

IBM Power Systems Scale Out Servers Internal Storage NVMe Options



Daniel R Sundt

Partner Technical Advocate

BP Technical Sales & Enablement

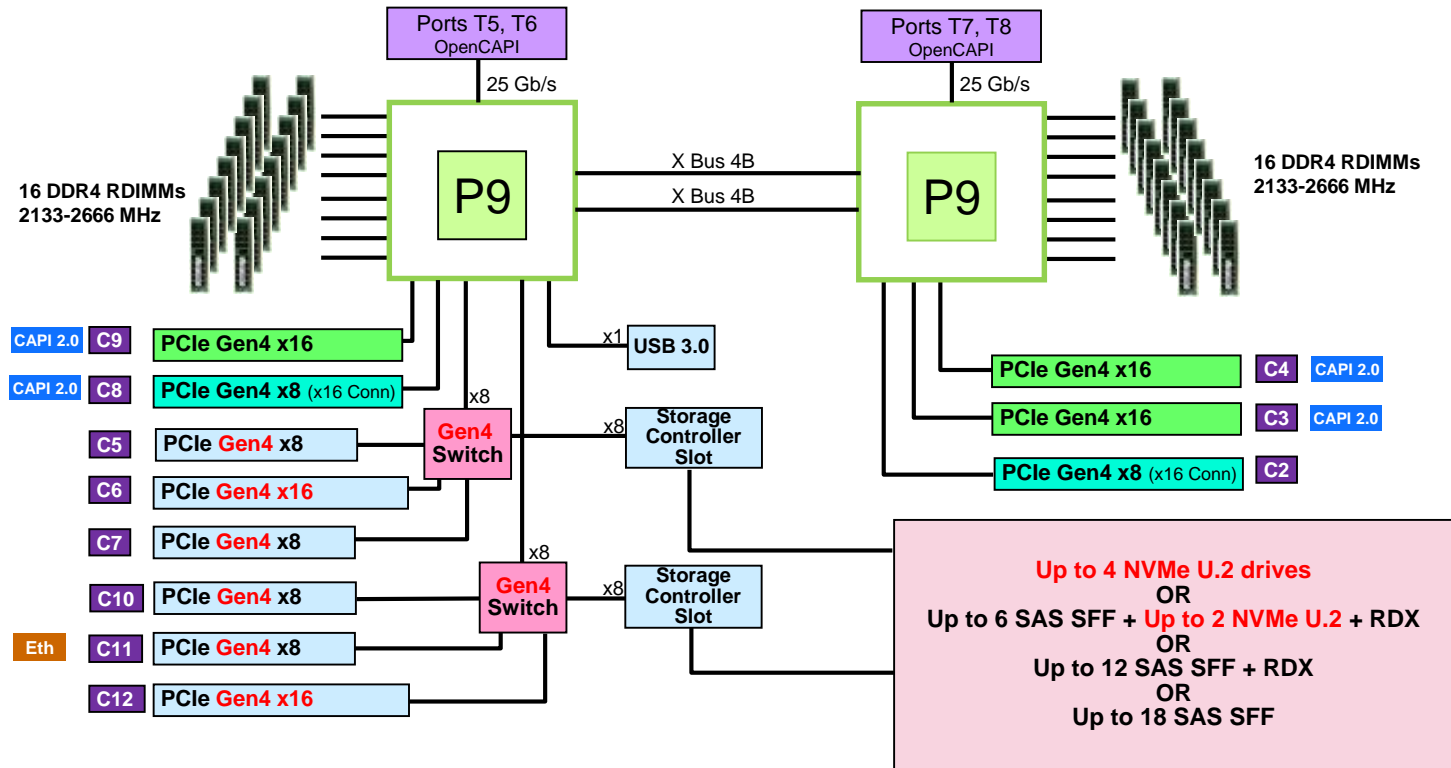
IBM Power Systems



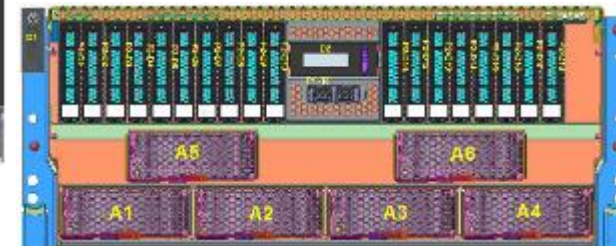
2800 37th ST NW
Rochester, MN 55901

Tel 507 253 3228
Mobile 507 261 5329
dansundt@us.ibm.com

Model S924 (9009-42G) System Topology/



C?? Indicates PCIe Slot Number



S924 Storage Options

NVMe drive slots, SAS bays, and storage backplane options

S924 or H924 Model	FC	Description
42A, 42H	EC59	NVMe Card with two M.2 connectors (withdrawn)
42A, 42H, 42G	EJ1C	12 SFF bays (Gen3-Carrier), 1 RDX bay / Single RAID 0,10,5,6
42A, 42H, 42G	EJ1E	6+6 SFF bays (Gen3-Carrier), 1 RDX bay / Split Backplane RAID 0,10,5,6
42A, 42H, 42G	EJ1M	12 SFF bays (Gen3-Carrier), 1 RDX bay / Dual Write Cache RAID 0,10,5,6,5T2,6T2
42A, 42H, 42G	EJ1D	18 SFF bays (Gen3-Carrier) / Dual Write Cache RAID 0,10,5,6,5T2,6T2
42A, 42H, 42G	EU00	RDX Docking Station (Internal)
42G only	EJ1S	Storage backplane with 6 SFF-3 Bays and 2 PCIe Gen4 capable NVMe U.2
42G only	EJ1T	Storage backplane with 2 PCIe Gen4 capable NVMe U.2 drive slots
42G only	EJ1U	Storage backplane with 4 PCIe Gen4 capable NVMe U.2 drive slots

External Storage Options

FC / MTM	Description
ESLL	19" Disk Expansion Drawer 12 LFF Gen2-Carrier Bays (Slider12)
ESLS	19" Disk Expansion Drawer 24 SFF Gen2-Carrier Bays (Slider24)
5887	19" Disk Expansion Drawer 24 SFF Gen2-Carrier Bays (EXP24S) Migrate
EUA5	USB DVD w/ Cable
EUA4	RDX USB Docking Station
7226-1U3	19" Media Drawer with 2 bays

Supported Media Overview

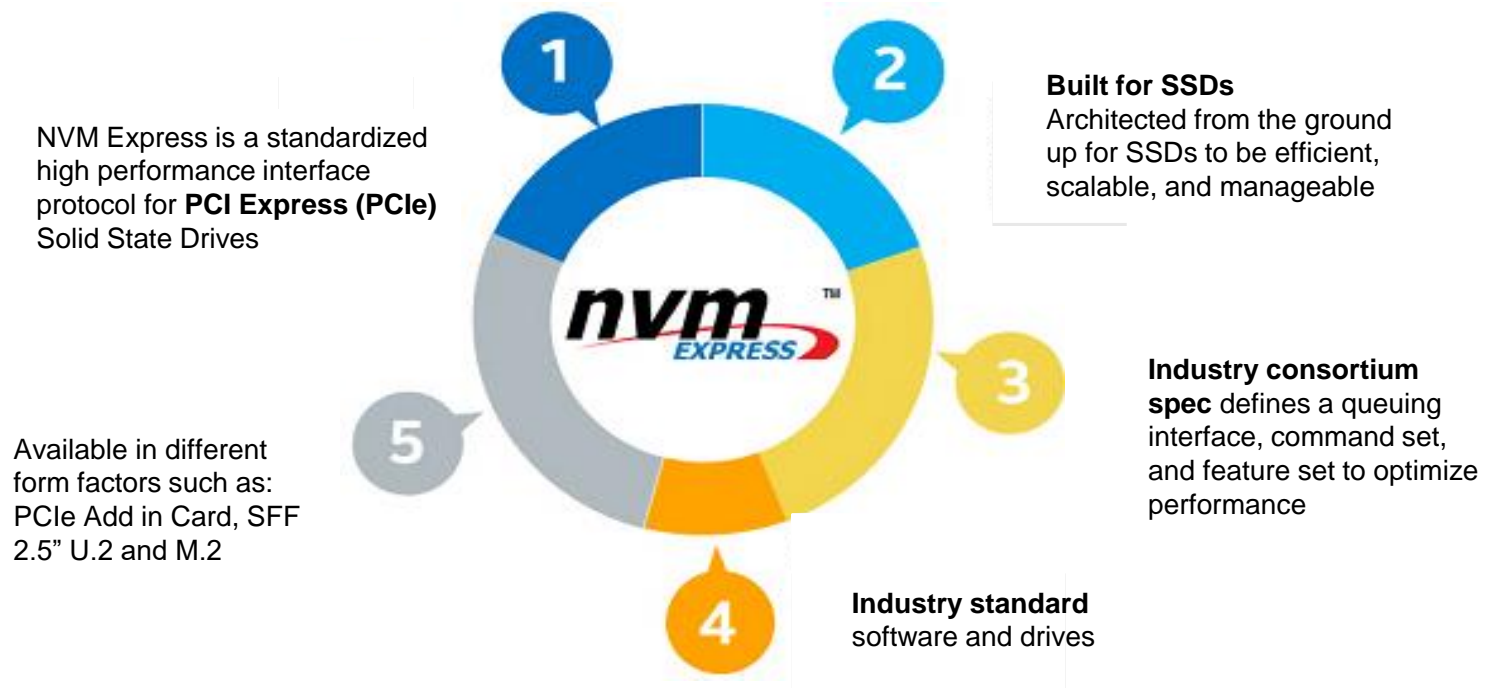
- **NVMe**
 - 42A and 42H only: 400GB 1.5 DWPD M.2 Flash devices (withdrawn)
 - All
 - Enterprise 1.6 TB PCIe3 x8 1.6 TB NVMe Flash Adapter
 - Enterprise 3.2 TB PCIe3 x8 1.6 TB NVMe Flash Adapter
 - Enterprise 6.4 TB PCIe3 x8 1.6 TB NVMe Flash Adapter
 - 42G only:
 - Mainstream 800 GB SSD PCIe3 NVMe U.2 module
 - Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module
 - Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module
 - Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module
- **SFF HDDs**
 - 571/600GB, 1100/1200GB, 1700/1800GB - 10K RPM
 - 283/300GB, 571/600GB – 15K RPM
- **SFF SSDs**
 - Enterprise
 - 387GB, 775GB, 1551GB – 10 DWPD
 - Mainstream
 - 931GB, 1860GB, 3720GB, 7450GB – 1 DWPD
- **RDX Disk Cartridge**
 - 500 GB Disk Cartridge (1107)
 - 1TB Disk Cartridge (EU01)
 - 2TB Disk Cartridge (EU2T)

nvm
EXPRESS®

The logo features the lowercase letters 'nvm' in a bold, black, italicized sans-serif font. Below this, the word 'EXPRESS' is written in a bold, blue, italicized sans-serif font. A red swoosh underline starts under the 'n', passes under the 'v' and 'm', and then curves upwards and to the right, ending under the 'S' of 'EXPRESS'. A registered trademark symbol (®) is located at the end of the 'EXPRESS' text.

What is NVMe?

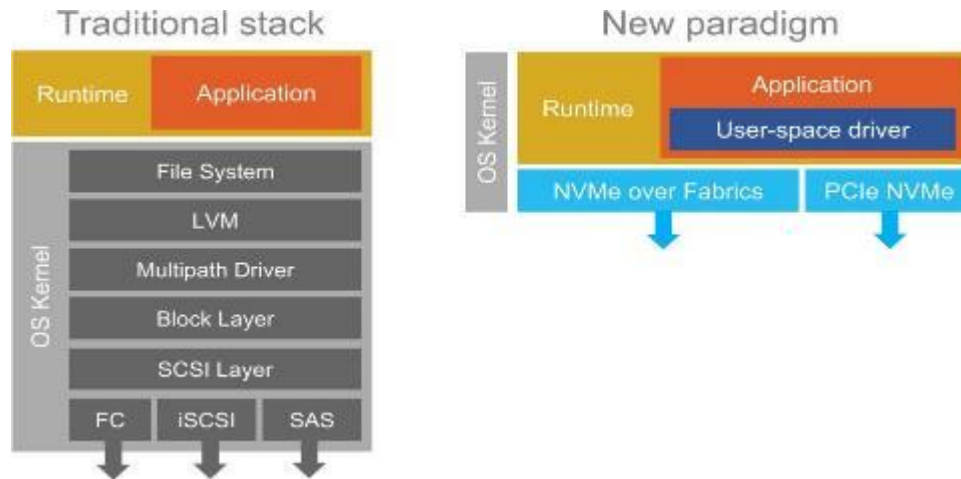
NVM Express® is an open collection of standards and information to fully expose the benefits of non-volatile memory in all types of computing environments from mobile to data center. NVMe™ is designed from the ground up to deliver high bandwidth and low latency storage access for current and future NVMe technologies.



<http://www.nvmexpress.org/>

What is NVMe?

SSDs are fast. **So fast in fact, their limiting factor is not their own hardware, but rather the SAS or SATA connection that hard drives** have traditionally used. **NVMe - “Non-Volatile Memory Express”** is an open standard developed to allow modern SSDs to operate at the read/write speeds their flash memory is capable of. Essentially, it allows flash memory to operate as an SSD directly through the PCIe interface rather than going through SATA and being limited by the slower SATA speeds.



Note that the diagram above is just a representation. Operating systems such as IBM i have a machine interface and single level storage, thus does not let applications write directly to hardware.

Is NVMe Fast?

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	Persistent Storage				15x	
		LL Flash (HMS)	TBD	TBD						
	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.

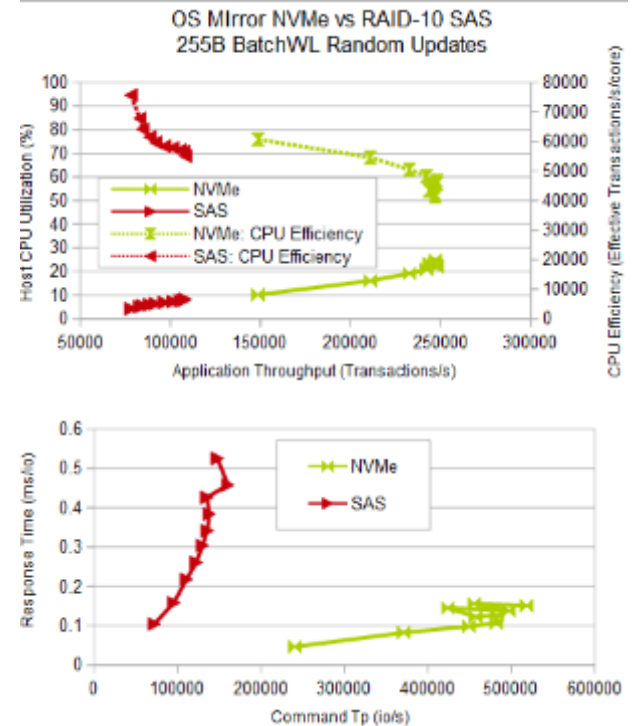
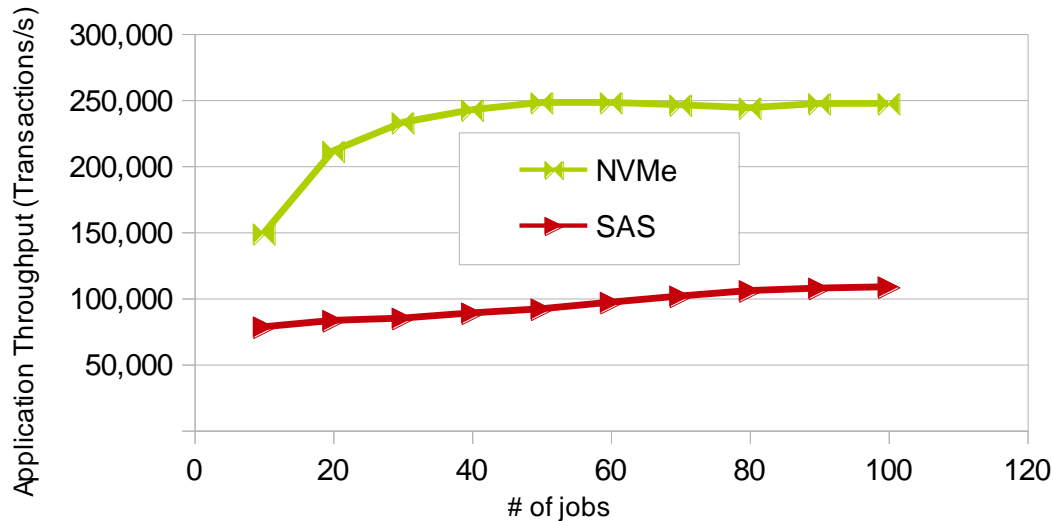
IOPs and Latencies: Normally measured on a random 4K ops. * <1us for 1K transfer utilizing Persistent Log Buffer feature

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

Is NVMe Fast?

Small Op Random Updates : Mirror vs Mirror

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



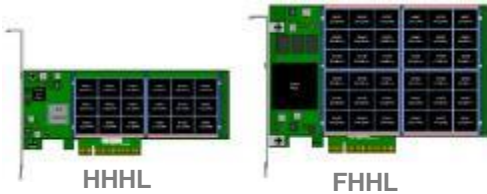
Take-aways:

1. Running SAS in a mirrored configuration does allow it to support more throughput than in RAID-5 mode, but still not nearly enough to keep up with the mirrored NVMe configuration
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

NVMe Product Form Types used in POWER9 Systems

Add-in Card (AIC)

- x4 / x8 PCIe Gen3
- 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

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



2.5" 15mm

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



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

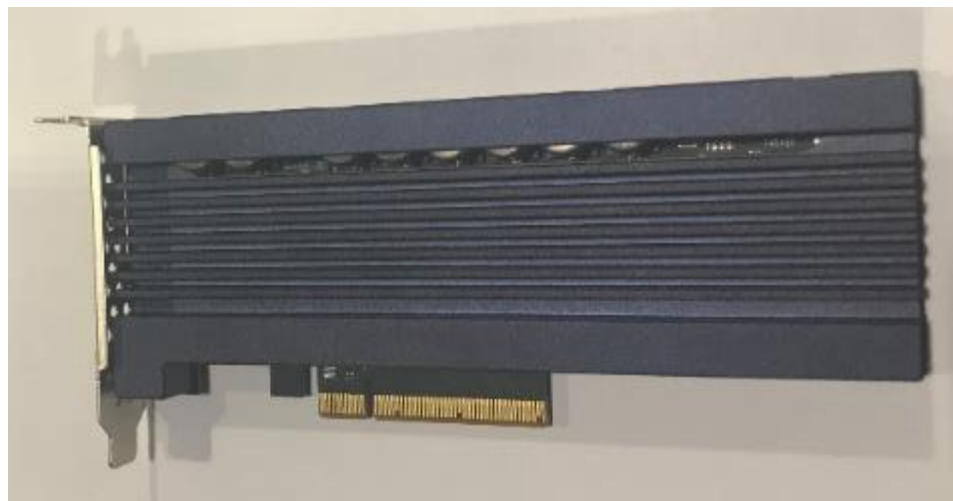


Scale Out Gen4 Refresh NVMe Performance Specs

FC	Capacity	Read BW GB/s	Write BW GB/s	Read IOPS	Write IOPS	DWPD (5 years)	PCIe Gen	Form Factor
EC5J	800 GB	1.6	1.05	380K	60K	2.4	3	U.2
EC5G, EC5B EC6U, EC6V	1.6 TB	4.7	1.9	700K	100K	5	3	AIC
EC5C, EC5D EC6W, EC6X	3.2 TB	6.0	3.0	910K	170K	5	3	AIC
EC5E, EC5F EC6Y, EC6Z	6.4 TB	6.0	3.0	910K	170K	5	3	AIC
ES1E, ES1F	1.6 TB	7.0	2.5	1000K	220K	3	4	U.2
ES1G, ES1H	3.2 TB	7.4	3.4	1500K	250K	3	4	U.2
EC5V, EC5W	6.4 TB	7.2	3.8	1500K	250K	3	4	U.2



IBM i Support of PCIe Card NVMe on POWER9 Servers



4Q19 Power Systems Hardware Announcement

- IBM i Support for Direct Attached NVMe devices (7.4 TR1)
- New feature codes for IBM i specific PCIe3 x8 SSD NVMe adapters (7.4 TR1)
 - 1.6 TB #EC6U/#EC6V
 - 3.2 TB #EC6W/#EC6X
 - 6.4 TB #EC6Y/#EC6Z
- POWER9 servers only - S914, S924, H924 and E980
- Announcement: 10/8/19
- GA: 11/22/19



POWER9 PCIe Add In Card NVMe Device

Hardware Features

- NVMe Specs. 1.2.1 Compliant
- NVMe Over Fabrics 1.0 Capable
- PCIe Gen 3 x 8
- **Multiple Namespace (32)**
 - **Namespace Granularity 16GB**
- Half Height Half Length (HH-HL)
- Power ≤ 25W
- Block Size 4096(Default), 512, **4160 (IBM i)**
- End-To-End Protection: T10 DIF & DIX
- Non Volatile Write Buffer
- **Endurance 5 DWPD for 1.6/3.2/6.4TB**
- 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)
- Non-TCG SED
- **No support for MEX Drawer**



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

Workload	Target (1.6 TB)	Target (3.2/6.4)
Read (IOPS)	700K	910K
Write (IOPS)	100K	170K
Mixed R/W (70/30)	250K	320K
Read Data Tp (GB/s)	4.7	6.0
Write Data Tp (GB/s)	1.9	3.0
Read Latency (us)	110	110
Write Latency (us)	30	30

Notes:

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

Software Support

- Linux
 - Power VM: RHEL 7.5LE, SLES 12 SP3 LE
 - Ubuntu 18.04
 - Power NV: RHEL 7.5LE, Ubuntu 18.04
- AIX (7.1Z & 7.2F), VIOS (2.2.6)
- **IBM i (7.4 TR1)**
- Load Source
- Software RAID 0, 1, 5 & 6 (Linux)
- **OS Mirroring (AIX, IBM i)**
- DIAG Support
- NVMe Over Fabrics (Linux Only)





PCIe Card NVMe/SSD Pricing Comparison

• 1.6 TB NVMe x 2

- #EC6V PCIe3 x8 1.6 TB NVMe Flash Adapter for IBM 2 \$5,198
- #ENS1 188 GB IBM i NVMe Load Source Name Space size
- $8 \times 188 = 1504 \times 2 = 3008 / 2 = 1504$ GB usable
- $\$5198 / 1504$ GB = $\$3.46/\text{GB}$

• 387 GB SAS SSD x 8

- #ESB9 387GB Enterprise SAS 4k SFF-3 SSD for IBM i 8 \$15,592
- Mirroring $8 \times 387 = 3096 / 2 = 1548$ GB usable
 - $\$15,592 / 1548$ GB = $\$10.07/\text{GB}$
- If the cost of a storage backplane (such as the #EJ1M \$4099) is included which is not required when only PCIe NVMe are used
 - $\$15,592 + \$4099 = \$19691 / 1548$ GB = $\$12.72/\text{GB}$
- NVMe have to be mirrored, but with SSDs could use RAID
 - RAID-5 $7 \times 387 = 2709$ GB usable
 - $\$15,592 + \$4099 = \$19691 / 2709$ GB = $\$7.27/\text{GB}$



PCIe Card NVMe/IBM i/IBM Power Systems Hardware

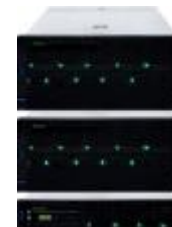


- NVMe is able to provide higher performance than SSDs. NVMe technology can provide significantly more read or write IOPS and significantly larger throughput (GB/sec) compared to SAS/SATA SSDs. Actual system or application performance differences will vary by customer and workload.
- NVMe provides additional virtualization capabilities since every device is a PCIe endpoint that can be dedicated to a partition/LPAR
- At least one identical NVMe adapter pair is required; subsequent NVMe adapter pairs can be different than the first pair. After an identical pair is on the order, one NVMe adapter of different capacity is allowed. Best practice would be to only order in pairs to make mirroring the most straightforward.
- NVMe devices require IBM i operating system mirroring as there isn't hardware RAID support. Mirrored pairs must be on different physical devices. NVMe can only be mirrored to NVMe and SAS drives can only be mirrored to SAS drives.
- Hot spare is not supported, however an extra NVMe could be on the system as a cold spare to speed up the repair process, and it is only a spare in the fact that a customer doesn't have to order/plug it in. IBM i development is aware of the desire for something more capable than a cold spare.

PCIe Card NVMe/IBM i/IBM Power Systems Hardware



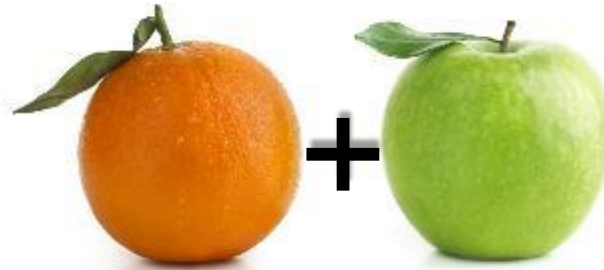
- NVMe is only supported in the system unit. It is not supported in a PCIe Gen3 I/O drawer due to both bandwidth issues and required testing.
- S914 supports up to 7 NVMe (A models 3). S924 supports up to 10 NVMe (A models 5). 7 with only one processor module (A models 3).
- S914 4-core P05 system is limited to 2 or one pair NVMe devices. No mixing of NVMe and SAS drives is allowed on the 4-core (10 maximum SAS drives).
- S922 1-core P05 system is limited to 2 or one pair NVMe devices. No mixing of NVMe and SAS drives is allowed on the 1-core (8 maximum SAS drives).
- E980 supports up to 8 NVMe per drawer (6 first drawer, 8 each drawer 2, 3, and 4 for a maximum total of 30)
- Note that using multiple PCIe card NVMe will affect available I/O slots which can affect system configurations with multiple partitions, and using external drawers may affect the maximum amount of PCIe NVMe.



PCIe Card NVMe/IBM i/IBM Power Systems Hardware



- Internal HDDs/SSDs are allowed in the system unit “with” PCIe card NVMe (except the S914 4-core).



- A storage backplane is “not” required when PCIe card NVMe adapters are used as the load source for IBM i, but may be nice to have for future storage expansion.

PCIe Card NVMe in an S9x4 Adapter Slot




PCIe Card NVMe IBM i HMC Screenshots



Hardware Management Console

ATSHMC4 Resources > All Systems > ATS_S924B > Partitions > ratsibmi4 > Physical I/O Adapters >



ratsibmi4

Running

- Capacity
- Partition Actions**
- Properties
- General Properties
- Processors
- Memory
- Persistent Memory
- Physical I/O Adapters**

Physical I/O Adapters

The table lists all the adapters that are connected to the logical partition. Click Add Adapter to add a physical I/O adapter to the logical partition. [Learn More](#) →

Assigned Physical I/O Adapters

Action ▾ Add Adapter

	Physical Location Code	Description
<input type="radio"/>	U78D2.001.WZS0ET2-P1-T3	Universal Serial Bus UHC Spec
<input type="radio"/>	U78D2.001.WZS0ET2-P1-C8	PCIe3 3.2TB NVMe Flash Adapter III x8
<input type="radio"/>	U78D2.001.WZS0ET2-P1-C2	PCIe3 3.2TB NVMe Flash Adapter III x8

Total : 3 Selected : 0

PCIe Card NVMe and IBM i



- IBM i supports virtualized NVMe (only for the PCIe Add In Card) via VIOS and requires the use of the VIOS LVM (Logical Volume Manager). This supports IBM i 7.2, 7.3 or 7.4. Since this has more "layers" between IBM i and the storage, it will not perform the same as native NVMe, thus would not be the recommended option if the best/most performance is required.
- PCIe card NVMe service/repair is similar to other PCIe card slot concurrent maintenance, but with extra steps such as described below:
https://www.ibm.com/support/knowledgecenter/ssw_ibm_i_74/rzaly/rzalnvmefailure.htm
- Encryption is not supported today. IBM i development is aware of the desire for hardware encryption support.
- IBM i treats NVMe as the same tier as SSDs, so currently there isn't a way to tier (say using the Trace ASP Balance (TRCASPBAL) command) between them.
- NVMe devices are now supported (4/14/20 announce) as direct attached devices for IBM Db2 Mirror for i.

PCIe Card NVMe Namespaces and IBM i

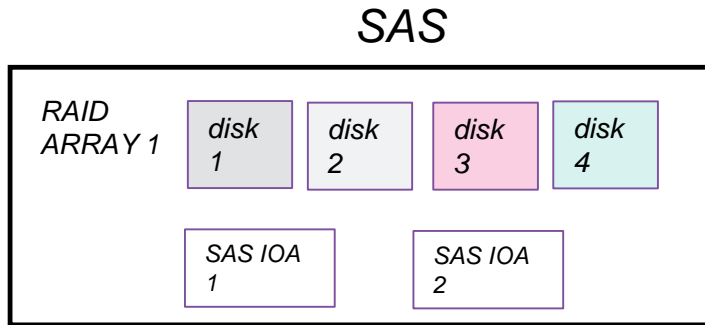


- NVMe architecture can use namespaces which is a collection of logical blocks whose logical block addresses range from 0 to the size of the namespace. A namespace ID (NSID) is an identifier used by a controller to provide access to a namespace.
- With NVMe, an ‘arm’ (logical unit) is a namespace. A namespace is a logical chunk of a physical NVMe device and multiple namespaces are allowed on one NVMe device.
- IBM i is the management interface used by a customer to create and manage namespaces
- IBM i’s use of NVMe architected multiple namespaces provides for many ‘arms’ on a small number of high capacity NVMe physical devices
- IBM i can use a NVMe device (up to 16 TB) with only a single namespace for the whole device. However, for almost all customers, this will cause sub-optimum performance since more (and smaller) ‘arms’ (logical units) are better than fewer and larger.
- Note that the word “namespace” is used in the industry and by IBM i in different ways and in different contexts, so for example, a NVMe namespace should not be confused with and has nothing to do with ASP namespaces.

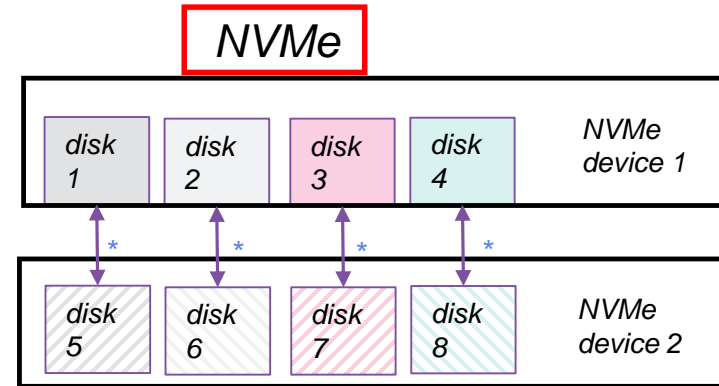
SAS Versus PCIe Card NVMe Storage with IBM i



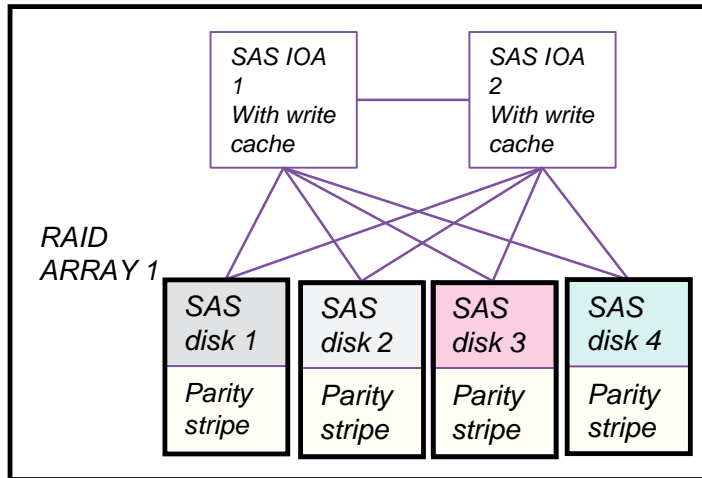
Customer and operating system view



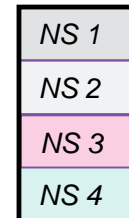
OS is aware of hardware RAID



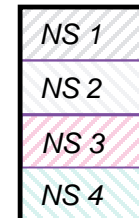
Device physical view



NVMe device 1



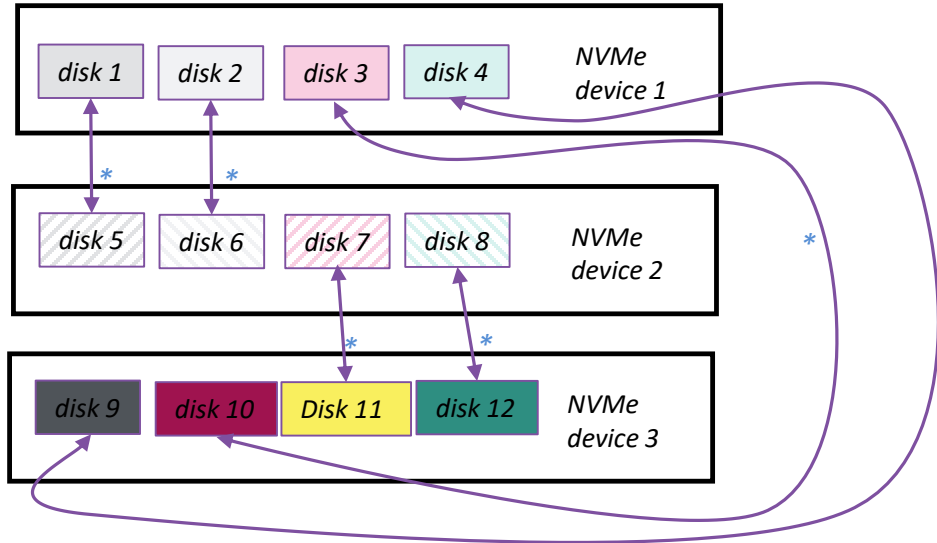
NVMe device 2



IBM i PCIe Card NVMe Mirroring (3 NVMe Option 1)



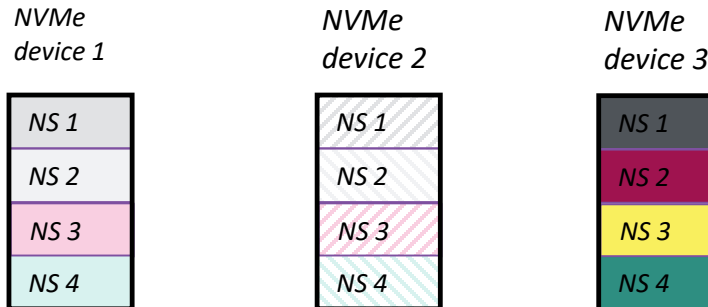
Customer and operating system view



All mirrored name spaces are the same size on all devices

**OS mirroring*

Device physical view

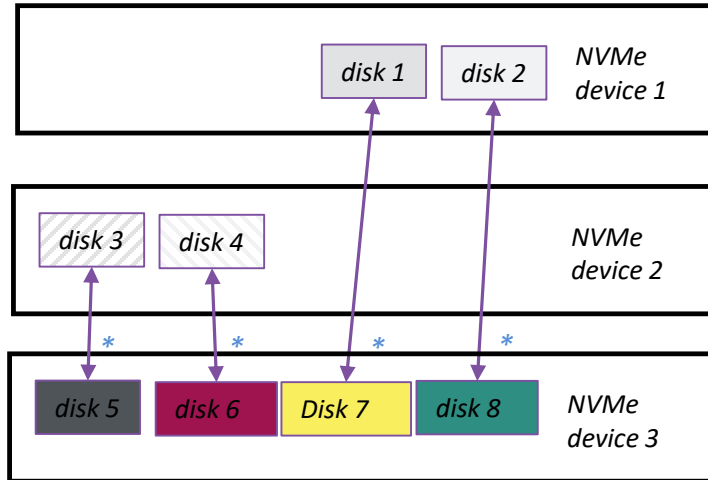


Device physical capacity does not have to be the same size

IBM i PCIe Card NVMe Mirroring (3 NVMe Option 2)



Customer and operating system view

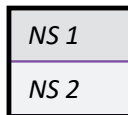


All mirrored name spaces are the same size on all devices

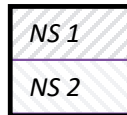
**OS mirroring*

Device physical view

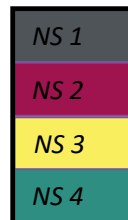
NVMe device 1



NVMe device 2



NVMe device 3



Device physical capacity does not have to be the same size

Recommended PCIe Card NVMe Namespace Sizes



- First generation NVMe devices have a hardware boundary of 16 GB for name spaces. Device capacity can be wasted/lost if name spaces are not multiples of 16. The maximum number of namespaces on a device is 32.
- IBM i screens show Capacities in ‘GB’ (1000**3 (GB), not 1024**3 (GiB))
- IBM recommends namespaces of 188 GB or 393 GB (the e-Config default)
- Consider only using 393 GB. On the 6.4 TB device there is a lot of unused space when using 188 because the maximum number of namespaces is 32. The choice of namespace size is a balance between number of “arms” for (storage management) performance and other individual customer factors such as scaling/growth of the system.

Device Nominal Size	Device Actual Size	Number of Namespaces	Namespace Size	Total User Capacity Used By Namespace	Remaining Space on the Device (unallocated)
1.6TB	1575	8	188	1448	87
3.2TB	3151	16	188	2977	174
6.4TB	6364	32	188	6016	348
1.6TB	1575	4	393	1556	19
3.2TB	3151	8	393	3112	38
6.4TB	6364	16	393	6288	76

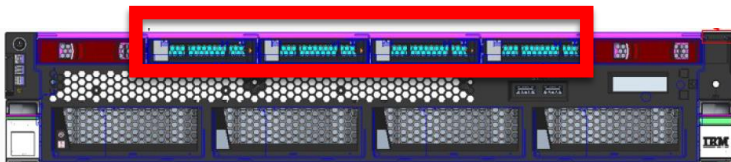
IBM i Support of U.2 NVMe on POWER9 Servers



U.2 NVMe



- New S922/S914/S924 G models added Enterprise PCIe GEN4 NVMe U.2 modules
- Same capacity points as PCIe card NVMe (1.6/3.2/6.4 TB). The 800GB is NOT supported by IBM i natively as it is meant as a VIOS/AIX/Linux boot drive.
- Up to 2 or 4 are supported based on the backplane
 - 6 SFF-3 bays & 2 NVMe U.2 drive slots
 - 4 NVMe U.2 drive slots
 - 2 NVMe U.2 drive slots
- Requires IBM i 7.4 TR2 or later
- Can mix PCIe card NVMe and U.2 NVMe on the same system
 - S914 – 7 PCIe card + 4 U.2 front drives (11 max)
 - S922 – 10 PCIe card + 4 U.2 front drives (14 max)
 - S924 – 10 PCIe card + 4 U.2 front drives (14 max)
- New IBM i load source name space sizes:
 - #ENSA 200GB IBM i NVMe LOAD SOURCE NAME SPACE SIZE
 - #ENSB 400GB IBM i NVMe LOAD SOURCE NAME SPACE SIZE



PCIe Card NVMe IBM i DST/SST Screenshots



Work with Disk Units

Select one of the following:

1. Save load source disk unit data
2. Copy load source disk unit data
3. Display/change page data
4. Analyze disk unit surface
5. Initialize and format disk unit
6. Reclaim I/OA cache storage
7. Stop device parity protection
8. Update system vital product data
9. Start device parity protection - RAID 5
10. Start device parity protection - RAID 6
11. Start device parity protection - RAID 5 with hot spare
12. Start device parity protection - RAID 6 with hot spare
13. Start device parity protection - RAID 10
14. Start device parity protection - RAID 10 with hot spare
15. Work with NVM Devices

Selection

15

F3=Exit

F12=Cancel

MA + A

21/007



Work with NVM Device

Select one of the following:

1. Display NVM namespaces
2. Display NVM devices
3. Create NVM namespaces
4. Delete existing NVM Namespaces
5. Sanitize/Erase NVM device
6. Format NVM device to prepare device for IBM i

Selection

—

F3=Exit

F12=Cancel

MA + A

21/007

PCIe Card NVMe IBM i DST/SST Screenshots



Display NVM Namespaces

NVM Device	ASP	Unit	Serial Number	Type	Model	Resource Name	Namespace Capacity in GB
1			Y0YACBYCB075	58FD		DC03	
	*	*	YCMHA6NF4E86	6600	205	DPH001	3151
2			Y0YACBYCB076	58FD		DC04	
	*	*	YUTUCYFW76AJ	6600	205	DPH002	3151

* - Non-configured disk unit

F3=Exit

F5=Refresh

F12=Cancel

PCIe Card NVMe IBM i DST/SST Screenshots



Display NVM Devices

Serial Number	Resource Name	Type	Capacity in GB			Namespaces		
			Used	Available	Total	Used	Avail	Total
Y0YACBYCB075	DC03	58FD	3151	0	3151	1	31	32
Y0YACBYCB076	DC04	58FD	3151	0	3151	1	31	32

F3=Exit

F5=Refresh

F12=Cancel

MA + A

01/001

PCIe Card NVMe IBM i DST/SST Screenshots



Select NVM Device

Type option, press Enter.
1=Select

Opt	Device	Serial Number	Resource Name	Type	Logical Address
<u>1</u>	NVM	Y0YACBYCB075	DC03	58FD	U78D2.001.WZS0ET2-P1-C8
-	NVM	Y0YACBYCB076	DC04	58FD	U78D2.001.WZS0ET2-P1-C2

F3=Exit

F12=Cancel

MA + A

09/002

PCIe Card NVMe IBM i DST/SST Screenshots



Delete Existing NVM Namespaces

Device	Serial Number	Resource Name	Type	Logical Address
NVM	Y0YACBYCB075	DC03	58FD	U78D2.001.WZS0ET2-P1-C8

Type option, press Enter.
4=Delete Namespace

Option	Serial Number	Type	Model	Resource Name	Capacity in GB
<u>4</u>	YCMHA6NF4E86	6600	205	DPH001	3151

F3=Exit

F7=Select All

F12=Cancel

MA + A

12/006

PCIe Card NVMe IBM i DST/SST Screenshots



Confirm Delete Existing NVM Namespaces

Device	Serial Number	Resource Name	Type	Logical Address
NVM	Y0YACBYCB075	DC03	58FD	U78D2.001.WZS0ET2-P1-C8

NVM Configuration	-----Capacity in GB-----			---Namespaces---		
	Used	Available	Total	Used	Avail	Total
Current :	3151	0	3151	1	31	32
Projected :	0	3151	3151	0	32	32

Attention: this function will destroy the data in the selected namespaces.

Option	Serial Number	Type	Model	Resource Name	Capacity in GB
4	YCMHA6NF4E86	6600	205	DPH001	3151

Press F10 to confirm the choice to delete the selected namespaces.
Press F12 to return to change your choice.

F10=Confirm F12=Cancel

PCIe Card NVMe IBM i DST/SST Screenshots



Work with NVM Device

Select one of the following:

1. Display NVM namespaces
2. Display NVM devices
3. Create NVM namespaces
4. Delete existing NVM Namespaces
5. Sanitize/Erase NVM device
6. Format NVM device to prepare device for IBM i

Selection

—

F3=Exit F12=Cancel

The Delete is complete.

MA + A

21/007

PCIe Card NVMe IBM i DST/SST Screenshots



Select NVM Device

Type option, press Enter.
1=Select

Opt	Device	Serial Number	Resource Name	Type	Logical Address
<u>1</u>	NVM	Y0YACBYCB075	DC03	58FD	U78D2.001.WZS0ET2-P1-C8
-	NVM	Y0YACBYCB076	DC04	58FD	U78D2.001.WZS0ET2-P1-C2

F3=Exit

F12=Cancel

MA + A

09/002

PCIe Card NVMe IBM i DST/SST Screenshots



Create NVM Namespaces

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

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

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



F3=Exit

F5=Refresh

F12=Cancel

MA + A

12/050

PCIe Card NVMe IBM i DST/SST Screenshots

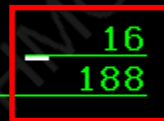


Create NVM Namespaces

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

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

Type choices, press Enter.
Quantity of namespaces to create . . . . . : 16 (1 - 32)
Capacity of each namespace . . . . . : 188 ( 64 - 3151) GB
```



F3=Exit

F5=Refresh

F12=Cancel

MA + A

12/050

PCIe Card NVMe IBM i DST/SST Screenshots



Confirm Create NVM Namespaces

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

Quantity of namespaces to create : 16
Capacity of each namespace to create : 188 GB

NVM Configuration	-----Capacity in GB-----			---Namespaces---		
	Used	Available	Total	Used	Avail	Total
Current	0	3151	3151	0	32	32
Projected	3008	143	3151	16	16	32

Note: Each namespace will be shown as a non-configured disk unit when the create operation completes.

Press F10 to confirm the choice to create namespaces.
Press F12 to return to change your choice.

F10=Confirm F12=Cancel

PCIe Card NVMe IBM i DST/SST Screenshots



Work with NVM Device

Select one of the following:

1. Display NVM namespaces
2. Display NVM devices
3. Create NVM namespaces
4. Delete existing NVM Namespaces
5. Sanitize/Erase NVM device
6. Format NVM device to prepare device for IBM i

Selection

—

F3=Exit F12=Cancel

The Create is complete.

MA + A

21/007

PCIe Card NVMe IBM i DST Screenshots



Display NVM Namespaces

NVM Device	ASP	Unit	Serial Number	Type	Model	Resource Name	Namespace Capacity in GB
1	*	*	Y0YACBYCB075	58FD		DC03	
	*	*	YMP9KAPZZW6G	6B7D	205	DD001	188
	*	*	YHKACXU3B3JB	6B7D	205	DD002	188
	*	*	Y9HQFUP5DPU2	6B7D	205	DD003	188
	*	*	YCMHA6NF4E86	6B7D	205	DD004	188
	*	*	YYQ7VVMTS8NH	6B7D	205	DD005	188
	*	*	YZHPQM9WD5DX	6B7D	205	DD006	188
	*	*	YNEM3WDHNQXS	6B7D	205	DD007	188
	*	*	YARDQBLT8VJG	6B7D	205	DD008	188
	*	*	YVDYFZCESX2Z	6B7D	205	DD009	188
	*	*	YCEN7SFBR7RC	6B7D	205	DD010	188
	*	*	YH6QQV83T562	6B7D	205	DD011	188
	*	*	YCCX3L6HZSQC	6B7D	205	DD012	188
	*	*	Y3QBNVUBAFYS	6B7D	205	DD013	188
	*	*	Y47PSM928RCJ	6B7D	205	DD014	188
	*	*	YASDH52XL3D9	6B7D	205	DD015	188

More...

* - Non-configured disk unit
 F3=Exit F5=Refresh F12=Cancel

PCIe Card NVMe IBM i WRKDSKSTS Screenshots



Work with Disk Status

IBMI4

12/03/19 12:49:08 CST

Elapsed time: 00:00:00

Unit	Type	Size (G)	% Used	I/O Rqs	Request Size (K)	Read Rqs	Write Rqs	Read (K)	Write (K)	% Busy
1	6B7D	188	5.5	.0	.0	.0	.0	.0	.0	0
1	6B7D	188	5.5	.0	.0	.0	.0	.0	.0	0
2	6B7D	188	1.5	.0						
2	6B7D	188	1.5	.0						
3	6B7D	188	1.5	.0						
3	6B7D	188	1.5	.0						
4	6B7D	188	1.5	.0						
4	6B7D	188	1.5	.0						
5	6B7D	188	1.5	.0						
5	6B7D	188	1.5	.0						
6	6B7D	188	1.5	.0						
6	6B7D	188	1.5	.0						
7	6B7D	188	1.5	.0						

Work with Disk Status

Elapsed time: 00:00:00

Unit	ASP	Type	Status	Protection	Compression
1	1	MRR	ACTIVE		
1	1	MRR	ACTIVE		
2	1	MRR	ACTIVE		
2	1	MRR	ACTIVE		
3	1	MRR	ACTIVE		
3	1	MRR	ACTIVE		
4	1	MRR	ACTIVE		
4	1	MRR	ACTIVE		
5	1	MRR	ACTIVE		
5	1	MRR	ACTIVE		
6	1	MRR	ACTIVE		
6	1	MRR	ACTIVE		
7	1	MRR	ACTIVE		

Command

===>

F3=Exit F5=Refresh F12=Cancel F24=More keys

PCIe Card NVMe IBM i DST/SST Screenshots



Display Disk Configuration Protection

ASP Unit	Serial Number	Type	Model	Resource Name	Protection	Hot Spare Protection
1	YMP9KAPZZW6G	6B7D	205	DD001	Mirrored planar	N
1	YRU8DUHFXXQF	6B7D	205	DD017	planar	N
2	YHKACXU3B3JB	6B7D	205	DD002	planar	N
2	YHDW8LMGBL8N	6B7D	205	DD018	planar	N
3	Y9HQFUP5DPU2	6B7D	205	DD003	planar	N
3	Y282DTK6GKDF	6B7D	205	DD019	planar	N
4	YCMHA6NF4E86	6B7D	205	DD004	planar	N
4	YUTUCYFW76AJ	6B7D	205	DD020	planar	N
5	YYQ7VVMTS8NH	6B7D	205	DD005	planar	N
5	YE9JSGR8CYWW	6B7D	205	DD021	planar	N
6	YZHPQM9WD5DX	6B7D	205	DD006	planar	N
6	YA9DEAKY2ZQM	6B7D	205	DD022	planar	N
7	YNEM3WDHNQXS	6B7D	205	DD007	planar	N

More...

Press Enter to continue.

F3=Exit
F11=Display non-configured units

F5=Refresh

F9=Display disk unit details
F12=Cancel



Checking the amount of remaining life in NVMe devices

☰ Table of Contents Change version or product ▾

Learn how to use operating system commands to find the amount of remaining life in an NVMe device.

Using the operating system commands

The AIX®, IBM® i, and Linux operating systems have commands that you can use to determine the amount of life that is remaining in an NVMe device.

- [Running the AIX command to check the amount of remaining life in NVMe devices](#)
Find information about using the AIX operating system to find the amount of remaining life in an NVMe device.
- [Running the IBM i command to check the amount of remaining life in NVMe devices](#)
Find information about using the IBM i operating system to find the amount of remaining life in an NVMe device.
- [Running the Linux smart-log command to check the amount of remaining life in NVMe devices](#)
Find information about using the Linux operating system smart-log command to find the amount of remaining life in an NVMe device.

https://www.ibm.com/support/knowledgecenter/en/9080-M9S/p9hak/pxhak_nvme_remaining_life_kickoff.htm

IBM i NVMe Fuel Gauge



```
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 : QPCSMPRT 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

MA + A

03/022

NVMe, IBM i, and SQL (new 2Q2020)



- The QSYS2.SYSDISKSTAT view contains information about disks
- The view is enhanced to recognize NVMe devices

```
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;
```

- Mixed SSD and NVMe example

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

NVMe, IBM i, and SQL (new 2Q2020)



- All NVMe system example (using 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	1.854
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

Done: 16 rows retrieved.



Modernize your platform with IBM i 7.4

Use the latest features to drive innovation

Home > IBM i 7.4 > Systems management > Disk management >

NVMe

Table of Contents Change version or product

>

NVMe (non-volatile memory express) is a host controller interface and storage protocol created to accelerate the transfer of data between enterprise and client systems and solid-state drives (SSD) over a computer's high-speed Peripheral Component Interconnect Express (PCIe) bus.

- **NVMe concepts**
NVMe (non-volatile memory express) is a host controller interface and storage protocol created to accelerate the transfer of data between enterprise and client systems and solid-state drives (SSD) over a computer's high-speed Peripheral Component Interconnect Express (PCIe) bus.
- **Managing NVMe**
This topic describes how to manage your NVMe devices and namespaces.

Parent topic:

→ [Disk management](#)

Related reference

- [NVMe failure with disk units mirror protected](#)
- [Installing a new partition using a NVMe as the load source disk unit](#)
- [Removing an operational NVMe device from a partition](#)
- [Upgrading the system ASP to NVMe - mirror protected](#)
- [Upgrading the system ASP to NVMe - not mirror protected](#)
- [Upgrading a user ASP to NVMe](#)

https://www.ibm.com/support/knowledgecenter/ssw_ibm_i_74/rzaly/rzalynvme.htm

IBM i I/O Support Summary – PCIe Card NVMe

IBM i I/O Support	Type of Configuration (Native, VIOS, iVirt, All)	IBM i 7.4	IBM i 7.3	IBM i 7.2
Enhancements from Nov 2019				
#EC6V, #EC6U - PCIe3 x8 NVMe 1.6 TB SSD NVMe Flash Adapter - IBM i	All	Tech Refresh 1	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#EC6X, #EC6W- PCIe3 x8 NVMe 3.2 TB SSD NVMe Flash Adapter - IBM i	All	Tech Refresh 1	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#EC6Z, #EC6Y - PCIe3 x8 NVMe 6.4 TB SSD NVMe Flash Adapter for IBM i	All	Tech Refresh 1	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#ESB9 - 387GB Enterprise SAS 4k SFF-3 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESBB - 387GB Enterprise SAS 4k SFF-2 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESBF - 775GB Enterprise SAS 4k SFF-3 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESBH - 775GB Enterprise SAS 4k SFF-2 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESBK - 1.55TB Enterprise SAS 4k SFF-3 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESBM - 1.55TB Enterprise SAS 4k SFF-2 SSD for IBM i	Native, iVirt	Base	Tech Refresh 3	Tech Refresh 7
#ESB8 - 387GB Enterprise SAS 4k SFF-3 SSD for AIX/Linux	VIOS	Base	Tech Refresh 3	Tech Refresh 7
#ESBA - 387GB Enterprise SAS 4k SFF-2 SSD for AIX/Linux	VIOS	Base	Tech Refresh 3	Tech Refresh 7
#ESBE - 775GB Enterprise SAS 4k SFF-3 SSD for AIX/Linux	VIOS	Base	Tech Refresh 3	Tech Refresh 7

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

IBM i I/O Support Summary – U.2 NVMe

IBM i I/O Support	Type of Configuration (Native, VIOS, iVirt, All)	IBM i 7.4	IBM i 7.3	IBM i 7.2
#EJ1Q - NVMe U.2 Passthru adapter Gen4 capable	All	Tech Refresh 2	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#EJ1S - Storage backplane with 6 SFF-3 Bays and 2 front PCIe Gen4 NVMe U.2 drive slots	All	Tech Refresh 2	Tech Refresh 8	Resave RS-720-Q
#EJ1T, #EJ1V - Storage backplane with 2 front PCIe Gen4 NVMe U.2 drive slots	All	Tech Refresh 2	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#EJ1U, #EJ1W - Storage backplane with 4 front PCIe Gen4 NVMe U.2 drive slots	All	Tech Refresh 2	Base (VIOS, iVirt)	Base (VIOS, iVirt)
#ES1F - Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i	Native, iVirt	Tech Refresh 2	Base (iVirt only)	Base (iVirt only)
#ES1H - Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i	Native, iVirt	Tech Refresh 2	Base (iVirt only)	Base (iVirt only)
#EC5W - Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i	Native, iVirt	Tech Refresh 2	Base (iVirt only)	Base (iVirt only)
#ES1E - Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	VIOS	Base	Base	Base
#ES1G - Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	VIOS	Base	Base	Base
#EC5V - Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	VIOS	Base	Base	Base

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

IBM i I/O Support Details – PCIe Card NVMe

November 2019 - IBM i 7.4 Technology Refresh 1 and IBM i 7.3 Technology Refresh 7 and IBM i 7.2

For more details on the new Power I/O features and enhancements listed below, see the October 8, 2019, announcement letter [IBM Power Systems enhancements](#).

Dedicated and VIOS support for selected PCIe3 x8 SSD NVMe adapters - IBM i 7.4 TR 1, IBM i 7.3, IBM i 7.2

IBM i 7.4 TR1 natively supports PCIe3 x8 SSD NVMe adapters that provide multiple capacity points for enterprise workloads on selected Power servers with POWER9™ technology. Support is for selected dedicated and VIOS VSCSI attached NVMe devices. These low latency devices can be used as IBM i load sources and are able to provide a high number of IOPS and enhanced virtualization capabilities. Mirroring is required. Pairs of these storage devices can be added to LPAR configurations as dedicated PCIe devices.

Each device is a partitionable endpoint that can be dedicated to an LPAR, meaning that multiple partitions may be configured for a single system unit without the need to use virtualization.

IBM i 7.4, 7.3, and 7.2 configurations with VIOS are also supported as virtual SCSI drives that are backed by NVMe devices. Other previously announced devices can be used for the VIOS boot disk but should not be virtualized to IBM i. For VIOS configurations, support is for VIOS VSCSI LVM client only, so block size is 4096.

	IBM i	IBM i with VIOS
1.6 TB	#EC6V, #EC6U	#EC5B, #EC5G
3.2 TB	#EC6X #EC6W	#EC5D, #EC5C
6.4 TB	#EC6Z, #EC6Y	#EC5F, #EC5E

Additional code levels required:

FW940, or later

VIOS 3.1.0 (for VIOS configurations)

New Enterprise SSDs – 387 GB, 775 GB, 1.55 TB - IBM i 7.4, IBM i 7.3 TR 3, and IBM i 7.2 TR 7

A new generation of enterprise 2.5-inch solid-state drives (SSD) improves enterprise-class reliability, endurance, and capacity characteristics.

<https://www.ibm.com/support/pages/node/1137508>

IBM i I/O Support Details – U.2 NVMe

July 2020 - IBM i 7.4 Technology Refresh 2 and IBM i 7.3 Technology Refresh 8 and IBM i 7.2 Resave RS-720-Q

For more details about the new I/O for Power Systems with POWER9 technology, see the features listed and described in the following July 14th announcement letters:

- IBM Power S914 model 9009-41G
- IBM Power S924 model 9009-42G
- IBM Power S922 model 9009-22G, which includes the Power S922 with a 1-core processor

Storage backplanes and Passtru Card for Power Systems Scale Out models with POWER9 technology and PCIe gen4 switches

There are several new options for storage backplanes for the new Power Systems Scale Out models with POWER9 technology and PCIe gen4 slots. There is also a passtru card that is required for cabling some of NVMe devices to the storage backplane in some configurations, such as with the #EJ1T and #EJ1V storage backplanes.

All options are supported natively for IBM i 7.4. However, only #EJ1S is supported natively for IBM i 7.3 and 7.2. All options are supported as IBM i virtual client for IBM i virtualization and VIOS configurations.

The storage backplane options provide different numbers of SFF-3 Bays for SAS disk and PCIe Gen4 U.2 bays for NVMe:

	Model 41G & 42G	Model 22G (8-, 10-, 11-cores)	Model 22G (1-core)
NVMe U.2 Passtru adapter Gen4 capable	#EJ1Q	#EJ1Q	#EJ1Q
6 SFF-3 bays & 2 NVMe U.2 drive slots	#EJ1S	N/A	N/A
4 NVMe U.2 drive slots	#EJ1U	#EJ1W	#EJ1W
2 NVMe U.2 drive slots	#EJ1T	#EJ1V	#EJ1V

Dedicated and VIOS support for Enterprise PCIe4 NVMe U.2 modules - IBM i 7.4 TR 2, IBM i 7.3 TR 8, IBM i 7.2

IBM i 7.4 TR2 natively supports selected Enterprise PCIe4 NVMe U.2 modules that provide multiple capacity points for enterprise workloads on selected Power servers with POWER9™ technology. These low latency devices can be used as IBM i load sources and are able to provide a high number of IOPS and enhanced virtualization capabilities. Mirroring is required. Mirrored sets of these storage devices can be added to LPAR configurations as dedicated PCIe devices. Support is also provided for VIOS VSCSI attached NVMe devices.

Each mirrored set of devices is a partitionable endpoint that can be dedicated to an LPAR, meaning that multiple partitions may be configured for a single system unit without the need to use virtualization.

IBM i 7.4, 7.3, and 7.2 configurations with VIOS are also supported as virtual SCSI drives that are backed by NVMe devices. For VIOS configurations, support is for VIOS VSCSI LVM client only, so block size is 4096.

	IBM i	IBM i with VIOS
1.6 TB	#ES1F	#ES1E
3.2 TB	#ES1H	#ES1G
6.4 TB	#EC5W	#EC5V

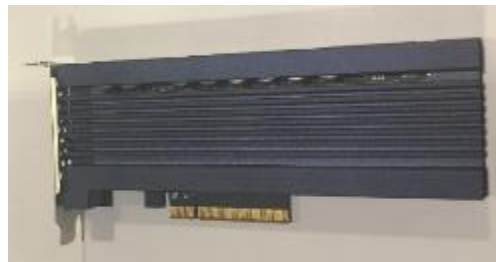
<https://www.ibm.com/support/pages/node/1137508>



Summary...



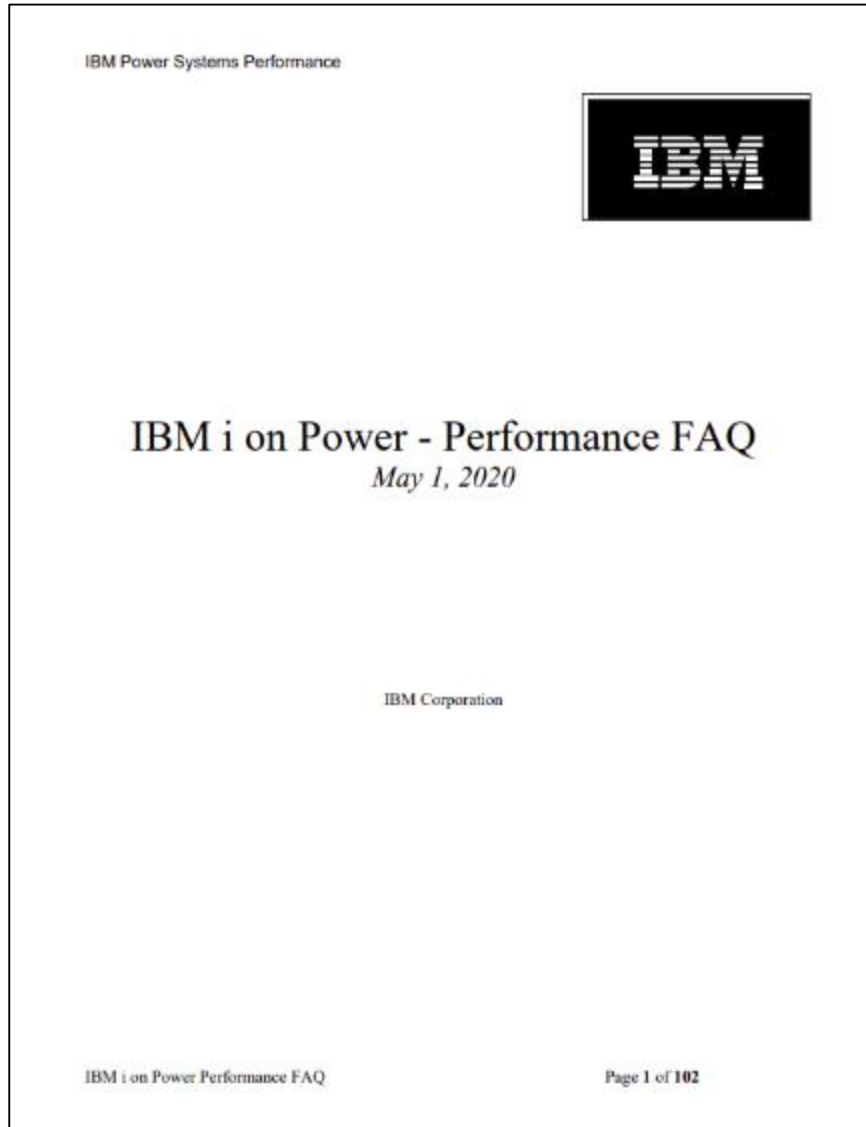
- Multiple NVMe devices are available on IBM POWER9 servers
- NVMe devices are supported on AIX, Linux, IBM i and VIOS
- NVMe devices are easy to install and configure
- NVMe devices are fast



Related Information



IBM i on Power - Performance FAQ



8.5 I/O QUESTIONS

- What are some common disk subsystem configuration changes that can hurt performance?
- How can I analyze I/O subsystem performance?
- I can't modify my applications. What can I do to improve my I/O performance?
- How do I virtualize IBM Serial-Attached SCSI (SAS) adapters for the best performance?
- How can I tell if my applications will benefit from SSDs?
- How many and what type of storage I/O products will meet performance requirements for an upgrade or new system or workload?
- How can I tell if a FlashSystem solution is the best storage configuration for my environment?
- How do I analyze I/O performance if I'm using external storage?
- What are the IBM i performance effects of using 4096 byte sector drives?
- Should I utilize NVMe drives?

E950/E980 NVMe Options



Hardware Features

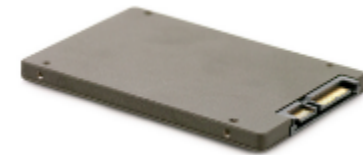
- NVMe Specs. 1.2.1 Compliant
- PCIe Gen 3 x 4
- SFF U.2 (2.5") 7mm z-Height
- Power < 12W
- Block Size 4096(Default), 512
- Non Volatile Write Buffer
- Endurance 2.4 DWPD
- Power Loss Protection
- RAIF: Tolerant of single flash die failures
- PCIe Vendor VPD Support
- Boot: Option ROM BAR 128KB
- Warranty 5 years
- Hot Plug capable
- MTBF ≥ 2 million hours
- End Of Life Data Retention ≥ 3 months
- EEH Support
- Live Firmware Update
- NVMe-MI Support

	PCIe3 800GB NVMe U.2 Slim SSD		
	800GB	1.6TB	3.2TB
FC	EC5J	EC5K	EC5L

Workload	Target (800GB)	Target (1.6/3.2TB)
Read (IOPS)	380K	500K
Write (IOPS)	60K	80K
Mixed R/W (70/30)	135K	150K
Read Data Tp (GB/s)	1.6	2.6/3.2
Write Data Tp (GB/s)	1.05	1.9
Read Latency (us)	80	80
Write Latency (us)	20	20

Software Support

- AIX, VIOS
- Linux



Supported POWER9 Platforms

- E950
- E980

Notes:

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

Where are NVMe Storage Devices in e-Config?

9009-42G: Server 1

Storage

Data Protection
 (0040) Mirrored System Disk Level, Specify Code (for IBM i use only)

[Filter Hard Drives](#)

Hard Drives	Proposed
(EC5W)-Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for IBM i	0
(ES1F)-Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for IBM i	0
(ES1H)-Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for IBM i	0
(ES91)-387GB Enterprise SAS 4k SFF-3 SSD for IBM i	0
(ES95)-387GB Enterprise SAS 4k SFF-2 SSD for IBM i	0
(ESB9)-387GB Enterprise SAS 4k SFF-3 SSD for IBM i	0
(ESBB)-387GB Enterprise SAS 4k SFF-2 SSD for IBM i	0

Additional Hardware	Proposed
(EJ1C)-Base Storage Backplane 12 SFF-3 Bays/RDX Bay	0
(EJ1D)-Expanded Function Storage Backplane 18 SFF-3 Bays/Dual IOA with Write Cache/Opt Ext SAS port	1
(EJ1M)-Expanded Function Storage Backplane 12 SFF-3 Bays/RDX Bay/Opt Ext SAS port	0
(EJ1S)-Storage Backplane Gen4 with 6 SFF-3 Bays and 2 NVMe U.2 drive slots	0
(EJ1T)-Storage Backplane Gen4 with 2 NVMe U.2 drive slots	0
(EJ1U)-Storage Backplane Gen4 with 4 NVMe U.2 drive slots	0

Where are NVMe Storage Devices in e-Config?

9009-42G: Server 1

Adapters

Category to View:
NVMe Adapter

Category Information:
Views are sorted by feature code

Form Factor
All

Adapters	Proposed
(EC6V)-PCIe3 x8 1.6 TB NVMe Flash Adapter for IBM i	0
(EC6X)-PCIe3 x8 3.2 TB NVMe Flash Adapter for IBM i	0
(EC6Z)-PCIe3 x8 6.4 TB NVMe Flash Adapter for IBM i	0

Configuration Information: (Additional features auto-configured to satisfy selections where possible)
PCI-E Slot: 1 of 8 used
Selections:
(1) (5899)-PCIe2 4-port 1GbE Adapter

Products ✓
Summary ✓
System ✓
Storage ✓
Adapters
Expansions
Power
Codes
CSP/LPAR
OS PCS
SOSWOS
Implementation Services
Integrated Supt Svcs

Load Source

(ESLV)-Load Source Specify for EC6V (NVMe 1.6 TB SSD for IBM i)

(Base)-Internal Load Source

Order Codes

(EB74)-IBM i 7.4 Indicator

(ECP0)-Cloud Private Solution

(ENS1)-188 GB IBM i NVMe Load Source Namespace size

(ENS2)-393 GB IBM i NVMe Load Source Namespace size

NVMe Devices in AIX – Listing NVMe Devices in AIX



```
TERMINAL — ttys000
[AIX 7.2 Server :> lscfg | grep nvme
+ nvme0          U78D5.ND3.CSS100A-P2-C1          PCIe3 x4 NVMe Flash Adapter
[AIX 7.2 Server :> lscfg -vp -l nvme0
nvme0          U78D5.ND3.CSS100A-P2-C1  PCIe3 x4 NVMe Flash Adapter

800GB NVMe Gen3 U.2 Slim SSD      :
Part Number.....@1CM501
EC Level.....P02092
FRU Number.....@1LU763
Product Specific.(AN).....@1LU760
Feature Code/Marketing ID...EC5J
Customer Card ID Number.....59B4
Serial Number.....84YAB0DV
Product Specific.(Z0).....10140637
Product Specific.(Z1).....2.4
Product Specific.(Z2).....800
Product Specific.(Z3).....3.50
Product Specific.(Z4).....0
Product Specific.(Z5).....0
Product Specific.(Z6).....A1800101
Product Specific.(Z7).....0YA
Product Specific.(Z8).....2637
Product Specific.(Z9).....08
Product Specific.(ZA).....5A
Product Specific.(ZB).....08
Manufacture ID.....34NY0K100518
Product Specific.(RV).....
ROM Level.(alterable).....REV.MN38
Hardware Location Code.....U78D5.ND3.CSS100A-P2-C1

PLATFORM SPECIFIC
```

NVMe Devices in AIX – Using SMIT



```
TERMINAL — ttys000
System Management

Move cursor to desired item and press Enter.

Software Installation and Maintenance
Software License Management
Manage Editions
Devices
System Storage Management (Physical & Logical Storage)
Security & Users
Communications Applications and Services
Workload Partition Administration
Print Spooling
Advanced Accounting
Problem Determination
Manage the AIX Cryptographic Framework
Performance & Resource Scheduling
System Environments
Processes & Subsystems
Applications
Installation Assistant
Electronic Service Agent
Using SMIT (information only)

F1=Help          F2=Refresh      F3=Cancel      Esc+8=Image
Esc+9=Shell      Esc+0=Exit      Enter=Do
```

NVMe Devices in AIX – Using SMIT



```
TERMINAL — ttys000
Devices

Move cursor to desired item and press Enter.

[MORE...13]
Cryptographic Adapters
Communications
Graphic Displays
Graphic Input Devices
Low Function Terminal (LFT)
Target Mode Initiator Device
SCSI Adapter
FC Adapter
IDE Adapters
iSCSI
Virtual SCSI Adapters
VirtIO SCSI Adapters
VirtIO SCSI Adapters
I/O Completion Ports
Multimedia
List Devices
Configure/Unconfigure Devices
Install Additional Device Software
PCI Hot Plug Manager
USB Adapter
SAS Adapter
SSD Cache Devices
SAS Flash Adapter
NVMe Manager
[BOTTOM]

F1=Help          F2=Refresh      F3=Cancel      Esc+8=Image
Esc+9=Shell      Esc+0=Exit     Enter=Do
```


NVMe Devices in AIX – Using SMIT



```
TERMINAL — ttys000
NVMe Manager

Move cursor to desired item and press Enter.
List NVMe Storage Configuration
Configure a Defined NVMe Controller
Change / Show Characteristics of a NVMe Controller
Namespace Management and Attachment

F1=Help          F2=Refresh      F3=Cancel      Esc+8=Image
Esc+9=Shell     Esc+0=Exit     Enter=Do
```

NVMe Devices in AIX – Using the nvmemgr Command



```
TERMINAL — ttys000
AIX 7.2 Server :> nvmemgr -?
Usage:
nvmemgr -A -l [ nvme# | hdisk# ]
nvmemgr -B -l nvme# -o opt#
nvmemgr -C -l nvme# -s size -b block_size
nvmemgr -D -l nvme# -p logpg_id
nvmemgr -E -l nvme# -f <feature id> -s <selection> [-v <int vect#>]
nvmemgr -H -l nvme# -c commit_action -s <slot number>
nvmemgr -I -l nvme# -z <FW_path> -c commit_action -s <slot number>
nvmemgr -J -l nvme# -k #
nvmemgr -S -l nvme#
nvmemgr -M -l [ nvme# | hdisk# ]
nvmemgr -Q -l nvme# -i <interval>
nvmemgr -R -l nvme# -d [hdisk# | ns_id]
nvmemgr -X -l nvme# -n ns_id
nvmemgr -U -l nvme# -d [hdisk# | ns_id]

-A Print Identify
  -l 'lname' - Adapter or disk logical name

-B Print list of namespaces attached to a controller/adapter
  -l 'lname' - Adapter logical name
  -o opt# - Option (0 - Active Namespaces, 1 - All Namespaces)

-C Create a Namespace
  -l 'lname' - Adapter logical name
  -s size - Namespace size (in MB, GB or TB)
  -b lba_format# - LBA Format
  -o opt# - Attach option (0 - Create Only, 1 - Attach after create)

-D Get Logpage
  -l 'lname' - Adapter logical name
```

NVMe Devices in AIX - Display nvmeX Device Details



```
TERMINAL — ttys000
AIX 7.2 Server :> nvmemgr -A -l nvme0
NVME Controller Identify
=====
PCI Vendor ID (pci_vid) = 0x144d
PCI Subsystem Vendor ID (pci_ssvid) = 0x1014
Serial number = "S434NY0K100518"
Model number = "800GB NVMe Gen3 U.2 Slim SSD"
Firmware Revision (fw_rev) = REV.MN38
Recommended Arbitration Burst (rab) = 0x2
IEEE OUI Identifier (IEEE) = 0x38250
Controller Multi-Path I/O and Namespace Sharing Capabilities (CMIC) = 0x0
    Controller is associated with PCI Function.
    NVM subsystem contains only a single controller.
    NVM subsystem contains only a single PCI Express port.

Maximum Data Transfer Size (MDTS) = 0x9
    i.e. Maximum Data Transfer Size is 0x200 minimum memory pages

Controller ID (cntlid) = 0x4
NVME Version supported = 1.2
RTD3 Resume Latency (RTD3R) = 0x7a1200
RTD3 Entry Latency (RTD3E) = 0x7a1200
Optional Asynchronous Events Supported (OAES) = 0x0
Optional Admin Command Support(OACS) = 0xf
    Controller supports the Namespace Management and Namespace Attachment commands.
    Controller supports the Firmware Commit and Firmware Image Download commands.
    Controller supports the Format NVM command.
    Controller supports the the Security Send and Security Receive commands.

Abort Command Limit (acl) = 0x7
Asynchronous Event Request Limit(aerl) = 0x3
Firmware Updates (frmw) = 0x16
```

NVMe Devices in AIX – Display NVMe hdiskX Device Details



```
TERMINAL — ttys000
[AIX 7.2 Server :> nvmemgr -A -l hdisk1
NVMe Namespace identify for nsid = 1
=====
Namespace Size (in LBAs) = 0xba4d9d6
Namespace capacity (in LBAs) = 0xba4d9d6
Namespace Utilization (in LBAs) = 0x95ffda8
Namespace Features = 0x2
Number of LBA Formats supported = 0x1
Formatted LBA Size = 0x0
    i.e. NS has been formatted with LBA format 0
Metadata Capabilities = 0x0
End-to-end Data Protection Capabilities = 0x0

End-to-end Data Protection Type Settings = 0x0
    Protection information is not enabled
Namespace Multi-path I/O and Namespace Sharing Capabilities = 0x0
    NS is private to controller

Reservation Capabilities = 0x0

Format Progress Indicator (FPI) = 0x80
    FPI is supported.
    Remaining percentage of NS to be formatted = 0x0.

Namespace Atomic Write Unit Normal (NAWUN) = 0x7f
Namespace Atomic Write Unit Power Fail (NAWUPF) = 0x0
Namespace Atomic Compare & Write Unit (NACWU) = 0x0
Namespace Atomic Boundary Size Normal (NABSN) = 0x7f
Namespace Atomic Boundary Offset (NABO) = 0x0
Namespace Atomic Boundary Size Power Fail (NABSPF) = 0x0
NVM Capacity (in bytes) (NVMCAP) = 800166076416
Namespace Globally Unique Identifier (NGUID) = 0x01000000593825001805104b30343334
```

NVMe Devices in AIX – Display Namespaces



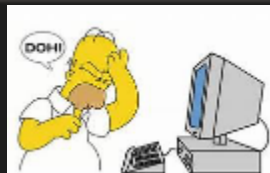
```
TERMINAL — ttys000
[AIX 7.2 Server :> nvmemgr -B -l nvme0
Name space identifier list
=====
NSID = 0x1
[AIX 7.2 Server :> nvmemgr -S -l nvme0
-----
Name          Location      State         Size
-----
nvme0         00-00        Available     800.17GB
  hdisk1      00-00        Available     800.17GB
AIX 7.2 Server :>
```

NVMe Devices in AIX – Delete a Namespace



```
TERMINAL — ttys000
AIX 7.2 Server :>
AIX 7.2 Server :>
AIX 7.2 Server :>
AIX 7.2 Server :> nvmemgr -U -l nvme0
Invalid Namespace or Format
AIX 7.2 Server :> nvmemgr -U -l nvme0 -d hdisk1
Namespace detached successfully.
AIX 7.2 Server :> nvmemgr -B -l nvme0

Name space identifier list
=====
AIX 7.2 Server :> █
```



NVMe Devices in AIX – Fuel Gauge



```
TERMINAL — ttys000
AIX 7.2 Server :> nvmmgr -M -l nvme0
Critical Warning ..... 0x0
Composite Temperature (Kelvin) ..... 307
Available Spare (%) ..... 100
Percentage of NVM subsystem life used ..... 0
Data Units Read (1000 units of 512 bytes) ..... 40608131
Data Units Written (1000 units of 512 bytes) ..... 19825240
Host Read Commands ..... 4071015568
Host Write Commands ..... 1896018759
Number of Power Cycles ..... 209
Power On Hours ..... 3829
Unsafe Shutdowns ..... 166
Media and Data Integrity Errors ..... 0
Number of Error Information Log Entries ..... 109
AIX 7.2 Server :>
```



- [NVMe Express Details on working with NVMe devices in Linux](#)
- NVM Express™ (NVMe™) technology has enabled a robust set of industry-standard software, drivers, and management tools that have been developed for storage. The tool to manage NVMe SSDs in Linux is called [NVMe Command Line Interface \(NVMe-CLI\)](#).

Command	Description
<code>nvme list</code>	Lists all the NVMe SSDs attached: name, serial number, size, LBA format, and serial
<code>nvme id-ctrl</code>	Discover information about NVMe controller and features it supports
<code>nvme id-ns</code>	Discover feature of NVMe namespaces, optimizations, features, and support
<code>nvme format</code>	Secure erase the data on an SSD, format an LBA size or protection information for end-to-end data protection
<code>nvme sanitize</code>	Securely erases all user data on the SSD
<code>nvme smart-log</code>	Outputs the NVMe SMART log page for health status, temp, endurance, and more
<code>nvme fw-log</code>	Outputs the firmware log page
<code>nvme error-log</code>	Outputs the NVMe error log page
<code>nvme reset</code>	Resets the NVMe controller / NVMe SSD
<code>nvme <vendor name> help</code>	e.g <code>nvme intel help</code> will display optional commands for Intel drives, this is the vendor plugins for <code>nvme-cli</code>
<code>nvme delete-ns</code>	Delete a namespace
<code>nvme create-ns</code>	Create a new namespace, e.g creating a smaller size namespace to overprovision an SSD for improved endurance, performance, and latency
<code>nvme fw-download</code>	Download a new firmware to the NVMe device
<code>nvme fw-commit</code>	Commit (activate) the firmware to run immediately or after the next reset



Daniel R Sundt

Partner Technical Advocate

BP Technical Sales & Enablement

IBM Power Systems



2800 37th ST NW
Rochester, MN 55901

Tel 507 253 3228
Mobile 507 261 5329
dansundt@us.ibm.com