# Vymezení požadavků na Malý cluster Příloha č. 1 Zadávací dokumentace veřejné zakázky „Malý cluster III – IT4I“

## Úvod

IT4Innovations národní superpočítačové centrum (dále jen „IT4Innovations“) poskytuje národní výzkumnou infrastrukturu v oblasti náročných výpočtů (HPC).

Výpočetní zdroje IT4Innovations jsou určeny pro řešení úloh ve výzkumu a vývoji, především pro akademická pracoviště a další výzkumné instituce v ČR, část kapacity je pak dedikována pro rozvoj spolupráce mezi akademickou sférou a průmyslovými partnery nebo pro samostatné využití průmyslovými podniky.

IT4Innovations je členem národní velké výzkumné infrastruktury e-INFRA CZ.

## Záměr zadavatele

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

Malý cluster III bude provozován IT4Innovations a bude využíván k poskytování služeb v oblasti náročných výpočtů (High Performance Computing).

Malý cluster III bude umístěn a provozován v datovém centru zadavatele, které se nachází v budově IT4Innovations, v areálu Vysoké školy báňské – Technické univerzity Ostrava.

Záměr zadavatele byl podrobně diskutován v rámci předběžné tržní konzultace.

## Předmět zakázky

Předmětem veřejné zakázky „Malý cluster III – IT4I“ je dodávka komplexního řešení systému pro náročné výpočty (High Performance Computing) tj. komplexu výpočetních, úložných, síťových a dalších systémů, softwarového řešení, včetně implementace, integrace do datového centra zadavatele, školení, servisních a dalších služeb.

## Legenda

V následujícím textu jsou uváděny následující značky:

**SPEC\_číslo** označuje pro snazší identifikaci jednotlivé požadavky zadavatele veřejné zakázky.

**SPEC\_číslo (I)** označuje požadavek zadavatele veřejné zakázky na informaci, kterou dodavatel musí uvést v nabídce.

## Malý cluster

### Obecné požadavky

1. Zadavatel požaduje komplexní řešení výpočetního systému určeného pro náročné výpočty (High Performance Computing) tj. komplexu výpočetních, úložných, síťových a dalších systémů, softwarového řešení, včetně implementace a integrace do datového centra zadavatele (dále označované též jako „řešení“, „Malý cluster“ nebo „superpočítač“).
2. Řešení musí umožňovat efektivní provádění mnoha současných výpočetních úloh všech fází životního cyklu úloh (příprava, výpočet, zpracování výsledků) různého typu (zejména paralelní, ale rovněž sériové; dávkové, interaktivní) mnoha uživatelů, bezpečné uložení dat uživatelů a rychlý přístup k datům, efektivní správu systému, komponent, zdrojů a služeb.
3. Řešení musí poskytovat výkonné výpočetní zdroje, které jsou dobře přístupné a využitelné uživateli systému a jejich úlohami. Řešení musí zajišťovat vlastnosti, služby a funkce potřebné pro efektivní provoz a efektivní správu systému zadavatelem. Řešení musí být vyvážené, parametry a vazby jednotlivých subsystémů musí zohledňovat ostatní subsystémy.
4. Nabídka musí obsahovat veškeré systémy, zařízení, komponenty, příslušenství, licence, dokumentace, projektové, implementační a další práce, školení atd. nezbytné k naplnění požadavků zadavatele.
5. Řešení musí respektovat dispozice a omezení vyplývající z prostředí a podmínek zadavatele.
6. Řešení musí komplexně (jako celek) splňovat požadavky zadavatele. Požadovaná funkcionalita a vlastnosti musí být reálně funkční a použitelné v provozu řešení, požadované parametry musí být reálně dosažitelné. Splnění požadavků zadavatele nesmí být nijak podmíněno.

Je nepřípustné, aby plnění požadavků zadavatele bylo postaveno pouze na funkcionalitě, vlastnostech či parametrech dílčích komponent a řešení jako celek požadavky zadavatele nesplňovalo.

1. Funkcionalita, vlastnosti a parametry řešení musí být uvedeny pro nabízenou/dodávanou konfiguraci určenou k běžnému, trvalému provozu.
2. Řešení musí splňovat všechny technické požadavky zadavatele současně. Všech požadovaných vlastností, funkcionalit a parametrů musí být dosaženo při použití jednoho, produkčního nastavení všech komponent řešení. Splnění požadavku nesmí být podmíněno změnou nastavení nebo změnou zapojení komponent.
3. Řešení nesmí obsahovat omezení, která by zabraňovala či omezovala užití Malého clusteru zadavatelem v požadovaném, nebo racionálním rozsahu. Malý cluster je určen pro přibližně tři tisíce uživatelů.
4. Řešení musí být v maximální míře autonomní, nezávislé na externích systémech a službách, soběstačné bez potřeby dalších zařízení, systémů či služeb.
5. Řešení musí být navrženo, dimenzováno a implementováno tak, aby zajistilo spolehlivý, bezpečný, výkonný a efektivní provoz v datovém centru zadavatele.
6. Dodávka musí obsahovat veškerá zařízení potřebná pro zajištění požadované funkcionality, pro zajištění požadovaného výkonu a pro efektivní provoz – a to i taková, která nejsou explicitně uvedena v tomto dokumentu.
7. Dodávka musí obsahovat veškeré potřebné softwarové vybavení a licence.
8. Dodavatel nesmí pro realizaci používat zařízení, systémy, infrastrukturu či služby zadavatele, pokud toto není explicitně uvedeno v tomto dokumentu, a to pouze k účelu a v rozsahu uvedeném v tomto dokumentu.

### Orientační schéma Malého clusteru

A diagram of a project

Description automatically generated

Obrázek 1 Orientační schéma Malého clusteru

Orientační schéma Malého clusteru je pouze zjednodušené ilustrativní znázornění Malého clusteru, nejedná se o úplné či přesné zapojení. Modré objekty představují komponenty, které jsou předmětem dodávky, oranžové objekty představují stávající infrastrukturu zadavatele.

### Komponenty Malého clusteru

1. Malý cluster musí obsahovat *Výpočetní cluster*. Výpočetní cluster je tvořen *Výpočetními servery* propojenými *Výpočetní sítí* – vysokorychlostní sítí s nízkou latencí. Výpočetní cluster je určen pro provádění výpočetních úloh uživatelů.
2. Výpočetní servery tvoří jeden typ serverů tzv. *Standardní výpočetní servery*.
3. Malý cluster musí obsahovat *Přístupové servery – servery* sloužící pro přístup uživatelů, pro přípravu úloh a dat, kompilaci a ladění kódů, pro zpracování výsledků a pro přenos dat.
4. Malý cluster musí obsahovat *Úložiště.* Úložiště slouží k ukládání a sdílení dat. Úložiště jsou realizována jako komplexní řešení úložných zařízení, I/O serverů (např. souborových serverů), sítí a potřebného softwarového vybavení. Úložiště musí poskytovat požadované datové služby.
5. Malý cluster musí obsahovat *Úložiště HOME*, *Úložiště SCRATCH* a *Úložiště infrastruktury*.
6. Malý cluster musí obsahovat *Úložiště HOME*. *Úložiště HOME* je souborové úložiště, které je určeno pro ukládání nastavení operačního systému a aplikací uživatelů na superpočítači.
7. Malý cluster musí obsahovat *Úložiště* *SCRATCH*. *Úložiště* *SCRATCH* je výkonné souborové úložiště, které je určeno pro krátkodobá data uživatelů pro výpočty a zpracování úloh, data jsou intenzivně používaná výpočetními servery superpočítače.
8. Malý cluster musí obsahovat *Úložiště infrastruktury. Úložiště infrastruktury* je souborové úložiště, které je určeno pro ukládání a sdílení dat infrastruktury superpočítače. Úložiště slouží pro uložení systémových obrazů (image) serverů, dat, uživatelského aplikačního vybavení, atp.
9. Malý cluster musí obsahovat infrastrukturní a management servery dále označované stručně jako *Infrastrukturní servery.* Infrastrukturní servery jsou určené pro správu superpočítače, zdrojů, úloh, licencí a poskytování infrastrukturních služeb (např. DHCP, DNS, LDAP, licenční servery, plánovače, monitoring, logování atd.), zdrojů.
10. Malý cluster musí obsahovat řešení zálohování dat tzv. *Zálohování*.
11. Malý cluster musí obsahovat *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 služeb.
12. Síťovou infrastrukturu tvoří zejména *Výpočetní síť*, *LAN síť*, *Integrace do WAN sítě zadavatele*, *Integrace do storage sítě zadavatele* a případně další sítě 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 realizována *Síťovými branami* realizujícími propojení storage sítě zadavatele a Výpočetní sítě clusteru.

1. Malý cluster musí obsahovat řešení a infrastrukturu pro instalaci a provoz superpočítače v datovém centru zadavatele (dále označováno jako *Infrastruktura pro provoz v datovém centru*), tj. zejména racky a příslušenství potřebné pro umístění zařízení superpočítače, řešení napájení a chlazení zařízení superpočítače, rozhraní a napojení na infrastrukturu datového centra zadavatele.
2. Malý cluster musí obsahovat veškeré potřebné softwarové vybavení a licence (*Software*).
3. Kromě uvedených komponent, které jsou nutnou součástí řešení, musí řešení obsahovat všechny další systémy potřebné pro zajištění požadované funkcionality a pro efektivní provoz Malého clusteru.

### Výpočetní cluster

#### Paměťová propustnost STREAM

1. Pro měření paměťové propustnosti STREAM bude použit benchmark STREAM https://www.cs.virginia.edu/stream/ verze 5.10.

Bude použit konkrétní zdrojový kód benchmarku STREAM v jazyce C zveřejněný na adrese <https://www.cs.virginia.edu/stream/FTP/Code/stream.c>

Zdrojový kód benchmarku STREAM je uveden v příloze č. 1 tohoto dokumentu.

1. Paměťová propustnost STREAM výpočetního serveru bude určena během benchmarku STREAM na serveru, za výsledek bude považována hodnota „Best Rate MB/s“ funkce TRIAD.
2. Paměťová propustnost STREAM výpočetního clusteru bude určena jako součet naměřených hodnot paměťové propustnosti STREAM všech Standardních výpočetních serverů. Měření bude prováděno na všech Standardních výpočetních serverech současně.
3. Pro měření paměťové propustnosti STREAM bude použit originální zdrojový kód benchmarku STREAM dle SPEC\_30 bez úprav.  
   Měření bude realizováno tak, aby využívalo a zatěžovalo paměti DDR serveru.  
   Pro nastavení STREAM\_ARRAY\_SIZE bude použita hodnota o velikosti minimálně 80% operační paměti RAM výpočetního serveru (v bytech).  
   Pro nastavení NTIMES bude použita hodnota 100.  
   Nastavení OFFSET je na zvážení dodavatele.  
   Pro nastavení STREAM\_TYPE bude použita hodnota double.  
   Optimalizace překladu zdrojového kódu a optimalizace běhu benchmarku nesmí negativně ovlivnit správnost provedení benchmarku – tj. zejména musí být reálně provedeny operace a přenosy dat v paměti v požadovaném rozsahu a musí být správně určeny časy běhu a výsledky měření. Pro měření bude spuštěna jedna instance programu STREAM na serveru a využije se OpenMP paralelizace.
4. (I) Dodavatel musí v nabídce uvést paměťovou propustnost STREAM výpočetního clusteru určenou dle požadavků v této kapitole.

#### Výpočetní výkon LINPACK

1. Pro měření výkonu Výpočetního clusteru musí být použity benchmarky High Performance LINPACK <http://www.netlib.org/benchmark/hpl/>
2. Výpočetní výkon LINPACK výpočetního clusteru bude určen během výpočetního benchmarku High Performance LINPACK spuštěným paralelně nad všemi CPU všech Standardních výpočetních serverů Výpočetního clusteru (jedna instance benchmarku na celém clusteru).
3. Pro realizaci měření není nutno použít referenční implementaci benchmarku LINPACK, lze použít optimalizovanou implementaci benchmarku. Použitá implementace benchmarku však musí zcela splňovat specifikaci benchmarku. Optimalizace překladu zdrojového kódu a optimalizace běhu benchmarku je možná za předpokladu, že negativně neovlivní správnost provedení benchmark – tj. zejména, že budou reálně provedeny výpočty v požadovaném rozsahu a že budou správně určeny časy běhu a výsledky měření.
4. (I) Dodavatel musí v nabídce uvést výpočetní výkon LINPACK výpočetního clusteru určený dle požadavků v této kapitole.

#### Měření – společné požadavky

1. Hodnoty paměťové propustnosti STREAM výpočetního clusteru a výpočetního výkonu LINPACK výpočetního clusteru (stejně tak jako další hodnoty v nabídce) musí být uvedeny pro nabízenou/dodávanou konfiguraci určenou k běžnému provozu. Dosažení hodnot výpočetních výkonů nesmí být nijak podmíněno např. specifickým režimem procesoru, ve kterém nelze systém dlouhodobě a bez dalších omezení provozovat nebo předpokládanou efektivitou, jejíž dosažení však dodavatel negarantuje.

### Výpočetní servery

1. Každý Výpočetní server musí splňovat následující požadavky

* Fyzický server
* Připojení do Výpočetní sítě
* Připojení do LAN sítě
* Podpora bootu operačního systému ze sítě
* 64-bitový operační systém Linux

1. Operační paměť RAM musí být rovnoměrně rozložena (kapacitou a rychlostí přístupu) na procesory a CPU jádra výpočetního serveru. Operační paměť RAM musí být složena z paměťových modulů stejného typu (velikost, rank, atd.) a rovnoměrně, se stejnou konfigurací rozložena na paměťové řadiče a na paměťové kanály výpočetního serveru. Musí být použity všechny paměťové kanály všech procesorů serveru.
2. Výpočetní servery musí používat přímé kapalinové chlazení (Direct Liquid Cooling - DLC).
3. Výpočetní servery jsou určeny výhradně pro výpočty, dodavatel nesmí využít žádný Výpočetní server pro zajištění jiné funkcionality.

#### Standardní výpočetní servery

1. Výpočetní cluster musí obsahovat minimálně 140 Standardních výpočetních serverů.
2. Každý Standardní výpočetní server musí splňovat následující požadavky:

* Architektura x86-64
* Dva procesory na server
* Každý procesor má minimálně 96 fyzických výpočetních jader
* Nebude použit/zapnutý simultaneous multithreading (někdy označovaný jako hyper-threading, SMT, HT)
* Operační paměť RAM typu DDR5 s ECC, minimálně 6000MT/s
* Kapacita operační paměti RAM minimálně 768GiB
* Připojení do LAN sítě rychlostí minimálně 1Gb/s
* Připojení do Výpočetní sítě rychlostí minimálně 200Gb/s

1. Všechny Standardní výpočetní servery musí mít shodnou hardwarovou konfiguraci a musí pracovat ve shodném provozním nastavení (frekvence, časování, nastavení vlastností).
2. (I) Dodavatel musí v nabídce uvést hardwarovou platformu řešení, detailní konfiguraci Výpočetních serverů včetně označení typu procesorů.
3. (I) Dodavatel musí v nabídce uvést přesné označení a verzi operačního systému.

### Výpočetní síť

1. Výpočetní síť musí používat technologii s propustností spoje minimálně 200Gb/s a s latencí spoje maximálně 10 mikrosekund. Připojení zařízení do výpočetní sítě a uvnitř výpočetní sítě musí splňovat požadavky na spoje uvedené v předchozí větě.
2. Výpočetní síť musí poskytovat neblokující ostrovy Výpočetních serverů.

Standardní výpočetní servery musí být zapojeny do neblokujících ostrovů v počtu minimálně 32 Standardních výpočetních serverů na neblokující ostrov.

Ostatní servery a systémy mohou být zapojeny do ostrovů s výpočetními servery (a je to preferováno), musí však být zachována neblokující komunikace výpočetních serverů.

1. Mezi neblokujícími ostrovy je pro výpočetní servery přípustný blokační faktor maximálně 1:2.
2. Řešení musí poskytovat efektivní MPI komunikaci výpočetních serverů.
3. Výpočetní síť musí podporovat a poskytovat IP protokol.
4. Výpočetní síť musí být vhodná a podporovaná pro řešení úložiště SCRATCH.
5. (I) Dodavatel musí v nabídce uvést technologii sítě, topologii sítě, propustnost sítě, počet a velikost neblokujících ostrovů, blokační faktor sítě a konfiguraci nabízených zařízení.

Výpočetní síť je použita pro zpřístupnění úložišť superpočítače na výpočetní uzly, požadavky jsou uvedeny v kapitole 5.8 Úložiště.

### Přístupové servery

1. Malý cluster musí obsahovat minimálně dva Přístupové servery.
2. Každý Přístupový server musí splňovat následující požadavky:

* Fyzický server
* Architektura x86-64
* Dva procesory na server, každý procesor minimálně 48 fyzických výpočetních jader
* Operační paměť RAM typu DDR5 s ECC, minimálně 6000MT/s
* Kapacita operační paměti RAM minimálně 384GiB
* Minimálně 2 lokální SSD disky o kapacitě minimálně 480GB v RAID1
* Připojení Výpočetní síť
* Připojení LAN síť
* Za provozu vyměnitelné (hot-swap) disky
* Redundantní, za provozu vyměnitelné (hot-swap) napájecí zdroje, redundantní napájení

1. Přístupové servery musí používat stejnou technologií procesorů, stejnou instrukční sadu procesorů a stejný operační systém jako Standardní výpočetní servery.
2. Přístupové servery musí mít stejnou hardwarovou konfiguraci a musí pracovat ve shodném provozním nastavení (frekvence, časování, nastavení vlastností).
3. Přístupové servery jsou určeny výhradně pro zajištění přístupu uživatelů a pro práci uživatelů, dodavatel nesmí využít žádný Přístupový server pro zajištění jiné funkcionality.
4. Přístupové servery musí poskytovat přístup uživatelům protokolem SSH2 a poskytovat služby pro přenos souborů SCP a SFTP.
5. (I) Dodavatel musí v nabídce uvést hardwarovou platformu řešení, detailní konfiguraci Přístupových serverů včetně označení typu procesorů.
6. (I) Dodavatel musí v nabídce uvést přesné označení a verzi operačního systému Přístupových serverů.

### Úložiště

#### Společné požadavky

Požadavky uvedené v této kapitole jsou požadavky společné pro všechna úložiště řešení – úložiště HOME, úložiště SCRATCH, Úložiště infrastruktury a případně další úložiště použitá v řešení (kromě lokálních disků v serverech).

##### Vysoká dostupnost

1. Řešení úložiště musí poskytovat vysokou dostupnost.
2. Výpadek či odstávka libovolného jednoho serveru řešení úložiště nesmí způsobit nefunkčnost služeb úložiště. Při výpadku či odstávce serveru řešení úložiště může být výkon úložiště nižší než požadovaný.
3. Řešení úložiště musí být odolné dlouhodobé vysoké zátěži.
4. Služby a provoz úložišť se nesmí vzájemně negativně ovlivňovat. Deklarované parametry agregované rychlosti a výkonu náhodných I/O operací musí být dosažitelné i při současném, paralelním zatížení všech úložišť.
5. (I) Dodavatel musí v nabídce uvést způsob zajištění vysoké dostupnosti a redundanci úložiště.

##### Redundance disků

1. Úložiště musí zajišťovat takové zabezpečení (redundanci) dat, že selhání libovolných dvou disků úložiště nezpůsobí ztrátu dat.
2. Úložiště musí zajištovat zotavení po selhání disku, tj. opětovné zajištění požadovaného zabezpečení (redundance) dat (např. rekonstrukce RAID skupiny za využití hot-spare disků). Zotavení po selhání disku musí probíhat automaticky, bez zásahu obsluhy.
3. Zotavení po selhání disku úložiště, tj. opětovné zajištění požadovaného zabezpečení (redundance) dat, musí být dokončeno do 24 hodin od selhání disku. Během zotavování po výpadku disku úložiště může být výkon úložiště dočasně nižší než požadovaný.
4. Úložiště musí mít takovou konfiguraci, že je možné zajištění požadovaného zabezpečení (redundance) dat po selhání libovolných dvou disků úložiště, a to bez zásahu obsluhy.
5. Každé diskové pole (nebo obdobné zařízení) řešení úložiště musí poskytovat rezervní kapacitu nebo náhradní disky v počtu či kapacitě minimálně disků.

Funkce max(a; b) vrací větší z čísel a, b. Výsledek vzorce se zaokrouhlí vždy nahoru na celé číslo. Celkový počet disků diskového pole zahrnuje jak disky obsahující samotná data a paritní data, tak disky realizující náhradní disky či rezervní kapacitu.

#### Souborová úložiště

Požadavky uvedené v této kapitole jsou společné požadavky pro úložiště HOME, úložiště SCRATCH a Úložiště infrastruktury, která jsou souborovými úložišti.

1. Souborové úložiště musí poskytovat služby síťového souborového systému.
2. Na všech klientech souborového úložiště musí být poskytována obvyklá funkcionalita souborového systému.
3. Souborové úložiště musí být na straně klientů transparentně integrováno do operačního systému, musí umožňovat obvyklé souborové operace a realizovat obvyklou sémantiku nativních souborových systémů a integrovat uživatele operačního systému jako uživatele souborového systému.
4. Souborové úložiště musí splňovat následující požadavky
5. Podpora Unicode ve jménech souborů
6. Podpora dlouhých jmen souborů
7. Řízení přístupu, přístupová práva na úrovni standardních Unixových práv (čtení, zápis, spuštění; uživatel, skupina, ostatní) a rozšířená ACL
8. Podpora souborů o velikosti větší než 1TB
9. Podpora symbolických linků
10. Podpora zamykání souborů
11. Data souborového úložiště musí být uložena na discích typu SSD nebo NVMe. Použité disky musí být vhodné pro jejich nasazení a charakter zátěže.
12. Požadavek, že souborové úložiště se musí z pohledu uživatele chovat jako jediná, souvislá oblast s jednotným prostorem jmen znamená, že uživatel úložiště pro přístup k souborům úložiště používá jednotný prostor jmen a v rámci tohoto jednotného prostoru jmen je bez omezení dostupná veškerá kapacita úložiště a vlastnosti úložiště.

#### Úložiště SCRATCH

1. Úložiště SCRATCH je souborové úložiště, které je určeno výhradně pro data uživatelů Malého clusteru.
2. Úložiště SCRATCH se musí z pohledu uživatele chovat jako jediná, souvislá oblast s jednotným prostorem jmen.
3. Úložiště SCRATCH musí být prostřednictvím Výpočetní sítě nativně zpřístupněné na všech Výpočetních serverech a Přístupových serverech.
4. Úložiště SCRATCH musí být na klientech dostupné v cestě /scratch.
5. Úložiště SCRATCH musí mít čistou dostupnou kapacitu na úrovni souborového systému minimálně **500TB** (500x10¹² byte).
6. Úložiště SCRATCH musí umožňovat uložení minimálně 1250 miliónů souborů.
7. Úložiště SCRATCH musí poskytovat dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB minimálně **162GB/s** (162x10⁹ byte/s). Požadovaná rychlost musí být reálně dosažitelná z Výpočetních serverů.
8. Úložiště SCRATCH musí poskytovat dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% minimálně **697 tisíc IOPs**. Požadovaný výkon musí být reálně dosažitelný z Výpočetních serverů.
9. Úložiště SCRATCH musí poskytovat projektové, nebo adresářové kvóty, nepřekročitelné omezení využité kapacity a počtu souborů nastavitelné individuálně buď pro projekt, nebo pro adresář úložiště.
10. (I) Dodavatel musí v nabídce uvést čistou dostupnou kapacitu, agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB a výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% a dlouhodobě udržitelný výkon metadata operací úložiště SCRATCH.
11. (I) Dodavatel musí v nabídce uvést architekturu řešení, typ a konfiguraci nabízených zařízení, počet a typ disků, úroveň RAID, počet disků v RAID, počet spare disků a popis softwarového řešení úložiště SCRATCH.

#### Úložiště HOME

1. Úložiště HOME je souborové úložiště, které je určeno výhradně pro data uživatelů Malého clusteru.
2. Úložiště HOME se musí z pohledu uživatele chovat jako jediná, souvislá oblast s jednotným prostorem jmen.
3. Úložiště HOME musí být nativně zpřístupněné prostřednictvím Výpočetní sítě na všech Výpočetních serverech a Přístupových serverech.
4. Úložiště HOME musí být na klientech dostupné v cestě /home.
5. Úložiště HOME musí mít čistou dostupnou kapacitu na úrovni souborového systému minimálně **25TB** (25x10¹²byte).
6. Úložiště HOME musí umožňovat uložení minimálně 500 miliónů souborů.
7. Úložiště HOME musí poskytovat dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB minimálně **1.5GB/s** (1.5x10⁹ byte/s). Požadovaná rychlost musí být reálně dosažitelná z Výpočetních serverů.
8. Úložiště HOME musí poskytovat dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% minimálně **50 tisíc IOPs**. Požadovaný výkon musí být reálně dosažitelný z Výpočetních serverů.
9. Úložiště HOME musí poskytovat uživatelské kvóty, nepřekročitelné omezení využité kapacity a počtu souborů nastavitelné individuálně pro každého uživatele.
10. (I) Dodavatel musí v nabídce uvést čistou dostupnou kapacitu, dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB a dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% úložiště HOME.
11. (I) Dodavatel musí v nabídce uvést architekturu řešení, typ a konfiguraci nabízených zařízení, počet a typ disků, úroveň RAID, počet disků v RAID, počet spare disků a popis softwarového řešení úložiště HOME.

#### Úložiště infrastruktury

1. Úložiště infrastruktury je souborové úložiště, které je určeno pro uložení dat infrastrukturních služeb superpočítače a pro uložení dat zadavatele potřebných pro realizaci a poskytování služeb uživatelům.
2. Zadavatel předpokládá následující použití Úložiště infrastruktury a pro ně stanovuje požadované parametry úložiště:

* instalační obrazy serverů (výpočetní a přístupové servery)
* aplikační software a data zadavatele

1. Dodavatel může využít Úložiště infrastruktury i k jiným účelům než uvedeným v SPEC\_103, musí však adekvátně navýšit parametry řešení a musí zajistit splnění zadavatelem požadovaných parametrů.
2. Úložiště infrastruktury musí mít čistou dostupnou kapacitu na úrovni souborového systému minimálně **25TB** (25x10¹²byte).
3. Úložiště infrastruktury musí umožňovat uložení minimálně 500 miliónů souborů.
4. Úložiště infrastruktury musí poskytovat dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB minimálně **1.5GB/s** (1.5x10⁹ byte/s). Požadovaná rychlost musí být reálně dosažitelná z Výpočetních serverů.
5. Úložiště infrastruktury musí poskytovat dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% minimálně **50 tisíc IOPs**. Požadovaný výkon musí být reálně dosažitelný z Výpočetních serverů.
6. Úložiště infrastruktury musí umožnovat rozdělení datové kapacity na logické části (souborové systémy či jeho části) požadované velikosti a zpřístupnění těchto logických částí výhradně na vybrané servery.
7. Úložiště infrastruktury musí být dostupné všem serverům řešení Malého clusteru.

Úložiště infrastruktury musí zpřístupňovat:

* instalační obrazy serverů – na Infrastrukturní servery určené pro běh vzdálené instalace, popř. běh síťového bootu serverů
* aplikační software a data zadavatele – zpřístupněný v cestě /apps
  + na Přístupové a Výpočetní servery prostřednictvím Výpočetní sítě
  + na vybrané Infrastrukturní servery; pro servery musí existovat datová cesta k úložišti infrastruktury s propustností minimálně 4Gb/s.

1. (I) Dodavatel musí v nabídce uvést čistou dostupnou kapacitu, dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB a dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80%/20% Úložiště infrastruktury.
2. (I) Dodavatel musí v nabídce uvést architekturu řešení, typ a konfiguraci nabízených zařízení, počet a typ disků, úroveň RAID, počet disků v RAID, počet spare disků a popis softwarového řešení Úložiště infrastruktury.

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

#### Vymezení pojmu – kapacita úložiště

1. Kapacita (velikost) úložiště je požadována a musí být uváděna jako čistá využitelná kapacita (není-li explicitně uvedeno jinak), tj. jako velikost úložiště reálně využitelná uživatelem na nejvyšší poskytované úrovni, tj. na úrovni poskytované služby.

Kapacita souborového úložiště je kapacita souborových systémů poskytovaných úložištěm.

Kapacita souborového systému je čistá využitelná kapacita souborového systému, to je reálně využitelná kapacita souborového systému.

1. Určení čisté využitelné kapacity úložiště musí být uvedeno pro nabízenou/dodávanou konfiguraci určenou k běžnému provozu.
2. Určení čisté využitelné kapacity nesmí kalkulovat, zohledňovat vlastnosti systému či jeho komponenty na potenciální možnost uložení většího množství dat za předpokladů, které nelze zajistit (komprese, deduplikace apod.).
3. Určení čisté využitelné kapacity nesmí kalkulovat, zohledňovat vlastnosti systému či jeho komponenty na alokování většího prostoru, než je fyzicky možné či reálně uskutečnitelné bez dalších úkonů (oversubscription).
4. Kapacity úložišť se uvádějí s využitím předpon dekadických násobků.

gigabyte (GB) 10⁹ byte

terabyte (TB) 10¹² byte

petabyte (TB) 10¹⁵ byte

1. Čistá dostupná kapacita datového úložiště se zjišťuje na vhodném klientském systému úložiště zápisem do úložiště do jeho zaplnění nebo čtením celého úložiště anebo vhodným systémovým nástrojem prezentujícím velikost – kapacitu úložiště.
2. Nástroje použité pro zjištění kapacity musí poskytovat důvěryhodné informace a musí pracovat se známou velikostí datového bloku nebo se známou a přesnou jednotkou.

Pro ilustraci, v OS Linux lze velikost souborového systému v bytech zjistit pomocí příkazu df ‑B1, hodnota Available.

#### Vymezení pojmu – rychlost úložiště

1. Rychlost úložiště (rychlost sekvenčních operací a výkon I/O operací) je požadována a musí být uváděna jako rychlost reálně dlouhodobě dosažitelná uživatelem na nejvyšší poskytované úrovni, tj. na úrovni poskytované služby z klientů úložiště.
2. Rychlost souborového úložiště je reálně dlouhodobě dosažitelná (sustained) rychlost operací prováděných na souborových systémech úložiště z klientů souborového úložiště.
3. Rychlost sekvenčních operací úložiště je menší z hodnot rychlost sekvenčního zápisu úložiště a rychlost sekvenčního čtení úložiště.
4. Určení rychlosti úložiště musí být uvedeno pro nabízenou/dodávanou konfiguraci určenou k běžnému provozu.
5. Určení rychlosti nesmí vycházet z předpokladu specifických výhodných podmínek či specifického výhodného režimu měření (např. operace z cache), pokud tyto podmínky nebo režim nejsou explicitně požadovány nebo uvedeny.
6. Rychlosti úložišť se uvádějí s využitím předpon dekadických násobků.

#### Měření rychlosti

1. Dodavatel musí prokázat splnění rychlosti úložišť provedením výkonových testů (benchmarků) v rámci akceptačních testů (měření rychlosti).
2. Běh výkonových testů úložišť bude realizován na Výpočetních serverech (testovací servery).
3. Před každým měřením rychlosti úložiště musí být zajištěno vymazání diskových cache pamětí testovacích serverů.
4. Měření rychlosti je nutno provádět postupy a za podmínek, které odpovídají běžnému provozu, běžnému poskytování služeb úložiště. Před a během měřením rychlosti úložiště se nesmí provádět žádné úkony, které by měly vliv na výsledek měření.

##### Použití nástroje fio

1. Měření rychlosti sekvenčních operací a měření výkonu I/O operací náhodného charakteru souborových úložišť se bude provádět nástrojem fio v klient-server režimu.

fio je open-source nástroj (licence GPL verze 2) pro benchmarkování a testování I/O dostupný na adrese <https://github.com/axboe/fio>.

Pro měření se použije fio verze 3.35 nebo novější, pro měření lze použít fio z distribuce dodávaného operačního systému.

Na každém serveru určeném pro realizaci měření souborového úložiště (testovacím servery dle SPEC\_127) se ověří dostupnost měřeného souborového úložiště a spustí program fio v serverovém režimu:

fio --server

Z vybraného serveru se pak iniciuje měření – spustí se na něm program fio v klientském režimu:

time fio \  
 --client=machinefile jobfile \  
 --output-format=normal,json |& \  
tee fio.out

V souboru machinefile jsou uvedena jména serverů určených pro realizaci měření úložiště, každé jméno je uvedeno na samostatném řádku.

jobfile je jméno konfiguračního souboru programu fio obsahujících popis testu dle požadavků konkrétního měření. Soubor jobfile je umístěn na serveru, který iniciuje měření.

Adresář souborového úložiště použitý pro vytváření testovacích souborů měření (v jobfile jde o nastavení directory) musí být před měřením prázdný.

Pro měření sekvenčního čtení souborového úložiště a pro měření výkonu I/O operací náhodného charakteru souborového úložiště se před měřením v adresáři připraví testovací soubory. Testovací soubory se vytvoří stejným postupem jako samotné měření – spuštěním programu fio v klientském režimu, ale v příslušném jobfile se uvede nastavení create\_only=1 a pro zkrácení doby běhu přípravy souborů je možno použít jinou velikost bloku (parametr bs).

Pro určení výsledku měření bude použit výstup programu fio na serveru, který inicioval měření, tj. soubor fio.out. Budou použity souhrnné informace o měření, které jsou uvedeny v části „All clients“, použije se JSON formát výstupu.

V JSON výstupu se použije poslední položka atributu „client\_stats“, atribut „jobname“ této položky má hodnotu „All clients“, tato položka je dále označovaná jako souhrn.

Pro měření rychlosti sekvenčních operací se použije hodnota atributu „bw\_bytes“ atributu „read“ nebo „write“ (podle typu měření) souhrnu. Získaná hodnota udává rychlost (propustnost) v bytech za sekundu.

Pro měření výkonu I/O operací se použije součet hodnot atributu „iops“ atributů „read“ a „write“ souhrnu. Získaná hodnota udává celkový počet I/O operací za sekundu (IOPS).

##### Měření rychlosti sekvenčních operací

1. Měření dlouhodobě udržitelné agregované rychlosti sekvenčního operací souborového úložiště se bude provádět nástrojem fio dle SPEC\_130.

Rychlost sekvenčních operací datového úložiště je menší z hodnot rychlost sekvenčního zápisu souborového úložiště a rychlost sekvenčního čtení souborového úložiště.

Pro měření sekvenčního zápisu souborového úložiště se použije následující konfigurační soubor – fio jobfile:

[global]

rw=write

bs=1M

create\_on\_open=1

time\_based

runtime=1h

numjobs=NUMJOBS

[FILESYSTEM]

directory=/MOUNTPOINT/test/fio

filesize=FILESIZE

Pro měření sekvenčního čtení souborového úložiště se použije následující konfigurační soubor – fio jobfile:

[global]

rw=read

bs=1M

time\_based

runtime=1h

numjobs=NUMJOBS

;create\_only=1

[FILESYSTEM]

directory=/MOUNTPOINT/test/fio

filesize=FILESIZE

Parametry FILESYSTEM, MOUNTPOINT, NUMJOBS a FILESIZE se nahradí vhodnými hodnotami.

Parametr FILESYSTEM udává jméno úložiště, např. HOME nebo SCRATCH.

Parametr MOUNTPOINT udává mountpoint měřeného úložiště na klientech.

Parametr NUMJOBS udává počet paralelně prováděných úloh na testovacím serveru, hodnotou je kladné přirozené číslo.

Parametr FILESIZE udává velikost jednotlivých souborů, se kterými úloha pracuje, hodnotou je velikost dle syntaxe programu fio (např. 3T).

Celková velikost souborů, se kterými měření pracuje z jednoho testovacího serveru, tj. hodnota FILESIZE \* NUMJOBS, musí být větší než dvacetinásobek velikosti operační paměti každého testovacího serveru.

Celková velikost souborů, se kterými měření pracuje ze všech použitých testovacích serverů, tj. hodnota FILESIZE \* NUMJOBS \* počet\_testovacích\_serverů musí být větší než

* 400TB pro Úložiště SCRATCH
* 20TB pro Úložiště HOME
* 20TB pro Úložiště infrastruktury

##### Měření výkonu I/O operací

1. Měření dlouhodobě udržitelného výkonu I/O operací náhodného charakteru souborového úložiště se bude provádět nástrojem fio dle SPEC\_130.

Použije následující konfigurační soubor – fio jobfile.

[global]

rw=randrw

rwmixread=80

bs=4k

time\_based

runtime=1h

numjobs=NUMJOBS

;create\_only=1

[FILESYSTEM]

directory=/MOUNTPOINT/test/fio

filesize=FILESIZE

Parametry FILESYSTEM, MOUNTPOINT, NUMJOBS a FILESIZE se nahradí vhodnými hodnotami.

Parametr FILESYSTEM udává jméno úložiště, např. HOME nebo SCRATCH.

Parametr MOUNTPOINT udává mountpoint měřeného úložiště na klientech.

Parametr NUMJOBS udává počet paralelně prováděných úloh na testovacím serveru, hodnotou je kladné přirozené číslo.

Parametr FILESIZE udává velikost jednotlivých souborů, se kterými úloha pracuje, hodnotou je velikost dle syntaxe programu fio (např. 3T).

Celková velikost souborů, se kterými úloha pracuje z jednoho testovacího serveru, tj. hodnota FILESIZE \* NUMJOBS, musí být větší než dvacetinásobek velikosti operační paměti každého testovacího serveru.

Celková velikost souborů, se kterými měření pracuje ze všech použitých testovacích serverů, tj. hodnota FILESIZE \* NUMJOBS \* počet\_testovacích\_serverů musí být větší než

* 400TB pro Úložiště SCRATCH
* 20TB pro Úložiště HOME
* 20TB pro Úložiště infrastruktury

##### Měření výkonu metadata operací

1. Za účelem ověření stability provádění metadata operací souborového úložiště a zjištění výkonu metadata operací souborového úložiště se v rámci akceptačních testů provede měření dlouhodobě udržitelného výkonu metadata operací souborového úložiště.
2. Měření dlouhodobě udržitelného výkonu metadata operací souborového úložiště se bude provádět nástrojem mdtest https://github.com/hpc/ior.

Pro měření bude použit následující příkaz:

time \

mpirun -n $NPROC -f machinefile \  
mdtest \  
-C -T -r \

-F \  
-d /MOUNTPOINT/test/mdtest \  
-I $FILES\_PER\_DIR \  
-i $ITERATIONS \  
-u \  
-z $TREE\_DEPTH -b $BRANCHING\_FACTOR \  
-L

kde machinefile je soubor obsahující jména serverů určených pro provádění testu, každé jméno uvedeno na samostatném řádku,

NPROC, FILES\_PER\_DIR ITERATIONS, TREE\_DEPTH, BRANCHING\_FACTOR jsou proměnné – kladná, přirozená čísla, stanovena tak, že

* že hodnota FILES\_PER\_DIR je větší nebo rovna 100
* počet souborů vytvořených v každé iteraci je minimálně 10⁸, tj.
* doba běhu testu je minimálně 1 hodina.

Za výsledek měření se považují hodnoty z řádků „File creation“ a „File stat“, ve sloupci Mean, tabulky „SUMMARY rate“.

### Infrastrukturní servery

1. Infrastrukturní servery a jejich infrastruktura musí být navrženy a dimenzovány tak, aby zajistily spolehlivý, bezpečný, rychlý a efektivní provoz Malého clusteru a jeho správu. Některé požadavky na infrastrukturní služby jsou uvedeny v kapitole 5.13 Software.
2. Malý cluster musí obsahovat dvě kategorie Infrastrukturních serverů:
   1. Infrastrukturní servery pro běh služeb dodavatele – Infrastrukturní servery určené pro běh služeb dodavatele, které zajišťují chod Superpočítače a komplexní naplnění požadavků zadavatele. Servery jsou mimo jiné určeny pro běh služeb uvedených v kapitole 5.13 Software.
   2. Infrastrukturní servery pro běh služeb zadavatele - Infrastrukturní servery určené pro běh dalších služeb zadavatele, které implementuje a provozuje zadavatel (další služby poskytované uživatelům, integrace superpočítače do prostředí zadavatele a další).
3. Malý cluster musí obsahovat minimálně tři fyzické Infrastrukturní servery pro běh služeb dodavatele (kategorie 1), servery musí poskytovat celkově (v souhrnu) minimálně 384GiB RAM a 48 CPU jader.
4. Malý cluster musí obsahovat minimálně tři fyzické Infrastrukturní servery pro běh služeb zadavatele (kategorie 2), každý server musí splňovat následující požadavky:

* Minimálně 16 CPU jader
* Paměť RAM minimálně 192GiB, provozovaná v režimu DDR5 s ECC
* Minimálně 2 lokální SSD disky o kapacitě minimálně 960GB v RAID1
* Připojení Výpočetní síť, min 100Gb/s
* Připojení LAN síť, min 100Gb/s
* Klient úložišť HOME (/home), SCRATCH (/scratch) a infrastruktury (/apps), je vyžadován nativní klient a dostupnost nativní funkcionality těchto úložišť.

1. Pro řešení infrastrukturních služeb může být použita virtualizační technologie. Použití virtualizace však nesmí mít negativní dopad na kvalitu poskytovaných služeb.
2. Každý fyzický Infrastrukturní server musí splňovat následující požadavky:

* Disky v RAID s redundancí dat
* Za provozu vyměnitelné (hot-swap) disky
* Redundantní, za provozu vyměnitelné (hot-swap) napájecí zdroje, redundantní napájení
* Připojení LAN síť

1. Klíčové služby musí být provozovány v režimu vysoké dostupnosti, preferovaně použitím nativního mechanismu dané služby.
2. Výpadek nebo odstávka libovolného jednoho Infrastrukturního serveru nesmí způsobit přerušení probíhajících úloh a provozu Výpočetního clusteru, Přístupových serverů, Úložišť, Síťové infrastruktury a správy superpočítače.
3. (I) Dodavatel musí v nabídce uvést počet Infrastrukturních serverů, jejich určení a detailní konfiguraci.

### Zálohování

1. Součástí dodávky musí být komplexní řešení zálohování dat – Zálohování.

1. Řešení Zálohování musí zajištovat zejména:

* zálohování všech dodávaných serverů s výjimkou Výpočetních serverů
* zálohování Úložiště HOME; Úložiště SCRATCH se nezálohuje.
* zálohování Úložiště infrastruktury

Pro Výpočetní servery se předpokládá bootování či instalace/reinstalace/obnova z jednotných instalačních obrazů. Zálohují se obrazy centrálního úložiště instalačních obrazů včetně instalačních obrazů Výpočetních serverů.

Pro úložiště HOME se předpokládá obsazení 80% celkové kapacity úložiště daty typickými pro takové úložiště v daném použití (tedy HOME v prostředí HPC centra obdobné velikosti), předpokládá se změna 4% celkové kapacity úložiště denně.

Úložiště infrastruktury se zálohuje na úrovni souborového systému realizovaném na logických částech Úložiště infrastruktury. Pro Úložiště infrastruktury se předpokládá obsazení 80% celkové kapacity a změna 4% celkové kapacity úložiště denně.

Není nutno zálohovat data serverů dočasného charakteru, která nejsou potřebná k obnově serveru.

1. Zálohování dat musí probíhat s periodou 1 den. Kompletní cyklus zálohování všech dat musí proběhnout v čase od 22:00 do 6:00 (tj. během maximálně 8 hodin).
2. Řešení Zálohování musí umožňovat/poskytovat obnovu dat z posledních 21 dnů, z posledních 21 denních záloh.
3. Řešení Zálohování musí poskytovat dostatečnou zálohovací kapacitu pro naplnění všech uvedených požadavků (s ohledem na použitou technologii a reálně dosažitelné parametry v daném nasazení).
4. Priority zálohování a obnovy dat (pořadí dle priority od nejvyšší k nejnižší):
5. Nejvyšší priorita: Servery řešení nezbytné pro poskytování služeb systému, Úložiště infrastruktury
6. Další servery řešení
7. Nejnižší priorita: Úložiště HOME
8. Řešení pro zálohování a obnovu dat musí splňovat následující základní vlastnosti:

* časový harmonogram záloh
* obnova individuálních souborů a složek
* obnova vlastníků, práv a atributů souborů a složek
* paralelní běh záloh a obnov

1. Řešení zálohování musí používat technologii zálohování do diskového úložiště. Předpokládá se použití technologie pro lepší efektivitu uložení dat (např. komprese nebo deduplikace).
2. Zálohování musí používat vyhrazené, nezávislé úložiště. Úložiště určené pro zálohování musí splňovat požadavky uvedené v kapitole 5.8.1.2 Redundance disků.
3. Zálohování musí být řešeno tak, aby mělo minimální negativní dopad na provoz a výkon Malého clusteru.
4. Licence zálohovacího systému musí pokrývat všechny uvedené potřeby v maximální konfiguraci.
5. (I) Dodavatel musí v nabídce uvést zálohovací kapacitu a propustnost řešení, která je dostatečná pro naplnění uvedených požadavků na zálohování.
6. (I) Dodavatel musí v nabídce detailně popsat výpočet potřebné kapacity pro naplnění požadavků zadavatele na zálohování.
7. (I) Dodavatel musí v nabídce uvést architekturu řešení, typ a konfiguraci nabízených zařízení, popis softwarového řešení zálohování, jeho vlastnosti a navrhovanou zálohovací politiku.

### Síťová infrastruktura

1. Malý cluster musí obsahovat veškerou síťovou infrastrukturu (aktivní prvky, moduly, kabeláž, atd.) potřebnou pro realizaci superpočítače a pro připojení superpočítače do WAN sítě zadavatele a do storage sítě zadavatele.

#### LAN síť

1. LAN síť musí zajišťovat komunikaci mezi zařízeními uvnitř superpočítače a poskytovat připojení do WAN sítě zadavatele.
2. LAN síť musí obsahovat veřejné a privátní části sítě.
3. Veřejné části sítě (dále jen veřejné sítě), ve kterých budou použity veřejné IPv4 a IPv6 adresy, budou sloužit pro poskytování služeb dostupných z internetu (přístupové servery, některé infrastrukturní servery zadavatele).
4. Privátní části sítě (dále jen privátní sítě), ve kterých budou použity privátní IPv4 adresy, budou sloužit pro vnitřní služby a pro management zařízení.
5. Zařízením v privátních sítích musí být umožněn přístup do internetu přes NAT. NAT bude poskytován síťovým prvkem zadavatele umístěným ve WAN síti zadavatele.
6. Řešení sítě musí podporovat IPv6.
7. LAN síť musí být rozdělena do různých L3 sítí. Pro každou L3 síť musí být použita jiná L2 síť (VLAN nebo jiný aktivní prvek). Je nežádoucí používat jednu L2 síť pro více L3 sítí. Rozdělení sítí musí být 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.)

##### Připojení zařízení

1. Připojení Přístupových serverů a Infrastrukturních serverů do LAN sítě musí být redundantní.
2. Minimální rychlost připojení zařízení do LAN sítě:

Minimální rychlost připojení každého Přístupového serveru je 100Gb/s.

Minimální rychlost připojení každého Infrastrukturní serveru pro běh služeb dodavatele je 25Gb/s.

Minimální rychlost připojení každého Infrastrukturní serveru pro běh služeb zadavatele je 100Gb/s.

Minimální rychlost připojení každého Výpočetního serveru je 1Gb/s.

1. Minimální rychlosti dle SPEC\_167 musí být dosažitelné při přístupu z WAN sítě zadavatele.

Pro systémy uvedené v požadavku SPEC\_166 musí být požadavek minimální rychlosti splněn i v případě výpadku jednoho libovolného aktivního síťového prvku.

1. Agregovaná rychlost připojení všech Přístupových serverů v LAN síti dosažitelná při přístupu z WAN sítě zadavatele musí být minimálně 200Gb/s.
2. Agregovaná rychlost připojení všech Infrastrukturních serverů pro běh služeb zadavatele v LAN síti dosažitelná při přístupu z WAN sítě zadavatele musí být minimálně 200Gb/s.
3. Agregovaná rychlost připojení všech Výpočetních serverů v LAN síti dosažitelná při přístupu z WAN sítě zadavatele musí být minimálně 25Gb/s.
4. Agregovaná rychlost připojení všech Výpočetních serverů v LAN síti dosažitelná při přístupu z WAN sítě zadavatele musí být minimálně 25Gb/s.

##### Hraniční prvky

1. Propojení LAN sítě Malého clusteru s WAN sítí zadavatele musí být na straně Malého clusteru řešena redundantními aktivními síťovými prvky (dále označovanými jako hraniční prvky, v obrázku označeno HP1, HP2). Hraniční prvky musí splňovat následující vlastnosti:

* duální napájení
* redundantní zapojení serverů pomocí multi-chassis ether channel
* propustnost propojení mezi hraničními prvky musí být minimálně 200Gb/s
* oddělené routovací instance pro privátní a veřejné sítě

1. Hraniční prvky musí umožňovat restrikce datového provozu pomocí access control listů (ACL). Hraniční prvky musí umožňovat konfigurace ACL pro každý port zařízení zvlášť. Každý hraniční prvek musí umožňovat konfiguraci ACL o minimálním počtu 2 tisíce vstupních a 2 tisíce výstupních pravidel.
2. V případě odděleného control-plane hraničních prvků, hraniční prvky musí umožňovat použití technologií HSRP nebo VRRP, kde keepalive komunikace probíhá po IP adresách z rozdílného subnetu (privátního) než je plovoucí IP adresa z veřejného IP rozsahu.
3. Hraniční prvky musí umožňovat autentizaci uživatelů protokolem RADIUS nebo TACACS+, definici různých rolí při správě sítě (operátor, administrátor, atd.) a logování použitých příkazů.
4. Hraniční prvky musí umožňovat export i import konfigurace na/ze serveru pomocí protokolu TFTP, FTP, SCP nebo SFTP. Konfigurace musí být uložena ve tvaru, který umožňuje její editaci.

##### Aktivní síťové prvky

Požadavky uvedené v této kapitole jsou společné požadavky pro všechny aktivní síťové prvky včetně Hraničních prvků.

1. Aktivní síťové prvky musí být vzdáleně spravovatelné použitím zabezpečeného management rozhraní, s použitím silných šifer, silných klíčů a silných hashovacích algoritmů.
2. Aktivní síťové prvky musí poskytovat vzdálený přístup pomocí protokolu SSH2.
3. Aktivní síťové prvky musí podporovat standard 802.1Q (VLAN Tagging).
4. Aktivní síťové prvky musí umožňovat provoz VLAN v počtu minimálně 100, s možností číslování VLAN od 1 do 4094.
5. Aktivní síťové prvky musí podporovat protokoly IGMPv2 a IGMPv3.
6. Aktivní síťové prvky, které pracují na L3, musí podporovat IPv4 multicast routing a protokoly PIM Sparse Mode and PIM Source-Specific Multicast.
7. Aktivní síťové prvky musí umožňovat čtení údajů o stavu a vytížení portů protokoly SNMPv2 a SNMPv3. Musí umožňovat:

* možnost definice omezení přístupu do vybraných větví SNMP stromu pro specifikovanou komunitu
* zasílání SNMP trapů pro definované události

1. Ethernetová management rozhraní aktivních prvků sítě budou připojena do OOB sítě zadavatele realizovaného OOB switchem zadavatele. Propojení bude realizováno fyzickým propojem dodaného síťového prvku, kde budou fyzicky zapojena výše zmíněná management rozhraní. Takové propojení bude fungovat jako L2 trunk. Dodavatel může využít maximálně dvou portů OOB switche zadavatele. Dodavatel dodá potřebnou kabeláž a zajistí připojení.

Sériová management rozhraní hraničních prvků sítě budou připojena do OOB routeru zadavatele. Dodavatel zajistí kabeláž mezi patchpanelem zadavatele ve WAN racku a rackem s dodaným hardware.

Pro ethernetové i sériové propoje bude použit ethernetový kabel typu drát kategorie 5E nebo vyšší, s pláštěm LSOH.

1. (I) Dodavatel musí v nabídce uvést schéma zapojení LAN sítě, počty spojů a rychlosti připojení.
2. (I) Dodavatel musí v nabídce uvést detailní specifikaci Hraničních prvků – jméno výrobce zařízení, celé modelové označení zařízení, počet zařízení, označení feature licencí.
3. (I) Dodavatel musí v nabídce uvést specifikaci aktivních síťových prvků.

#### Adresace, jmenné služby

1. Správu veškerých adresních rozsahů, jmenných rozsahů, uživatelských účtů a skupin realizuje výhradně zadavatel.
2. Adresace IP sítí řešení úložiště musí být ve shodě s adresní politikou a adresním plánem zadavatele, použité adresní rozsahy (nejsou-li již definovány v jiné části zadávací dokumentace) budou stanoveny vždy v jednání se zadavatelem.
3. Zadavatel požaduje možnost ovlivnění konfigurace DNS služeb dodávaných dodavatelem a možnost integrace vlastních DNS služeb (pohledů).
4. Zařízení a systémy úložiště musí používat přesný čas. Pro synchronizaci času se budou používat NTP servery zadavatele v síti zadavatele.

#### Integrace do WAN sítě

1. Připojení do WAN sítě musí poskytovat konektivitu 4x100Gb/s.
2. Připojení do WAN prvků zadavatele musí být redundantní. Výpadek či odstávka libovolného jednoho hraničního prvku WAN sítě zadavatele nesmí způsobit nedostupnost služeb superpočítače.
3. Každý Hraniční prvek Malého clusteru musí být připojen na každý ze dvou hraničních prvků WAN sítě zadavatele pomocí jednoho spoje o rychlosti 100Gb/s. Pro připojení Malého clusteru do WAN sítě zadavatele jsou ve dvou hraničních prvcích WAN sítě zadavatele (CISCO Nexus 9336C-FX2) vyhrazeny celkem čtyři porty QSFP28, po dvou v každém hraničním prvku. Součástí dodávky jsou moduly a optické kabely potřebné pro připojení Hraničních prvků Malého clusteru do hraničních prvků WAN sítě zadavatele. Propojení musí využívat moduly kompatibilní s propojovanými zařízeními. Je požadováno natažení síťové kabeláže v podhledech datového sálu.

WAN síť zadavatele je popsána v kapitole 8.5 WAN síť, kde jsou rovněž znázorněny kabelové trasy.

1. Propojení Hraničních prvků Malého clusteru s hraničními prvky WAN sítě zadavatele bude realizováno přes Ethernet a privátní IPv4 adresy vyhrazené pro tyto účely zadavatelem.
2. (I) Dodavatel musí v nabídce uvést popis řešení integrace do WAN sítě.

#### Integrace do storage sítě

1. Pro realizaci integrace do storage sítě zadavatele dodávka musí obsahovat minimálně dvě Síťové brány. Síťové brány propojují Výpočetní síť Malého clusteru se storage sítí zadavatele, zajišťují konektivitu těchto dvou sítí, a tak síťově zpřístupňují serverům Malého clusteru úložiště zadavatele připojená do storage síti zadavatele
2. Každá Síťová brána musí splňovat následující požadavky:

* Fyzický server
* Architektura x86-64
* Operační systém Linux 64-bit
* Minimálně jeden procesor, celkem minimálně 16 CPU jader
* Paměť RAM minimálně 32GiB, DDR s ECC
* Paměť RAM musí být tvořena paměťovými moduly shodných parametrů
* Teoretická propustnost procesoru/procesorů do paměti RAM minimálně 160GB/s (v nabízené konfiguraci serveru)
* Minimálně 2 lokální SSD disky o kapacitě minimálně 240GB v RAID1
* Za provozu vyměnitelné (hot-swap) disky
* Redundantní, za provozu vyměnitelné (hot-swap) napájecí zdroje, redundantní napájení

1. Každá Síťová brána musí být připojena do Výpočetní sítě Malého Clusteru s celkovou, agregovanou rychlostí minimálně 100Gb/s.
2. Každá Síťová brána musí být připojena do storage sítě zadavatele s celkovou, agregovanou rychlostí minimálně 100Gb/s. Pro připojení Malého clusteru do storage sítě zadavatele jsou ve dvou hraničních prvcích storage sítě zadavatele (CISCO Nexus 9336C-FX2) vyhrazeny celkem dva porty QSFP28, po jednom v každém hraničním prvku. Součástí dodávky jsou moduly a optické kabely potřebné pro připojení do hraničních prvků storage sítě zadavatele. Propojení musí využívat moduly kompatibilní s propojovanými zařízeními. Je požadováno natažení síťové kabeláže v podhledech datového sálu.

Storage síť zadavatele je popsána v kapitole 8.6 Storage síť.

1. Konektivita mezi Výpočetní sítí Malého clusteru a storage sítí zadavatele musí být zajištěna i v případě výpadku či odstávky libovolné jedné Síťové brány.
2. Součástí dodávky je implementace softwarového řešení Síťových bran, v řešení bude použit software keepalived.
3. (I) Dodavatel musí v nabídce uvést popis řešení integrace do storage sítě zadavatele, počet a konfiguraci Síťových bran.

### Obecné technické požadavky

1. Paměti RAM všech serverů a řadičů diskových polí musí používat mechanismus detekce a opravy chyby – Error-correcting code memory (paměť ECC).
2. Běžný provoz a dostupnost deklarovaných kapacit zařízení nesmí vyžadovat zásah obsluhy.
3. Všechna zařízení a systémy musí být spravovatelné vzdáleně.
4. Servery musí mít vzdálený síťový management nezávislý na provozu operačního systému serveru poskytující ovládání napájení, reset, grafickou konzoli a připojení virtuálních médií.
5. Požaduje-li se podpora protokolu SSH, SCP a/nebo SFTP musí být podporována verze 2 protokolu a je požadováno použití silných šifer, silných klíčů a silných hashovacích algoritmů.
6. Všechna zařízení Malého clusteru musí být fyzicky označena jednoznačnou identifikací, která musí být na zařízeních snadno dostupná a čitelná, a vhodným způsobem evidována.

### Software

1. Součástí Malého clusteru musí být komplexní softwarové řešení provozu a správy Malého clusteru. Pro veškeré softwarové vybavení musí být zajištěna možnost bezplatného používání po neomezenou dobu.

#### OS a aplikace

1. Všechny servery musí používat operační systém Red Hat Enterprise Linux nebo Rocky Linux, verze 9 nebo novější. Servery určené výhradně na řešení Zálohování mohou používat operační systém Microsoft Windows Server verze 2022 nebo novější.
2. Všechny Výpočetní a Přístupové servery musí používat stejnou distribuci a stejnou verzi operačního systému.
3. Licence operačního systému musí pokrývat všechny dodávané servery.
4. Dodané operační systémy a softwarové vybavení musí mít dobrou dostupnost aktualizací, zejména těch, které jsou určeny ke zvýšení bezpečnosti, spolehlivosti, odstranění funkčních či výkonových nedostatků.
5. Součástí dodávky je poskytování a provádění aktualizací softwarového vybavení, podmínky jsou uvedeny ve smlouvě.
6. Dodávka musí obsahovat softwarové vybavení potřebné pro provedení akceptačních testů a pro možnost opakování těchto testů v záruční době (s výjimkou software poskytnutého zadavatelem). Zadavatel poskytne kompilátory a MPI/MKL knihovny Intel v poslední dostupné stabilní verzi.
7. Komplexní správu uživatelského aplikačního vybavení (tj. software pro výpočty), jeho získání, sestavení, instalaci, konfiguraci a nastavení bude zajišťovat zadavatel.
8. (I) Dodavatel musí v nabídce uvést řešení požadavků zadavatele na software, názvy a počet licencí navrhovaného software.

Zadavatel preferuje použití software:

* Pro správu a provádění konfigurací – configuration management software Ansible
* Pro správu verzí souborů – GIT

#### Prostředí

1. Malý cluster musí poskytovat koncovým uživatelům transparentní, jednotné, sdílené uživatelské prostředí.
2. Malý cluster musí poskytovat koncovým uživatelům v maximální míře shodné prostředí na všech uživatelsky dostupných serverech – tj. na Přístupových serverech a Výpočetních serverech.

Uživatelské prostředí na serverech stejného typu musí být shodné.

Uživatelské prostředí serverů různých typů musí být v maximální účelné míře shodné, rozdíly vyplývají pouze z rozdílů povahy a služeb serverů.

Servery musí poskytovat en\_US a C locale.

1. Přístupové servery musí uživatelům poskytovat koncovým uživatelům řádkové (command-line) rozhraní a grafické rozhraní.

#### Uživatelské účty

1. Malý cluster musí poskytovat koncovým uživatelům jednotný účet a jednotnou autentizaci. Účet koncového uživatele musí být dostupný a shodný na všech Přístupových a Výpočetních serverech.
2. Realizace uživatelských účtů musí 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. Zadavatel provádí centralizovanou správu všech uživatelů centra IT4Innovations, informace o uživatelích a skupinách Malého clusteru budou zadavatelem synchronizovány z centrální repository zadavatele do LDAP databáze Malého clusteru. Správu domovských adresářů uživatelů bude rovněž realizovat výhradně zadavatel.

LDAP služba musí být realizována dvojicí LDAP serverů s replikovanou databází. LDAP služba bude realizována na vhodných Infrastrukturních serverech, které nemusí být vyhrazeny pouze pro běh této služby. Pro řešení služby LDAP zadavatel preferuje software OpenLDAP.

Na serverech poskytujících uživatelské účty bude použita služba SSSD využívající LDAP službu zabezpečeným LDAP protokolem. Na Přístupových serverech bude použita SSSD LDAP enumerace.

1. Realizace jednotné autentizace uživatelů bude používat SSH klíče. SSH klíče budou spravovány zadavatele nebo uživatelem a budou uloženy v LDAP databázi nebo v domovském adresáři uživatele. Služba SSH musí umožňovat ověřování autorizovaných SSH klíčů uživatele proti LDAP záznamu uživatelského účtu, předpokládaná realizace je konfigurací vlastností sshd a sssd. Pro autentizaci se nepředpokládá použití hesel.

#### Plánovač

1. Malý cluster musí poskytovat jednotný přístup k výpočetním zdrojům. Malý cluster musí pro přístup k výpočetním zdrojům, pro vykonávání výpočetních úloh na Výpočetních serverech používat pokročilý plánovač úloh a manažer zdrojů (dále jen Plánovač).
2. Jako Plánovač Malého clusteru bude použit Slurm <https://slurm.schedmd.com/>.
3. Serverová část Plánovače musí být provozována na (vybraných) Infrastrukturních serverech. V případě virtuálních serverů servery musí být vyhrazené pro běh této služby, v případě fyzických serverů servery nemusí být vyhrazeny pouze pro běh této služby.
4. Dodavatel zajistí prvotní, základní instalaci a konfiguraci Plánovače, tak aby předvedl základní funkčnost clusteru a provedl akceptační testy. Není požadována HA konfigurace plánovače. Zadavatel si poté řešení Plánovače komplexně upraví a bude spravovat a upravovat dle svých představ a potřeb.
5. V případě, že jako plánovač bude dodavatelem použit standardní Slurm bez úprav a specifických rozšíření, dodavatel není povinen poskytovat podporu a aktualizace sw Plánovače. V ostatních případech dodavatel je povinen poskytovat podporu a aktualizace sw Plánovače. Zadavatel bude na clusteru provozovat Slurm v provedení, které nejvíce vyhovuje jeho komplexním potřebám.

#### Správa

1. Malý cluster musí obsahovat nástroje a řešení pro správu (management) všech systémů a služeb, které jsou součástí dodávky.
2. Malý cluster musí obsahovat vzdálenou správu a monitoring všech hardwarových zařízení (servery, disková pole, switche, atd.) poskytující zejména konfiguraci a ovládání zařízení, detekci závažných stavů a událostí a jejich oznamování prostřednictvím SMTP (email) nebo SNMP.
3. Malý cluster musí poskytovat vzdálené ovládání napájení a reset serverů bez závislosti na operačním systému ovládaného serveru (předpokládáme použití protokolu IPMI).

Funkcionalita musí být realizovatelná command-line nástrojem.

Funkcionalita musí být dostupná na vybraných Infrastrukturních serverech (určených pro správu).

1. Malý cluster musí poskytovat zobrazení a ovládání systémových konzolí serverů po síti v terminálovém režimu bez závislosti na operačním systému ovládaného serveru (předpokládáme použití protokolu IPMI).

Funkcionalita musí být realizovatelná command-line nástrojem.

Funkcionalita musí být dostupná na vybraných Infrastrukturních serverech (určených pro správu).

1. Malý cluster musí poskytovat efektivní centralizovanou vzdálenou správu operačních systémů serverů, a to jak jednotlivě (jeden server) tak hromadně (po skupinách serverů nebo všechny servery). Systém musí poskytovat zejména vzdálené vykonávání příkazů, přenos, modifikaci a odstraňování souborů, porovnávání výstupu příkazů a návratových hodnot příkazů. Systém musí umožňovat paralelní provádění akcí. Vzdálená správa musí efektivně pracovat i v případech, kdy některé spravované servery nekomunikují nebo nepracují korektně, musí identifikovat chyby prováděných akcí a oznamovat je.

Dodavatel zajistí vhodné pojmenování serverů a definici skupin serverů pro efektivní používání centralizované vzdálené správy.

Funkcionalita musí být realizovatelná command-line nástrojem.

Funkcionalita musí být dostupná na vybraných Infrastrukturních serverech (určených pro správu).

Pro řešení bude použit software ClusterShell https://github.com/cea-hpc/clustershell.

1. Funkcionalita popsaná v SPEC\_233 až SPEC\_235 musí být komplexně (a současně) dostupná na minimálně jednom Infrastrukturním serveru (určeném pro správu).

#### Vzdálená instalace a bootování

1. Malý cluster musí poskytovat efektivní vzdálenou instalaci serverů a bootování serverů po síti z centrálního úložiště obrazů. Řešení musí umožňovat instalaci/bootování jednoho serveru, tak hromadnou instalaci/bootování více serverů současně (paralelně). Řešení musí poskytovat správu (instalačních a bootovacích) obrazů (image) serverů - zejména vytváření, modifikaci a odstraňování obrazů. Serverová část řešení musí být provozována na (vybraných) Infrastrukturních serverech, které nemusí být vyhrazeny pouze pro běh této služby.
2. Řešení vzdálené instalace a bootování serverů dle SPEC\_237 musí být použito pro bootování Výpočetních serverů a pro instalaci Přístupových serverů. Pro servery shodného typu se použije shodný instalační obraz.
3. Centrální úložiště obrazů serverů dle SPEC\_237 musí pojmout obrazy o celkové velikosti minimálně 1TB.

Centrální úložiště obrazů může být realizováno na Úložišti infrastruktury.

1. Servery využívající bootování po síti musí po nastartování běžet z operační paměti, nesmí používat pro provoz operačního systému úložiště poskytované jiným serverem (tj. pro běh operačního systému se nesmí použít např. NFS share).
2. Realizace serverů dle SPEC\_238 musí probíhat ve dvou fázích:
   1. Instalace, resp. bootování serveru z centrálního úložiště obrazů za využití vybraného obrazu pro daný typ serveru. Obraz serveru obsahuje zejména aplikační balíky (package) potřebné pro chod serveru. Součástí obrazu je jen minimální konfigurace (zejména konfigurace sítě) tak, aby bylo možno realizovat následující fázi.
   2. Konfigurace systému serveru aplikací Ansible předpisů. Konfigurace systému bude popsána pomocí Ansible předpisů, které se budou aplikovat po instalaci (resp. bootu) operačního systému a dále dle potřeby. Ansible předpisy budou rozděleny do dvou skupin:
3. Ansible předpisy dodavatele – komplexní konfigurace serveru tak, aby byl zajištěn správný a efektivní chod serveru a naplnění požadavků zadavatele. Vytvoření a údržbu těchto předpisů zajišťuje dodavatel.
4. Ansible předpisy zadavatele – úprava konfigurace, nastavení prostředí a služeb zadavatele. Vytvoření a údržbu těchto předpisů zajišťuje zadavatel.

Na servery budou aplikovány aktuální verze Ansible předpisů, Ansible předpisy nejsou obsaženy v obrazech serverů. Ansible předpisy budou organizovány v GIT repozitáři.

1. Malý cluster musí poskytovat nastartování všech Výpočetních serverů (všech najednou) z vypnutého stavu do stavu, kdy je server schopen vykonávat úlohy, za méně než 15 minut.

#### Monitoring

1. Malý cluster musí obsahovat monitoring dostupnosti komponent a služeb systému. Monitoring musí poskytovat informace o dostupnosti a stavu relevantních služeb všech dodávaných serverů (včetně výpočetních serverů), uložišť, atd.

Pro řešení bude použit software Icinga2 provozovaný na vybraném Infrastrukturním serveru, který nemusí být vyhrazen pouze pro běh této služby.

Instalaci a konfiguraci Icinga2 zajistí dodavatel. Zadavatel doplní řešení sadou svých testů a začlení Icinga2 instanci do svého Icinga2 clusteru.

#### Logování

1. Systémy Malého clusteru musí zaznamenávat a centralizovaně uchovávat záznamy o aktivitách, činnostech, změnách stavu, událostech apod. (logování). Záznamy musí obsahovat časové razítko, identifikaci systému, služby, uživatele a popis události. Řešení musí poskytovat efektivní prohledávání velkého množství záznamů.

Záznamy (logy) musí být uchovávány po dobu minimálně 3 měsíců.

Řešení logování musí poskytovat kapacitu pro sbíraná data o celkové velikosti minimálně 1TB.

Řešení logování musí být provozováno na (vybraných) Infrastrukturních serverech, které nemusí být vyhrazeny pouze pro běh této služby.

#### Mail

1. Malý cluster musí veškerou mailovou komunikaci realizovat výhradně prostřednictvím SMTP serverů zadavatele.

Zadavatel poskytne SMTP servery.

#### Bezpečnost

IT4Innovations národní superpočítačové centrum je držitelem certifikátů systému managementu bezpečnosti informací podle normy ISO 27001 (ISO/IEC 27001:2013, ČSN ISO/IEC 27001:2014). Implementace Malého clusteru bude realizována v souladu s vnitřními předpisy zadavatele.

1. Malý cluster musí poskytovat přístup a služby pouze oprávněným uživatelům a systémům. Systém nesmí umožnit či poskytovat přístup a služby neoprávněným uživatelům a systémům. Systém musí být zabezpečen proti úniku dat, proti zneužití služeb, proti kompromitaci služeb a systémů.
2. Systémy a služby musí používat bezpečná, silná hesla, bezpečné klíče, bezpečné šifrování a protokoly. Je nepřípustné použití výchozích (default) hesel a klíčů a slabých hesel. Je nepřípustné používat stejné autentizační údaje pro různé účty či služby.
3. Řešení Malého clusteru musí poskytovat nastavení různých úrovní oprávnění uživatelů, tak aby byl poskytován přístup jen k nezbytným službám a datům. Systém musí zejména rozlišovat koncové uživatele, kteří využívají poskytované služby a mají přístup k některým (zejména vlastním) datům, a uživatele správy systému, kteří dohlížejí a spravují celý systém.
4. Služby, které nejsou pro provoz řešení a zajištění funkcionality potřeba, nebudou na serverech spouštěny/provozovány, nejlépe ani instalovány.
5. Malý cluster nesmí bez souhlasu zadavatele komunikovat s jinými systémy.

### Integrace do datového centra zadavatele

1. Dodavatel je povinen realizovat a zprovoznit Malý cluster v infrastruktuře zadavatele – v datovém centru IT4Innovations (tzv. integrace do datového centra zadavatele).
2. Integrací do datového centra zadavatele se rozumí veškeré dodávky a činnosti, jejichž výsledkem bude zprovoznění Malého clusteru v prostorách a infrastruktuře datového centra IT4Innovations.
3. Integrace systémů do datového centra zadavatele musí být provedena dle platné legislativy a předpisů a dle požadavků a doporučení výrobců jednotlivých systémů. Pro instalované systémy musí být provedeny revize požadované legislativou a předpisy.
4. Dodavatel je povinen respektovat infrastrukturu zadavatele. Infrastruktura zadavatele je popsána v kapitole 8  Infrastruktura zadavatele.

Nabídka nesmí používat či předpokládat parametry infrastruktury datového centra, které se liší od těch, které zadavatel specifikoval v zadávací dokumentaci.

1. Rozšíření infrastruktury zadavatele, která jsou nezbytná pro realizaci zakázky, musí být součástí dodávky.
2. Dodavatel nesmí měnit provozní parametry datového centra zadavatele. Nabídka nesmí předpokládat parametry datového centra zadavatele, které se různí od parametrů specifikovaných v zadávací dokumentaci.
3. Veškeré dodávky a činnosti řešení integrace do datového centra zadavatele musí být projednány se zadavatelem a zadavatelem schváleny. Veškeré použití a případné úpravy zařízení či technologií datového centra zadavatele (potřebné pro realizaci integrace do datového centra zadavatele) musí být projednány se zadavatelem a se smluvní společností zadavatele poskytující servis datového centra a schváleny zadavatelem. Zadavatel si vyhrazuje právo na změnu, úpravu technických návrhů integrace (vedení tras, přípojná místa, použité materiály, apod.) s ohledem na provozní podmínky a best-practices provozu datového centra.

#### Umístění

1. Umístění zařízení musí respektovat dispozice prostor datového centra zadavatele. Zadavatel explicitně upozorňuje na přítomnost sloupů v prostoru datového sálu.
2. Všechna dodávaná ICT zařízení musí být určena pro instalaci do racku a musí být nainstalována do racků, nebo musí jít o zařízení, jejichž konstrukce má charakter racku.
3. Zařízení musí být umístěna do racků dodaných dodavatelem.
4. Součástí dodávky je připojení racků potřebných pro realizaci na infrastrukturu zadavatele.
5. Dodavatel zajistí veškerou údržbu a servis dodaných racků po záruční dobu.
6. Dodávaná zařízení budou umístěna ve vymezeném prostoru vyznačeném na následujícím obrázku červenou linií, předpokládané standardní umístění racků je vyznačeno žlutou barvou a popisy „MC III“, světle zelenou barvou je vyznačen širší prostor, který je možno využít. Prostor se nachází v blízkosti stávajícího clusteru Barbora (SEQ, SRV1 a SRV2) a datového úložiště PROJECT (Project storage). Umístění může být předmětem změny ze strany zadavatele, nebo po dohodě se zadavatelem.



Obrázek 3 Umístění Malého clusteru

1. Vzdálenost racků od překážek (zdí, sloupů, řad dalších racků) a způsob instalace zařízení v racích musí umožňovat bezproblémový návoz, instalaci, výměnu a servis všech zařízení.
2. Umístění veškerého instalovaného zařízení dodaného řešení nesmí omezit možnosti revizních zásahů na elektroinstalaci, chladících rozvodech, vzduchotechnice a bezpečnostních systémech (čidla apod.).
3. Postavení racků na podlaze datového sálu musí být stabilní a bezpečné, hmotnost racků musí být vhodně rozložena. Musí být respektována konstrukce a únosnost podlahy (25kN/m2). V případě, že by rack působil zátěž bodově nebo na malou plochu, musí být použit roznášecí rám. Úpravy podlahových dlaždic (prostupy do prostoru zdvojené podlahy) musí být provedeny způsobem, kterým nedojde ke snížení únosnosti podlahy. Úpravy podlahových dlaždic musí být provedeny smluvní společností zadavatele poskytující servis datového centra.
4. Provedení kabeláže (uvnitř i vně racků) musí zajišťovat spolehlivé připojení, umožňovat manipulaci se zařízeními (např. při servisu zařízení) a bránit poškození kabelů při manipulaci. Prostupy do stropního podhledu datového sálu musí být osazeny průchodkami vhodnými do plechových stropnic. Kabeláž do storage sítě a do WAN racku bude vedena kabelovými žlaby.
5. Zařízení umístěná v racku, musí být orientována tak, aby všechna nasávala vzduch na stejné straně racku.
6. (I) Dodavatel musí v nabídce uvést schéma navrhovaného umístění zařízení Malého clusteru v datovém sálu datového centra zadavatele.

#### Napájení

1. Řešení napájení a provoz Malého clusteru musí respektovat omezení zadavatele, zejména parametry napájecích okruhů.
2. Malý cluster bude současně napájen z nezávislých napájecích větví datového centra (A a B).
3. Výpadek či odstávka napájecího okruhu nebo napájecí větve nesmí způsobit poškození žádného zařízení a nesmí způsobit ohrožení zdraví osob či majetku.
4. Výpadek či odstávka libovolného jednoho napájecího okruhu nebo napájecí větve nesmí způsobit:

* nedostupnost či výpadek služeb superpočítače s výjimkou služeb Výpočetních serverů
* nedostupnost více než 67 % Standardních výpočetních serverů

Po obnovení napájení napájecího okruhu nebo napájecí větve musí být automaticky zajištěna redundance napájení všech zařízení.

1. Pro řešení redundance napájení zařízení nesmí být použity ATS switche.
2. Systémy Malého clusteru musí rovnoměrně zatěžovat napájecí větve a jednotlivé fáze elektrického napájení.
3. Řešení musí počítat s maximálním provozním příkonem všech dodaných zařízení.
4. Malý cluster musí umožňovat vypnutí celého systému. Malý cluster musí umožnit korektní vypnutí celého systému nejdéle za 30 minut.
5. (I) Dodavatel musí v nabídce uvést energetickou kalkulaci. Energetická kalkulace musí obsahovat:

* maximální elektrický příkon celého řešení
* elektrický příkon pro každý osazený rack či samostatně umístěné zařízení

1. (I) Dodavatel musí v nabídce uvést schéma a parametry napojení Malého clusteru na elektrické rozvody datového centra.

#### Chlazení

1. Řešení chlazení a provoz Malého clusteru musí respektovat omezení zadavatele, zejména parametry chladících okruhů datového centra.
2. Řešení musí zajistit dostatečné chlazení všech dodaných zařízení.
3. Řešení musí zajistit chlazení (odvod) veškerého tepla uvolněného zařízeními Malého clusteru.
4. Chlazení Malého clusteru musí jako zdroj chladu používat výhradně chladících kapalinových okruhů (studené a teplé vody) datového sálu.

Vzduch datového sálu lze jako zdroj chladu (tj. bez jeho následného ochlazení) použít pouze v havarijní situaci, kdy dojde k výpadku zdroje chladu – chladícího kapalinového okruhu.

1. Chlazení Malého clusteru bude napájeno z nezávislých okruhů vodního chlazení datového centra.

Pro chlazení studenou vodou musí být použity okruhy studené vody SV1 a SV3.

Pro chlazení teplou vodou musí být použity okruhy teplé vody TV1 a TV2.

Uvedené chladící okruhy jsou v tomto dokumentu považovány za primární. Vnitřní chladicí okruhy zařízení dodavatele (např. CDU) jsou považovány za sekundární.

1. Řešení Malého clusteru musí poskytovat dlouhodobý provoz pod plnou zátěží při vstupní teplotě primárních chladících okruhů teplé vody (okruhů chlazení datového centra zadavatele) 38°C.
2. Chlazení výpočetních serverů musí využívat jako převažující zdroj chlazení chladicí okruhy teplé vody (TV1 a TV2).

Základní požadavek, jeho splnění je povinné: Minimálně 60% tepla generovaného výpočetními servery je chlazeno využitím chladících okruhů teplé vody.

Nadstandardní požadavek, jeho splnění není povinné: Minimálně 95% tepla generovaného výpočetními servery je chlazeno využitím chladících okruhů teplé vody.

Uvede-li dodavatel splnění nadstandardního požadavku, toto musí být prokazatelné.

1. Výpadek či odstávka chladícího okruhu nesmí způsobit poškození žádného zařízení a nesmí způsobit ohrožení zdraví osob či majetku.
2. Výpadek či odstávka libovolného jednoho chladícího okruhu datového sálu nesmí způsobit nedostupnost či výpadek služeb superpočítače.
3. V případě rizika překročení maximální provozní teploty zařízení či komponent zařízení (např. při výpadku chladícího okruhu), musí automaticky a včas proběhnout odstávka postižených zařízení tak, aby nedošlo k přehřátí nebo poškození zařízení či komponent.
4. V případě, že řešení chlazení nesplňuje nadstandardní požadavek na chlazení dle SPEC\_286, pak řešení chlazení musí využívat oba chladící okruhy studené vody S1 a S3 současně a zátěž musí být rozložena na tyto dva okruhy.
5. Každé připojení Malého clusteru k primárním chladicím okruhům musí umožňovat samostatné vzdáleně řízené přepínání mezi dvěma primárními chladicími okruhy. (Pro vyloučení pochybností - takové připojení lze použít pro chlazení více zařízení/racků.)

Přepínání okruhů nesmí způsobit nedostupnost či výpadek služeb či poškození žádného zařízení. Přepínání okruhů nesmí způsobit nežádoucí propojení primárních okruhů. Přepnutí mezi chladícími okruhy musí proběhnout v čase kratším než 3 minuty.

Vzdáleně řízené přepínání chladících okruhů musí být integrováno do systému Měření a regulace (MaR) zadavatele. Integraci provede dodavatel, na náklady dodavatele, ve spolupráci se zadavatelem. Řešení musí být kompatibilní s technologií MP-Bus, musí poskytovat rozhraní MP-Bus a fungovat jako zařízení MP-Bus Slave řízené MaR MP-Bus Masterem.

1. Úpravy rozdělovačů chlazení datového centra a připojení na rozdělovače chlazení datového centra musí být provedeny smluvní společností zadavatele poskytující servis datového centra na náklady dodavatele.
2. (I) Dodavatel musí v nabídce uvést řešení chlazení dodávaných technologických celků a způsob chlazení položek uvedených v energetické kalkulaci dle SPEC\_278, včetně uvedení jaká část příkonu (potažmo generovaného tepla) je chlazena teplou vodou a jaká studenou vodou.
3. (I) Dodavatel musí v nabídce uvést schéma a parametry napojení Malého clusteru na chladící okruhy datového centra.

#### Transport

1. Při transportu a instalaci zařízení nesmí být překročena únosnost podlahy přepravní trasy.

Únosnost podlahy datového sálu a jeho přístupové chodby (místnosti 219 a 223) je 25kN/m2.

Únosnost podlahy přístupových prostor (místnosti 217 a 218) je 5kN/m2. Tento úsek musí být pro transport materiálu dodavatelem dočasně ošetřen instalací ochranných desek pro rozklad zatížení (kupříkladu položením překližkových desek), tak aby výsledné zatížení podlahy v průběhu fyzické dodávky nepřekročilo 5kN/m2 a předešlo se tak poškození podlahy.

#### Prostředí

1. Při integraci a provozu musí být zajištěny tyto vlastnosti a parametry prostředí:
   * Teplota vzduchu 22 °C až 25 °C
   * Nekondenzující vlhkost
   * Relativní vlhkost 20-60%

## Implementace a další aktivity

### Implementace

1. Součástí dodávky musí být komplexní implementace Malého clusteru, tak aby byly splněny všechny požadavky zadavatele.
2. Součástí dodávky musí být návrh, doprava, instalace, zprovoznění, konfigurace, ladění, testování všech systémů a provedení akceptačních testů.

### Školení

1. Součástí dodávky musí být školení v rozsahu a detailu dostatečném pro získání a osvojení znalostí potřebných pro samostatné provozování a správu Malého clusteru.
2. Školení musí poskytnout informace potřebné pro pochopení vnitřního fungování systémů, funkčních celků, hardware a software. Školení musí zahrnovat důkladné seznámení s provozními postupy a správou systému.
3. Školení musí být v rozsahu minimálně 16 hodin, počet účastníku školení nesmí být omezen na méně než 16 účastníku.

Účastníky školení budou pracovníci IT zadavatele se zkušenostmi s IT problematikou na různé úrovni v různých oblastech.

1. Školení musí zahrnovat představení technologií, postupů a nástrojů pro správu systému, praktické ukázky a práci s reálným systémem. Praktické ukázky a práce s reálným systémem musí probíhat na Malém clusteru.
2. Školení musí vést erudovaní školitelé. Školení musí probíhat v českém nebo v anglickém jazyce.
3. Součástí dodávky musí být i poskytnutí výukových materiálů a prezentací v anglickém jazyce.
4. Časový harmonogram a detailní plán školení bude vypracován ve spolupráci s projektovým manažerem zadavatele.
5. Školení musí probíhat v místě zadavatele. Pro účely školení je možno bezplatně využít prostor zadavatele.
6. (I) Dodavatel musí v nabídce uvést základní harmonogram a obsah školení.

### Dokumentace

1. Součástí dodávky musí být vypracování a dodání komplexní dokumentace Malého clusteru.
2. Dokumentace musí komplexně pokrývat všechny dodávané systémy a musí být logicky navržena a strukturována. Součástí dokumentace musí být dokumentace skutečného stavu a dokumentace provozních postupů (provozní manuály).
3. Dokumentace musí pokrývat rovněž procesy a postupy dohledu systémů, pravidelné údržby systémů, řešení závažných stavů a obnovy.
4. Součástí dokumentace musí být též poskytnutí dokumentace dodávaného hardware a software (manuály) v anglickém jazyce. Dokumentace musí být poskytnuta v elektronické podobě umožňující kopírování textu.

### Prohlášení o shodě

1. Ke všem dodaným systémům a zařízením musí být doloženo prohlášení o shodě.

### Likvidace odpadů

1. Součástí dodávky musí být likvidace veškerých odpadů vzniklých realizací dodávky.

Zadavatel není povinen a nebude uchovávat obaly, obalový materiál.

## Požadavky na obsah Návrhu technického řešení

Návrh Malého clusteru musí obsahovat detailní popis architektury řešení, použitých technologií, funkcionalitu a vlastností řešení, uvedení konkrétního počtu a konkrétních konfigurací zařízení, konkrétního počtu licencí a konkrétních názvů software a způsob naplnění požadavků zadavatele.

Návrh Malého clusteru musí být poskytnut v elektronické formě umožňující kopírování textu.

Licenční podmínky budou vloženy do příslušné přílohy závazného vzoru smlouvy.

## Infrastruktura zadavatele

Systém Malý cluster bude instalován a provozován v datovém sálu datového centra zadavatele v budově IT4Innovations. Budova IT4Innovations se nachází v areálu kolejí Vysoké školy báňské – Technické univerzity Ostrava, na adrese Studentská 6231/1B, 708 00 Ostrava-Poruba.

V datovém sále jsou umístěny stávající systémy zadavatele, včetně superpočítačů Barbora a Karolina.

Datový sál není vybaven racky pro umístění systému Malý cluster.

### Dispozice sálu

Ve druhém nadzemním podlaží budovy IT4Innovations se nachází datový sál (místnost č. 223) o výměře 511,1m2, rozměrech 24,97 x 20,47m a světlé stavební výšce 4,5m. Sál je koncipován jako samostatný, stavebně nedělený prostor.

V datovém sále je instalována systémová antistatická podlaha rastru 600x600mm a výšky 900mm a dále těsný stropní podhled rastru 600x600mm, podvěšený 700mm pod stavebním stropem. Světlá výška mezi systémovou podlahou a stropním podhledem činí 2,9m.

#### Podlaha

Zdvojená podlaha výška 980 mm – systémové řešení. Zatížení 2500kg/m2 ( 25kN/m2).

Rámová konstrukce z ocelových "C" profilů v rastru 600x600 mm. Podpůrné sloupky tvořeny výškově rektifikovanými stojkami s montážně nastavitelnou hlavou. Sloupky a vodorovné rámy jsou propojeny pevným šroubovaným spojením.

Zakrytí je provedeno z panelů s jádrem na minerální bázi (třída reakce na oheň dle EN 13501 A2).

Spodní líc panelu s AL folií. Boky panelů jsou opatřeny plastovou hranou. Horní povrch s nalepeným PVC. Povrch v antistatické úpravě.

#### Stropní podhled

Pod stropem je podvěšený kovový podhled z galvanizované oceli 600x600x33mm, zkosená hrana o 3mm zaklapnutá to skryté konstrukce, hladký povrch bez perforace, UV stabilní elektrostaticky nanášený polyesterový lak. Revizní kazety doplněny klipy pro vyklápění kazet směrem dolů dle výkresu podhledů.

Skrytá závěsná kovová konstrukce s U-profilem a kolmým DP-profilem zavěšená pomocí závitových tyčí.

### Napájení

**Hlavní napájecí větev:**

Topologie: Dvě nezávislé napájecí větve, redundance 1+1

Rozvodná soustava NN: 400/ 230V;3+N+PE; 50Hz; TN-S

**Distribuční přípojnice:**

Počet distribučních přípojnic: 5ks pro každou napájecí větev

Jmenovitý proud distribučního systému: 800A

Napájecí soustava distribuční přípojnice: TN-S 3x400/230V 50Hz

Jmenovitý výkon distribuční přípojnice: 552kVA

Umístění distribučního systému : pod zdvojenou podlahou

Způsob napojení na distribuční systém: pomocí rozvaděče PDB

Specifikace vývodů rozvaděče PDB 6x3f/32A; možnost jiné konfigurace;   
jištění pojistkami s charakteristikou gG

Způsob připojení IT technologie: pomocí průmyslových zásuvek IEC60309

Přípojná místa jsou vybavena v konfiguraci 3f/32A a jsou osazena koncovkami PKY32G435. V případě potřeby rozdělení na 1f musí být toto realizováno dodavatelem v prostoru nad podlahou.



Obrázek 4 Rozvody napájení

### Chlazení

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

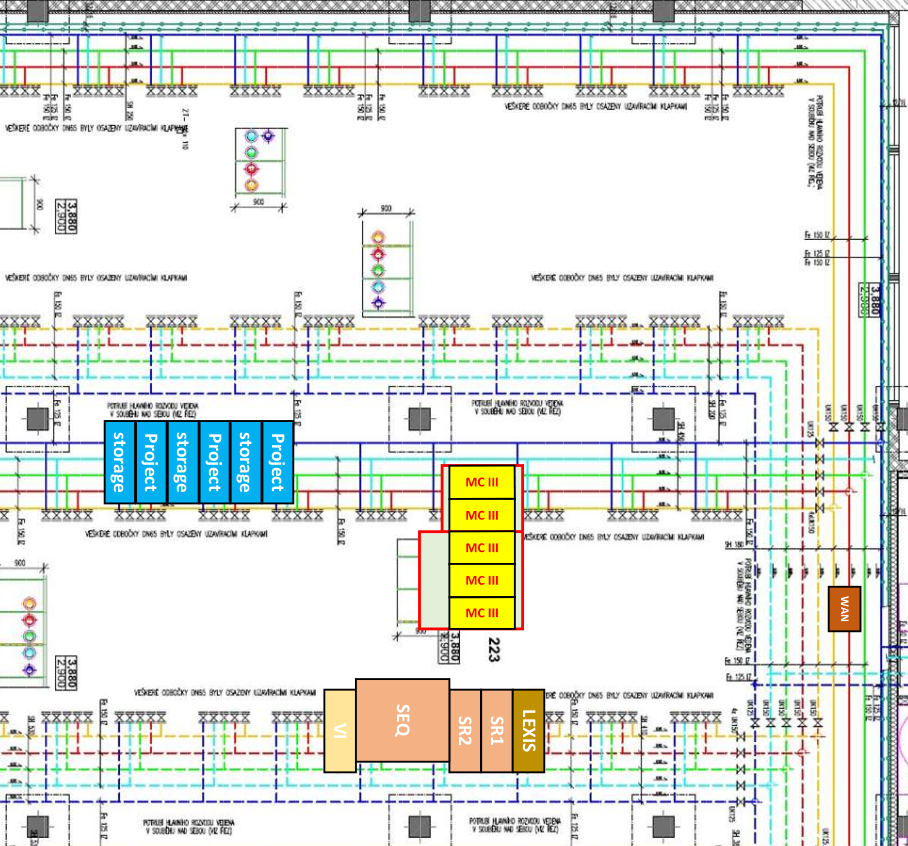
 SV2 – okruh Studené Vody – modrý:

* 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

Jedná se o stávající hodnoty parametrů chladících okruhů, zadavatel předpokládá, že dojde k navýšení teploty okruhů teplé vody na 38°C.



Obrázek 5 Rozvody chlazení

### Přístupová cesta do datového sálu

Návoz technologií do budovy IT4Innovations je možný ze severovýchodní strany objektu, kde je k tomuto účelu připravena rampa. Rampa je široká 2850 mm a vysoká 1030 mm. Pro transport je výhodné použití nákladních aut s hydraulickým čelem. Z prostoru rampy je vstup do budovy, resp. místnosti č. 218.

Vstupními dveřmi vedoucími do místnosti č. 218 je možné transportovat předměty o rozměrech 2410x1540 mm (výška x šířka).

Místnost č. 218 o rozměrech 5,3 x 5,6m je možné použít pro sejmutí transportních obalů či jako malý mezisklad v době transportu.

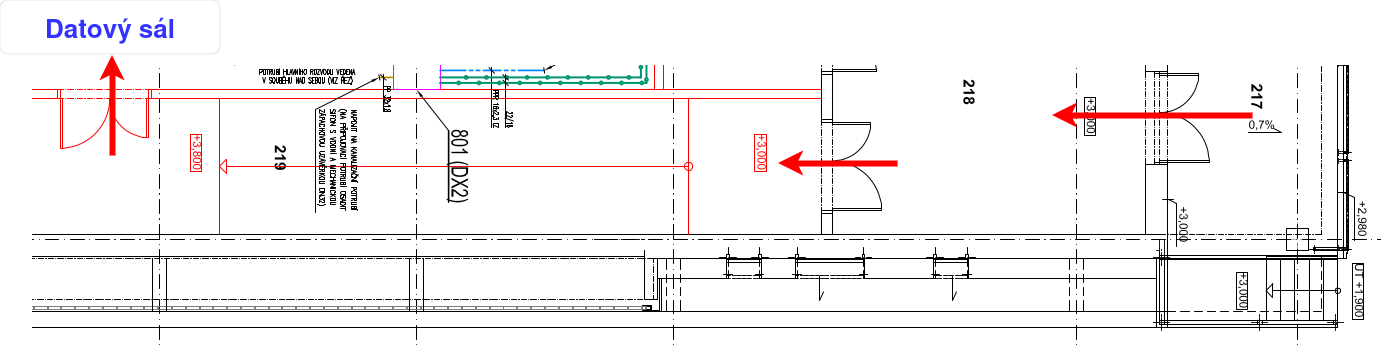
Z místnosti č. 218 vede přístupová chodba (místnost č. 219) k datovému sálu. Mezi místností č. 218 a místností č. 219 jsou dveře, jimiž je možné transportovat předměty o rozměrech 2340x1600 mm (výška x šířka). Část této chodby, konkrétně v délce 8,5m, je v provedení šikminy o sklonu 6,5°.

Vstup na datový sál (místnost č. 223) je v horní části přístupové chodby. Vstupními dveřmi je možné transportovat předměty o rozměrech 2360x1520 mm (výška x šířka).

Předmět, který se projde všemi třemi dveřmi, může mít max. rozměry 2340x1520mm (výška x šířka).

Pro transport předmětů v datovém sále je možné využít více variant transportních cest, žádná z nich však není překážkou pro předměty maximálních rozměrů danými dveřmi, viz výše.

Přípustné zatížení podlah v přístupové chodbě (místnost č.219) a na datovém sále (místnost č.223) je 2500 kg/m2 (25 kN/m2).



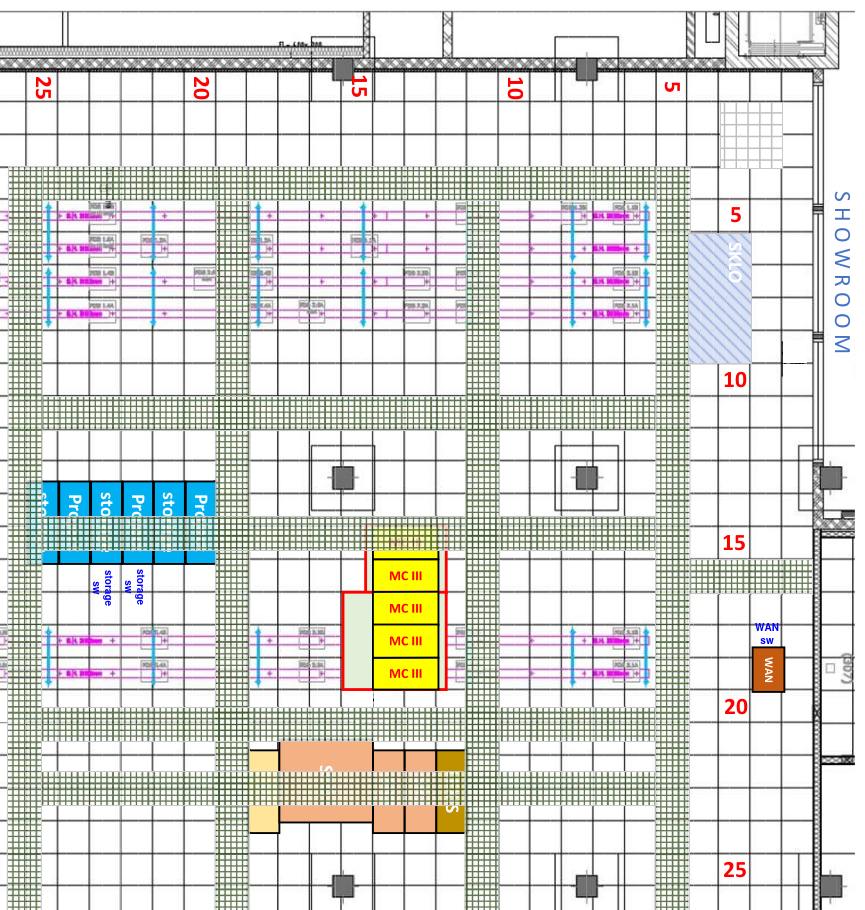
Obrázek 6 Přístupová cesta

### 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 zadavatele jsou umístěny ve WAN racku.

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



Obrázek 7 Kabelové trasy a síťové racky

### Storage síť

Storage sítí zadavatele je nazývána ethernetová síť, která připojuje centrální úložiště zadavatele a zpřístupňuje je (pomocí specializovaných síťových bran a připojením do WAN/LAN sítě) na výpočetní a další systémy zadavatele. Do storage sítě zadavatele jsou v době přípravy dokumentu připojeny čtyři bloky úložiště PROJECT.

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 zadavatele jsou umístěny ve dvou racích úložiště PROJECT, v horní části racku.

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

### OOB síť

Stávající OOB síť zadavatele je tvořena prvky:

* OOB router Cisco 2901/K9, konsolové moduly HWIC-16A (kabely typu CAB-HD8-ASYNC)
* OOB switch Cisco C3750X-48TS-S, porty 10/100/1000BaseTX

OOB zařízení jsou umístěna ve WAN racku.

## Přílohy

Příloha č. 1 Zdrojový kód benchmarku STREAM