**„****Cloudové a archivní úložiště“**

**Základní požadavky zadavatele na předmět veřejné zakázky**

**Obsah**

[Základní požadavky zadavatele na předmět veřejné zakázky 1](#_Toc1882243122)

[1. Úvod 1](#_Toc1484765431)

[2. Parametry zakázky 1](#_Toc758164252)

[3. Struktura předmětu zakázky 1](#_Toc1394436699)

[4. Účel a plánované užití úložišť 1](#_Toc47521190)

[5. Požadavky na cloudové úložiště 2](#_Toc2117908162)

[5.1. Obecné 2](#_Toc866392627)

[5.2. Komponenty 2](#_Toc1772403837)

[5.3. Parametry 3](#_Toc1532966742)

[5.4. Redundance, dostupnost 3](#_Toc498642275)

[5.5. Rozšiřitelnost a kontinuita řešení úložiště 3](#_Toc1039831024)

[5.6. Nezávislost na konkrétním dodavateli 3](#_Toc1534695858)

[5.7. Podpora operačních systémů a technologií 4](#_Toc1129955201)

[6. Požadavky na archivní úložiště 4](#_Toc2026076504)

[6.1. Obecné 4](#_Toc880120160)

[6.2. Komponenty 4](#_Toc576479536)

[6.3. Parametry 5](#_Toc1565241279)

[7. Zpřístupnění úložišť 5](#_Toc1207459048)

[8. Rozsah zakázky 5](#_Toc1711245524)

[9. Požadavky na záruku a servisní služby 6](#_Toc499138502)

[10. Infrastruktura zadavatele 7](#_Toc329074459)

[10.1. Datové centrum 7](#_Toc574539046)

## Úvod

Obsahem tohoto dokumentu jsou informace pro připravovanou veřejnou zakázku na pořízení cloudového a archivního úložiště a předběžné požadavky na vlastnosti, parametry a funkcionality těchto úložišť.

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

Pozornost prosím věnujte rovněž požadavkům a pokynům k Návrhu úložišť obsaženým v textu tohoto dokumentu. Požadavky jsou vyznačeny touto modrou barvou písma.

## Parametry zakázky

Předpokládaný termín realizace nákupu úložišť je **rok 2023**.

Předpokládaný finanční objem zakázky je **cca 20 mil. Kč bez DPH**.

Předpokládáme, že hodnotícím kritériem bude ekonomická výhodnost zakázky, kde jako hlavní kritérium bude zvolena kapacita úložišť při stanovení pevné ceny za předmět veřejné zakázky.

## Struktura předmětu zakázky

Předmět veřejné zakázky zahrnuje dvě hlavní části:

1. **cloudové úložiště** pro ukládání dat a
2. **archivní úložiště** pro archivaci dat

jejich vzájemnou integraci a jejich zprovoznění v prostředí a v prostorách zadavatele.

## Účel a plánované užití úložišť

Cloudové úložiště a archivní úložiště jsou pořizována se záměrem **rozšířit stávající portfolio úložišť** zadavatele, které tvoří zejména uživatelská souborová úložiště clusterů HOME a SCRATCH, centrální úložiště projektových dat PROJECT, disková pole virtualizační infrastruktury, disková pole řešení pro zálohování a další, o úložiště odlišných typů a charakteristik pro pokrytí nadcházejících potřeb zadavatele na ukládání a archivaci dat a pro poskytování a realizaci nových služeb zadavatele.

Cloudové úložiště a archivní úložiště budou provozovány na hardwarových prostředcích a softwaru ve vlastních prostorách zadavatele – on**-premises**.

Cloudové úložiště bude sloužit jako centrální úložiště dat superpočítačového centra IT4Innovations a bude poskytovat diskové zdroje nově budované cloud infrastruktuře (několika instancím), virtuálním serverům ve virtualizační infrastruktuře zadavatele, **k ukládání objektových dat** laboratoří a výzkumných projektů.

Archivní úložiště bude sloužit jako centrální úložiště dat superpočítačového centra IT4Innovations, **k dlouhodobé archivaci dat** laboratoří a výzkumných projektů.

## Požadavky na cloudové úložiště

### Obecné

Cloudové úložiště bude poskytovat:

* objektové přístup
* blokový přístup (bloková zařízení) a
* souborové služby (POSIX compliant).

Všechny uvedené služby/přístupy bude úložiště poskytovat (dynamicky, flexibilně) nad jednou množinou diskových zdrojů.

Předpokládáme řešení postavené na **technologii** **Ceph**. V aktuální fázi předběžné tržní konzultace není tento požadavek nastaven jako mandatorní a dodavatel může navrhnout řešení postavená na jiné softwarové technologii, vzhledem ke komplexu požadavků technického i netechnického charakteru však předjímáme, že úložiště bude postaveno právě na technologii Ceph, a již v tomto dokumentu předpokládáme vlastnosti a funkcionality dostupné v technologii Ceph.

V případě, že dodavatel navrhne řešení na jiné softwarové technologii než Ceph, v Návrhu uvede rovněž vzájemné porovnání technologií (objektivní – popisující silné i slabé stránky) a poskytne dokumentaci, nebo odkaz na dokumentaci navrženého software.

Cloudové úložiště bude poskytovat protokoly:

* Pro objektový přístup S3 a OpenStack Swift
* Pro blokový přístup nativní Ceph block device a iSCSI
* Pro souborové služby CephFS, alternativně NFSv4

Pro bloková zařízení požadujeme funkcionality thin provisioning, resize a snapshoty.

Pro souborové služby požadujeme funkcionalitu snapshotů.

Disková kapacita bude realizována pomocí

* kapacitních rotačních disků NL-SAS – „HDD oblast“
* SSD disků – „SSD oblast“

Zvážíme/předpokládáme použití SSD disků rovněž pro další specifické potřeby (WAL, DB, CephFS metadata).

SSD oblast cloudového úložiště bude určena zejména pro systémové image virtuálních serverů v cloudové infastruktuře (využití blokového přístupu).

HDD oblast cloudové úložiště bude určena pro ukládání dat, bude optimalizovaná na kapacitu (designem, použitými zařízeními, konfigurací), bude však rovněž disponovat dostatečným agregovaným výkonem.

Zadáním tedy je dosáhnout efektivním způsobem požadovanou kapacitu a přiměřeným způsobem využít potenciál instalovaných úložných zařízení/disků.

### Komponenty

Předpokládáme, že cloudové úložiště bude sestaveno z komoditních zařízení – x86 serverů, Ethernet síťových switchů, a že pro object storage servery budou použity x86 servery určené pro software-defined storage nasazení s větším počtem diskových slotů.

### Parametry

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

#### Kapacita

Požadovaná celková čistá kapacita úložiště (pro počáteční instalaci) je

* **minimálně 5PB pro HDD oblast**
* **minimálně 100TB pro SSD oblast**

Jde o kapacitu dostupnou z klientů úložiště.

Řešení musí umožňovat rozšíření úložiště na celkovou čistou kapacitu HDD oblasti **minimálně 10PB** navýšením počtu object storage serverů bez nutnosti nahrazování zařízení pořizovaných pro počáteční instalaci.

Určete kapacity vašeho návrhu.

#### Propustnost, výkon

Prosím, navrhněte/doporučte vhodné výkonové metriky pro charakterizaci zvažovaného úložiště a určete očekávané hodnoty pro váš návrh (kvalifikovaným odhadem).

Prosím, určete výkonové hodnoty vašeho návrhu (kvalifikovaným odhadem) pro blokový přístup úložiště:

* agregovaná rychlost sekvenčních operací (čtení, zápis) pro velikost bloku 1MiB
* výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 80/20

Propustnost a I/O výkon musí být dosažitelné z klientů úložiště a musí být dlouhodobě dosažitelné.

### Redundance, dostupnost

Cloudové úložiště musí poskytovat dostatečnou redundanci dat, spolehlivost a vysokou dostupnost.

Cloudové úložiště nesmí obsahovat komponentu, jejíž výpadek by způsobil nefunkčnost služeb úložiště – nesmí existovat „single point of failure“. Komponenty řešení úložiště – napájecí zdroje, switche, servery musí být redundantní a vyměnitelné za provozu bez výpadku služeb úložiště. Dostupnost služeb a dat musí být zajištěna rovněž v případě selhání libovolných dvou disků úložiště a při odstávce či výpadku libovolných dvou serverů.

### Rozšiřitelnost a kontinuita řešení úložiště

Řešení cloudového úložiště musí poskytovat rozšiřitelnost a kontinuitu úložiště – možnost rozvoje úložiště v čase. Požadujeme, aby počáteční řešení bylo možno (postupně) rozšiřovat, tak aby byla dosažena vyšší agregovaná kapacita a vyšší agregovaný výkon úložiště. První plánované rozšíření je uvedeno v kapitole 5.3.1 Kapacita.

Požadujeme, aby řešení úložiště bylo možno nejenom rozšiřovat o nové komponenty, ale aby bylo možné staré či jinak nevyhovující komponenty odstraňovat anebo nahrazovat novými.

### Nezávislost na konkrétním dodavateli

Požadujeme, aby rozšiřitelnost a kontinuita řešení cloudového úložiště nebyly závislé na komponentě/technologii (hardware či software) jednoho dodavatele/výrobce.

Řešení cloudového úložiště musí umožňovat rozšíření úložiště o další object storage servery použitím hardware jiného výrobce/dodavatele.

Řešení cloudového úložiště musí být postaveno na open-source software.

### Podpora operačních systémů a technologií

Cloudové úložiště (jeho klientská část) musí být podporováno na operačním systému OS RHEL (a ideálně jeho derivátech CENTOS, RockyLinux) verze 7, 8 a 9 na platformě x86-64. Cloudové úložiště musí být podporované a dobře integrovatelné v technologii OpenStack ve verzi alespoň Zed.

V návrhu prosím uveďte:

* Orientační schéma řešení
* Počty a typ/roli použitých serverů
* Předpokládanou konfiguraci serverů
* Počty, kapacity a typ/třídu disků v serveru a jejich určení
* Předpokládané užití SSD disků
* Design sítí úložiště
* Počty a parametry síťových switchů
* Síťové připojení serverů
* Softwarové řešení
* Redundanci a dostupnost dat, navrhovaný způsob zabezpečení dat (repliky, erasure coding)
* Předpokládané kapacitní a výkonové parametry

## Požadavky na archivní úložiště

### Obecné

Archivní úložiště je určeno pro archivaci dat – dlouhodobé uložení dat, přičemž zdrojem dat budou:

1. Cloudové úložiště – objektový přístup (S3)
2. Cloudové úložiště – souborové služby
3. Souborové úložiště projektových dat zadavatele PROJECT – souborový přístup NFSv4

Archivní úložiště bude komplexně realizovat funkcionalitu archivu – implementovat procesy a politiky archivace a poskytovat úložnou kapacitu pro archivaci dat. Součástí dodávky je softwarové řešení archivace.

Předpokládáme úzkou integraci archivního úložiště na cloudové úložiště – jakožto zdroj dat k archivaci. Předpokládáme, že operace s archivem budou nastavovat/provádět pověření uživatelé.

Očekáváme, že úložná kapacita bude realizována páskovou technologií LTO aktuální generace (tj. LTO8 nebo LTO9). Důvodem pro páskovou technologii je předpokládaná nižší pořizovací cena na jednotku kapacity a nižší náklady na provoz v porovnání s diskovou technologií a delší životnost.

### Komponenty

Předpokládáme, že archivní úložiště bude sestávat z páskové knihovny, či páskových knihoven, obslužných serverů a potřebných sítí. Ze síťového pohledu se předpokládá silné propojení na cloudové úložiště. Součástí dodávky jsou potřebná média – datové a čistící pásky.

### Parametry

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

#### Kapacita

Požadovaná celková kapacita úložiště – kapacita datových pásek je **10PB**.

Prosím, navrhněte technické řešení archivního úložiště a popište jeho vlastnosti a funkcionality.

V návrhu prosím uveďte:

* Orientační schéma řešení
* Navrhovaná zařízení a jejich počet a parametry
* Předpokládané kapacitní a výkonové parametry
* Navrhované softwarové řešení, jeho funkcionality a vlastnosti
* Způsob práce s archivem
* Možnosti organizace archivu
* Možnosti autentizace a autorizace uživatelů, přístupová práva
* Způsob integrace s cloudovým úložištěm

## Zpřístupnění úložišť

Služby úložišť budou systémům zadavatele poskytovány prostřednictvím LAN sítě zadavatele využívající technologii Ethernet.

Na straně LAN sítě zadavatele budou k dispozici dva přístupové Ethernet switche podporující multi-chassis etherchannel vybavené konektory typu QSFP28. Úložiště budou na přístupové switche zadavatele připojena redundantními linkami o rychlosti 100Gbps. Přístupové switche zadavatele nejsou osazeny transceivery, potřebné transceivery jsou předmětem dodávky, stejně tak jako potřebná kabeláž.

Prosím určete počet 100Gbps linek potřebných pro zapojení úložišť.

## Rozsah zakázky

Součástí zakázky bude:

* dodávka hardware řešení úložišť (infrastrukturní/management servery, I/O servery a další potřebné servery, pásková knihovna, datové, management a přístupové sítě, a další)
* dodávka software řešení úložišť včetně potřebných licencí
* dodávka infrastruktury pro instalaci a provoz úložišť v datovém centru zadavatele – racky a příslušenství potřebné pro umístění zařízení úložišť, řešení napájení a chlazení zařízení úložišť, napojení na infrastrukturu datového centra zadavatele (chlazení vodou)
* připojení úložišť do sítě zadavatele
* 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 softwarových upgradů po dobu záruční doby
* konzultace v průběhu provozu řešení

## Požadavky na záruku a servisní služby

Předpokládáme obdobné požadavky a formulace záručních podmínek jako u dříve realizovaných zakázek. Dále požadujeme, aby podmínky a finanční náklady na prodloužení užívání softwarových produktů a podpory softwarových produktů byly předvídatelné a akceptovatelné.

Předpokládáme následující parametry záruky a servisních služeb:

* záruční doba 5 let
* 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 Úložiště,
* Vada kategorie B – vada, která nebrání užívání Úložiště, ale podstatným způsobem jej omezuje, nebo která vytváří riziko znemožnění užívání Úložiště,
* 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 – cloudové úložiště:

* Vada kategorie A – do 1 pracovního dne od nahlášení vady
* Vada kategorie B – do 3 pracovních dnů od nahlášení vady
* Vada kategorie C – do 7 pracovních dnů od nahlášení vady

Odstranění vady – archivní úložiště:

* Vada kategorie A – do 5 pracovních dnů od nahlášení vady
* Vada kategorie B – do 10 pracovních dnů od nahlášení vady
* Vada kategorie C – do 20 pracovních dnů 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.

V návrhu prosím uveďte odhadované finanční náklady ve struktuře:

* Cloudové úložiště
  + Hardware
    - Servery (bez OSD disků)
    - Disky – HDD oblast
    - Disky – SSD oblast
    - Sítě
  + Software
  + Implementace
  + Podpora
* Archivní úložiště
  + Hardware
  + Software
  + Implementace
  + Podpora
* Racky a instalace v datovém centru zadavatele

## Infrastruktura zadavatele

### Datové centrum

IT4Innovations provozuje veškeré systémy ve svém datovém centru umístěném v budově IT4Innovations.

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

Chlazení pro datový sál poskytuje pět samostatných chladících kapalinových okruhů – dva okruhy teplé vody teploty (cca 30°C) a tři okruhy studené vody (cca 11,5°C), je použita směs 35% propylenglykol a 65% voda.

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