

Příloha č. 1 Zadávací dokumentace

Technická dokumentace

Technická dokumentace – Popis stávající úložné infrastruktury v Plzni a požadavky na předmět plnění

Informace a údaje uvedené v jednotlivých částech této technické dokumentace vymezují závazné požadavky zadavatele na plnění veřejné zakázky. Tyto požadavky je uchazeč povinen plně respektovat při zpracování nabídky.

Pokud není uvedeno jinak, veškeré kapacity jsou uvedeny v dekadických násobcích, tj. 1TB = 10^{12} B, 1PB = 10^{15} B.

1. Popis stávající úložné infrastruktury v Plzni

Stávající úložná infrastruktura sdružení CESNET umístěná v prostorách Západočeské univerzity v Plzni byla dodána na sklonku roku 2011. Jedná se o HSM (Hierarchical storage management) systém o celkové kapacitě cca 3,8 PB. Logicky se úložiště skládá z vrstvy (tieru) rychlých FC disků o hrubé kapacitě cca 76,8 TB, z vrstvy pomalejších SATA disků o hrubé kapacitě 480 TB a páskové knihovny o stávající hrubé kapacitě 3300 TB. Data jsou na servery nahrávány skrze 6 frontend serverů zapojených v HA režimu. Migraci dat mezi jednotlivými tiery je zajištěna pomocí softwaru DMF, všechny frontendy využívají clusterový souborový systém CXFS.

Bližší popis hardwaru a softwarového nastavení následuje níže.

1.1. Hardwarová struktura úložiště

Hardwarově se úložiště skládá:

- Z dvojice diskových polí InfiniteStorage 4600, kdy každé pole obsahuje:
 - 2x řadič s konektivitou 8xFC 4Gbit a 4x 10Gbit iSCSI, 32GB cache
 - 4x 16 diskovou polici s 600GB 15k rpm FC disky
 - 2x 60 diskovou polici s 2TB 7.2k rpm SATA disky

- Páskové knihovny Spectralogic T-finity v konfiguraci
 - 1x základní rack umístěný uprostřed s kapacitou 12xLTO5 mechanika, 950 slotů na média (z toho 50 nakonfigurováno pro import/export pásek)
 - 2x rozšiřující media only rack, každý s kapacitou 1300 LTO5 slotů
 - 2x krajový servisní rack každý s kapacitou 500 slotů
 - 2x nezávislý redundantní robot pro obsluhu knihovny
 - 11x LTO5 mechanika 8Gbit FC
 - 2200 kusů LTO5 pásek s čipem pro správu životního cyklu, 50 kusů čistících médií

- o licence na 3670 páskových slotů

Celkově lze knihovnu ve stávající konfiguraci osadit až 4500 LTO páskami (přitom je nutno doplnit licence, neboť 830 slotů knihovny licenci nemá).

- 2x servisní uzel pro zajištění služeb clusterového souborového systému (CXFS) a migrace (DMF)
 - o 2x Intel X5650
 - o 48GB RAM
 - o 2x 500GB 7.2k rpm disk v RAID1
 - o 2x 1Gbit síťové rozhraní
 - o 2x 8Gbit FC rozhraní (jedna karta)
 - o SLES 11 SP1 + HA rozšíření + SGI Foundation Software
- 6x frontend serverů zajišťujících vnější konektivitu, každý v konfiguraci
 - o 2x Intel X5650
 - o 48GB RAM
 - o 2x 500GB 7.2k rpm disk v RAID1
 - o 2x 10Gbit a 2x 1Gbit síťové rozhraní
 - o 4x 8Gbit FC rozhraní (dvě karty po dvou portech)
 - o SLES 11 SP1 + HA rozšíření + SGI Foundation Software
- 2x virtualizační servery
 - o 2x Intel X5650
 - o 48GB RAM
 - o 2x 300GB 15k rpm disky v RAID1
 - o 2x 10Gbit a 2x 1Gbit síťové rozhraní
 - o 2x 8Gbit FC rozhraní (jedna karta)
 - o VMWare ESXi 5.0
- 2x speciální aplikační servery SGI C3108-TY11
 - o 2x Intel X5650
 - o 48GB RAM
 - o 2x 500GB 7.2k rpm disky v RAID1
 - o 2x 10Gbit a 2x 1Gbit síťové rozhraní
 - o 2x 8Gbit FC rozhraní (jedna karta)
- 1x centrální administrační stroj
 - o 1x Intel E5620
 - o 12GB RAM
 - o 2x 1TB 7.2k rpm disky v RAID1
 - o 2x 10Gbit a 2x 1Gbit síťové rozhraní

Úložiště dále obsahuje potřebnou propojovací infrastrukturu:

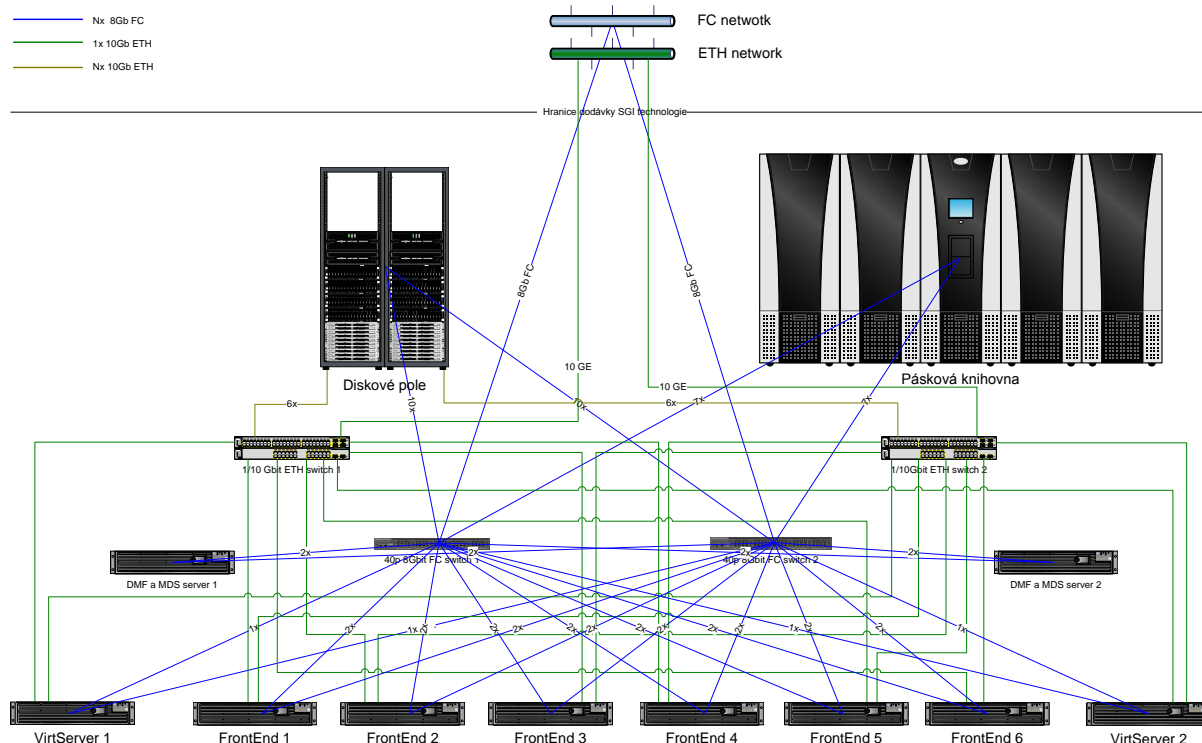
- 2x 40 portový 8Gbit FC switch Brocade 510

- všechny porty osazeny SFP a licencemi
- neobsazeno 8 portů v každém switchi (16 celkem)
- 2x 48 portový 10Gbit Cisco Nexus 5548, s vlastnostmi
 - ke každému připojen modul obsahující 24x1Gbit porty
 - každý 32x10Gbit portů zalicencováno, 8 portů s SFP typu 10GBASE-SR volných (celkem 16)
 - 5 respektive 2 1Gbit porty volné (celkem 7 portů)
- 1x 24portový 1Gbit SMC 8824M
 - 1 volný 1Gbit port.

1.2. Propojení komponent

Propojení jednotlivých komponent pomocí 10Gbit Ethernetu a 8Gbit FC zobrazuje Obrázek 1.

Na všech frontend serverech a speciálních aplikačních serverech jsou jejich 10Gbit rozhraní zapojeny do obou switchů, na strojích je konfigurován active-active bonding, na straně switchů pak vPC. Kvůli podpoře vPC jsou oba switche propojeny pomocí (na obrázku není znázorněno).



Obrázek 1 Propojení jednotlivých částí úložiště v sítích 10Gbit Ethernet a 8Gbit FC

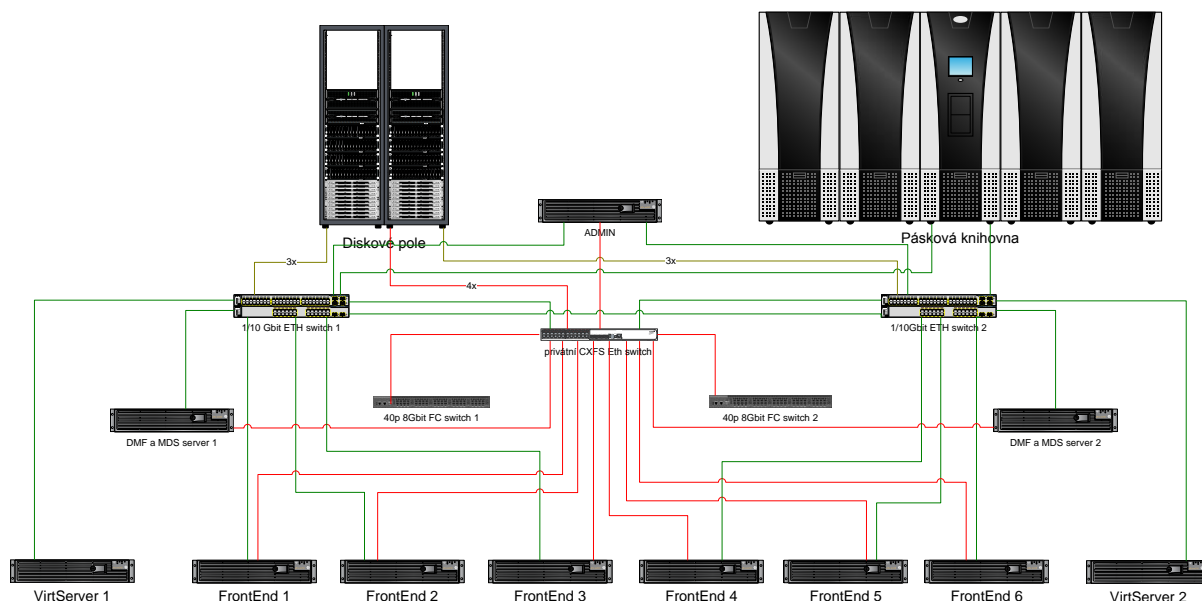
Všechna zařízení obsahující FC porty jsou pak propojena do obou FC switchů minimálně jedním rozhraním, úložiště se skládá ze dvou nezávislých fabric sítí. Výjimku tvoří jen jednotlivé drivy páskové knihovny, které jsou zapojené vždy jen do jednoho switche.

1Gbit Ethernet síť slouží k několika účelům. Část sítě je vyhrazena pro primární a záložní okruh CXFS clusteru sloužící pro výměnu metadat a heartbeat zpráv. Po této síti jsou také posílány heartbeat zprávy HA řešení jiných komponent. Jiná část 1Gbit sítě pak slouží pro přístup na management rozhraní jednotlivých zařízení. Zapojení 1Gbit Ethernet sítě znázorňuje Obrázek 2.

— 1x 1Gb privátní CXFS ETH
— 1x 1Gb ETH
— Nx 1Gb ETH

— FC network
— ETH network

Hranice dodávky SGI technologie



Obrázek 2 Propojení jednotlivých částí úložiště pomocí sítě 1Gbit Ethernet

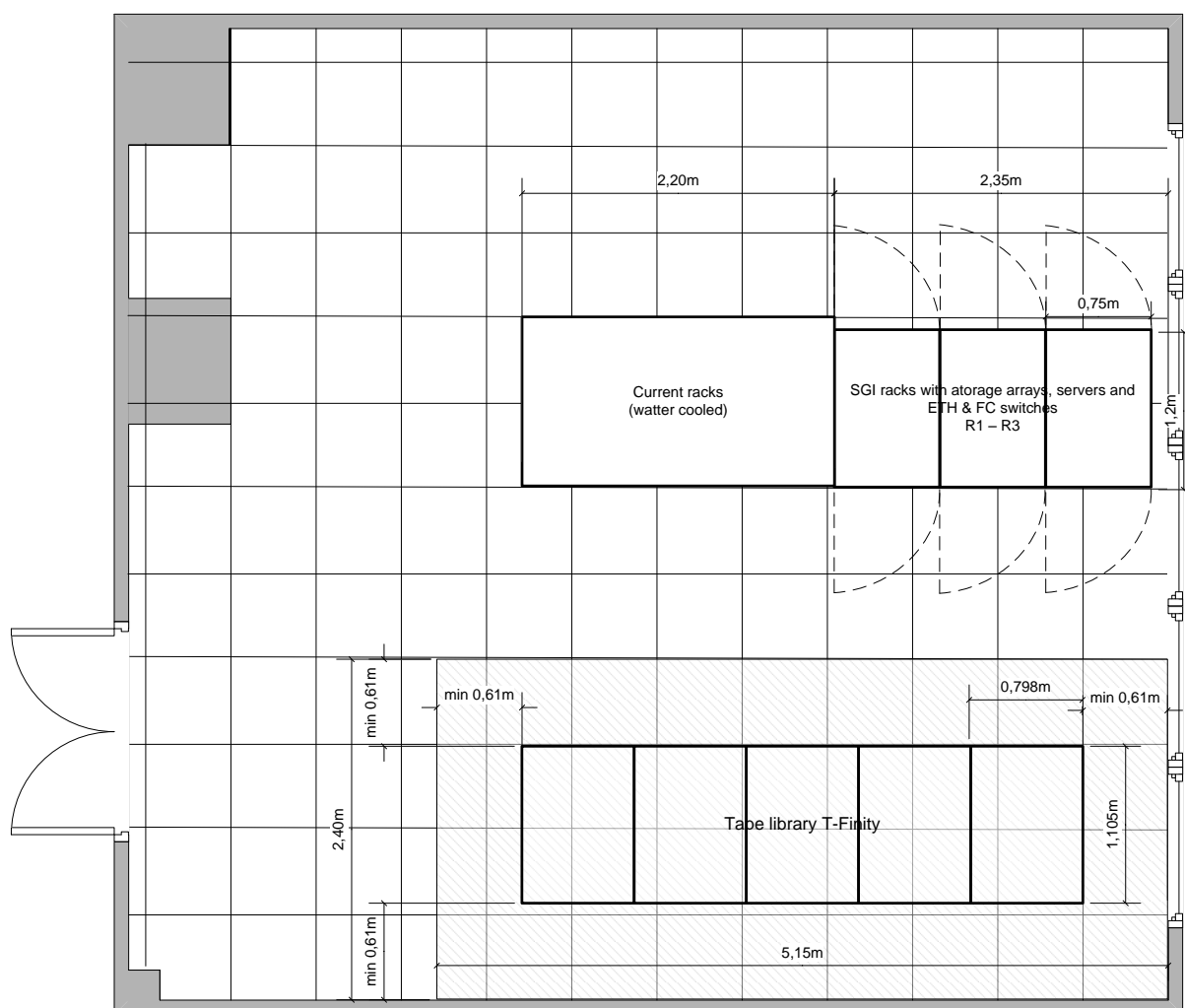
1.3. Fyzické umístění

Současná technologie je umístěna ve třech 19palcových 80 cm širokých 42 U vysokých racích. Pásková knihovna se skládá z 5 racků, které tvoří jeden funkční celek. Rozmístění jednotlivých komponent v racích znázorňuje Obrázek 3. Pozn.: Na místě pro „souborový řadičový modul diskového pole“ je ve skutečnosti umístěn 2x speciální aplikační server.



Obrázek 3 Rozmístění technologií do jednotlivých racků

Celkové rozmístění v serverovně pak zobrazuje Obrázek 4.



Obrázek 4 Schéma rozmístění zařízení v serverovně

1.4. Softwarová struktura úložiště

CXFS metadata servery/DMF servery, všechny frontendy a administrační server jsou provozovány na operačním systému SLES 11 SP1 s HA rozšířeními. Na úložišti je provozován software SGI ISSP v2.5 skládající se z distribuovaného souborového systému CXFS v 6.5.0.4 a DMF ve verzi 5.5.0.

Softwarové konfigurace speciálních aplikačních serverů a virtualizovaných serverů nejsou pro účel tohoto výběrového řízení podstatné.

1.4.1. Stávající softwarové licence

Součástí stávajícího úložiště jsou tyto licence:

- 11x SLES 11 SP1
- SGI DMF server + HA rozšíření, s kapacitou 6PB
- CXFS Server + 12x CXFS client

1.4.2. Poskytované úložné služby

Úložiště je nakonfigurováno tak, že frontendy pro uživatele poskytují stejný jmenný prostor pomocí následujících protokolů:

- NFSv4 s Kerberos autentizací, provozujeme sgi-enhanced-nfs, které je součástí ISSP v2.5.
- FTPS implementovaným pomocí vsftpd serveru
- SFTP/SCP pomocí openssh
- Rsync

Pomocí všech protokolů lze přistupovat na stejná data.

1.4.3. HA konfigurace úložiště

Všechny služby úložiště jsou konfigurovány jako vysoce dostupné. Toho je dosahováno pomocí softwaru Pacemaker. Úložiště je rozděleno na tři HA clustery:

- 3 frontendy tvoří produkční HA cluster zajišťující služby pro uživatele. Všechny protokoly jsou provozovány v režimu active-active s výjimkou NFSv4, které je provozováno v režimu active-passive.
- 3 frontendy tvoří vývojový HA cluster, který je nakonfigurován velice podobně jako produkční HA cluster a slouží k ověřování změn nastavení
- Oba MDS/DMF servery tvoří HA cluster zajišťující funkci metadata serveru CXFS a DMF serveru. Do toho clusteru je přiřazen i uzel admin, který však slouží jen jako tie breaker a neběží na něm žádné služby.

1.4.4. Rozdělení úložiště

Úložiště je momentálně rozděleno na cca 20 souborových systémů s velikostí od 1 TB do 50 TB. Migrace většiny souborových systémů je nastavená v dvoutierovém režimu, kdy uživatelé pracují na SATA discích, ze kterých jsou data migrována na jednu nebo více pásek. Jeden ze souborového systému je nakonfigurován ve třítierovém režimu, kdy uživatelé pracují na FC discích, odkud jsou data paralelně migrována na SATA disky a pásy.

2. Požadavky na předmět plnění

Předmětem plnění zakázky jsou následující komponenty:

- 2.1. Software SGI DMF Parallel Data Mover option včetně příslušných licencí tak, aby bylo možné tento software provozovat na stroji SGI C3108-TY11 (jeden ze speciálních aplikačních serverů) a začlenit ho do celého HSM popsaného v sekci 1.
- 2.2. 1000 kusů LTO5 pásek s čipem pro správu životního cyklu, pásky musí být kompatibilní a osaditelné do knihovny uvedené v sekci 1. Součástí je také instalace všech dodaných pásek do této knihovny.
- 2.3. Diskové pole typu MAID s následujícími vlastnostmi:
 - 2.3.1. Polem typu MAID rozumíme takové pole, jehož disky lze vypnout a to vč. elektroniky. Samotné zastavení rotace disků nestačí.
 - 2.3.2. Diskové pole zaujímá nejvýše 1 rack o rozměrech maximálně (š x d x v) 85 x 135 x 240 cm.
 - 2.3.3. Disky v poli jsou organizované v RAID 5 v konfiguraci s maximálně čtyřmi datovými disky (4+1) nebo v RAID 6 v konfiguraci s maximálně 12 datovými disky (12+2).
 - 2.3.4. Na každých i započatých 60 disků pole požadujeme alespoň 1 samostatný hot-spare disk (např. při počtu 300 disků zapojených v RAID svazcích požadujeme navíc 5 hot-spare disků, celkem tedy 305 disků).
 - 2.3.5. Diskové pole musí být plně redundantní, výpadek jakékoliv jedné komponenty nesmí způsobit nedostupnost pole, může ale vést k dočasné degradaci výkonu.
 - 2.3.6. Diskové pole 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 chybná komponenta v lidsky čitelné podobě.
 - 2.3.7. Čistá kapacita (tj. kapacita dostupná uživateli po režii RAIDu a hotspare, před režii souborového systému) diskového pole je alespoň 2,4 PB.
 - 2.3.8. Diskové pole je připojitelné ke stávající infrastruktuře alespoň 8x 8Gbit FC do každého z obou switchů uvedených v sekci 1, tj. celkem minimálně 16x.
 - 2.3.9. Diskové pole bude umístěno buď v serverovně se zbytkem HSM viz obrázek 4, nebo v serverovně ve stejné budově, tato serverovna je od serverovny s HSM vzdálená naproti přes chodbu o šířce přibližně 3,5 m (délka kabelů bude větší).
 - 2.3.10. Veškerá nezbytná kabeláž (napájení, datové kabely, management) je součástí dodávky pole. Maximální délka FC kabelů je 40 metrů, přesnou délku kabeláže je nutné zjistit při instalaci.
 - 2.3.11. Maximální provozní spotřeba diskového pole je 7 kW.

- 2.4. Rozšíření licencí SGI Data Migration Facility. V současnosti zadavatel používá licenci na 6 PB kapacity. Předmětem dodávky je rozšíření licence tak, aby licence celkem pokryla 1,5násobek čisté kapacity celého zařízení (tj. jak stávajícího systému zadavatele, tak i zařízení dodaného v rámci tohoto rozšíření). Čistou kapacitou rozumíme a. pro disková pole kapacitu dostupnou uživateli po režii RAIDu a hotspare, před režii souborového systému, b. nominální nekomprimovanou kapacitu pro pásky.
- 2.5. Součástí dodávky bude rekonfigurace stávajícího SGI DMF a souvisejících softwarových komponent pro začlenění všech komponent předmětu plnění do stávajícího systému zadavatele.
- 2.6. Součástí dodávky je rovněž rekonfigurace HW (zejména SAN) pro začlenění všech komponent předmětu plnění do stávajícího systému zadavatele.
- 2.7. Zatížení podlahy systémem úložiště po přidání nových prvků pořízených v rámci tohoto výběrového řízení nesmí překročit 1700 kg/m^2 , bodové zatížené podlahy nesmí překročit 300 kg.
- 2.8. Všechna dodaná zařízení a technologie musí být možné transportovat stavebním otvorem s maximálními rozměry 145cm x 190cm, za kterým jsou ve vzdálenosti 70cm tři schody směrem vzhůru s celkovou výškou 56cm.

3. Související požadavky

Uchazeč je povinen poskytnout nezbytnou součinnost s dodavateli dotčených zařízení. Rozšíření datového úložiště nesmí ovlivnit jeho funkci a dostupnost jeho služeb, s výjimkou prací nutných k jeho rozšíření s maximální dobou trvání 24 hodin. Uchazeč garantuje dodávku jednotlivých komponent tak, aby byly certifikovány výrobcem daného zařízení, aby nedošlo ke ztrátě garance funkčnosti celého zařízení.