

## TECHNICKÁ DOKUMENTACE - SPECIFIKACE POŽADOVANÉHO PLNĚNÍ

### A. Požadavky zadavatele na dodávku úložiště jako celku

1. Předmětem veřejné zakázky je kompletní řešení, sestávající se z dodávky, instalace a zprovoznění uzlů úložiště. Součástí předmětu veřejné zakázky je poskytnutí rozšířené záruky včetně technické podpory (dále jen „rozšířená záruka“) ve formě reakce Next Business Day On-site v lokalitě BioPharmaHub v datovém sále 1S35, v budově G62 BioPharma Hub na ulici Studentská, Brno.
  - a) Instalací a zprovozněním se rozumí instalace hardware do stávajících rack skříní, zapojení všech síťových rozhraní, zapojení do elektrické sítě, spuštění hardware a ověření bezchybného chodu všech komponent.
  - b) Všechny komponenty dodávky se musí vejít do standardního racku a celkové velikosti 42U.
2. Úložiště je tvořeno alespoň čtyřmi úložnými uzly s celkovou kapacitou alespoň 2240 TB, vše dle níže uvedené specifikace.

### B. Požadavky zadavatele na úložné uzly

1. Provedení do standardního 19" racku, velikost uzlu musí být maximálně 4 U.

#### Uzel provedení case 4U, do 19" racku, včetně ližin

2. V případě sdílení některých komponent mezi více uzly clusteru musí být všechny sdílené aktivní komponenty redundantní.

#### 4 samostatné uzly/výpočetní jednotky

3. Každý uzel (výpočetní jednotka se samostatnou pamětí, chipsetem, procesory, diskem, atd.) musí mít jeden procesor se sdílenou pamětí v architektuře x86\_64.

#### Každý uzel/výpočetní jednotka osazena 1×CPU, AMD EPYC

4. Výkon celého uzlu měřený nástrojem SPECfp2017 ve variantě "FP rate, baseline" musí být alespoň 500 bodů. Dále musí platit, že SPECfp2017 na jedno jádro ve variantě "FP rate, baseline" (tj. spec celého uzlu děleno počtem fyzických jader v celém uzlu bez hyperthreadingu) musí být alespoň 10. (Účastník uvede hodnotu SPEC výkonu pro jeden výpočetní uzel, jejíž splnění **garantuje** a ověření této hodnoty **doloží** protokolem s naměřenou stejnou nebo vyšší hodnotou v souladu s postupem uvedeným v bodu D.1.).

#### Výkon celého uzlu = 535 bodů, CPU má 48jader, 535/48 = 11,1 na jedno jádro

5. Nabízený procesor i další komponenty (motherboard, BIOS) musí podporovat virtualizaci, včetně virtualizace I/O (v terminologii firmy Intel VT-d, v terminologii firmy AMD AMD-Vi).

#### AMD EPYC

požadavky (včetně I/O virtualizace) splněny díky AMD IOMMU na desce a Intel E810 NIC s podporou SR-IOV; Intel I350-T2v2 plně splňuje požadavky na virtualizaci I/O včetně SR-IOV a VMware integrace

6. Operační paměť alespoň 768 GB ECC. Rychlost paměti nesmí být horší než rychlost paměti použité ve SPEC benchmarku v bodu B.4). Na všech použitých paměťových kanálech musí být stejný počet DIMMů. Všechny použité paměťové kanály musí být osazeny stejně.

**12× 64GB 4800MHz DDR5 ECC Registered = 768 GB ECC**

7. Každý server musí být osazen alespoň dvěma SSD disky s celkovou kapacitou alespoň 480 GB. Rychlost lineárního čtení/zápisu každého SSD disku musí být alespoň 500/200 MB/s, každý SSD disk musí nabízet alespoň 50000/10000 IOPS pro náhodné čtení/zápis, každý disk musí mít TBW (TeraBytes Written) alespoň 1 PB. Disky budou použity jako mirror a musí být vyměnitelné za chodu.

**2× ██████████ 960GB SATA3 2,5" 560/530MBs 94/65kIOPS, 3D TLC, 1DWPD, TBW=1,75 PB**

8. Každý uzel clusteru musí mít přístup k lokálnímu SSD prostoru pro úložná data, velikost prostoru musí být alespoň 60 TB a musí být realizován alespoň 4 NVMe SSD disky (datové NVMe). Pro celý prostor musí dodavatel garantovat životnost alespoň 100000 TBW (TeraBytes Written). Všechny datové NVMe disky musí být totožného typu a kapacity.

**4× ██████████ 15,36TB NVMe5 U.2 (2,5"/15mm), PCI-E4(g5), 2800/410kIOPS, 13,5/9,5GB/s, 1DWPD, TBW=28 PB; Celý prostor 4×28= 112 PBW , velikost diskového prostoru = 61,44 TB / uzel**

9. Každý uzel clusteru musí mít přístup k lokálnímu prostoru pro úložná data, velikost prostoru musí být alespoň 500 TB a může být realizován maximálně 24 rotačními disky (datové rotační disky). Všechny datové rotační disky musí být totožného typu a kapacity.

**21× 24TB ██████████  
██████████ 3,5" Celý prostor 21×24= 504 TB / uzel**

10. Pro datové NVMe disky (bod B.8) je požadována rychlost čtení: počet\_disků krát 1500 MiB/s a rychlost zápisu počet\_disků: krát 750 MiB/s, kde počet\_disků je počet datových NVMe disků v serveru.

**Požadavek – čtení: 4× 1500 = 6000 Mib/s      Naměřeno - čtení: 18 909,49 MiB/s**  
**Požadavek – zápis: 4× 750 = 3000 MiB/s      Naměřeno - zápis: 10 349,42 MiB/s**

11. Pro datové rotační disky (bod B.9) je požadována rychlost čtení: počet\_disků krát 80 MiB/s a rychlost zápisu: počet\_disků krát 60 MiB/s, kde počet\_disků je počet datových rotačních disků v serveru.

**Požadavek – čtení: 21× 80 = 1 680 Mib/s      Naměřeno - čtení: 3 449,79 MiB/s**  
**Požadavek – zápis: 21× 60 = 1 260 MiB/s      Naměřeno - zápis: 2 665,76 MiB/s**

12. Všechny disky musí být typu enterprise, tedy určené pro použití v serverech nebo diskových polích a navržené pro nepřetržitý provoz.

**Všechny SSD/HDD jsou určeny pro provoz 24/7 servery/diskové pole**

13. Všechny osazené NVMe disky musí být vybaveny technologií PLP (Power Loss Protection).

**Všechny osazené NVMe SSD mají technologii PLP**

14. Všechny disky musí být samostatně vyměnitelné za chodu. Výměna jednoho disku nesmí vést k dočasné nedostupnosti jiných disků nebo serveru samotného.

**Všechny disky jsou v HotSwap pozicích, vyměnitelné za chodu**

15. Všechny disky musí být prezentovány operačnímu systému jako jednotlivá bloková zařízení, přitom musí podporovat vyčtení provozních stavů a statistik pomocí standardu SMART.

**Všechny disky jsou prezentovány OS jako bloková zařízení a podporují standard SMART**

16. U všech disků musí být zajištěno jejich dostatečné chlazení umožňující jejich trvalý provoz pod plnou

zátěží. Disky (zejm. NVMe) mohou být osazeny chladiči, pokud to dodavatel považuje za nezbytné pro dosažení stabilního provozu.

**Všechny disky jsou dostatečně chlazeny.**

17. Všechny disky všech typů musí být zapojeny přímo do serveru, nepřipouští se jejich připojení přes expanzi. Expanzí rozumíme samostatný fyzický box pro umístění disků, typicky se samostatným napájením a chlazením.

**Všechny disky zapojeny přímo do serveru, bez přídavné expanze.**

18. Osazení HDD s použitou technologií SMR (Shingled magnetic recording) není přípustné.

**Osazené HDD 24TB nemají technologii SMR, používají CMR**

19. Každý uzel clusteru musí mít dvě síťová rozhraní 100Gb Ethernet a alespoň jedno rozhraní 1Gb Ethernet. Rozhraní 1Gb Ethernet musí podporovat bootování přes PXE, 1Gb rozhraní musí být typu BaseT s RJ45 konektorem.

**Každý uzel/výpočetní jednotka obsahuje:**

**1× Ethernet [redacted] [redacted] Dual port 100GbE (QSFP28) PCI-E16(g4);  
1GbE podporuje bootování přes PXE (konektor RJ45, [redacted])**

20. 100Gb Ethernet rozhraní musí podporovat UDP TNL offload v Linuxovém ovladači.

**LAN [redacted] podporuje UDP tunnel offload (UDP TNL offload), včetně VXLAN/GENEVE/GRE**

21. Každý uzel clusteru umožňuje centralizovaný přístup ke konzoli (klávesnice + monitor), a to jak lokálně (KVM switch), tak po síti (síťový KVM nebo BMC).

**Každý uzel/výpočetní jednotka umožňuje lokální KVM: přes VGA/DP + USB, vzdálený KVM: přes BMC (AST2600)**

22. Každý uzel clusteru podporuje bootování z externího zařízení, a to jak lokálně (boot z USB – CD-ROM, flash disk, harddisk), tak po síti (boot z virtuálního média implementovaného pomocí síťového KVM nebo BMC). Je-li na provoz síťového KVM nutná licence, musí být tato licence součástí dodávky.

**Každý uzel/výpočetní zařízení podporuje bootování z externího zařízení (lokálně i po síti) bez přídavné licence**

23. Základní deska musí umožňovat změnu pořadí bootovacích zařízení.

**Základní deska umožňuje změnu bootovacích zařízení**

24. Základní deska musí obsahovat management controller (BMC) kompatibilní se specifikací Redfish nebo IPMI 2.0 nebo vyšší. BMC musí umět monitorovat minimálně funkčnost ventilátorů a zdroje, teplotu CPU a základní desky; dále musí BMC poskytovat základní vzdálený power management (vypnout, zapnout, reset). Požadujeme možnost změny bootovacího zařízení vzdáleně pomocí BMC nebo KVM. Funkcionalita BMC musí být přístupná po dedikovaném portu (100BaseT nebo 1000BaseT). Ve spojení s bodem B.19) musí mít každý uzel alespoň **4 Ethernet** rozhraní.

**Základní deska obsahuje ASPEED AST2600 BMC (podpora IPMI 2.0, Redfish API), Základní deska podporuje: Monitoring CPU / MB teplot, monitoring ventilátorů, power ON/OFF/Reset, dedikovaný 1GbE management port, možnost změny bootovacích zařízení (BMC/KVM)**

25. Každý uzel musí mít duální napájení. Zdroje musí být vyměnitelné za chodu.

**Každý uzel/výpočetní zařízení disponuje 2× HotSwap rPS 1k3W (80+Titanium)**

## C. Další společné požadavky

1. Délka záruční doby musí být minimálně 36 měsíců s reakční dobou nejpozději následující pracovní den (NBD) v místě instalace (On-site).

### **Typ SLA - Oprava 3r NBD on-site pro 1xCPU, Příjem hlášení 8-22 hod. PO-PA**

2. Veškeré zařízení by mělo být možno koupit bez jakéhokoliv software. Pokud je programové vybavení nutnou součástí nabídky (například SW pro vzdálenou správu), musí být jasně specifikovány důvody a cena za takový SW musí být zahrnuta do ceny dodávky (na dobu neurčitou; pokud autor / výrobce / dodavatel SW neposkytuje licenci na dobu neurčitou, je účastník povinen tuto skutečnost zadavateli prokázat a zajistit licenci nejméně po dobu poskytování záruky).

### **Každý uzel/výpočetní jednotka je bez přídavného SW**

3. Všechny souborové servery musí mít nastaveno IPMI na získání DHCP adresy a musí mít nastaveno bootování ze sítě přes PXE z 100Gb interfacu. Úložné servery musí mít nastaveno IPMI na získání DHCP adresy a musí mít nastaveno bootování ze sítě přes PXE z 1Gb interfacu.

### **Všechny uzly/výpočetní jednotky budou nastaveny dle výše uvedeného požadavku. IPMI přes DHCP a nezávislé PXE bootování jak přes 100GbE [REDACTED], tak přes 1GbE NIC**

4. Všechny komponenty, které jsou touto technickou specifikací požadovány, musí být použitelné v prostředí operačního systému Linux (zejména, ale nikoliv výhradně 64bit Rocky Linux), tj. musí být podporovány distribučním nebo originálním jádrem nebo s využitím externích ovladačů dostupných ve zdrojovém kódu.

### **Všechny komponenty jsou použitelné v prostředí operačního systému Linux**

5. Součástí nabídky musí být celková maximální spotřeba sestavy (maximální spotřeba odpovídá spotřebě při plném zatížení všech komponent, tedy uzlů).

### **Maximální teoretická spotřeba uzlu/výpočetní jednotky při 100% zátěži je: 983 W**

6. Zařízení budou uchycena ve standardním 19" rámu, montážní sady je potřeba dodat spolu se zařízením v potřebném počtu, doporučené provedení šroubů M6 philips (PH).

### **Všechny uzly/výpočetní jednotky budou dodány včetně potřebného spojovacího/installačního materiálu.**

7. Napájení v rozvaděčích je jednofázové 230V~. PDU lišty zadavatele obsahují kulaté zásuvky CEE 7/5.

### **Napájecí kabeláž bude korespondovat s připravenými zásuvkami CEE 7/5.**

8. Při montáži zařízení musí být proudění vzduchu v rozvaděči orientováno vodorovným podélným směrem, od předních dveří rozvaděče do teplé uličky.

### **Všechny uzly/výpočetní jednotky jsou v provedení chlazení/proudění vzduchu od předních dveří rozvaděče do teplé uličky.**

9. Součástí dodávky musí být veškerá propojovací kabeláž pro připojení uzlů do síťových switchů zadavatele pro obě 1 Gb a všechna 100 Gb rozhraní, všechny o délce až 6 m, po dohodě se zadavatelem budou dodány kabely na míru.

### **Součástí dodávky je i veškerá propojovací kabeláž uzel – switch. Konkrétní délky a typy kabeláže budou upřesněny po dohodě se zadavatelem.**

10. Pro účely posouzení nabídky je účastník povinen uvést SPEC výkon všech serverů.

### **SPEC výkon uzlu je doložen dle specifikace bodu B/4.**

11. Přesné umístění serverů a diskových polí v racku bude zadavatelem upřesněno během dodání.

### **Dle požadavku**

12. Pro účely posouzení splnění technických parametrů je účastník povinen popsat technické parametry nabízené sestavy. Popis lze realizovat formou komentářů k jednotlivým bodům

Technické dokumentace, zpravidla však nepostačují odpovědi typu ANO/NE, je nutné konkrétně popsat konfiguraci navrženého řešení. Nabídky bez technického popisu nejsou přípustné.

Navrženou konfiguraci ve formě uvedení výrobce a jednoznačného typového označení každé položky předmětu veřejné zakázky včetně řádného ocenění všech položek rozpočtu přiloží účastník k formuláři nabídky jako samostatnou přílohu Položkový rozpočet.

#### D. Měření výkonu úložiště

Součástí nabídky budou výkonnostní testy diskového subsystému dle následujícího popisu:

1. Součástí nabídky budou výkonnostní testy výpočetních uzlů, kterými dodavatel demonstruje dosažení požadovaných parametrů procesorů dle bodu B.4. (Spec2017). Výkon lze prokázat předložením oficiálního výsledku z webu [www.spec.org](http://www.spec.org) dosaženého na ekvivalentním stroji (typ a počet procesorů totožný s dodanými procesory; počet a frekvence paměťových DIMMů a organizace paměti totožné s dodanou pamětí, velikost paměti se může lišit) nebo výsledkem spuštění benchmarku na uzlu nakonfigurovaném dle uvedené technické specifikace. Zadavatel je schopen poskytnout kompilátor Intel verze 17.1 a CPU2017 SPEC. Případné licence pro komerční distribuce Linuxu si dodavatel musí zajistit sám.
2. Účastníkem dodané výsledky výkonnostního měření musí být provedeny na účastníkem navržené konfiguraci vyhovující tomuto zadání (není tedy možné dodat výkonnostní charakteristiky pouze pro RAID 0 nebo pro jinou RAID konfiguraci nesplňující uvedené požadavky).
3. Měření propustnosti datových NVMe a datových HDD disků (ve smyslu bodů B.10 a B.11) bude provedeno pomocí nástroje iozone nad souborovým systémem vytvořeným nad všemi osazenými disky příslušného typu spojenými do softwarového RAID 0. Měření bude provedeno samostatně pro datové HDD a samostatně pro NVMe disky. Měření datových rotačních disků (B.11) bude provedeno příkazem `iozone -Mce -t64 -s50G -r256 -i0 -i1 -F soubor1 ... soubor64` kde soubory soubor1 až soubor64 leží na souborovém systému vytvořeném nad testovanými disky. Měření datových NVMe disků (B.10) bude provedeno příkazem `iozone -Mce -l -t64 -s50G -r256 -i0 -i1 -F soubor1 ... soubor64` kde soubory soubor1 až soubor64 leží na souborovém systému vytvořeném nad testovanými disky. Jako výsledek testu pro zápis, respektive pro čtení je brána průměrná hodnota tří testů udaná výstupem programu iozone jako „Children see throughput for X initial writers“, respektive, „Children see throughput for X readers“. Program iozone používá jednotky v dvojkových násobcích (KiB, MiB) apod.
4. Testy dodané pro účely posouzení nemusejí být pořízeny na stejném hardware, který bude dodán, případně v dodávané konfiguraci. Dodavatel nicméně odpovídá za to, že případně skutečně naměřené hodnoty během možného ověření na skutečně dodané konfiguraci nebudou horší, než jaké přikládá k nabídce. Nevadí, budou-li skutečně naměřené hodnoty lepší.
5. Pro rychlost úložiště je pro zadavatele podstatná rychlost naměřená programem iozone (<http://www.iozone.org>). **Výstup programu iozone je nutné přiložit k nabídce.**