exploitation de Power10 et Power9.

■ Question

- Pour VIOS, combien d'écritures par jour en général ? Les périphériques Mainstream conviennent-ils à une utilisation avec VIOS ?

■ Réponse

- La réponse à cette question est que tout dépend de la configuration. Si le VIOS est initialisé uniquement pour fournir Ethernet virtuel et/ou une autre virtualisation de périphérique de stockage, les lectures/écritures seront très faibles. Les lectures/écritures peuvent être élevées si le VIOS héberge un stockage virtuel sur le même stockage physique. Privilégiez un NVMe Enterprise.

■ Question

- Le S1024 dispose de 10 emplacements d'E/S et le S924 de 11 emplacements d'E/S. Si je configure le S1024 avec un fond de panier NVMe, obtiendrai-je 10 emplacements d'E/S ou 9 emplacements d'E/S ?

■ Réponse

- Tous les fonds de panier NVMe de serveur scale-out Power10 doivent utiliser un emplacement d'E/S PCIe. Cela signifie que 9 emplacements d'E/S sont disponibles avec un fond de panier et 8 emplacements d'E/S sont disponibles avec 2 fonds de panier.

■ Question

- Configuration d'une paire de miroirs NVMe de 3,2 To sur IBM i. Si une unité NVMe tombe en panne, le système continuera-t-il à fonctionner avec les deux capacités de lecture/écriture sur les périphériques restants ? Ou le client l'utilisera-t-il en mode "lecture seule" jusqu'à ce que l'appareil soit réparé ?

■ Réponse

- Si un périphérique NVMe d'une paire miroir tombe en panne, le volume/LUN restera en mode lecture/écriture normal. Peu importe le système d'exploitation que vous utilisez.

■ Question

- La configuration nécessite 4 disques NVMe sur S1024 (2 disques pour VIOS1 et 2 disques pour VIOS2). De combien d'emplacements d'E/S auriez-vous besoin dans ce cas ? Aurai-je jamais besoin de plus de deux emplacements d'E/S pour mon fond de panier NVMe ?

■ Réponse

- La plate-forme évolutive Power10 a la possibilité de configurer un ou deux fonds de panier NVMe. Chaque fond de panier utilise un logement d'E/S PCIe.

■ Question

- Supposons que vous configurez un certain nombre de partitions logiques nécessitant des disques internes (sans tiroirs d'extension de disque). Quel est le nombre maximal de LPAR qui peuvent avoir des disques NVMe comme disques de démarrage du système d'exploitation si seulement deux disques NVMe sont requis par LPAR ?

■ Réponse

- Avec 2 NVMe U.2 dédiés par LPAR, un total de 4 LPAR peut être créé sur un P10 2U (1022, S1022s), 8 LPAR sur une P10 4U (S1024, S1014). Si vos unités NVMe sont virtualisées (par VIOS), vous pouvez créer davantage de LPAR en utilisant des espaces de noms configurés en tant que volumes logiques utilisés en tant que cibles vscsi.

Questions & Réponses

■ Question

- VIOS fournit-il une configuration d'espace de noms dans le programme d'installation du système d'exploitation ? , ou dois-je continuer à utiliser le diag autonome ?

■ Rponse

- L'installation initiale du système d'exploitation utilise un espace de noms unique par défaut qui utilise l'ensemble du périphérique. Si vous souhaitez modifier cela, vous devez utiliser le diag boot autonome pour supprimer l'espace de noms par défaut et créer l'espace de noms souhaité.

■ Question

- Existe-t-il des bonnes pratiques pour dimensionner l'espace de noms NVMe ?
Quelle taille de périphérique NVMe est recommandée lors de l'utilisation d'une taille d'espace de noms de 400 Go au lieu de 200 Go ?

■ Réponse

- Cela dépend de la taille de votre appareil NVMe. Ces tailles d'espace de noms recommandées sont utilisées pour optimiser l'espace disponible pour un maximum de 64 espaces de noms et optimiser les performances d'E/S IBM i. Les appareils de 800 Go et 1,6 To doivent cibler 200 Go, et les appareils de 3,2 To et 6,4 To doivent cibler 400 Go.

Questions & Réponses

■ Question

- NVMe over FC est-il déjà pris en charge sur IBM Power ?

■ Réponse

- FC-NVMe n'est actuellement pris en charge que sur les adaptateurs Fibre Channel EN1A/EN1B 2x32 Gb sur IBM AIX et Linux. Il est disponible sur AIX version 7.2 niveau technologique 5 et supérieur et RHEL 8, POWER LE version 8.3 et supérieur. Pas de support sur IBM i car les performances ne sont pas énormes par rapport au Fibre Channel

Questions & Réponses

■ Question

- Pour IBM i, après avoir préparé NVMe pour IBM i, dois-je également initialiser le namespace créé ?

■ Réponse

- Une fois qu'un namespace est créé, il est immédiatement initialisé et prêt à l'emploi. Ce processus est nettement plus rapide pour les unités NVMe que pour les unités SAS, quelque soit le système d'exploitation.

Questions & Réponses

■ Question

- Puis-je augmenter ou diminuer la taille d'un namespace?

■ Répondre

- Les namespaces ne peuvent pas être redimensionné une fois créé. S'il vous reste suffisamment d'espace sur 2 unités NVMe, vous pouvez:
 - créer un nouveau namespace de la taille souhaitée sur les 2 NVMe
 - Lancez le miroir sur les 2 namespaces.
 - Ajoutez les 2 nouveaux namespaces à l'ASP
 - Enlever la paire miroir (2 anciens namespaces) NVMe de l'ASP
 - Lancez un étalement des données sur l'ASP (STRASPBAL).

IBM