**Příloha č. 1 zadávací dokumentace**

Dodávka hierarchického datového úložiště – Ostrava

**Technická dokumentace a specifikace požadovaného plnění**

1. **Popis požadované sestavy datového úložiště**

Předmětem plnění veřejné zakázky je dodávka, instalace a zprovoznění (uvedení do řádného provozu) sestavy hierarchického datového úložiště (Hierarchical Storage Management, HSM; dále také jen „datové úložiště“), včetně potřebného SW, dalšího potřebného příslušenství a poskytnutí rozšířené záruky včetně technické podpory v místě instalace (popř. vzdáleně), a to v souladu s technickými požadavky uvedenými níže.  
Hlavními částmi požadovaného systému jsou:

* 1. Tier-1: diskové pole
  2. Tier-2: pásková knihovna
  3. servery pro řídící software (pro realizaci HSM), správu a zabezpečení provozu datového úložiště a servery uživatelského rozhraní (front-end)
  4. prvky síťové infrastruktury pro zajištění SAN a LAN komunikace
  5. řídící software a operační systémy nezbytné pro jeho provoz
  6. další potřebné příslušenství ke zprovoznění sestavy datového úložiště (kabely, adaptéry atd.)
  7. licence na všechny dodané programové produkty
  8. pomocné systémy: UPS, chlazení a zastřešení studené uličky

K veškeré funkcionalitě požadované v této zadávací dokumentaci musí v době podání nabídky existovat oficiální dokumentace příslušných komponent nabízeného řešení, kterou je uchazeč schopen na vyžádání zadavateli předložit, a která tuto funkcionalitu jednoznačně popisuje a prokazuje.

Není-li explicitně určeno jinak, všechny požadavky této zadávací dokumentace musí být splněny v jediné provozní konfiguraci, tj. současně.

1. **Předpokládané využití datového úložiště**

Předpokládáme využití úložiště v následujících hlavních kategoriích:

* 1. Úložiště pro zálohy uživatelů

Datové úložiště bude sloužit jako back-end pro již existující zálohovací SW uživatelů jako např. Tivoli TSM, Legato Networker a podobně. Zálohovací SW není předmětem této veřejné zakázky.

* 1. Souborový systém

Předpokládáme nasazení protokolů NFSv4, NFSv3, SMB 2.0, HTTP(S), SFTP, FTP(S), SCP a dále protokolu rsync. Všechny tyto protokoly budou provozovány v konfiguraci vysoké dostupnosti.

Nepožadujeme, aby kategorie 2.1 a 2.2 sdílely mezi sebou data.

Musí být možno úložiště logicky rozdělit na alespoň dvě logicky nezávislé části (s dedikovanými front-end servery, exkluzivním přiřazením LUNů diskových polí těmto serverům a virtuálních části páskové knihovny). Vyžaduje-li souborový systém servisní uzly (např. metadata servery), musí být pro každou logickou část samostatný.

Systém bude provozován jako dostupný po síti.

1. **Požadavky kladené na datové úložiště**

Pokud není uvedeno jinak, veškeré kapacity jsou uvedeny v dekadických násobcích, tj. 1TB = 1012B, 1PB = 1015B.

Využitelná kapacita diskových polí je definována jako kapacita po režii RAID, hot-spare disků či jiných případných režijních kapacit pole, ale před režií souborového systému.

Využitelná kapacita páskových knihoven je definována jako součet nativních kapacit (bez komprese) všech využitelných (umístěných v knihovnách v licencovaných slotech) dodávaných páskových médií.

V následujícím textu jsou použity následující zkratky a pojmy:

IB - Infiniband

FC - Fibre channel

1GE - 1Gbit Ethernet

10GE - 10 Gbit Ethernet

V textu je rozlišeno několik druhů příkonů sestavy, ty jsou vždy sázeny *pomocí kurzívy*, aby bylo zdůrazněno využití definice. Typy příkonů jsou následující:

*Peak příkon:* Příkon zařízení dosažitelný v řádu několika sekund. U diskových polí se jedná typicky o příkon při roztáčení pevných disků. Na tuto hodnotu je třeba dimenzovat elektrické rozvody. Nejedná se o krátkodobý příkon v řádu tisícin až desetin sekundy způsobený náběhem zdrojů.

*Maximální příkon:* Průměrný hodinový příkon zařízení při jeho plné zátěži. U diskových polí se plná zátěž měří spuštěním zátěžového testu (benchmarku) využívajícím všechny disky. U serverů je to pak příkon při spuštění několika benchmarků využívající všechny komponenty serveru (CPU, paměti, lokální disky, síť, SAN, ...). U páskové knihovny pak spotřeba při využití všech páskových jednotek. Na tuto hodnotu je potřeba mít dimenzované chlazení.

Všechny uváděné typy příkonů nesmí být při akceptaci, kdy budou zadavatelem měřeny, překročeny.

* 1. Datové úložiště typu HSM vychází z modelu hierarchicky uspořádaných úložných vrstev, které mají různou kapacitu úložného prostoru a zároveň poskytují vzájemně odlišný výkon. Úložiště bude obsahovat níže uvedené tiery (vrstvy):
     1. Tier-1: Diskové pole osazené disky minimálně o využitelné kapacitě 2700 TB. Nejméně 0,2 % z dodané kapacity disků musí být realizována prostřednictvím SSD technologie. Tato úložná vrstva může být případně sestavena i z více samostatných diskových polí, ta však musí být možné vhodným nástrojem (např. distribuovaným souborovým systémem) prezentovat front-end serverům jako jeden svazek. Není-li řečeno v textu jinak, veškeré požadavky na Tier-1 se týkají každého diskového pole v Tier-1.
     2. Tier-2: Pásková knihovna osazená páskami o minimální využitelné kapacitě 8 000 TB, kapacitou pásek rozumíme kapacitu bez komprese.

Součástí dodávky musí být také adekvátní počet čistících pásek. Pásková média musí být dodána s čárovým identifikačním kódem (labelem), musí také obsahovat čip umožňující provádět Media Lifecycle Management, potřebný software musí být součástí dodávky.

* + 1. Využitelná kapacita Tier-1 musí být větší než 1/3 využitelné kapacity Tier-2.
    2. Vedle úložných kapacit musí sestava obsahovat všechny servery nezbytné pro zajištění HSM funkcionality včetně zajištění chodu sestavy datového úložiště (např. správa knihovny apod. – dále jen obslužné servery).
  1. Vedle diskových polí a serverů požadujeme i odpovídající síťovou infrastrukturu včetně propojů diskových polí, páskové knihovny a řídících serverů. V případě použití FC nebo IB nebo 10GE pro propojení pole a front-end serverů musí být toto propojení realizováno přes switche, které jsou nutnou součástí dodávky. Každý switch propojovací infrastruktury musí po konečném zapojení všech prvků celé dodávky obsahovat navíc minimálně n/2 volných portů, kde n je počet obsazených portů, přitom se zaokrouhluje nahoru. Navíc počet volných portů v každém switchi musí být alespoň 6. (Např. je-li použito 7 10GE portů na daném switchi, celkový počet portů na tomto switchi musí být alespoň 13.) Dále musí platit, že na switchích, do kterých jsou připojena disková pole a páskové mechaniky, musí být v součtu nejméně 16 volných portů. Všechny obsazené i volné sloty musí být osazeny transceiverem a zalicencovány. Dodávka musí obsahovat kabeláž pro propojení jednotlivých částí úložiště, tato kabeláž nesmí být typu Direct Attach (Direct Attach kabelem rozumíme metalický kabel pro 10GE, který je zakončen XFP nebo SFP/SFP+, tj. připojuje se do XFP, SFP/SFP+ housingu. Tato kabeláž se vyskytuje pod různými názvy, např. 10GSFP+Cu, 10GBase-CR, 10GBase-CX1, 10GbE Cu SFP.). Zároveň je nutné dodat navíc 2 kabely každého typu jako rezervu. Optické kabely pro připojení do vnější sítě musí mít délku dostačující k propojení do rozvaděče – optické vany umístěné ve stejné serverovně (lze předpokládat, že nepřesáhne 20 metrů, při realizaci lze dodat délku na míru).
  2. HSM musí poskytovat souborový systém (dle normy POSIX) připojitelný na front-end servery, tento souborový systém musí být možno reexportovat ve smyslu sekce 2 této dokumentace. HSM musí umožnit vytvoření alespoň 100 souborových systémů (oddílů), každý oddíl musí být možno nezávisle na ostatních zvětšit (přidáním disků) a zmenšit (odebráním disků).
  3. Diskové pole v Tier-1 musí být zabezpečeno technologií RAID (či ekvivalentní technologií poskytující stejné či lepší zabezpečení, dále jen RAID) proti ztrátě dat při současném výpadku libovolných dvou rotačních disků. Počet rotačních disků (včetně paritních disků) v jedné takové RAID skupině nesmí být vyšší než 10. Přitom rebuild libovolné takto vytvořené RAID skupiny nesmí trvat déle než 48 hodin. Diskové pole v Tier-1 musí být zabezpečeno proti ztrátě dat při výpadku libovolného jednoho SSD disku. Pro SSD disky je požadována RAID konfigurace typu MIRROR. Požadovaný výkon musí být dosažitelný v této konfiguraci, měřen bude při stabilní konfiguraci pole (ne při rebuildu).
  4. Na každých započatých 30 rotačních disků v Tier-1 požadujeme alespoň 1 další samostatný hot-spare disk (např. při počtu 300 rotačních disků zapojených v RAID svazcích požadujeme navíc 10 hot-spare disků, celkem tedy 310 disků). Na každé započaté 4 SSD disky požadujeme alespoň 1 další samostatný SSD hot-spare disk.
  5. Pro použité SSD disky požadujeme DWPD 3.
  6. Diskové pole v Tier-1 musí podporovat a být vybaveno licencí na funkcionalitu LUN maskingu pro maximální možnou konfiguraci (maximální konfiguraci pole, maximální počet storage partitions). Je-li takto vypočtený počet licencí větší než 64, bude součástí dodávky 64 licencí.
  7. Write-back cache řadičů diskového pole musí být zabezpečena proti všem následujícím jevům: ztrátě dat, poškození dat při výpadku napájení (např. baterií) a poruše řadiče (např. zrcadlením cache redundantních řadičů). Požadovaný výkon musí být dosažitelný se zapnutou funkcionalitou zabezpečení dle tohoto bodu.
  8. Čistá velikost (po započtení případné režie redundance) Write-back cache řadičů diskového pole v Tier-1 musí být minimálně 8 GB.
  9. Pásková knihovna musí obsahovat dva páskové roboty (tj. zařízení pro přenášení pásek mezi sloty knihovny a páskovými mechanikami). Ovládání knihovny musí být redundantní (control path failover). Minimální počet mechanik je 12. Páskovou knihovnu musí být možno rozdělit na dvě logicky nezávislé knihovny (partitioning).
  10. Pásková knihovna nesmí zabrat podlahovou plochu větší než prostor sedmi racků šířky 80 cm.
  11. Pro zpřístupnění úložiště požadujeme minimálně 8 plně zastupitelných front-end serverů. Souborový systém poskytovaný HSM musí být možno sdílet pro čtení i zápis mezi všemi front-endy najednou. Zadavatel požaduje přístup s plnými administrátorskými právy na všechny front-end a obslužné servery. Front-end servery musí sloužit výhradně tomuto účelu (exportu svazků), nepřipouští se kombinace s jinou funkcionalitou (není např. možné používat front-end současně jako TSM či databázový server).
  12. Každý front-end musí mít připojení k Tier-1 úložiště technologií 8Gbps FC nebo 10GE nebo IB a to minimálně dvěma aktivními cestami (tj. 2x 8Gbps FC nebo 2x 10GE nebo 2x IB). Každý řadič pole Tier-1 musí mít alespoň 4 odpovídající porty.
  13. Všechny front-end servery musí být provozuschopné pod 64bitovým operačním systémem na bázi Unix (např. Linux, Solaris, AIX) na architektuře x86\_64.
  14. Pokud je na front-endech nutné provozovat jakýkoli komerční software, musí být všechny nutné licence pro všechny front-endy součástí dodávky (například operační systém). Jsou-li součástí dodávky instalace komerčních distribucí Linuxu (např. RHEL, SLES), požadujeme, aby v nich byly nakonfigurovány zdroje pro development balíky a všechny závislosti pro rekompilaci libovolného balíčku, který dodavatel OS poskytuje.
  15. Každý front-end server musí mít minimálně 256 GB RAM ECC.
  16. Všechny front-end servery musí být z hlediska operačního systému nakonfigurovány jako plně zastupitelné (High Availability, realizováno např. pomocí nástroje Pacemaker, RHCS, či dalšího obdobného nástroje). Předpokládáme, že HSM bude nabízet alespoň 4 nezávislé souborové systémy. Je-li High Availability režim realizován pomocí komerčního rozšíření základu operačního systému (např. Red Hat Enterprise Linux + HA add on), požadujeme všechny potřebné licence v rámci dodávky.
      1. Všechny front-end servery budou exportovat všechna HSM data v režimu active-active pomocí protokolů: SCP, FTP(S), CIFS/SAMBA. Export CIFS/SAMBA musí být nakonfigurován jako klastrový.
      2. Dále musí front-end servery exportovat HSM data protokolem NFSv4.0 se zapnutou Kerberos autentizací. Čtyři front-endy budou exportovat všechny HSM souborové systémy tímto protokolem tak, že žádné dva front-endy neexportují stejný HSM souborový systém. Pátý front-end bude nakonfigurován jako pasivní NFSv4.0 server. V případě výpadku libovolného z front-endů exportujícího NFSv4.0 musí pasivní front-end plně převzít veškerou NFSv4.0 funkcionalitu bez nutnosti administrativního zásahu jak na serveru, tak na klientech. Implementace NFSv4 pro klienta může být dodána dodavatelem.
      3. Funkcionalita všech požadavků v tomto bodě bude ověřena v akceptačních testech.
  17. Součástí dodávky musí být switche/routery pro připojení front-end serverů do vnější sítě a to: 4x typu 10GE, 2x typu 1GE. Pro každý tento typ připojení požadujeme 2 plně zastupitelné switche (fail-over). I pro tyto switche platí požadavek na volné porty z bodu 3.2. Každý front-end musí být přímo připojený do každého switche každým rozhraním. Každý tento switch bude připojen do hraničního routeru (uplink) pomocí 1x10Gbit a dále 1x10Gbit do vnitřní sítě. Všechny Ethernet switche musí splňovat požadavky z následující tabulky.

|  |
| --- |
| HW |
| Neblokující architektura. Možnost využití plné kapacity všech portů v obou směrech. |
| Redundantní hot-swap AC zdroje. |
| Hot-swap větráky. |
| Standardní rackové provedení. Switche budou instalovány do racků, které jsou součástí této dodávky. Switche musí být schopny trvalého provozu v systému teplé a studené uličky. |
|  |
|  |
| L2 funkcionalita |
| Velikost tabulek MAC adres: minimálně 20 000 záznamů. |
| Možnost použití minimálně 100 VLAN s číslováním od 1 do 4094. |
| Podpora rapid STP – MST podle 802.1s a 802.1w minimálně pro 16 instancí. Nezbytná je možnost filtrování BPDU, root guard a loop guard. |
| Podpora 802.1Q na všech portech. |
| Podpora jumbo rámců na všech portech minimálně 9000 bytů. |
| Možnost agregace až 8 portů do jednoho kanálu podle 802.3ad staticky i se signalizací LACP. Při použití LACP je nutné porty zablokovat pokud protější strana nepoužívá LACP také. Možnost agregace portů přes dva fyzické switche. |
| IGMP snooping v2 a v3. |
| MLD snooping v2. |
|  |
|  |
| L3 Funkcionalita |
| IPv4 i IPv6 unicast a multicast routing. |
| Velikost tabulek pro IPv4: minimálně 8 000 záznamů. |
| Velikost tabulek pro IPv6: minimálně 5 000 záznamů. |
| Plná podpora IPv4 i IPv6 protokolu. Nutná je podpora pro použití více směrovacích tabulek u obou protokolů. Minimálně je nutné oddělit datový provoz a management prvků. |
| Směrovací protokol BGPv4 pro IPv4 i IPv6. Nezbytná je možnost filtrování, nastavování parametrů (local-preference, metriky, komunity, ...) přijímaných i propagovaných prefixů podle IPv4/IPv6 adres, čísla AS a komunity. |
| Podpora čísel autonomních systémů (ASN) o velikosti 4 byte. |
| Podpora některého FHRP (First Hop Redundancy Protocol – HSRP, VRRP, GLBP, …) pro IPv4 i IPv6. |
| Podpora DHCP pro IPv4 i IPv6. Možnost přeposílání DHCP rámců do jiné IPv4/IPv6 sítě. |
| Podpora MTU na L3 rozhraních do velikosti minimálně 9000 bytů. |
| Možnost filtrování protékajícího IPv4 i IPv6 provozu na vstupu i na výstupu. |
| FHS (First Hop Security). Pro IPv4 minimálně DHCP snooping, Dynamic ARP inspection a IP Source Guard. Pro IPv6 minimálně Router Advertisement Guard a DHCPv6 guard. |
| IPv4 i IPv6 PIM. |
| IPv4 IGMP. |
| IPv6 MLD. |
| QoS umožňující upřednostnění určitého typu provozu, definice šířky pásma pro určité typy provozu a zajištění dostupnosti managementu i při zcela vytížených linkách. |
| Kontrola unicast RPF (reverse-path-forwarding) pro IPv4 i IPv6. |
|  |
|  |
| Management |
| Správa z příkazové řádky a vzdálená správa konfigurace přes grafické rozhraní bez nutnosti instalace zvláštního SW, se zabezpečeným přístupem (SSH, SSL, …) s možností definovat seznam IPv4/IPv6 adres, ze kterých bude povolen přístup. |
| Možnost správy přes lokální konzoli. |
| Podpora SNMP v2c i v3 s možností definice seznamu IP adres pro použití komunity nebo uživatelského jména. Přes SNMP musí být dostupné informace o systému a všech rozhraních. U rozhraní musí být dostupné informace o stavu rozhraní. Dále o přenesených bytech, přenesených paketech, zahozených paketech a chybovosti v obou směrech. |
| Možnost exportovat informace o přenesených datech (IPFIX, NetFlow v9 nebo vyšší, SFlow, …). |
| Možnost uložení konfigurace v editovatelné formě na server. Možnost načtení připravené nebo zazálohované konfigurace ze serveru. |
| Ukládání informací o událostech na vzdálený syslog server a lokálně do paměti nebo na lokální médium. |
| Možnost zrcadlení provozu lokálně i vzdáleně. |
| Ochrana proti přetížení procesoru nežádoucím provozem. |
| Podpora LLDP (Link Layer Discovery Protokolu). |

* 1. 10GE rozhraní switchů i serverů použité pro vnitřní propoje komponent úložiště musí být stejného optického typu, a to LR (long range). Porty pro uplink do hraničního routeru musí být typu LR. Požadované volné porty switchů musí být osazeny LR transceivery.
  2. Všechna datová (ne management porty) síťová Ethernet rozhraní front-end serverů musí podporovat jumbo rámce (alespoň 9000 bytů).
  3. Diskové pole v Tier-1 musí podporovat vytváření LUN o velikosti více než 16 TB.
  4. Systém úložiště (diskové a páskové kapacity, jejich řadiče, SAN infrastruktura, servisní stroje HSM a páskové knihovny a vzájemné propoje těchto komponent) musí být plně redundantní, výpadek jakékoliv jedné komponenty nesmí způsobit nedostupnost úložiště, může ale vést k dočasné degradaci výkonu. Tento požadavek se týká i jednoho samostatného front-end či obslužného serveru, kdy při jejich výpadku nesmí dojít k nedostupnosti funkcionality, kterou poskytují. Všechny typy serverů musí mít redundantní napájecí zdroje, disky a karty zajišťující SAN a LAN připojení (alespoň dual port, tj. alespoň dva porty stejného typu na jedné dvouportové kartě). Pásková knihovna, SAN a 10GE switche musí mít také redundantní napájení.
  5. Výměna jakékoliv části HW musí být možná za chodu (nesmí být nutné převést úložiště do stavu, kdy jsou některá data nedostupná). Upgrade SW (firmware) pole a páskové knihovny musí být možný taktéž za chodu. V případě poruchy jednoho z robotů páskové knihovny připouštíme následné plánované odstavení páskové knihovny na dobu nejvýše 60 minut.
  6. Z hlediska zajištění provozu musí být všechny prvky datového úložiště vybaveny managementem kontroly funkčnosti a provozních parametrů (teplota, napájení, …) a možností vzdálené správy. U všech dodaných serverů požadujeme možnost vzdáleného managementu včetně grafické konzole, možnosti využití virtuálních médií pro boot serverů a vzdáleného přístupu do BIOS/UEFI. Veškerý management musí být možný z prostředí OS Linux. BMC kontrolery serverů musí být připojeny samostatným kabelem, není možné sdílet fyzické porty s datovými rozhraními serverů.
  7. Všechny servery musí být stejného typu.
  8. Datové úložiště musí mít automatický systém hlášení poruch na bázi protokolu SNMP. Zprávy systému hlášení poruch musí být možno zpracovat na stroji s operačním systémem Linux, z těchto zpráv musí být rozpoznatelná chybující komponenta v lidsky čitelné podobě.
  9. Veškerý management HSM musí být ovladatelný ze stroje s operačním systémem Linux.
  10. HSM musí být schopno pojmout alespoň miliardu souborů.
  11. Exporty HSM z front-endů ve smyslu sekce 2 musí být realizovatelné přes všechna rozhraní, tj. 10GE, 1GE.
  12. Systém datového úložiště musí obsahovat nástroje pro zálohování dat v Tier-1, které zajistí v případě hardwarového nebo softwarového výpadku Tier-1 nebo kterékoliv jeho části (např. výpadek pole, volumu, rozpad souborového systému) plnou obnovu dat (dostupnost všech dat v systému) ve stavu maximálně 24 hodin před výpadkem.
  13. Není-li požadováno touto technickou dokumentací jinak, všechny dodané komponenty musí být osazeny v systému. Dále musí být zalicencovány, je-li k jejich provozu nutná nebo požadovaná licence.

1. **Požadavky na HSM**
   1. Systém pro správu migrací dat mezi tiery HSM musí správci umožňovat nastavit minimálně níže popsané typy migračních politik.
   2. Akce přesunu souborů mezi tiery mohou být nastaveny na základě:
      1. vzorů na cesty a jména souborů (včetně možnosti používat zástupné znaky alespoň pro „libovolný znak“ a to minimálně na začátku a konci vzoru a dále samotného znaku pro „libovolný znak“) nebo dle umístění souborů ve svazku (tj. pravidlo se týká všech souborů na svazku),
      2. času vzniku, poslední modifikace, posledního použití souborů,
      3. velikosti souborů,
      4. logické funkce vytvořené z pravidel 1. až 3. alespoň pomocí and,
      5. pravidla podle zaplnění jednotlivých vrstev (např. „začni přesouvat na pásky nejdéle nepoužité soubory, pokud zaplnění disků je větší než 70 procent“) a explicitního příkazu správce nebo autorizovaného uživatele.

Pro soubory odpovídající výše uvedeným pravidlům musí být možné provést alespoň následující akce:

* + 1. přesun na zvolený tier,
    2. vytvoření kopie na zvoleném tieru (např. soubor se zapíše na pásky, zpoždění zápisu na pásky dané technologií je tolerováno),
    3. řízení počtu kopií na archivním tieru, alespoň tři kopie. Musí být možná konfigurace, kdy se soubor zapisuje do archivního tieru na dvě samostatné sady pásek. Vytváření kopií nelze realizovat vytvořením snapshotu celého svazku.
  1. HSM musí poskytovat API nebo CLI nástroje (lépe obojí), které umožňují explicitně řídit migraci a zjišťovat aktuální umístění a stav souborů a adresářů. Je vhodné, aby statistické souhrny stavu migrací a shody s nastavenou politikou byly dostupné přes GUI.
  2. Systém nastavování pravidel musí poskytovat CLI rozhraní (například konfigurací uloženou v souboru dokumentovaného formátu), je vhodné, aby pravidla bylo možno nastavovat přes GUI.
  3. HSM software musí obsahovat funkcionalitu pro automatické periodické provádění kontrol konzistence dat na páskách. Kontrola dat nesmí vyžadovat migraci dat z pásek na diskové pole.
  4. Systém musí podporovat kvóty na velikost uživatelských dat na základě identifikace uživatelů a jejich skupin.
  5. Součástí dodávky je licence HSM software tak, aby pokryla veškerou dodanou kapacitu Tier-1 a dvojnásobek dodané kapacity Tier-2 (platí pro páskovou knihovnu).
  6. Je-li HSM software licencován na fyzickou dodanou kapacitu Tier-2 a neomezuje licenčně skutečné množství nekomprimovaných dat uložitelných (po kompresi) do Tier-2, považujeme požadavek „dvojnásobek dodané kapacity Tier-2“ za splněný dodáním licence na jednonásobek dodané kapacity Tier-2. Není-li HSM software licencován na kapacitu, je požadavek na dodání licence na dvojnásobek dodané kapacity Tier-2 logicky splněn.

1. **Výkonové požadavky**
   1. Výkony disků uveďte ve dvojkových násobcích, tj. 1MiB = 220B, 1TiB = 240B, výkony u páskové knihovny uveďte v desítkových násobcích, tj. 1MB = 106B, 1TB = 1012B.
   2. Rychlosti čtení a zápisu dat na datové úložiště musí být měřeny na identické sestavě, která je předmětem nabídky, a musí být v našich podmínkách reprodukovatelné. V případě, že nepůjde v nabídce deklarovaná čísla reprodukovat, dostane dodavatel možnost provést optimalizaci zařízení. V případě, že se ani tak nepodaří výkonu dosáhnout, vyhrazujeme si právo odstoupení od smlouvy.
   3. Výkonové požadavky pro Tier-1:
      1. Test bude prováděn ze všech dodaných front-end serverů. Velikost testovaného oddílu není omezena, musí být však umístěn na HSM. V případě využití více diskových polí je nutné test spouštět na takovém svazku či svazcích, který zahrnuje všechna disková pole najednou. SSD disky smí být použity pouze pro metadata svazku.
      2. Rychlost čtení a zápisu dat u disků na Tier-1 bude měřena nástrojem iozone pomocí příkazu:

**iozone -Mce –t50 -sMEMg -r512k -i0 -i1 -+mNODES cesta\_k\_souborům**  
MEM je velikost paměti jednoho front-end serveru vynásobená dvěma a NODES je soubor obsahující:

hostnames všech front-end serverů a cesty dle dokumentace programu iozone (popis volby -+m),

Rozložení souborů na jednotlivé servery, ze kterých budou prováděny testy, musí být rovnoměrné (případně vhodně zaokrouhleno, nelze-li rozložit soubory zcela rovnoměrně).

* + 1. Jako výsledek testu pro zápis respektive pro čtení je brána průměrná hodnota tří testů udaná výstupem programu iozone jako „Children see throughput for X initial writers“, respektive, „Children see throughput for X readers“.
    2. Požadované rychlosti pro čtení a zápisu jsou minimálně 6000 MiB/s. Program iozone používá jednotky v dvojkových násobcích (KiB, MiB) apod.
  1. Výkonové požadavky pro páskovou knihovnu v Tier-2:
     1. Souhrnná rychlost lineárního zápisu dat bez zapnutí komprese na všechny mechaniky páskové knihovny a stejně tak i souhrnná rychlost lineárního čtení ze všech mechanik páskové knihovny bez komprese musí dosáhnout alespoň počet\_dodaných\_mechanik krát 280MB/s.
  2. Výkonové požadavky pro front-end servery:
     1. Pro všechny front-end servery požadujeme minimální skóre získané aplikací SPEC2006 ve variantě int (integer), rate, no-autoparalel, base 350 bodů. Hodnota SPEC2006 pro servery musí být v nabídce uvedena.

1. **Požadavky na energetickou zátěž a UPS**
   1. Součástí dodávky úložiště je zapojení všech nezbytných silnoproudých rozvodů pro napájení všech dodaných technologií (jak samotného úložiště, tak všech pomocných systémů UPS a chlazení), a to do rozvodů připravených v serverovně (včetně připojení do rozvaděče). V této dokumentaci se o rozvodech, které jsou v serverovně připraveny a je možno je použít, hovoří v přítomném čase.
   2. V servervovně je k dispozici zdroj třífázového napětí 230 V z distribuční sítě, který je v případě výpadku do cca 30 sekund přepojen na napájení z diesel agregátu (dále jen „napájení z diesel agregátu“).

Takto zálohovaný přívod je vyveden do prvního pole el. rozvaděče v serverovně. Z tohoto pole je přívod veden do druhého pole el. rozvaděče k manuálnímu bypassu, ze kterého je napájení přivedeno do UPS. UPS je součástí dodávky a bude umístěna v serverovně. UPS musí podporovat plné odstavení/přemostění použitím manuálního bypassu bez výpadku zálohovaných okruhů. Připojení UPS z rozvaděče a do rozvaděče je součástí dodávky.

Výstup z této UPS bude přiveden zpět do druhého pole el. rozvaděče a dále veden do třetího pole el. rozvaděče, ze kterého budou napájeny další komponenty úložiště (zejména disková pole, servery, pásková knihovna a síťové prvky), jednozdrojové i dvouzdrojové.

Komponenty chlazení úložiště budou připojeny přímo na napájení z diesel agregátu (tj. budou připojeny „před“ UPS).

* 1. Z třetího pole el. rozvaděče je do prostoru pod každým rackem (umístění racků je naznačeno v doporučeném schématu využití serverovny v následující sekci) vyvedeno celkem 6 samostatně jištěných vývodů C16/1P, jednotlivé fáze jsou označeny přímo na zásuvkách. Vývody jsou ukončeny ve zdvojené podlaze pod racky, každý vývod je osazen dvěma dvojzásuvkami v provedení „na omítku“, tj. pro každý rack je k dispozici 24x zásuvek.
  2. V zálohovaném i nezálohovaném poli el. rozvaděče jsou instalovány rezervní 1F i 3F jističe a svorkovnice vč. místa na případnou dodatečnou instalaci kabeláže a případně i dalších jističů. Instalace dalších jističů vč. kabeláže, revize atd. bude součástí této dodávky, pokud to bude vyžadovat připojení technologií.
  3. Silnoproudé kabely budou vedeny pod dvojitou podlahou. Součástí dodávky je rovněž prořezání dlaždic zdvojené podlahy za účelem protažení kabeláže včetně případných nezbytných úprav pro zajištění stabilního umístění racků na takto upravenou podlahu (doplnění stojky pod dlaždice a podobně).
  4. Na připojení elektroinstalace nechá vybraný dodavatel svým nákladem zpracovat revizní zprávu, kterou předá zadavateli.
  5. *Maximální příkon* všech dodaných technologií (včetně systému chlazení a ztrát UPS) nesmí překročit 50 kW. *Peak příkon* všech dodaných technologií však může být po dobu maximálně 60 vteřin až 60 kW. Pokud sestava úložiště bude obsahovat takové technické prostředky, které zamezí vyššímu *peak příkonu* (např. nedovolení roztáčení všech disků v jeden okamžik), může být čistý součet *peak příkonů* dodaných zařízení vyšší, výše uvedené podmínky však musí být při provozu splněny.
  6. *Maximální příkon* každého jednotlivého racku datového úložiště nesmí překročit 15 kW. Zadavatel preferuje energeticky vyvážené rozmístění komponent do racků a rovnoměrné zatížení jednotlivých fází.
  7. V nabídce musí být uveden celkový deklarovaný *peak příkon* a *maximální příkon* sestavy.
  8. Součástí dodávky musí být online UPS s následujícími vlastnostmi.
     1. Rack-mount provedení, instalace v jednom dodaném racku. UPS musí být schopna trvalého provozu při chlazení v systému teplé a studené uličky.
     2. Redundance N+1, výpadek jednoho modulu nezpůsobí výpadek celé UPS.
     3. Třífázový vstup napětí 230 V/400 V o frekvenci 50 Hz, vstupů může být více. Vstupy UPS budou připojeny do připraveného rozvaděče v serverovně.
     4. Výstupní napětí 230 V/400 V, frekvence 50 Hz, třífázové.
     5. Možnost monitoringu přes SNMP.
     6. Výdrž baterií alespoň 10 minut při zátěži *maximálním příkonem* úložiště.
     7. Životnost baterií min. 5 let, bezúdržbové baterie. Možnost výměny bateriových modulů při provozu UPS bez výpadku zálohovaných systémů.
     8. Podpora automatického a manuálního bypassu. Manuální bypass je nainstalován v poli el. rozvaděče. UPS musí podporovat plné odstavení/přemostění použitím manuálního bypassu bez výpadku zálohovaných okruhů.
     9. UPS musí mít účinnost alespoň 95 % při zatížení v intervalu 50 až 90 %. Pokud UPS podporuje i účinnější provoz mimo režim dvojité konverze, musí být doba přepnutí dostatečně krátká pro zajištění nepřetržitého napájení spínaných zdrojů.
  9. Všechna zařízení musí být k elektrické síti připojena tak, aby platilo:
     1. Napájení musí být realizováno tak, že výpadek UPS nesmí způsobit výpadek datového úložiště či výpadek poskytované funkcionality (může dojít k degradaci výkonu).
     2. Výpadek napájení v době do spuštění diesel agregátu nezpůsobí výpadek zařízení či výpadek poskytované funkcionality (může dojít k degradaci výkonu).

1. **Prostorové, hmotnostní a hygienické požadavky, chlazení**

Sestava datového úložiště musí splňovat a respektovat následující omezení. Pozice v serverovně (nahoře, dole, vlevo, vpravo) jsou odkazovány ve vztahu k obrázku uvedenému na poslední straně této dokumentace.

* 1. Všechny racky sestavy datového úložiště a racky páskové knihovny mohou mít výšku nejvýše 42 U. Racky budou součástí dodávky. Racky musí být připraveny na umístění do systému studené a teplé uličky (viz schéma serverovny), zejména nesmí mít skleněné dveře a podobně. Všechny dodané racky budou identické. Neobsazené pozice v racích musí být osazeny záslepkami výšky 1U. S ohledem na provedení přívodů napájení k rackům se doporučuje uchazeči ponechat dostatek volného prostoru ve spodní části racku pro protažení silnoproudé kabeláže. Pokud nebudou dostačovat v serverovně předchystané zásuvky, může dodavatel k rozvodům napájení použít PDU do racku.
  2. Rozměry jednotlivých dále nedělitelných technologických dílů sestavy datového úložiště musí umožnit transport zařízení do serverovny takovým způsobem, který neporuší záruční podmínky výrobce těchto zařízení. Vstupní dveře serverovny mají šířku 110 cm a výšku 193 cm. Zadavatel doporučuje uchazečům využít možnosti prohlídky místa plnění pro přesné zaměření a prohlídku přístupových tras k serverovně.
  3. Plošná nosnost podlahy v serverovně pod racky je 1500 kg/m2. Podlaha serverovny je zdvojená.
  4. Dispozice serverovny a možné umístění komponent je zachyceno na obrázku na konci této dokumentace. Řada v pravé části obrázku umožňuje umístění nejvýše sedmi racků šířky 80 cm. Řada v levé části obrázku je určena pro umístění až pěti racků, mezi nimiž se nachází až čtyři (dle počtu potřebného pro dodržení podmínek dokumentace pro chlazení) mezirackové (in-row) chladící jednotky, které jsou součástí požadované dodávky. Zadavatel preferuje uspořádání, kdy se v dolní části (nejblíže rozvaděči) bude nacházet rack s dodanými UPS, v dalších čtyřech racích budou ostatní komponenty úložiště, tj. zejména disková pole a servery, a v řadě sedmi racků se bude nacházet pásková knihovna.
  5. Umístění zařízení musí dovolovat stabilní trvalý provoz dodaného zařízení. Dále musí respektovat nezbytnost zachování únikových tras ze sálu naznačenými dveřmi. Mezi řadami racků se nachází studená ulička.
  6. Umístění komponent v serverovně musí respektovat polohu vyznačené studené uličky, zejména zachovat linie předních hran obou řad racků. Celá sestava úložiště a pomocných technologií smí v serverovně zabrat nejvýše prostor vyznačený ve schématu plochou racků a chladících jednotek.
  7. Součástí nabídky musí být předběžné schéma rozmístění komponent. Detailní umístění komponent bude nicméně upřesněno před realizací dohodou zadavatele a vybraného uchazeče.
  8. Součástí dodávky je zastřešení studené uličky splňující následující podmínky:
     1. Shora bude ulička překryta vhodně uchycenými průsvitnými nebo průhlednými polykarbonátovými deskami.
     2. Vstupy do uličky z obou stran budou realizovány buď posuvnými dveřmi v kovovém rámu osazenými polykarbonátovou nebo skleněnou průhlednou výplní.
     3. Racky budou v neosazených pozicích vybaveny záslepkami proti míchání teplého a studeného vzduchu.
     4. Pokud pro umístění komponent úložiště bude použito méně racků nebo kratší pásková knihovna, než předpokládá schéma z bodu 7.4, budou součástí dodávky další prázdné racky v počtu, který postačí k doplnění studené uličky do délky nejméně 520 cm (tj. 6,5 racku[[1]](#footnote-2)). Prázdné racky budou použity k dobudování obou stran studené uličky. Je-li to nezbytné pro zachování manipulačních prostorů z boku páskové knihovny, dodavatel navrhne vhodné technické řešení pro uzavření studené uličky mezi páskovou knihovnou a sousedícím rackem či dveřmi uličky.
  9. Součástí dodávky bude systém chlazení úložiště splňující následující podmínky:
     1. Pro chlazení prostor jsou požadovány přesné mezirackové chladicí jednotky plně vhodné do datových center. Mezirackové jednotky musí plně zajišťovat chlazení vzduchu z teplé uličky datového centra a jeho distribuce do uzavřené chladné uličky před IT zařízení. Jednotky musí být dodány vč. venkovních částí systému, zejména kompresoru, výparníku a dalšího příslušenství, které jsou potřeba pro jejich provoz. Systém chlazení musí tvořit funkční celek.
     2. Mezirackové chladící jednotky vč. veškerého příslušenství musí být kompatibilní s použitými datovými racky a systémem uzavřené studené uličky.
     3. Systém chlazení musí být schopen trvale uchladit tepelný výkon všech do serverovny dodaných technologií, tj. zejména samotného úložiště při *maximálním příkonu a* UPS, a dále mít rezervu pro dalších 10 kW instalovaných do sálu. Na tuto hodnotu musí být konfigurován v redundanci N+1.
     4. Komponenty systému chlazení budou napájeny 400V AC 3f nebo 230V AC 1f.
     5. Chladícím médiem je směs HFC typu, náplně média budou součástí dodávky.
     6. Všechny komponenty nutné pro správný běh systému chlazení musí být odolné proti výpadku napájení, tzn. po obnovení napájení po výpadku se musí automaticky spustit do plně provozního stavu.
     7. Technické parametry každé mezirackové chladící jednotky musí být následující:
        + výška jednotky musí být ve standardních rozměrech rackových skříní 2000 mm, hloubka max. 1200 mm a šířka max. 400 mm a to z důvodu využití maximálního prostoru v místě instalace,
        + ovládací displej každé chladicí jednotky nebo min. dvou jednotek kompletu,
        + přístup ke správě a monitoringu prostřednictvím protokolu SNMP a to i z nástrojů monitoringu zadavatele,
        + čidla vlhkosti i teploty v přední i zadní části jednotky integrovaná do správy a přístupné prostřednictvím SNMP protokolu,
        + jednotka musí umět úpravu vlhkosti vzduchu dle nastavení uživatelem (zvlhčování/odvlhčování),
        + připojení na přívod vody (je součástí této dodávky) a možnost osazení čerpadlem pro odvod kondenzátu,
        + za provozu vyměnitelné EC ventilátory,
        + kompresor jednotky bude umístěn mimo datový sál z důvodu minimalizace vibrací.
     8. Zadavatel požaduje v nabídce uvést položkový technický rozpis dodávaného zařízení, např. vnitřní a vnější klimatizační jednotky a další komponenty dodávaného systému vč. veškeré projektové i jiné dokumentace.
     9. Funkční zapojení dodaných klimatizačních jednotek, oživení a nastavení klimatizačních jednotek a zaškolení obsluhy v nutném rozsahu provede vybraný dodavatel.
     10. Instalace kompletu musí být zajištěna vč. všech dalších navazujících prací, jako např. stropní průrazy pro vedení potrubí, požární ucpávky, zpracování požárně-bezpečnostního řešení, omítky, štuky, výmalby, úklid apod.

1. **Akceptační testy**

Po dodávce a instalaci datového úložiště požaduje zadavatel v rámci zkušebního provozu provést akceptační testy (viz článek 3.6 zadávací dokumentace). Tyto testy budou minimálně zahrnovat:

1. Ověření funkcí a vlastností dodaných zařízení a komponent v souladu s deklarovanými parametry v nabídce vybraného dodavatele;
2. Ověření funkčnosti managementu SW, komunikačních protokolů a přístupových rozhraní. Detailní popis akceptačních testů funkčnosti přístupových protokolů je uveden níže.
3. Výkonové testy podle specifikace v části 5. V rámci akceptačních testů budou požadované rychlosti zápisu/čtení páskové knihovny dle bodu 5.4.1 kontrolovány oproti dokumentaci dodané konfigurace páskové knihovny.

Test active-passive režimu front-end serverů proběhne následovně.

Předpokládáme, že klient má připojen souborový systém přes NFSv4.0 z front-endu A. Během akceptačního testu musí úspěšně proběhnout následující testy:

1. Z klienta se na svazek připojený přes NFSv4 rozbalí archiv zdrojových souborů linuxového jádra verze 3.x čtyřikrát (do čtyř samostatných adresářů).
2. Na klientu bude spuštěn iozone test podobně jako při měření výkonu na plně sestaveném a funkčním úložišti, bude spouštěn opakovaně v nekonečné smyčce.
3. Z front-endu A budou odpojeny všechny napájecí kabely (IPMI v tomto případě bude nedostupné a je tedy nezbytné zajistit funkcionalitu failoveru i takto, např. pomocí sekundárního fencingu na switchích). Funkci front-endu A musí automaticky (bez administrativního zásahu) převzít pasivní front-end. Po převzetí funkce pasivním front-endem musí klient pokračovat v měření výkonu bez administrativního zásahu (tj. nepřipouští se ruční rekonfigurace serveru ani klienta).
4. Funkčnost High Availability konfigurace bude dále ověřena následovně a v každém případě musí být klient schopen připojit a používat NFSv4.0 export bez přerušení a zásahu administrátora
   1. vypnutí libovolného jednoho switche
   2. shození libovolného síťového rozhraní na front-endu (ifdown ethX, ibX)
   3. odpojení libovolného jednoho FC, IB, GE, 10GE kabelu
   4. reboot front-endu používaného klientem
   5. power off front-endu používaného klientem (přes IPMI rozhraní)
   6. power off libovolného služebního serveru (např. metadatového serveru souborového systému).

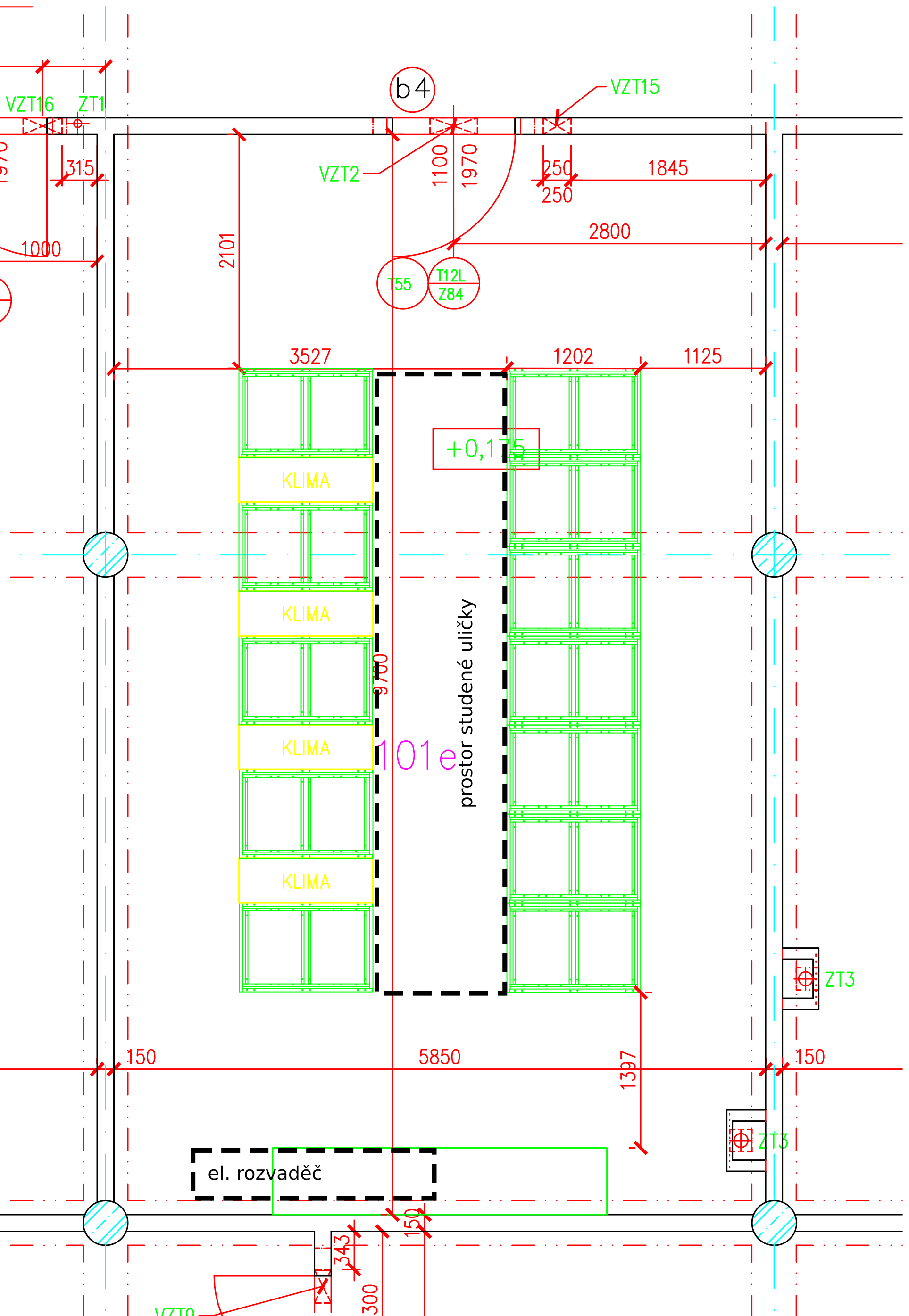
Ve všech těchto případech musí běžící test dle popisu v bodu 1. pokračovat v činnosti, tj. nesmí skončit chybou. Nebyl-li klient připojen v okamžiku provedení akcí 4.1 až 4.6, musí být schopen se připojit na funkční server a pracovat.

1. Pro NFSv4.0 export musí dále uspět test zámků dostupný na:  
   <http://nfsv4.bullopensource.org/tools/tests/locktest.php>  
   a to jak v rámci jednoho klienta, tak i pro více klientů (test je k dispozici zde: <http://nfsv4.bullopensource.org/tools/tests_tools/locktests-net.tar.gz>).

Test bude spuštěn na klientovi, který má připojený NFSv4.0 svazek z některého z front-endů lokálně do /mnt/nfs4. Spouštěný příkaz bude:  
./locktests -n 10 -f /mnt/nfs4/lockfile.  
Žádná část testu nesmí skončit chybou.

1. Síťový test bude spuštěn na dvou klientech, kteří mají připojený stejný NFSv4.0 svazek z některého z front-endů lokálně do /mnt/nfs4. Spouštěný příkaz bude:  
   Klient A: ./locktests -n 10 -f /mnt/nfs4/lockfile -c 1  
   Klient B: ./locktests --server Klient\_A\_FQDN  
   Žádná část testu nesmí skončit chybou.

Celá sekvence testů musí být provedena bezprostředně za sebou. Přitom se do průběhu testů nesmí administrativně zasahovat (např. jinými než popsanými restarty služeb nebo strojů a podobně).



1. To umožňuje osadit lichý počet mezirackových chladících jednotek. Při sudém počtu chladících jednotek z tohoto požadavku plyne, že je nutno studenou uličku vybudovat v délce celých sedmi racků. [↑](#footnote-ref-2)