**„Výpočetní cluster VC III“ a „HPCQC – AI výpočetní cluster“**

**Základní požadavky zadavatele na předmět veřejné zakázky**

**Obsah**

[**Základní požadavky zadavatele na předmět veřejné zakázky** 1](#_Toc198733118)

[1 Úvod 1](#_Toc198733119)

[2 Parametry zakázky 1](#_Toc198733120)

[3 Záměr 2](#_Toc198733121)

[4 Výpočetní cluster VC III 2](#_Toc198733122)

[4.1 Komponenty 3](#_Toc198733123)

[4.2 Řešení superpočítače 4](#_Toc198733124)

[4.3 Požadavky 5](#_Toc198733125)

[5 HPCQC – AI výpočetní cluster 8](#_Toc198733126)

[5.1 Komponenty 8](#_Toc198733127)

[5.2 Řešení superpočítače 10](#_Toc198733128)

[5.3 Požadavky 11](#_Toc198733129)

[6 Rozsah zakázky 14](#_Toc198733130)

[7 Závazný obsah Návrhu 14](#_Toc198733131)

[8 Požadavky na záruku a servisní služby 16](#_Toc198733132)

[9 Infrastruktura zadavatele 17](#_Toc198733133)

[9.1 Datové centrum 17](#_Toc198733134)

[9.2 WAN síť 18](#_Toc198733135)

[9.3 Storage síť 18](#_Toc198733136)

## Úvod

Obsahem tohoto dokumentu jsou informace pro dvě připravované veřejné zakázky „**Výpočetní cluster VC III“ a** „**HPCQC – AI výpočetní cluster“,** předběžné požadavky na rozsah, požadované vlastnosti, parametry a funkcionality.

***Všechny požadavky a parametry uvedené v tomto dokumentu jsou předběžné.***

## Parametry zakázky

Předpokládaný termín realizace zakázky je rok 2026.

Předpokládaná hodnota navazujícího zadávacího řízení „**Výpočetní cluster VC III“** je 195 500 000 Kč bez DPH.

Předpokládaná hodnota navazujícího zadávacího řízení „**HPCQC – AI výpočetní cluster“** je 177 000 000 Kč bez DPH.

Zakázka „**Výpočetní cluster VC III“** je financována z prostředků operačního programu Jan Amos Komenský (OP JAK), projekt Modernizace e-INFRA CZ II (CZ.02.01.01/00/23\_016/0008329) a vlastních prostředků.

Zakázka „**HPCQC – AI výpočetní cluster“** je financována z prostředků operačního programu Jan Amos Komenský (OP JAK), projekt Centrum pro umělou inteligenci a kvantové výpočty v systémovém výzkumu mozku (registrační číslo: CZ.02.01.01/00/23\_029/0008437), akronym CLARA, a z vlastních zdrojů

Zadavatel předpokládá, že v rámci hodnocení stanoví pevnou cenu a bude hodnotit pouze kvalitu nabízeného plnění. V rámci kritérií kvality bude hodnocena zejména technická úroveň řešení s důrazem na hodnotu výpočetního výkonu.

## Záměr

Záměrem zadavatele je modernizace a posílení kapacit ICT infrastruktury Národního superpočítačového centra IT4Innovations.

V rámci modernizace a posílení kapacit ICT infrastruktury centra budou pořízeny, instalovány a zpřístupněny uživatelům centra superpočítače pro řešení nejnáročnějších vědeckých a inženýrských výpočetních úloh. Tyto nové superpočítače postupně nahradí stávající superpočítač Karolína a dále poslouží vědeckým cílům projektu CLARA.

V roce 2026 budou realizovány oba superpočítače, kde superpočítač VC III nahradí v tom čase již morálně zastaralý stávající superpočítač Karolína. Superpočítač VC III bude výkonem převyšovat stávající superpočítač Karolína a bude používat v době realizace nejmodernější technologie. Hlavním cílem je poskytnout moderní technologie uživatelům centra pro široké spektrum úloh, od tradičních HPC simulací až po nejmodernější úlohy z oblasti umělé inteligence (AI). **Realizace superpočítače VC III je předmětem této veřejné zakázky.**

Ve stejném roce dojde taktéž k instalaci nového superpočítače HPCQC – AI výpočetní cluster, který bude primárně sloužit potřebám projektu CLARA, který je zaměřen na výzkum neurodegenerativních onemocnění mozku. **Realizace HPCQC – AI výpočetní clusteru je předmětem této veřejné zakázky.**

Ačkoliv z technického hlediska mohou být výsledná řešení do jisté míry podobná, bude každý výpočetní systém pořizován na základě samostatného zadávacího řízení, jelikož se jedná o dva samostatné funkční celky.

Oba superpočítače budou využívány taktéž k simulacím kvantových počítačů jakožto doplnění a předstupeň kvantového počítače VLQ, konsorcia LUMI-Q, který je instalován a provozován v datovém centru IT4Innovations (tedy stejném jako předchozí i nově pořizované superpočítače).

Nově realizované superpočítače VC III a HPCQC – AI výpočetní cluster budou integrovány do infrastruktury superpočítačového centra IT4Innovations, budou instalovány na datovém sále zadavatele, budou integrovány do síťové infrastruktury zadavatele a budou zpřístupňovat centralizovaná úložiště zadavatele.

Pozn. V textu jsou uváděny rychlosti a kapacity za použití dekadických jednotek (násobky deseti), pokud není explicitně uvedena binární jednotka (např. KiB, MiB).

## Výpočetní cluster VC III

Pro superpočítač VC III jsou v současné době na straně zadavatele zajištěny prostředky financování z národních zdrojů (OP JAK) ve výši 195 500 000 Kč bez DPH.

### Komponenty

Superpočítač je komplexní systém sestávající z výpočetních, úložných, síťových a dalších systémů   
a softwarového řešení určený pro provádění náročných výpočtů (HPC).



Obrázek 1 Orientační schéma řešení superpočítače VC III

**Výpočetní cluster** je tvořen **Výpočetními servery** propojenými **Výpočetní sítí** – vysokorychlostní sítí s nízkou latencí. Výpočetní cluster je určen pro provádění výpočetních úloh uživatelů. Výpočetní servery poskytují výpočetní výkon, Výpočetní síť umožňuje rychlou komunikaci mezi servery a efektivní provádění paralelních úloh. Pro servery clusteru je dále používáno též označení uzel.

**Přístupové servery** *–* servery sloužící pro přístup uživatelů, pro přípravu úloh a dat, kompilaci a ladění kódů, pro zpracování výsledků a pro přenos dat.

**Úložiště** slouží k ukládání a sdílení dat. Úložiště jsou realizována jako komplexní řešení úložných zařízení, I/O serverů (např. souborových serverů), datových sítí a potřebného softwarového vybavení na straně serverů i klientů. Úložiště poskytující souborové služby jsou označována jako **Souborová úložiště**.

IT4Innovations strukturuje uživatelská úložiště výpočetních clusterů takto:

* úložiště HOME – souborové úložiště pro uživatelská nastavení operačního systému a aplikací na superpočítači
* úložiště SCRATCH – souborové úložiště pro krátkodobá data uživatelů, intenzivně používaná výpočetními uzly superpočítače
* úložiště PROJECT – souborové úložiště pro střednědobá a dlouhodobá projektová data uživatelů, sdílená mezi superpočítači IT4Innovations (existující úložiště zadavatele, není součástí zakázky)

**Úložiště** infrastruktury je určeno pro ukládání a sdílení dat infrastruktury superpočítače, tj. pro uložení systémových obrazů (image) serverů, logů, dat infrastrukturních služeb, uživatelského aplikačního vybavení, plánovače atd.

**Infrastrukturní a management servery** jsou určené pro správu superpočítače, zdrojů, úloh, licencí a poskytování infrastrukturních služeb (např. DHCP, DNS, LDAP, licenční servery, plánovače, monitoring, logování atd.), zdrojů.

Řešení **Zálohování** zajišťuje ochranu dat před úplnou ztrátou v případě vážného selhání primárního úložiště dat, chyby apod.

Superpočítač obsahuje **Síťovou infrastrukturu,**tj. síťové propojení komponent, systémů tak, aby bylo dosaženo požadované funkcionality, byl zajištěn přístup na jednotlivé služby, byl zajištěn výkon, dostupnost a bezpečnost.

Síťovou infrastrukturu tvoří zejména Výpočetní síť, LAN síť, integrace do WAN sítě zadavatele, integrace do storage sítě zadavatele a případně další sítě (například datové sítě SAN) dle návrhu dodavatele*.*

**Výpočetní síť**propojuje výpočetní servery, přístupové servery, je použita pro zpřístupnění některých úložišť.

**LAN síť** zajišťuje komunikaci mezi zařízeními uvnitř superpočítače, slouží ke správě systému, poskytování infrastrukturních služeb atp.

**Integrace do WAN sítě** zadavatele zajišťuje propojení do WAN sítě zadavatele a zprostředkovaně služby internetu v superpočítači.

**Integrace do storage sítě** zadavatele zajišťuje propojení do storage sítě zadavatele, kde jsou připojena centralizovaná úložiště zadavatele (např. úložiště PROJECT), tak aby byla zajištěna dostupnost úložišť na uzlech superpočítače. Integrace je obvykle realizována síťovými branami realizujícími propojení storage sítě zadavatele (100 Gb/s Ethernet) a výpočetní sítě clusteru.

**Infrastrukturu pro provoz v datovém centru** tvoří. zejména racky potřebné pro umístění zařízení, řešení chlazení zařízení superpočítače, napájení, rozhraní a napojení na infrastrukturu datového centra.

**Softwarové vybavení** je nezbytnou součástí řešení superpočítače. Softwarové vybavení zahrnuje mj. operační systémy všech serverů, firmware a ovladače hardware, knihovny, infrastrukturní software, software pro realizaci výpočetního prostředí superpočítače a výpočetní aplikace.

### Řešení superpočítače

Řešení superpočítače musí umožňovat efektivní provádění mnoha současných výpočetních úloh všech fází životního cyklu úloh (příprava, výpočet, zpracování výsledků) různého typu (paralelní, sériové; dávkové, interaktivní) mnoha uživatelů, bude zajišťovat bezpečné uložení dat uživatelů a rychlý přístup k datům, efektivní správu systému, komponent, zdrojů a uživatelů.

Řešení musí poskytovat transparentní, jednotné, sdílené uživatelské prostředí a jednotný přístup k výpočetním zdrojům.

Redundance řešení musí být v takovém rozsahu, aby při selhání fyzické komponenty nedošlo k významnému omezení poskytování služeb (aby neselhala klíčová služba nutná pro běh řešení) nebo aby nedošlo ke ztrátě důležitých dat. Zabezpečení významných dat bude dále zajištěno zálohováním.

Všechny servery řešení budou umožňovat vzdálenou správu nezávislou na běhu operačního systému.

Řešení musí umožňovat efektivní provoz a nízké provozní náklady.

Superpočítač VC III bude obsahovat tyto komponenty:

* **Výpočetní servery**
* **Výpočetní síť**
* **Přístupové servery**
* **Úložiště HOME**
* **Úložiště SCRATCH**
* **Úložiště infrastruktury**
* **Infrastrukturní a management servery**
* **Zálohování**
* **Síťová infrastruktura**
  + **LAN síť**
  + **Integrace do WAN sítě**
  + **Integraci do storage sítě**
* **Infrastruktura pro provoz v datovém centru**
* **Softwarové vybavení**

### Požadavky

#### Výpočetní servery

* minimálně 110 výpočetních serverů
* architektura x86\_64, nebo ARM
* OS Linux
* jeden nebo dva procesory na výpočetní server
* minimálně 960 GB standardní paměti DDR RAM na výpočetní server
* min. 450 GB/s propustnost do paměti CPU/socket
* min. čtyři GPU akcelerátory
* výkon min. 4000 TFlop/s na server v FP32 (tensorflow/sparsity AI maximální teoretický výkon)
* minimálně 2 NIC pro vysokorychlostní síť s agregovanou šířkou pásma min. 800 Gb/s
* chlazení pomocí DLC, celkově min. 90% tepla do vody o vstupní teplotě min. 38 °C

#### Přístupové servery

* 2 přístupové (login) servery
* CPU, GPU a paměť serveru – shodná konfigurace jako u výpočetních serverů

#### Úložiště SCRATCH

Úložiště SCRATCH slouží pro uložení krátkodobých dat uživatelů, data jsou obvykle intenzivně používaná výpočetními uzly superpočítače.

* Úložiště SCRATCH bude mít čistou kapacitu minimálně 2PB, propustnost minimálně 1000 GB/s a bude těsně (s vysokou propustností a s nízkou latencí) připojeno na všechny uživatelsky dostupné uzly (výpočetní, přístupové).
* Předpokládáme použití technologie SSD disků (technologie NVMe) pro zajištěné vysoké propustnosti a I/O výkonu (5000 kIOPS).
* Jako souborový systém bude použit vysoce výkonný paralelní clusterový souborový systém „enterprise“ kvality, zajišťující kromě vysokého výkonu i vysokou bezpečnost přístupu k datům ve sdíleném prostředí.

#### Výpočetní síť

* Výpočetní síť bude mít rychlost portu minimálně 400Gb/s a latenci menší než 10 mikrosekund.
* Topologie sítě musí poskytovat neblokující ostrovy výpočetních uzlů o velikosti minimálně 32 uzlů.
* Síť musí podporovat hardwarovou akceleraci MPI komunikace.
* Síť musí podporovat IP protokol.
* Síť musí být vhodná a podporovaná pro řešení úložiště SCRATCH.

#### Úložiště HOME

Úložiště HOME slouží pro uložení uživatelských nastavení operačního systému a aplikací v rámci superpočítače, předpokládáme tedy potřebu relativně nízké kapacity, přiměřené propustnosti a výkonu.

* Úložiště HOME bude mít čistou kapacitu minimálně 25TB a bude sdíleno na všechny uživatelsky dostupné uzly (výpočetní, přístupové).
* Úložiště bude postaveno na SSD discích.
* Předpokládáme použití technologie NFS s vysokou dostupností (HA).

#### Úložiště infrastruktury

Úložiště infrastruktury je určeno pro ukládání a sdílení dat infrastruktury superpočítače, tj. pro uložení systémových obrazů (image) serverů, logů, dat infrastrukturních služeb, uživatelského aplikačního vybavení, plánovače atd.

Úložiště bude mít čistou kapacitu minimálně 25TB.

Úložiště bude postaveno na SSD discích.

Konfigurace bude v režimu redundance/vysoké dostupnosti minimálně na úrovni disků, řadičů diskového pole, připojení k serverům a souborových serverů.

Úložiště infrastruktury bude připojeno k serverům (přímo či vhodnou sítí) blokovým způsobem anebo pomocí síťového souborového systému (např. NFS). Předpokládané použití ze strany zadavatele:

* APPS – aplikační software sdílený pomocí NFS na všechny uživatelsky dostupné servery

Pro realizaci úložiště infrastruktury a úložiště HOME lze použít jedno společné diskového pole.

#### Infrastrukturní a management servery

Systém bude obsahovat minimálně dva fyzické servery pro infrastrukturní služby a management výpočetního systému. Klíčové infrastrukturní služby budou realizovány v konfiguraci vysoké dostupnosti (HA).

Několik infrastrukturních serverů (předpokládaný počet cca 5ks) bude určeno pro běh služeb zadavatele.

Preferujeme – v případě použití malého počtu fyzických serverů – oddělení služeb na virtuální servery (použití serverové virtualizace).

#### Zálohování

Zálohování bude poskytovat zabezpečení dat infrastrukturních a management serverů, úložiště infrastruktury a úložiště HOME.

Zálohování bude řešeno bez využití pásek (na diskové úložiště) a bude využívat technologii pro vyšší efektivitu uložení dat (např. kompresi nebo deduplikaci). Zálohování bude využívat vyhrazené, nezávislé úložiště dat.

#### LAN síť

LAN síť bude zajišťovat komunikaci mezi zařízeními uvnitř superpočítače a poskytovat připojení do WAN sítě IT4Innovations.

LAN síť bude obsahovat veřejné a privátní části sítě.

Veřejné části sítě (dále jen veřejné sítě), ve kterých budou použity veřejné IPv4 a IPv6 adresy, budou sloužit pro poskytování služeb dostupných z internetu.

Privátní části sítě (dále jen privátní sítě), ve kterých budou použity privátní IPv4 adresy, budou sloužit pro vnitřní služby a pro management zařízení.

Bude zajištěna konektivita výpočetních serverů do LAN sítě a internetu.

Zařízením v privátních sítích musí být umožněn přístup do internetu přes NAT.

LAN síť bude rozdělena do různých L3 sítí. Pro každou L3 síť bude použita jiná L2 síť (VLAN nebo jiný aktivní prvek). Je nežádoucí používat jednu L2 síť pro více L3 sítí. Rozdělení sítí bude zajišťovat zejména oddělení těchto provozů:

* služby dostupné z internetu
* datová komunikace mezi servery (služby)
* management síťových aktivních prvků
* management diskových polí a storage zařízení
* management serverů (BMC, IPMI, apod.)
* management non-IT infrastruktury (napájení, chlazení apod.)

LAN síť nezbytná pro propojení přístupových serverů do WAN sítě bude řešena redundantními prvky. Aktivní prvky sítě budou vzdáleně řízené a centrálně spravovatelné.

#### Integrace do WAN sítě

Součástí zakázky je připojení do WAN sítě zadavatele.

Připojení do WAN sítě zadavatele musí poskytovat konektivitu min. 2x100 Gb/s, pro připojení jsou vyhrazeny dva porty QSFP28 na prvcích WAN sítě zadavatele.

#### Integrace do storage sítě

Součástí zakázky je připojení do storage sítě zadavatele. Pro realizaci se předpokládá použití minimálně dvou vyhrazených serverů implementujících síťové brány mezi storage sítí zadavatele (na technologii 100Gb/s Ethernet) a výpočetní sítí clusteru

Připojení do storage sítě zadavatele musí poskytovat konektivitu min. 2x100 Gb/s, pro připojení jsou vyhrazeny dva porty QSFP28 na prvcích storage sítě zadavatele.

#### Softwarové vybavení

Požadovanou součástí řešení superpočítače je softwarové vybavení.

Softwarové vybavení zahrnuje operační systémy všech serverů, firmware a ovladače hardware, systémové, komunikační a další knihovny, infrastrukturní software a software pro realizaci výpočetního prostředí superpočítače.

Infrastrukturní software a software pro realizaci výpočetního prostředí superpočítače zahrnuje software pro správu a provisioning serverů/clusteru, management software zařízení, software řešení datových úložišť a bran, sdílené a/nebo clusterové, paralelní souborové systémy, software infrastrukturních služeb, zálohovací software, software pro serverovou virtualizace (bude-li použita v řešení), software pro monitoring infrastruktury, sběr a analýza logů, optimalizovaný softwarový stack pro komunikaci ve výpočetní síti, optimalizované MPI a další.

Předpokládáme použití OS Linux na všech serverech řešení superpočítače.

Na clusteru bude provozován plánovač úloh Slurm.

#### Bezpečnost

Vzhledem k otevřenému charakteru centra IT4Innovations, které poskytuje služby širokému spektru uživatelů, usilujeme o bezpečnost našich systémů. IT4Innovations je držitelem certifikátů systému managementu bezpečnosti informací podle normy ISO 27001 (ISO/IEC 27001:2013, ČSN ISO/IEC 27001:2014). Preferuje použití operačního systému a softwarového vybavení s dobrou dostupností bezpečnostních aktualizací a jejich rychlou aplikaci. Dostupnost bezpečnostních aktualizací by měla být zajištěna po celou dobu provozování celého komplexního softwarového stacku.

#### Identity management

IT4Innovations provádí centralizovanou správu všech uživatelů centra. Informace o uživatelích jsou propagovány/synchronizovány z centrální repository na všechny provozované systémy/superpočítače. Realizace uživatelských účtů superpočítače bude používat technologie LDAP a SSSD. Účty uživatelů budou uloženy a spravovány v LDAP databázi. Pro realizaci uživatelských účtů a skupin budou použita LDAP schémata posixAccount a posixGroup. Správa uživatelských účtů a skupin bude realizována zabezpečeným protokolem LDAP. Správu uživatelských účtů a skupin bude realizovat výhradně zadavatel.

## HPCQC – AI výpočetní cluster

### Komponenty

Superpočítač je komplexní systém sestávající z výpočetních, úložných, síťových a dalších systémů   
a softwarového řešení určený pro provádění náročných výpočtů (HPC).



Obrázek 1 Orientační schéma řešení superpočítače VC III

**Výpočetní cluster** je tvořen **Výpočetními servery** propojenými **Výpočetní sítí** – vysokorychlostní sítí s nízkou latencí. Výpočetní cluster je určen pro provádění výpočetních úloh uživatelů. Výpočetní servery poskytují výpočetní výkon, Výpočetní síť umožňuje rychlou komunikaci mezi servery a efektivní provádění paralelních úloh. Pro servery clusteru je dále používáno též označení uzel.

**Přístupové servery** *–* servery sloužící pro přístup uživatelů, pro přípravu úloh a dat, kompilaci a ladění kódů, pro zpracování výsledků a pro přenos dat.

**Úložiště** slouží k ukládání a sdílení dat. Úložiště jsou realizována jako komplexní řešení úložných zařízení, I/O serverů (např. souborových serverů), datových sítí a potřebného softwarového vybavení na straně serverů i klientů. Úložiště poskytující souborové služby jsou označována jako **Souborová úložiště**.

IT4Innovations strukturuje uživatelská úložiště výpočetních clusterů takto:

* úložiště HOME – souborové úložiště pro uživatelská nastavení operačního systému a aplikací na superpočítači
* úložiště SCRATCH – souborové úložiště pro krátkodobá data uživatelů, intenzivně používaná výpočetními uzly superpočítače
* úložiště PROJECT – souborové úložiště pro střednědobá a dlouhodobá projektová data uživatelů, sdílená mezi superpočítači IT4Innovations (existující úložiště zadavatele, není součástí zakázky)

**Úložiště** infrastruktury je určeno pro ukládání a sdílení dat infrastruktury superpočítače, tj. pro uložení systémových obrazů (image) serverů, logů, dat infrastrukturních služeb, uživatelského aplikačního vybavení, plánovače atd.

**Infrastrukturní a management servery** jsou určené pro správu superpočítače, zdrojů, úloh, licencí a poskytování infrastrukturních služeb (např. DHCP, DNS, LDAP, licenční servery, plánovače, monitoring, logování atd.), zdrojů.

Řešení **Zálohování** zajišťuje ochranu dat před úplnou ztrátou v případě vážného selhání primárního úložiště dat, chyby apod.

Superpočítač obsahuje **Síťovou infrastrukturu,**tj. síťové propojení komponent, systémů tak, aby bylo dosaženo požadované funkcionality, byl zajištěn přístup na jednotlivé služby, byl zajištěn výkon, dostupnost a bezpečnost.

Síťovou infrastrukturu tvoří zejména Výpočetní síť, LAN síť, integrace do WAN sítě zadavatele, integrace do storage sítě zadavatele a případně další sítě (například datové sítě SAN) dle návrhu dodavatele*.*

**Výpočetní síť**propojuje výpočetní servery, přístupové servery, je použita pro zpřístupnění některých úložišť.

**LAN síť** zajišťuje komunikaci mezi zařízeními uvnitř superpočítače, slouží ke správě systému, poskytování infrastrukturních služeb atp.

**Integrace do WAN sítě** zadavatele zajišťuje propojení do WAN sítě zadavatele a zprostředkovaně služby internetu v superpočítači.

**Integrace do storage sítě** zadavatele zajišťuje propojení do storage sítě zadavatele, kde jsou připojena centralizovaná úložiště zadavatele (např. úložiště PROJECT), tak aby byla zajištěna dostupnost úložišť na uzlech superpočítače. Integrace je obvykle realizována síťovými branami realizujícími propojení storage sítě zadavatele (100 Gb/s Ethernet) a výpočetní sítě clusteru.

**Infrastrukturu pro provoz v datovém centru** tvoří. zejména racky potřebné pro umístění zařízení, řešení chlazení zařízení superpočítače, napájení, rozhraní a napojení na infrastrukturu datového centra.

**Softwarové vybavení** je nezbytnou součástí řešení superpočítače. Softwarové vybavení zahrnuje mj. operační systémy všech serverů, firmware a ovladače hardware, knihovny, infrastrukturní software, software pro realizaci výpočetního prostředí superpočítače a výpočetní aplikace.

### Řešení superpočítače

Řešení superpočítače musí umožňovat efektivní provádění mnoha současných výpočetních úloh všech fází životního cyklu úloh (příprava, výpočet, zpracování výsledků) různého typu (paralelní, sériové; dávkové, interaktivní) mnoha uživatelů, bude zajišťovat bezpečné uložení dat uživatelů a rychlý přístup k datům, efektivní správu systému, komponent, zdrojů a uživatelů.

Řešení musí poskytovat transparentní, jednotné, sdílené uživatelské prostředí a jednotný přístup k výpočetním zdrojům.

Redundance řešení musí být v takovém rozsahu, aby při selhání fyzické komponenty nedošlo k významnému omezení poskytování služeb (aby neselhala klíčová služba nutná pro běh řešení) nebo aby nedošlo ke ztrátě důležitých dat. Zabezpečení významných dat bude dále zajištěno zálohováním.

Všechny servery řešení budou umožňovat vzdálenou správu nezávislou na běhu operačního systému.

Řešení musí umožňovat efektivní provoz a nízké provozní náklady.

Superpočítač HPCQC – AI bude obsahovat tyto komponenty:

* **Výpočetní servery**
* **Výpočetní síť**
* **Přístupové servery**
* **Úložiště HOME**
* **Úložiště SCRATCH**
* **Úložiště infrastruktury**
* **Infrastrukturní a management servery**
* **Zálohování**
* **Síťová infrastruktura**
  + **LAN síť**
  + **Integrace do WAN sítě**
  + **Integraci do storage sítě**
* **Infrastruktura pro provoz v datovém centru**
* **Softwarové vybavení**

### Požadavky

#### Výpočetní servery

* minimálně 100 výpočetních serverů
* architektura x86\_64, nebo ARM
* OS Linux
* jeden nebo dva procesory na výpočetní server
* minimálně 960 GB standardní paměti DDR RAM na výpočetní server
* min. 450 GB/s propustnost do paměti CPU/socket
* min. čtyři GPU akcelerátory
* výkon min. 4000 TFlop/s na server v FP32 (tensorflow/sparsity AI maximální teoretický výkon)
* minimálně 2 NIC pro vysokorychlostní síť s agregovanou šířkou pásma min. 800 Gb/s
* chlazení pomocí DLC, celkově min. 90% tepla do vody o vstupní teplotě min. 38 °C

#### Přístupové servery

* 2 přístupové (login) servery
* CPU, GPU a paměť serveru – shodná konfigurace jako u výpočetních serverů

#### Úložiště SCRATCH

Úložiště SCRATCH slouží pro uložení krátkodobých dat uživatelů, data jsou obvykle intenzivně používaná výpočetními uzly superpočítače.

* Úložiště SCRATCH bude mít čistou kapacitu minimálně 2PB, propustnost minimálně 1000 GB/s a bude těsně (s vysokou propustností a s nízkou latencí) připojeno na všechny uživatelsky dostupné uzly (výpočetní, přístupové).
* Předpokládáme použití technologie SSD disků (technologie NVMe) pro zajištěné vysoké propustnosti a I/O výkonu (5000 kIOPS).
* Jako souborový systém bude použit vysoce výkonný paralelní clusterový souborový systém „enterprise“ kvality, zajišťující kromě vysokého výkonu i vysokou bezpečnost přístupu k datům ve sdíleném prostředí.

#### Výpočetní síť

* Výpočetní síť bude mít rychlost portu minimálně 400Gb/s a latenci menší než 10 mikrosekund.
* Topologie sítě musí poskytovat neblokující ostrovy výpočetních uzlů o velikosti minimálně 32 uzlů.
* Síť musí podporovat hardwarovou akceleraci MPI komunikace.
* Síť musí podporovat IP protokol.
* Síť musí být vhodná a podporovaná pro řešení úložiště SCRATCH.

#### Úložiště HOME

Úložiště HOME slouží pro uložení uživatelských nastavení operačního systému a aplikací v rámci superpočítače, předpokládáme tedy potřebu relativně nízké kapacity, přiměřené propustnosti a výkonu.

* Úložiště HOME bude mít čistou kapacitu minimálně 25TB a bude sdíleno na všechny uživatelsky dostupné uzly (výpočetní, přístupové).
* Úložiště bude postaveno na SSD discích.
* Předpokládáme použití technologie NFS s vysokou dostupností (HA).

#### Úložiště infrastruktury

Úložiště infrastruktury je určeno pro ukládání a sdílení dat infrastruktury superpočítače, tj. pro uložení systémových obrazů (image) serverů, logů, dat infrastrukturních služeb, uživatelského aplikačního vybavení, plánovače atd.

Úložiště bude mít čistou kapacitu minimálně 25TB.

Úložiště bude postaveno na SSD discích.

Konfigurace bude v režimu redundance/vysoké dostupnosti minimálně na úrovni disků, řadičů diskového pole, připojení k serverům a souborových serverů.

Úložiště infrastruktury bude připojeno k serverům (přímo či vhodnou sítí) blokovým způsobem anebo pomocí síťového souborového systému (např. NFS). Předpokládané použití ze strany zadavatele:

* APPS – aplikační software sdílený pomocí NFS na všechny uživatelsky dostupné servery

Pro realizaci úložiště infrastruktury a úložiště HOME lze použít jedno společné diskového pole.

#### Infrastrukturní a management servery

Systém bude obsahovat minimálně dva fyzické servery pro infrastrukturní služby a management výpočetního systému. Klíčové infrastrukturní služby budou realizovány v konfiguraci vysoké dostupnosti (HA).

Několik infrastrukturních serverů (předpokládaný počet cca 5ks) bude určeno pro běh služeb zadavatele.

Preferujeme – v případě použití malého počtu fyzických serverů – oddělení služeb na virtuální servery (použití serverové virtualizace).

#### Zálohování

Zálohování bude poskytovat zabezpečení dat infrastrukturních a management serverů, úložiště infrastruktury a úložiště HOME.

Zálohování bude řešeno bez využití pásek (na diskové úložiště) a bude využívat technologii pro vyšší efektivitu uložení dat (např. kompresi nebo deduplikaci). Zálohování bude využívat vyhrazené, nezávislé úložiště dat.

#### LAN síť

LAN síť bude zajišťovat komunikaci mezi zařízeními uvnitř superpočítače a poskytovat připojení do WAN sítě IT4Innovations.

LAN síť bude obsahovat veřejné a privátní části sítě.

Veřejné části sítě (dále jen veřejné sítě), ve kterých budou použity veřejné IPv4 a IPv6 adresy, budou sloužit pro poskytování služeb dostupných z internetu.

Privátní části sítě (dále jen privátní sítě), ve kterých budou použity privátní IPv4 adresy, budou sloužit pro vnitřní služby a pro management zařízení.

Bude zajištěna konektivita výpočetních serverů do LAN sítě a internetu.

Zařízením v privátních sítích musí být umožněn přístup do internetu přes NAT.

LAN síť bude rozdělena do různých L3 sítí. Pro každou L3 síť bude použita jiná L2 síť (VLAN nebo jiný aktivní prvek). Je nežádoucí používat jednu L2 síť pro více L3 sítí. Rozdělení sítí bude zajišťovat zejména oddělení těchto provozů:

* služby dostupné z internetu
* datová komunikace mezi servery (služby)
* management síťových aktivních prvků
* management diskových polí a storage zařízení
* management serverů (BMC, IPMI, apod.)
* management non-IT infrastruktury (napájení, chlazení apod.)

LAN síť nezbytná pro propojení přístupových serverů do WAN sítě bude řešena redundantními prvky. Aktivní prvky sítě budou vzdáleně řízené a centrálně spravovatelné.

#### Integrace do WAN sítě

Součástí zakázky je připojení do WAN sítě zadavatele.

Připojení do WAN sítě zadavatele musí poskytovat konektivitu min. 2x100 Gb/s, pro připojení jsou vyhrazeny dva porty QSFP28 na prvcích WAN sítě zadavatele.

#### Integrace do storage sítě

Součástí zakázky je připojení do storage sítě zadavatele. Pro realizaci se předpokládá použití minimálně dvou vyhrazených serverů implementujících síťové brány mezi storage sítí zadavatele (na technologii 100Gb/s Ethernet) a výpočetní sítí clusteru

Připojení do storage sítě zadavatele musí poskytovat konektivitu min. 2x100 Gb/s, pro připojení jsou vyhrazeny dva porty QSFP28 na prvcích storage sítě zadavatele.

#### Softwarové vybavení

Požadovanou součástí řešení superpočítače je softwarové vybavení.

Softwarové vybavení zahrnuje operační systémy všech serverů, firmware a ovladače hardware, systémové, komunikační a další knihovny, infrastrukturní software a software pro realizaci výpočetního prostředí superpočítače.

Infrastrukturní software a software pro realizaci výpočetního prostředí superpočítače zahrnuje software pro správu a provisioning serverů/clusteru, management software zařízení, software řešení datových úložišť a bran, sdílené a/nebo clusterové, paralelní souborové systémy, software infrastrukturních služeb, zálohovací software, software pro serverovou virtualizace (bude-li použita v řešení), software pro monitoring infrastruktury, sběr a analýza logů, optimalizovaný softwarový stack pro komunikaci ve výpočetní síti, optimalizované MPI a další.

Předpokládáme použití OS Linux na všech serverech řešení superpočítače.

Na clusteru bude provozován plánovač úloh Slurm.

#### Bezpečnost

Vzhledem k otevřenému charakteru centra IT4Innovations, které poskytuje služby širokému spektru uživatelů, usilujeme o bezpečnost našich systémů. IT4Innovations je držitelem certifikátů systému managementu bezpečnosti informací podle normy ISO 27001 (ISO/IEC 27001:2013, ČSN ISO/IEC 27001:2014). Preferuje použití operačního systému a softwarového vybavení s dobrou dostupností bezpečnostních aktualizací a jejich rychlou aplikaci. Dostupnost bezpečnostních aktualizací by měla být zajištěna po celou dobu provozování celého komplexního softwarového stacku.

#### Identity management

IT4Innovations provádí centralizovanou správu všech uživatelů centra. Informace o uživatelích jsou propagovány/synchronizovány z centrální repository na všechny provozované systémy/superpočítače. Realizace uživatelských účtů superpočítače bude používat technologie LDAP a SSSD. Účty uživatelů budou uloženy a spravovány v LDAP databázi. Pro realizaci uživatelských účtů a skupin budou použita LDAP schémata posixAccount a posixGroup. Správa uživatelských účtů a skupin bude realizována zabezpečeným protokolem LDAP. Správu uživatelských účtů a skupin bude realizovat výhradně zadavatel.

## Rozsah zakázky

Součástí zakázky je:

* komplexní implementace celého systému v datovém centru zadavatele (doprava, integrace do datového centra, instalace, konfigurace, ladění, testování, likvidace obalů)
* příprava a provedení akceptačních testů pro prokázání funkčnosti, vlastností a parametrů díla
* zaškolení obsluhy (školení)
* poskytování servisu, podpory a provádění pravidelných sw upgradů po dobu záruční doby
* konzultace v průběhu provozu řešení

## Závazný obsah Návrhu

Návrh musí obsahovat tyto informace:

1. **Návrh technického řešení** 
   * návrh hardwarového řešení
     + uvedení zvažovaných komponent, jejich počty, názvy výrobců a modelů
   * návrh softwarového řešení
     + uvedení zvažovaného SW vybavení, názvy výrobců, produktů a verzí
   * výkonové a kapacitní parametry řešení
     + výpočetní výkon v FP64 a FP32 na úrovni komponent i agregovaný
     + paměťová propustnost
     + výkon úložných zařízení (jak agregovaná propustnost, tak výkon IOPS)
   * energetická náročnost řešení
     + specifikace proudového připojení [A]
   * popis chlazení systému
     + rámcové uvedení kapacity [kW] odvodu tepla do okruhů s teplou a studenou vodou
     + uvedení souladu návrhu chlazení s technickými parametry chladících okruhů, především objemu průtoků a způsobených tlakových ztrát při chlazení plně zatíženého systému kapalinou o teplotě 38°C v okruhu zadavatele a s chemickým složením uvedeným v popisu vlastností datového sálu (směs voda/glykol)
     + v případě nesouladu řešení s vlastnostmi chladících okruhů prosím uveďte možnosti jak chlazení zajistit, ať již snížením teploty chladící kapaliny, či snížení tlakových ztrát za pomocí snížení objemu průtoku chladící kapaliny v okruhu zadavatele
   * prostorová náročnost řešení
     + použité racky, zabraná plocha, plošná i bodová zátěž [kg/m2] na podlahovou plochu
   * předpokládaná časová dostupnost navrhovaných technologií

*Zadavatel požaduje, aby každý z účastníků PTK ve svém návrhu potvrdil, že zvažovaná technická řešení jsou realizovatelná včetně jejich finální akceptace do konce října roku 2026.*

* + popis funkcionality a vlastností řešení
    - Explicitní uvedení jakým způsobem lze na navrženém HW provozovat akcelerovaným způsobem kvantový simulátor. Je potřeba uvést produktový název SW, verzi a referenci využití ve výzkumné instituci nebo soukromé firmě

*Funkcionalita kvantového simulátoru bude mandatorním požadavkem u systému HPCQC – AI a proto je potřeba volit komponenty především výpočetní části tak, aby akceleraci umožňovala.*

* + spolehlivost a dostupnost řešení
  + soulad technického řešení (především výkonových a kapacitních parametrů) a rozpočtu
    - v případě, že požadované kapacitní a výkonové parametry zadavatele přesahují předpokládaný rozpočet, uveďte, prosím, varianty řešení v původním požadovaném rozsahu a taktéž adekvátně redukovaný návrh s dodržením daného rozpočtu. Varianty s původními požadovanými parametry prosím označte jako „Původní varianta“ a redukovanou verzi jako „Redukovaná varianta“
    - dále prosíme o explicitní potvrzení adekvátního poměru výpočetního výkonu výpočetních uzlů a kapacit síťového připojení výpočetních uzlů do vysokorychlostní sítě za účelem zpracování náročných úloh z oblasti umělé inteligence jako např. trénování velkých jazykových modelů jak z pohledu komunikace při paralelním trénování, tak přístupu na úložiště SCRATCH
    - taktéž prosíme o explicitní potvrzení adekvátního poměru kapacit a výkonu výpočetní a SCRATCH úložné části, především tedy celkový výpočetní výkon k agregované propustnosti a agregovanému výkonu IOPS operací
    - v případě, že byste doporučili jiné kapacity v rámci vysokorychlostní sítě či kapacity a výkon úložiště SCRATCH, uveďte toto prosím jako samostatnou pod-variantu varianty Původní a/nebo Redukované s označením „Optimalizovaná varianta“
  + návrh musí být k dispozici rovněž v elektronické formě umožňující editaci

1. **Návrh záruky a servisních služeb**
   * reaktivní služby
   * proaktivní služby
2. **Návrh licenčních podmínek** upravujících práva a závazky zadavatele k navrhovanému řešení
3. **Orientační cenová kalkulace navrhovaného řešení Výpočetního clusteru VC III / HPCQC – AI výpočetního clusteru** (indikativní nabídka) včetně rozkladu cen na logické a funkční celky a služby v členění minimálně:

Výpočetní servery, Výpočetní síť, Přístupové servery, Úložiště HOME, Úložiště SCRATCH, Úložiště infrastruktury, Infrastrukturní a management servery, Zálohování, LAN síť, Integrace do WAN sítě, Integrace do storage sítě, Racky, Chlazení, Softwarové vybavení, Slurm s podporou, Napojení na infrastrukturu DC, Implementace, Školení, Servisní služby.

Indikativní nabídka bude rovněž obsahovat výhled ohledně stability cen. Zadavatel cílí na možnost/vhodnost využit inflační doložku nebo vhodné mechanismy pro narovnání kritických cenových rozdílů v cenách vzniklých v průběhu plnění navazující veřejné zakázky.

## Požadavky na záruku a servisní služby

Zadavatel požaduje zohlednit následující parametry záruky a servisních služeb:

* záruční doba 5 let
* autorizovaný servis
* servis v místě instalace
* jediné kontaktní místo pro hlášení poruch
* možnost hlášení poruch a vad 24 hodin denně

Kategorie vad:

* Vada kategorie A – vada, která zcela nebo podstatným způsobem znemožňuje užívání superpočítače,
* Vada kategorie B – vada, která nebrání užívání superpočítače, ale podstatným způsobem jej omezuje, nebo která vytváří riziko znemožnění užívání superpočítače,
* Vada kategorie C – jiná vada, která není vadou kategorie A ani B

Zahájení řešení odstranění vady:

* Vada kategorie A – okamžitě po nahlášení vady
* Vada kategorie B – maximálně do 2 hodin od nahlášení vady
* Vada kategorie C – maximálně do 24 hodin od nahlášení vady

Odstranění vady:

* Vada kategorie A – do 24 hodin od nahlášení vady
* Vada kategorie B – do 72 hodin od nahlášení vady
* Vada kategorie C – do 7 pracovních dní od nahlášení vady

Za odstranění vady kategorie A nebo B lze považovat poskytnutí akceptovatelného náhradního řešení. Náhradní řešení vady kategorie A se považuje za nahlášenou vadu kategorie B a náhradní řešení vady kategorie B se považuje za nahlášenou vadu kategorie C; přípustné je jen to náhradní řešení, které skutečně umožňuje změnu kategorizace vady.

## Infrastruktura zadavatele

### Datové centrum

IT4Innovations národní superpočítačové centrum provozuje veškeré systémy ve svém datovém sále umístěném v budově IT4Innovations.

Datové centrum IT4Innovations disponuje redundantními zdroji zálohovaného napájení (dvě nezávislé napájecí větve, redundance 1+1) rozvodné soustavy 400/230 V;3+N+PE; 50 Hz; TN-S.

Datový sál o výměře 511 m2 poskytuje světlou výšku mezi systémovou podlahou a stropním podhledem 2.9 m.

#### Chlazení na datovém sále — současný stav

Chlazení pro datový sál poskytuje pět samostatných chladících kapalinových okruhů – dva okruhy teplé vody teploty a tři okruhy studené vody.

V datovém sále jsou ve zdvojené podlaze vyhotoveny tři větve z každého z pěti chladících okruhů. Na větvích jsou pro napojení technologie připraveny odbočky s bezúkapovými uzávěry DN65. Odbočky jsou umístěné pod technologickou podlahou datového sálu a jsou přístupné po vyjmutí některé z dlaždic technologické podlahy. V okruzích je použita směs 35% propylenglykol a 65% voda.

Základní parametry chladících okruhů, včetně údajů pro připojení:

TV1 – okruh Teplé Vody – červený:

* Hydraulické zakončení: Victaulic UK 76,1
* Pracovní tlak: 4,5bar
* diferenční tlak pro datový sál: 50kPa
* Průtok: 40m3/hod
* Vstupní teplota chladícího média: 30°C

 TV2 – okruh Teplé Vody – žlutý:

* Hydraulické zakončení: Victaulic UK 76,1
* Pracovní tlak: 4,5bar
* diferenční tlak pro datový sál: 50kPa
* Průtok: 40m3/hod
* Vstupní teplota chladícího média: 30°C

 SV1 – okruh Studené Vody – zelený:

* Hydraulické zakončení: Victaulic UK 76,1
* Pracovní tlak: 4,5bar
* diferenční tlak pro datový sál: 100kPa
* Průtok: 20m3/hod
* Vstupní teplota chladícího média 11,5°C

 SV2 – okruh Studené Vody – modrý:

* Hydraulické zakončení: Victaulic UK 76,1
* Pracovní tlak: 4,5bar
* diferenční tlak pro datový sál: 100kPa
* Průtok: 20m3/hod
* Vstupní teplota chladícího média 11,5°C

 SV3 – okruh Studené Vody – tyrkysový:

* Hydraulické zakončení: Victaulic UK 76,1
* Pracovní tlak: 4,5bar
* diferenční tlak pro datový sál: 100kPa
* Průtok: 20m3/hod
* Vstupní teplota chladícího média 11,5°C

#### Chlazení na datovém sále — plánovaný stav

Zadavatel předpokládá, že v rámci plánované modernizace datového sálu dojde k navýšení teploty okruhů teplé vody na 38°C a navýšení kapacity o další (třetí okruh teplé vody). Předpokládaný termín realizace je do konce listopadu 2026. Z tohoto důvodu již zadavatel očekává technická řešení výpočetní části clusteru (plus další, je-li to technický možné) umožňující odvod tepla do okruhů s provozní teplotou cca. 38°C.

### WAN síť

WAN síť zadavatele je postavena na technologii 100Gb/s Ethernet a je tvořena dvěma hraničními prvky – L3 switchi Cisco Nexus 9336C-FX2 s porty QSFP28. Hraniční prvky jsou vzájemně propojeny čtyřmi 100Gb/s propoji.

Hraniční prvky WAN sítě nejsou osazeny transceivery, je potřeba je dodat.

### Storage síť

Storage síť zadavatele je postavena na technologii 100Gb/s Ethernet a je tvořena dvěma hraničními prvky – L3 switchi Cisco Nexus 9336C-FX2 s porty QSFP28. Hraniční prvky jsou vzájemně propojeny čtyřmi 100Gb/s propoji. Storage síť je připojena do centrální 100Gb/s WAN/LAN sítě zadavatele čtyřmi 100Gb/s propoji.

Hraniční prvky storage sítě nejsou osazeny transceivery, je potřeba je dodat.