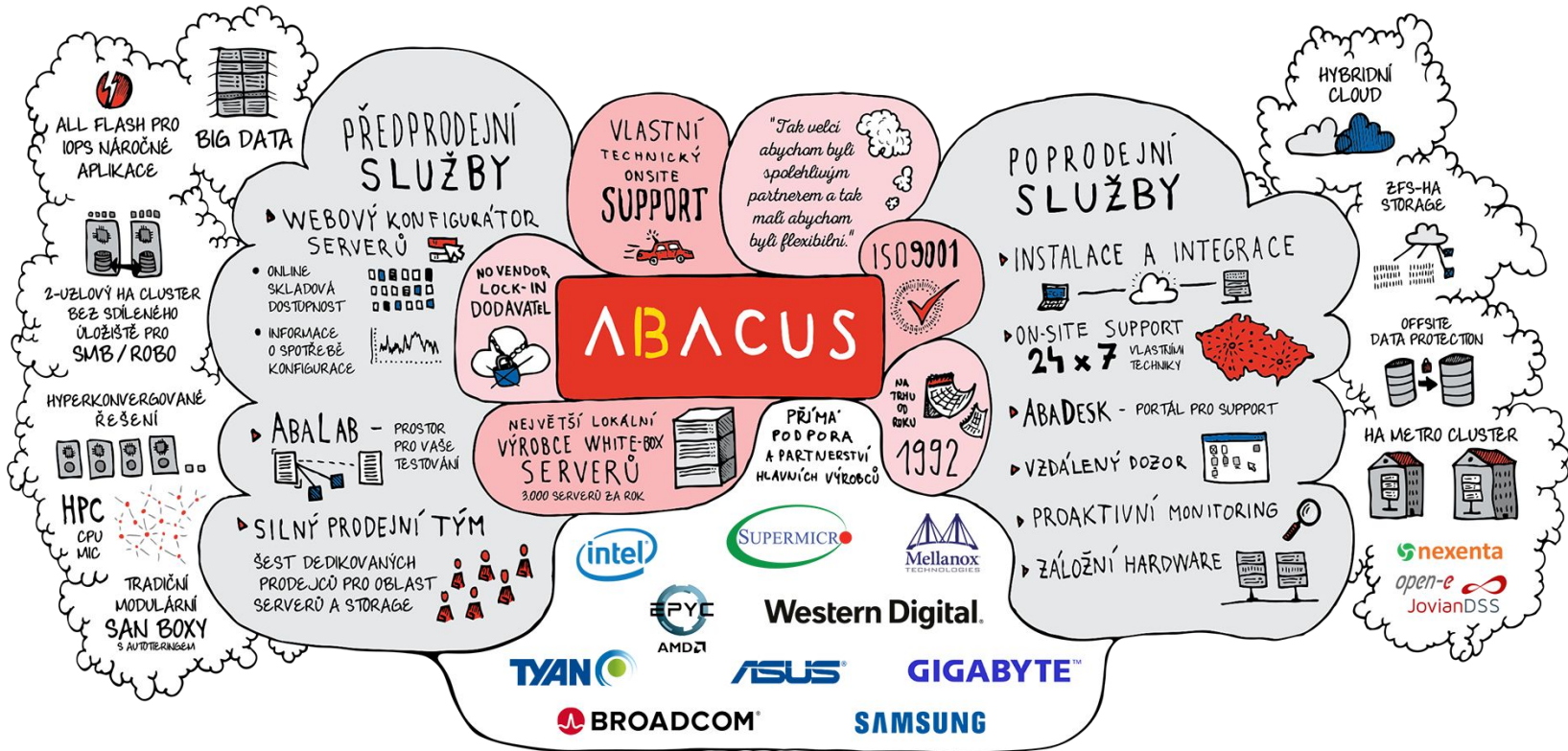


» Abacus 2022 servery a storage

Jan Petrák
jp@abacus.cz

ABACUS



- » Představení společnosti Abacus
- » Procesory Intel® Xeon® Scalable třetí generace (IceLake) a co dál
- » Procesory AMD EPYC™ 7003 třetí generace (Milan) a co dál
- » SSD – NVMe, Optane, formáty
- » NVMe & RAID
- » Servery
- » NVMe over Fabric
- » 25G a 100G Ethernet
- » SDS (Software-Defined-Storage) – přehled řešení
- » Co jsme nestihli a kde se dozvíte více

- » Největší výrobce serverů v CZ
- » Největší dodavatel „white-box“ serverů a příslušenství
 - » AIC, AMD, Asus, Areca, Broadcom, Gigabyte, Intel, Mellanox Supermicro, QSAN...
- » ISO 9001
- » Servery dodáváme od roku 2003
- » Téměř 3000 serverů v roce 2021
- » Certifikovaná řešení pro Open-e Jovian (ZFS)

IDC	1Q 2022	2Q 2022
Dell	1680	1244
HPe	1508	1499
Abacus	506	620
Fujitsu	152	90
Lenovo	136	237
Cisco	28	38



Co děláme (a jsme v tom dobří)

- » **Build-to-Order** - asemblace serverů
 - » Level 6 - Level 12
 - » OEM služby (zákaznické firmware, branding)
- » **Návrh a dodávky storage řešení (SDS)**
 - » Scale up (ZFS): Open-e, Nexenta
 - » Scale out: Ceph
- » **Zajištění podpory a SLA**
 - » Vlastní technici
 - » 8h fix
 - » Česká republika a Slovensko
- » **Kompletní dodávky pro SMB**
 - » servery, ukládání dat, networking, koncová zařízení, zálohování, dozor ...



» **AbaKonfigurátor**

- » Rozsáhlý webový konfigurátor serverů
- » 150 upravitelných konfigurací
- » Online kalkulačtor spotřeby (power budget)
- » Online informace o skladové dostupnosti

» **AbaLab**

- » Hands-on Lab – lokálně nebo vzdáleně
- » Try&Buy program

» **Abadesk**

- » Portál podpory
- » Výpis konfigurace
- » Aktuální verze firmware
- » Sledování a eskalace ticketů



- » **Nezávislo**
- » **Nejen x86**
- » **Vlastní se**
- » **Konfigura**
- » **Otevřeno**

Intel Technology Provider


PARTNER WITH AN HIGH PERFORMANCE COMPUTING (HPC) DATA CENTER SPECIALIST

Engage with an Intel® Technology Provider high performance computing (HPC) Data Center Specialist to architect your next HPC solution.



HPC Data Center Specialist

Intel® Technology Provider high performance computing (HPC) Data Center Specialist demonstrated deep experience in implementing and value-added benefits from Intel that differentiate the HPC Data Center Specialist.

Learn why you should partner with an HPC Data Center Specialist.

GOT AN HPC OPPORTUNITY?

Partner with an HPC Specialist today.

Americas
Asia Pacific and Japan
China



2CRSI
France, United Arab Emirates



Abacus Electric
Czech Republic



AMAX
Ireland

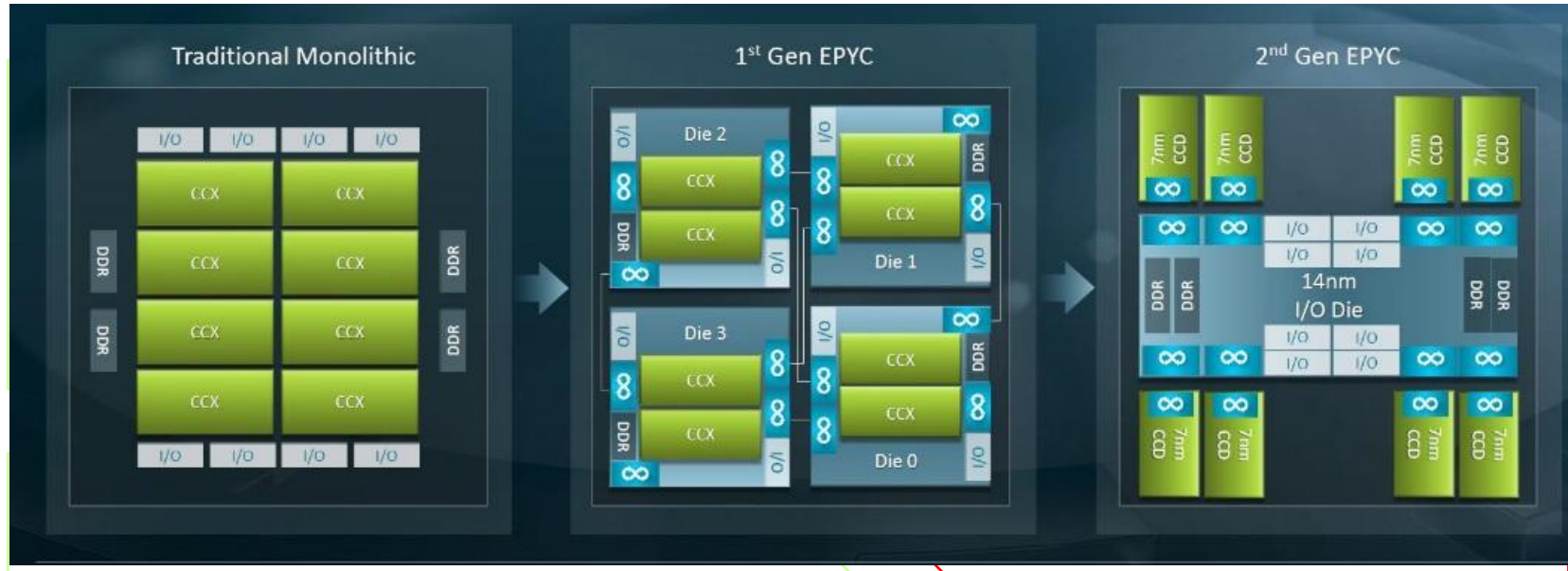
Proč „No Vendor Lock-in“?

- » **Dostupné HDD hotswap rámečky**
- » **Volně dostupný firmware**
Pro aktuální verze firmware není nutné mít zaplacený support
- » **Bez omezení na úrovni firmware:**
Lze osadit běžné CPU, SSD, HDD, RAM ...
- » **Dodržování standardů a jejich názvů:**
 - IPMI 2.0 (iLO, DRAC)
 - SDDC (ChipKill, Chipspare)
 - SAS expander
 - ...



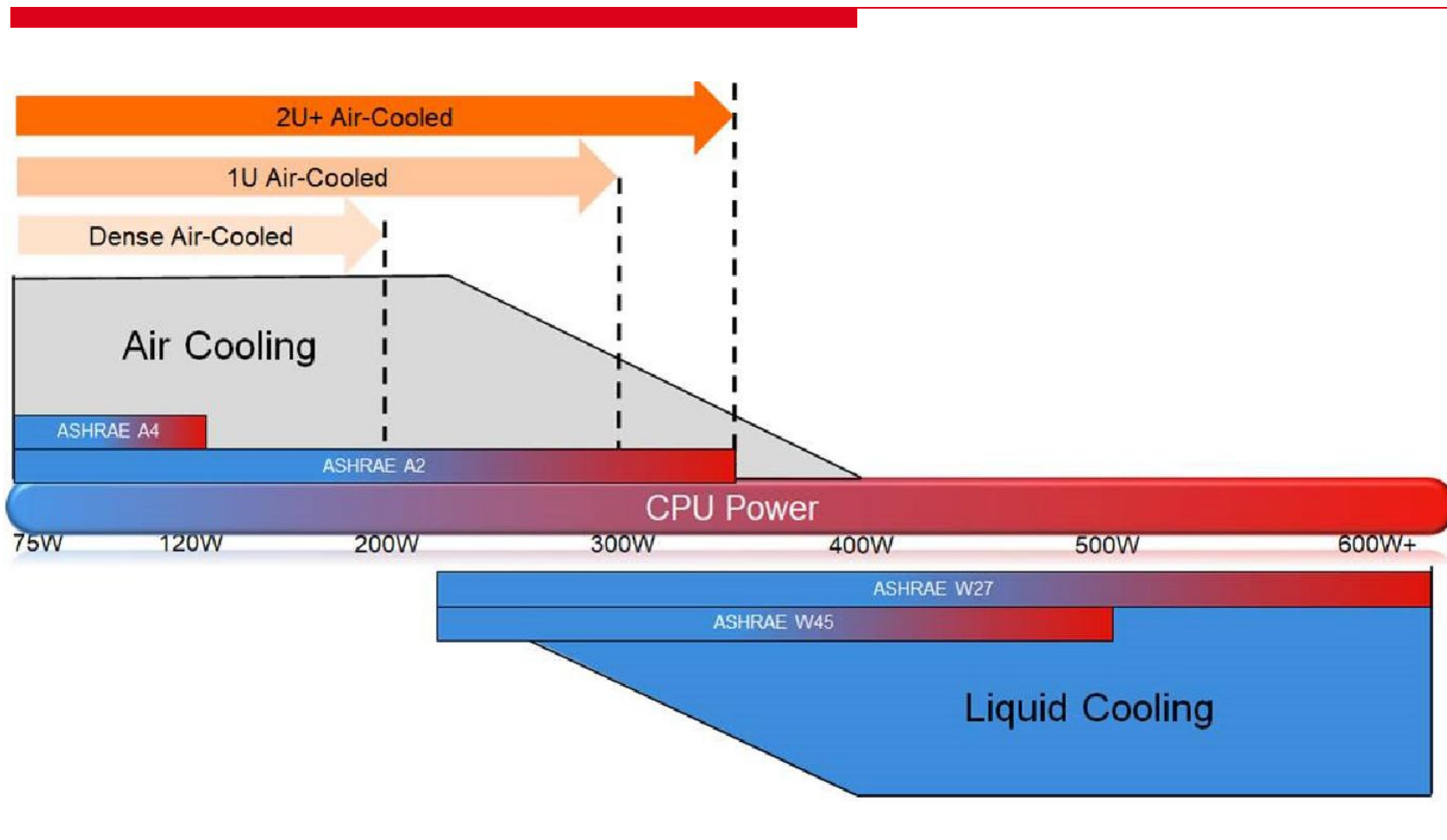
NO VENDOR LOCK-IN

Procesory (a jak zvýšit výkon)



Multi-die „dlaždičky“

Chlazení



Intel® Xeon® Scalable

	SkyLake (x1xx)	CascadeLake (x2xx)	IceLake (x3xx)	Sapphire Rapids
Tech.	14nm, monolithic	14nm, monolithic	10nm, monolithic	7nm, multie die
Max.jader	28 jader/max.8P	28 jader/max.8P	40 jader/max.2P	
RAM/CPU	DDR4-2666, 6ch. 768GB 1,5TB(+\$2800)	DDR4-2933, 6ch. 1TB, 2TB(+\$2800) 4,5TB(+\$7500)	DDR4-3200, 8ch	DDR5
PCIe/CPU	48 linek G3	48 linek G3	64 linek G4	G5
\$ 12core	\$1000 (4116/2,1GHz)	\$700 (4214/2,2GHz)	\$550 (4310/2,1GHz)	2023?
\$ 16core	\$1900 (6130/2,2GHz)	\$1000 (4216 / 2,1GHz)	\$700 (4314/2,4GHz)	
\$ 20core	\$2600 (6138/2GHz)	\$1900 (6230 / 2,1GHz)	\$1000 (4316/2,3GHz)	
\$ 22core	\$3650 (6152/2,1GHz)	\$2650 (6238/2,1GHz)	\$1300 (5318/24c@2,1GHz)	
\$ 32core			\$2500 (6338 / 2GHz)	

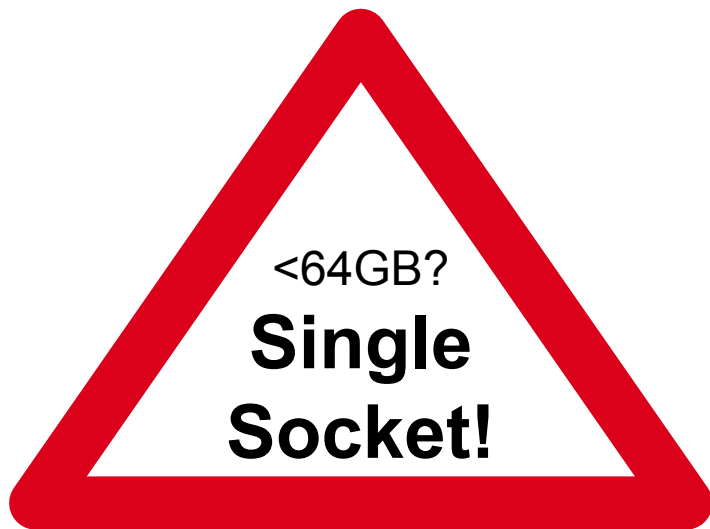
	Naples(Zen) 7xx1	Rome(Zen2) 7xx2	Milan/-X (Zen3) 7xx3/7xx3-X	Genoa/-X (Zen4) 9xx4	Bergamo (Zen4c) 9xx4
Tech	14nm, multiedie	7nm, multiedie	7nm+, multiedie	5nm, multiedie	4nm, multiedie
Max.jader	32 jader/2P	64 jader/2P	64 jader/2P	96 jader/2P	128 jader/2P
L3 cache	64MB	256MB	256/786MB	384/1152 MB	256.1B
RAM/CPU	DDR4-2666, 8ch., 2TB	DDR4-3200, 8ch., 4TB	DDR4-3200, 8ch., 4TB	DDR5, 12ch.	DDR5, 12ch.
PCIe/CPU	128 l., G3	128 l., G3	128 l., G4	128 l., G5	128 l., G5
TDP (max)	180W	240W	280W	400W	400W

AMD EPYC™ 7003 vs. Intel® Xeon® Scalable

		AMD EPYC2/3 (Rome/Milan)	Intel Xeon (IceLake x3xx)
Výrobní technologie		7nm, multi-die hybrid, SoC	10nm, monolitický čip
Max.jader/CPU		64 jader/2P	40 jader/max.2P
RAM/socket (8channel, 2DPC)		3200MHz, 16DIMM 4TB (bez příplatku)	3200MHz, 16DIMM 4TB + 2TB Optane PM200
PCI-Express		1P/2P: 128/162 linek G4	1P/2P: 64/128 linek G4
Cena procesoru	\$900	1P: 7302P (16c@3GHz, 4TB)	1P: 5317 (12c@3GHz, 4TB) 2P: 4310 (24c@2,1GHz, 8TB)
	\$1.300	1P: 7402P (24c@2,8GHz, 4TB)	1P: 5320 (26c@2,2GHz, 4TB) 2P: 4314 (32c@2,4GHz, 8TB)
	\$2.500	1P: 7502P (32c@2,5GHz, 4TB) 2P: 7352 (48c@2,3GHz, 8TB)	1P: 6314U (32c@2,3GHz, 4TB) 2P: 6326 (32c@2,4GHz, 8TB)
	\$4.500 <i>(CPU2017 base_rate_INT/FP)</i>	1P: 7702P (64c@2GHz, 4TB) 1P: 7643 (48c@2,3GHz, 4TB, 2P: 7452 (64c@2,35GHz, 8TB, 400/350) 2P: 7443 (48c@2,85GHz, 4TB, 420/410)	1P: 8360Y(36c@2,4GHz, 4TB) 2P: 6346 (32c@3,1GHz, 8TB, 280/300)
	\$15.000 <i>(CPU2017 base_rate_INT/FP)</i>	2P: 7742 (128c@2,25GHz, 8TB, 680/520) 2P: 7763 (128c@2,45GHz, 8TB, 800/650)	2P: 8380 (80c@2,3GHz, 8TB, 560/470)

Optimální počet DIMM modulů

	Optimální počet DIMMů/CPU	Minimální kapacity RAM/CPU (GB)
AMD EPYC 7001/7002/7003	8 / 16, tedy 16 / 32 v 2P serveru	64GB (8×8GB), dále pak 128GB/256GB/512GB
Intel Xeon E-2100/2200/2300	2 / 4	16GB (2×8GB)
Intel Xeon Scalable 3.gen.	8 / 16, tedy 16 / 32 v 2P serveru	64GB (8×8GB), dále pak 128GB/256GB/512GB
AMD EPYC 9004	12, tedy 24 v 2P serveru	
Optimální poměr GB/\$ = 64GB DIMMy		



NVMe × sATA šířka pásma a latence

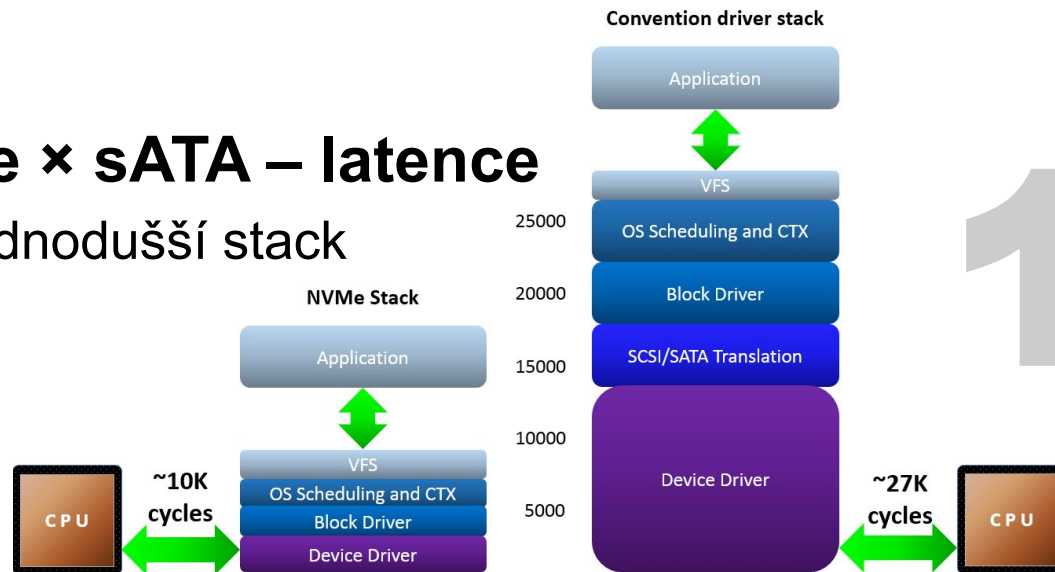
» NVMe × sATA – šířka pásma

- » NVMe3 = PCI-Expres ×4 g3 = 4GB/s
- » NVMe4 = PCI-Expres ×4 g4 = 8GB/s
- » sATA 6Gbs = 768MB/s

10x

» NVMe × sATA – latence

- » Jednodušší stack



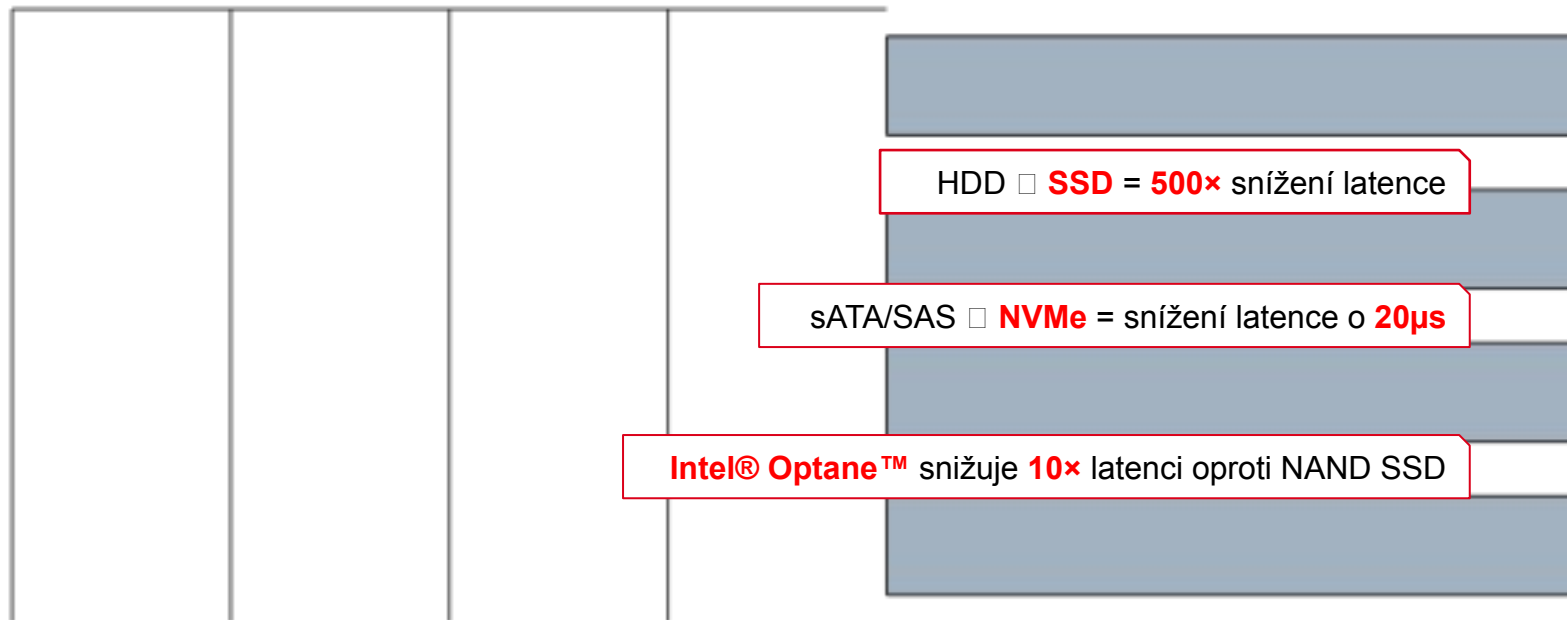
1/6



Latence

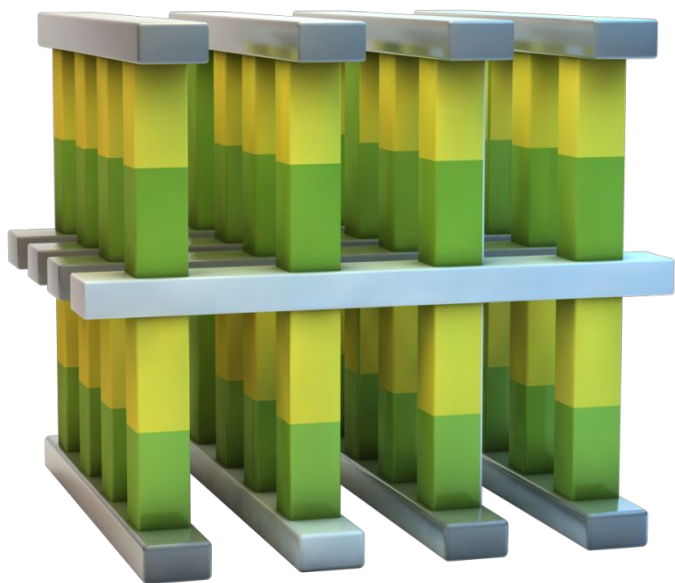
Latence (μs)

1 10 100 1,000 10,000



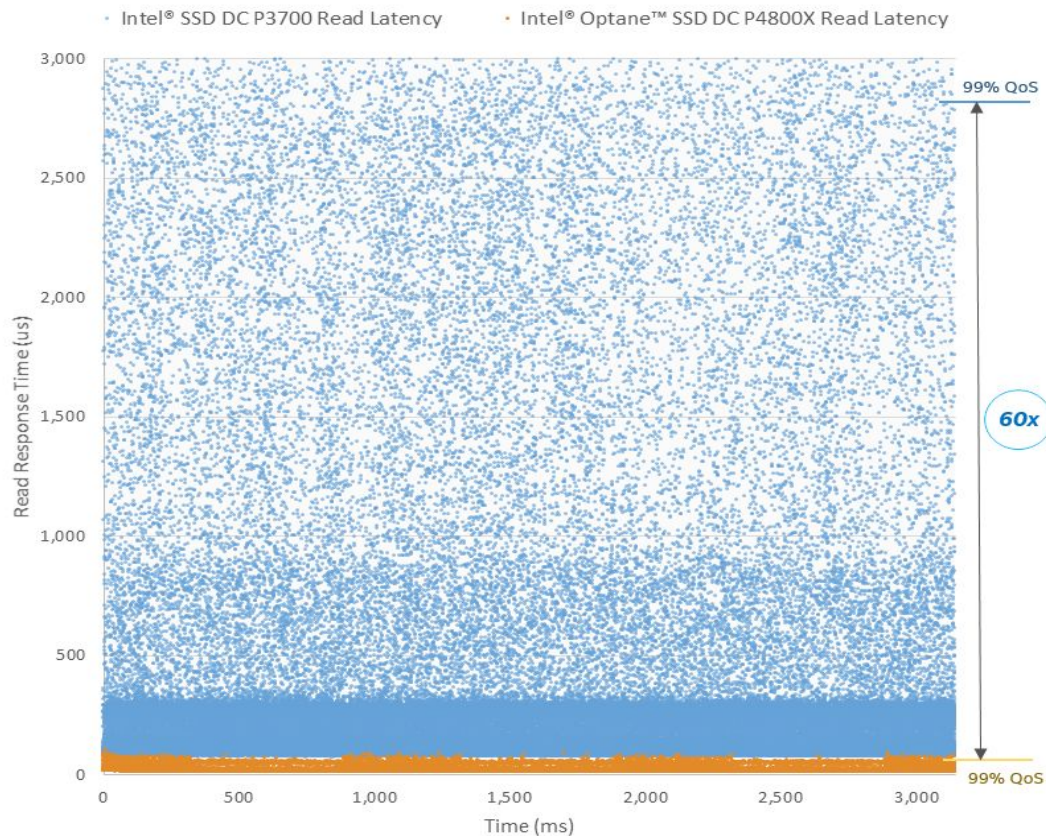
■ Software ■ Řadič ■ Disk

Intel® Optane™ Technologie

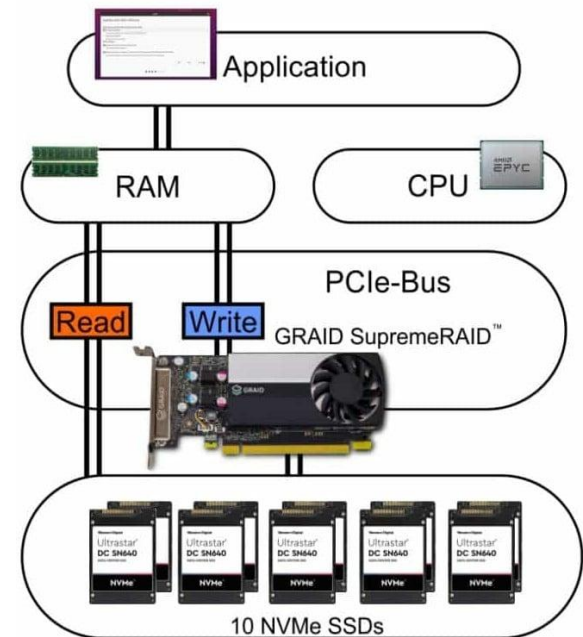


6.000 Kč/100GB

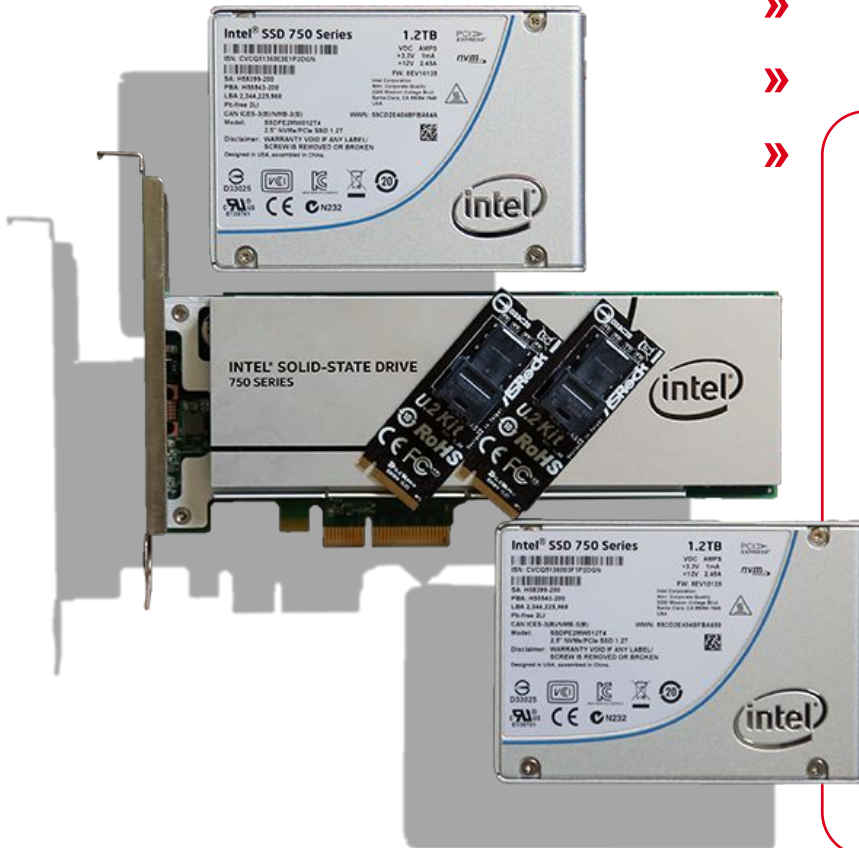
Read QoS in Mixed Workload



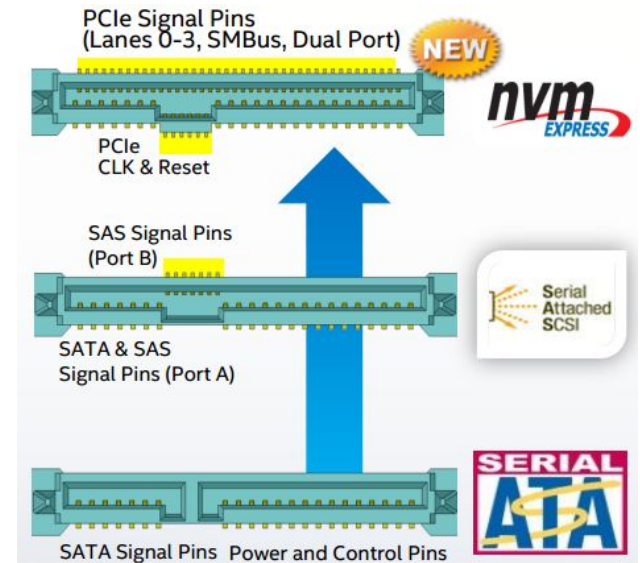
- » **HW RAID**
 - » Bootleneck na PCIe sběrnici i na ROC
 - » VMw kompatibilní
 - » Reálně max.4 disky v R10
- » **Intel VROC (Virtual RAID on CPU)**
 - » Není podporováno v VMware
 - » Bez bootlenecku na PCIe sběrnici
 - » Pouze Intel a ne všechny desky
- » **SW RAID**
 - » Linux, Windows
 - » Dovedávána jediný způsob jak „nezabít“ NVMe výkon
- » **GRAID – RAID na GPU**
 - » 5-10× výkon proti SW RAID (MIOPS)
 - » „Citlivé“ na OS – nejlépe v Linuxu, bude VMw
 - » Až 32 SSD, podpora NVMe-o-F



NVMe SSD mechanické provedení

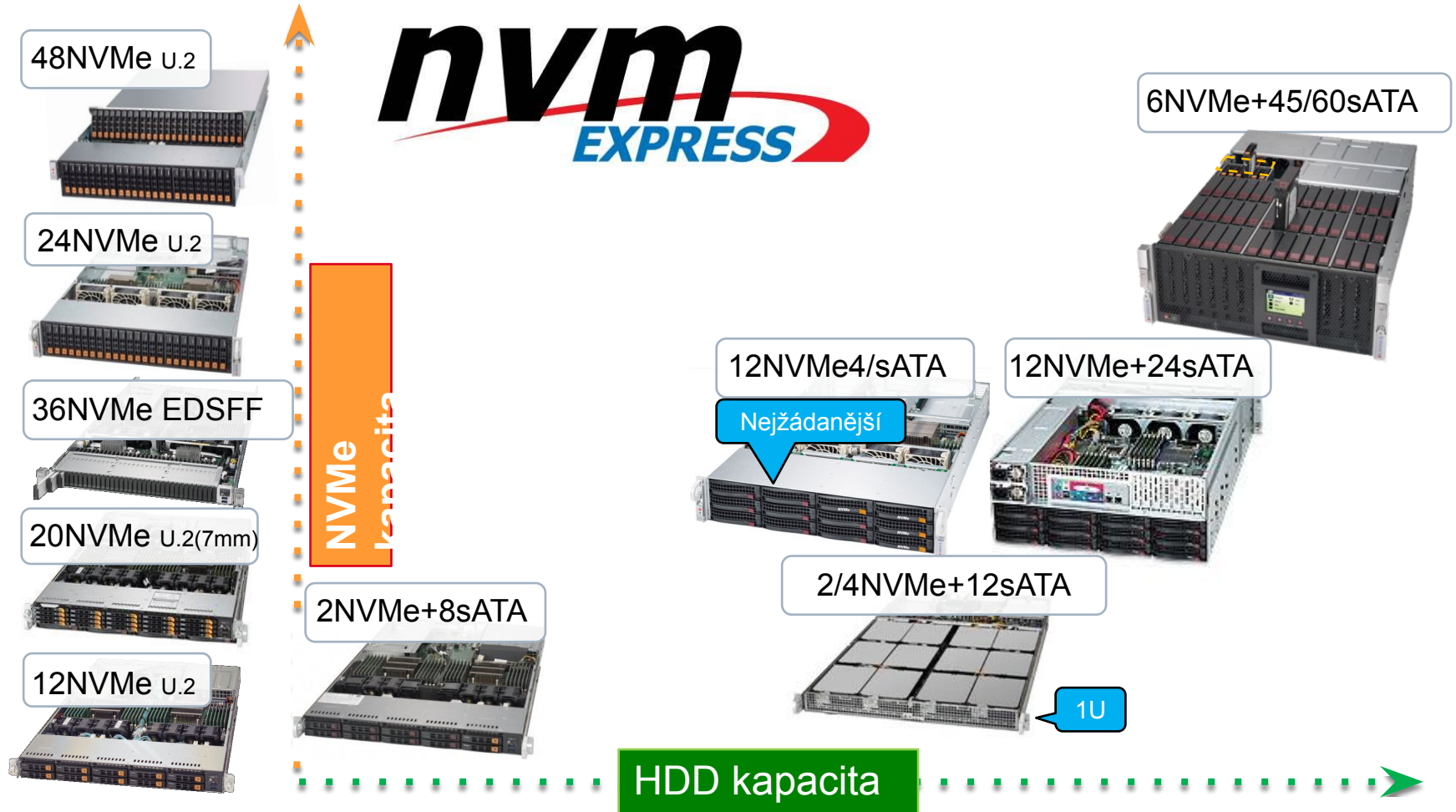


- » PCI-Express x4 karta (AIC/AOC)
- » M.2: 2260 / 2280 / 22110 (22×60mm)
- » U.2/U.3 (7mm a 15mm)
 - = SFF-8639
 - = Hybrid port
 - = 2,5" SFF
 - = hotswap



	Hot Swap	Dual path	Low latency	Počet SSD jednotek	HW RAID	Poznámka
sATA 6Gbs (2,5")	Ano	Ne	Ne	24/2U	Ano	
SAS 12Gbs (2,5")	Ano	Ano, vždy	Ne	24/2U	Ano	
NVMe3/4 M.2	Ne	Ne	Ano	1-2 na základní desce Přídavné karty	Omezeně, pomocí Intel VROC GRAID	
NVMe3/4 U.2/U.3 (2,5") 7/15 mm	Ano	Ano	Ano	24/2U (15mm) 12/1U (15mm) 32/1U JBOF	GRAID TriMODE MegaRAID	<i>Dual path pouze u některých řad SSD Dvě tloušťky: 7mm a 15mm Velmi omezené TDP</i>
NVMe3/4 AOC	Ne	Ne	Ano	10/2U	GRAID	
NVMe EDSFF E1.S	Ano	Ano	Ano	32/1U (9,5mm) 24/1U (15mm)	GRAID	<i>Zřejmě bude převažovat. „M.2 s hotswapem“ Tři tloušťky: 9,5mm (20), 15mm (+20W), 25mm(+25W)</i>
NVMe EDSFF E1.L	Ano	Ano	Ano	32/1U	GRAID	<i>Maximální kapacita, typicky s QLC Dvě tloušťky: 9,5mm (25W) a 18mm (40W) Ale 32cm hloubka</i>
NVMe EDSFF E3.S/L 1T/2T	Ano	Ano	Ano	20/1U	GRAID	<i>70W, do budoucna i GPU a OCP LAN, FPGA, CXL ×4/×8/×16 Zpětně kompatibilní s 2,5" U.2/U.3 .S lze do .L slotu, 1T do 2T slotu, zřejmě i E1 do E3</i>

Hybridní a all-NVMe storage servery



4U toploaded JBODy a servery

» Supermicro

- » Ucelená řada (45/60/90) [+]
- » Servery i JBODy [+]
- » NVMe [+]
- » Proprietární základní deska [-]
- » L9

» AIC

- » 60×3,5"
- » 4NVMe + 6 sATA SSD
- » ATX/eATX deska [+]

1,3PB/4U

45×3,5" Server



60×3,5" Server/JBOD



» WDC

- » Pouze JBOD
- » Inband management
- » JBOD téměř „zdarma“ [+]

102×3,5" JBOD



Dual port NVMe3 SSD Cluster-in-Box

48NVMe



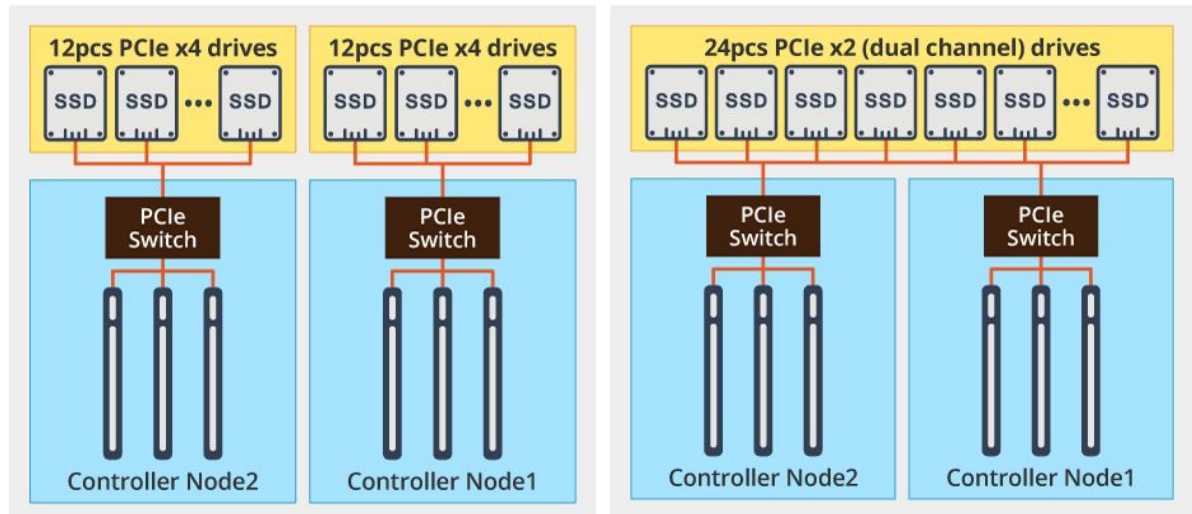
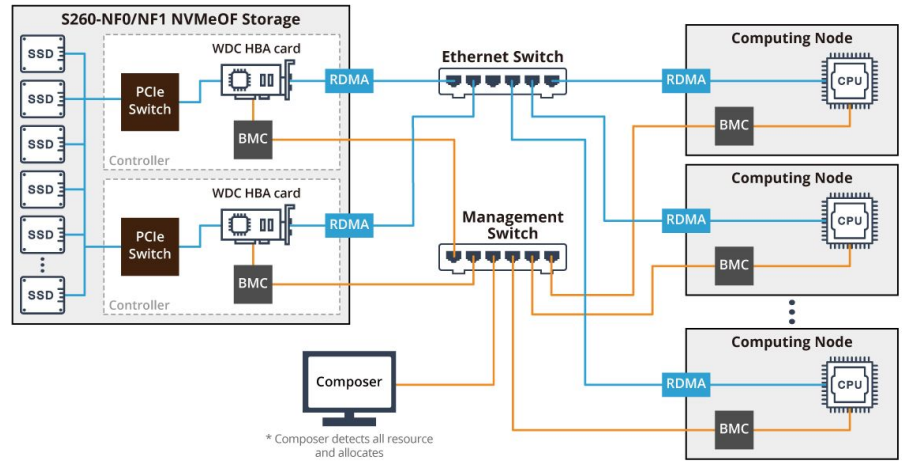
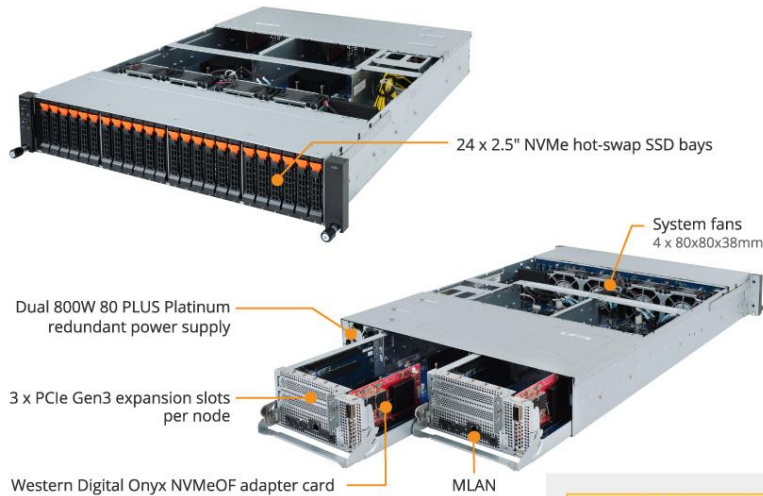
24NVMe



- » NVMe Dual Port SSD
 - » Samsung 1733, Kioxia, WDC SN840
 - » Intel® Optane™ Dual port
- » CiB platforma:
 - » Dual controller
 - » 2U – 24× dual port NVMe 2,5“
 - » 2U – 48× dual port NVMe 2,5“
- » SDS software
 - » Open-e Jovian
 - » MS Windows Server

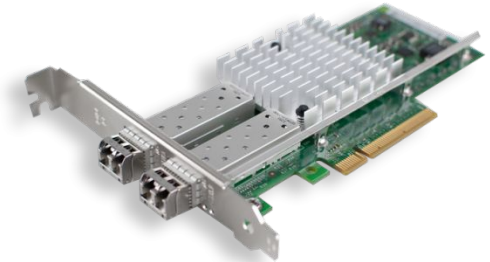
nvm
EXPRESS

NVMe over Fabric (JBOF, NVMe-o-F)



Ethernet: 10G/25G/40G/50G/100G

- » **10GbE-T**
- » **10GbE SFP+**
 - » Optika a GBIC
 - » Ale také DAC kabely
(Direct-Attached metalické kabely bez GBICu, 5m max)
- » **40GbE QSFP+** (původně Infiniband QDR)
 - » QSFP+ = quad SFP+
 - » 40GbE = 4× 10GbE SFP+
- » **25GbE SFP28**
 - » SFP28 = „vyladěné“ 10GbE SFP+
 - » Zpětně kompatibilní s 10GbE SFP+
 - » Obdobné ceny jako 10G
- » **100GbE QSFP28**
 - »
 - » 4×25GbE SFP28



Scale up



Scale out



Proprietární

QSAN

- » Hybridní i AFA
- » Podpora WDC 60/102 JBODů

.. ale stejně je tam Linux -)



Software Defined Storage (SDS)

Open-e Jovian

- » ZFS based (snapshots)
- » NVMe cache
- » Metro i shared-JBOD klastr
- » Disaster Recovery
- » Také jako VSA
- » VAAI

CEPH

- » Jako private cloud

- » Oddělení HW a SW
- » Komoditní x86 HW
- » Vlastnosti do SW
- » Neomezená škálovatelnost



Co jsme nestihli

- » GPU
 - » MicroCloud, Twin a Blade servery
 - » Edge a IoT servery
 - » Mechanické disky
 - » SANky
 - » Infiniband
 - » CXL
 - » DPU
 - » ZFS
 - »
-

