Předmět zakázky

Předmětem nabídky na veřejnou zakázku „Superpočítač - Velký cluster" 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.

Předkládaná nabídka řešení dodávky Velkého clusteru byla vypracována s maximálním úsilím a pečlivostí a ve všech bodech naplňuje požadavky zadavatele stanovené a popsané v zadávací dokumentaci.

1 Velký cluster – popis funkcionality a vlastností navrženého řešení

Navržené řešení Velkého clusteru je komplexní ICT řešení výpočetního systému pro náročné vědecko-technické výpočty – tj. komplex výpočetních, úložných a infrastrukturních propojovacích / síťových a dalších systémů a SW řešení. Navržené řešení umožňuje efektivní běh mnoha současných výpočetních úloh ve všech fázích jejich životního cyklu (příprava, výpočet, zpracování výsledků) pro různé typy úloh (paralelní, sériové; dávkové, interaktivní) od mnoha uživatelů, bezpečné uložení dat uživatelů a rychlý a snadný lokální i vzdálený přístup k datům, efektivní správu systému, komponent, zdrojů a uživatelů.

Řešení poskytuje výkonné výpočetní zdroje, které jsou dobře přístupné a využitelné uživateli systému a jejich úlohami. Řešení zajišťuje vlastnosti, služby a funkce potřebné pro efektivní provoz a efektivní správu systému zadavatelem. Řešení taktéž zajišťuje dostupnost a kvalitu služeb v čase. Řešení je navrženo jako vyvážené, parametry a vazby jednotlivých subsystému zohledňují ostatním subsystémy a tvoří jeden efektivní celek.

Nabídka obsahuje veškeré systémy, zařízení, komponenty, příslušenství, licence, dokumentaci, projektové, implementační a další práce, školení atd. nezbytné k naplnění požadavků zadavatele vzhledem k definici předmětu zakázky.

Navržené řešení respektuje dispozice a omezení vyplývající z prostředí a podmínek zadavatele.

Řešení neobsahuje omezení a limity, které by zabraňovaly či omezovaly užití Velkého clusteru zadavatelem v požadovaném, očekávaném nebo racionálním rozsahu.

Řešení je v maximální míře autonomní, nezávislé na externích systémech a službách, soběstačné bez potřeby dalších zařízení, systémů či služeb.

Řešení je navrženo, dimenzováno a implementováno tak, aby zajistilo spolehlivý, bezpečný, výkonný a efektivní provoz Velkého clusteru v datovém centru zadavatele.

1.1 Komponenty Velkého clusteru

Velký cluster obsahuje Výpočetní cluster. Výpočetní cluster je tvořen Výpočetními servery propojenými Výpočetní sítí - vysokorychlostní FDR InfiniBand sítí s nízkou latencí. Výpočetní cluster je určen pro provádění výpočetních úloh uživatelů.
Výpočetní servery tvoří standardní výpočetní servery tzv. *Výpočetní servery bez akcelerace* a výpočetní servery osazené specializovanou akcelerační výpočetní kartou s velkým množstvím výpočetních jader tzv. *Výpočetní servery s akcelerací.* Řešení využívá tzv. MIC akceleraci - jsou osazeny specializované výpočetní karty s mnoha výpočetními jádry architektury x86 (MIC akcelerátory).

Velký cluster obsahuje čtyři *Přístupové servery* - servery sloužící pro přístup uživatelů, pro přípravu úloh a dat, kompilaci a ladění kódů, pro zpracování výsledků a pro přenos dat.

Velký cluster obsahuje *Vizualizační servery* - servery pro vzdálenou vizualizaci dat uživatelů.

Velký cluster obsahuje *Datová úložiště.* Datová úložiště jsou realizována jako komplexní řešení úložných zařízení, I/O serverů (např. souborových serverů), sítí a potřebného softwarového vybavení. Datová úložiště poskytují požadované datové služby.
Datová úložiště zahrnují *Souborová datová úložiště* – výkonná úložiště poskytující služby souborového systému a *Datové úložiště infrastruktury.*
Velký cluster obsahuje dvě Souborová datová úložiště - *Souborové datové úložiště HOME* určené pro středně a dlouhodobá data uživatelů a *Souborové datové úložiště SCRATCH* určené pro krátkodobá data uživatelů po dobu výpočtu úloh a střednědobá data úloh a projektů.
Velký cluster obsahuje *Datové úložiště infrastruktury* určené pro ukládání a sdílení dat infrastruktury clusteru. Slouží pro uložení systémových obrazů (image) serverů, logů, dat infrastrukturních služeb, uživatelského aplikačního vybavení, databázových systémů a rovněž jako datové úložiště *Virtualizační infrastruktury.*
Velký cluster obsahuje řešení zálohování dat tzv. *Zálohovací systém.*

Velký cluster obsahuje *Infrastrukturní servery. Infrastrukturní servery* jsou určené pro řízení clusteru, zdrojů, úloh, licencí a poskytování infrastrukturních služeb Velkého clusteru (např. DHCP, DNS, LDAP, licenční servery, plánovače, monitoring, logování, atd.)

Velký cluster obsahuje *Management servery. Management servery* jsou servery určené pro správu, administraci zařízení, úloh, uživatelů, zdrojů a služeb.

Velký cluster obsahuje infrastrukturu serverové virtualizace poskytující virtuální servery tzv. *Virtualizační infrastrukturu, která jako* datové úložiště využívá *Datové úložiště infrastruktury.*

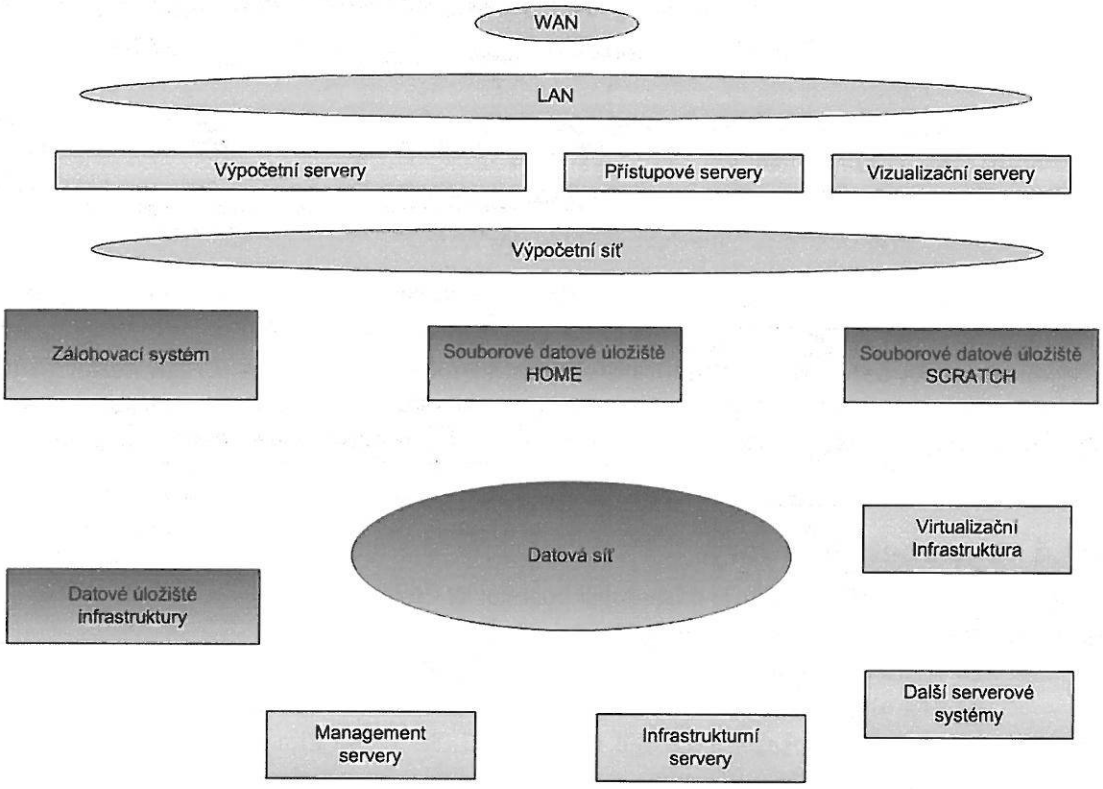
Velký cluster obsahuje další fyzické servery určené výhradně pro běh specifických služeb zadavatele (databázových, portálových a dalších) tzv. *Další serverové systémy.* Těmto Dalším serverovým systémům je rovněž dostupné *Datové úložiště infrastruktury.*

Velký cluster obsahuje *Síťovou infrastrukturu* tj. síťové propojení komponent, systémů, tak aby bylo dosaženo požadované funkcionality, byl zajištěn přístup na jednotlivé služby, byl zajištěn výkon, dostupnost a bezpečnost.
Síťovou infrastrukturu tvoří zejména *Výpočetní síť* clusteru a *Ethernetová síť.*
Výpočetní síť, založená na technologi FDR InfiniBand, propojuje Výpočetní servery Výpočetního clusteru, Přístupové servery, Vizualizační servery, servery kategorie Další serverové systémy a datová úložiště a další servery viz dále v popisu technického řešení.
Ethernetová síť zajišťuje komunikaci mezi zařízeními uvnitř Velkého clusteru, mezi Velkým clusterem, dalšími systémy provozovanými zadavatelem (Malý cluster, Specializovaný systém, atd.) a internetem. Řeší připojení do internetu a ostatních vzdálených sítí (např. CESNET a GEÁNT) a zabezpečení takového připojení.
Datové sítě propojují zařízení Datových úložišť, propojují Datová úložiště a konzumenty jejich služeb, propojují zařízení řešení Zálohování, atp.

Velký cluster obsahuje řešení a infrastrukturu pro instalaci a provoz Velkého clusteru v datovém centru, tj. racky a příslušenství potřebné pro umístění, řešení napájení a chlazení zařízení Velkého clusteru a rozhraní a napojení na infrastrukturu datového centra zadavatele.

Kromě uvedených komponent, které jsou nutnou součástí řešení, řešení obsahuje všechny další systémy potřebné pro zajištění požadované funkcionality a pro efektivní provoz Velkého clusteru.

1.2 Orientační schéma Velkého clusteru



Obrázek - Orientační schéma Velkého clusteru

Orientační schéma Velkého clusteru je pouze zjednodušené ilustrativní, schématické znázornění Velkého clusteru, v žádném případě nejde o úplné či přesné zapojení.

1.3 Výpočetní cluster

Pro měření výpočetního výkonu Výpočetního clusteru bude použit High Performance LINPACK benchmark (http://www.netlib.org/benchmark/hpl/)

Výpočetní výkon Rmax Výpočetního clusteru při využití pouze CPU bude určen během výpočetního benchmarku High Performance LINPACK spuštěného paralelně nad všemi CPU všech Výpočetních serverů Výpočetního clusteru (jedna instance benchmarku na celém clusteru).

Výpočetní výkon Rmax akcelerační karty bude určen během výpočetního benchmarku High Performance LINPACK spuštěného v hybridní konfiguraci (na procesorech a akceleračních kartách serveru současně), od naměřené hodnoty benchmarku bude odečten výpočetní výkon procesorů a poté bude výsledná hodnota podělena počtem akceleračních karet v serveru.

Výpočetní výkon Rmax akceleračních karet Výpočetních serverů s akcelerací bude určen jako součet výpočetních výkonů Rmax všech akceleračních karet Výpočetních serverů s akcelerací.

Agregovaný výpočetní výkon Rmax Výpočetního clusteru bude určen jako součet Výpočetního výkonu Rmax Výpočetního clusteru při využití pouze CPU Výpočetních serverů a výpočetního výkonu Rmax akceleračních karet Výpočetních serverů s akcelerací.

Hodnoty výpočetních výkonů jsou uvedeny pro nabízenou konfiguraci určenou k běžnému provozu. Dosažení hodnot výpočetních výkonů není nijak podmíněno např. specifickým režimem procesoru, ve kterém nelze systém dlouhodobě a bez dalších omezení provozovat nebo předpokládanou efektivitou, jejíž dosažení však uchazeč negarantuje.

Agregovaný výpočetní výkon Rmax Výpočetního clusteru překračuje minimálně požadovanou hodnotu 1000 Tflop/s, dosahuje 1546.56 Tflop/s.
Výpočetní výkon Rmax Výpočetního clusteru při využití pouze CPU překračuje minimálně požadovanou hodnotu 650 Tflop/s, dosahuje 756 Tflop/s.

Výkony Rpeak, výkony Rmax (Výpočetní výkon Rmax Výpočetního clusteru při využití pouze CPU, Výpočetní výkon Rmax akceleračních karet Výpočetních serverů s akcelerací, Agregovaný výpočetní výkon Rmax Výpočetního clusteru) a efektivita Výpočetního clusteru v High Performance LINPACK benchmarku při měření výpočetního výkonu pouze CPU jsou uvedeny v tabulce (zpracované dle vzoru v příloze číslo 4 Zadávací dokumentace), která je přílohou tohoto dokumentu.

1.4 Výpočetní servery

Každý Výpočetní server splňuje následující požadavky definované zadávací dokumentací:

- Fyzický server, architektura x86-64
- 24 fyzických CPU jader na server - osazeny jsou 2 procesory Intel Xeon E5-2680v3 12 core 2.5GHz na server
- Paměť RAM - osazeno 128GiB DDR4 s ECC (8x 16GiB dimm) na server, tj. 5.3GiB na jedno fyzické jádro CPU
- Konektivita Výpočetní síť - 1x FDR 56Gb/s
- Konektivita Ethernetová síť
- Podpora bootu operačního systému ze sítě
- 64-bitový operační systém s jádrem Linux - CentOS 6.5

Operační paměť je rovnoměrně rozložena (kapacitou a rychlostí přístupu) na procesory a CPU jádra výpočetního serveru (osazeno 8x16GiB DIMM, 4xDIMM na jeden procesor). Operační paměť je složena z paměťových modulů stejné velikosti a rovnoměrně rozložena na paměťové řadiče a paměťové kanály výpočetního serveru.
Všechny Výpočetní servery bez akcelerace mají shodnou hardwarovou konfiguraci.
Nabídka neobsahuje výpočetní servery s GPU akcelerací.
Všechny Výpočetní servery s MIC akcelerací mají shodnou hardwarovou konfiguraci.
Všechny Výpočetní servery mají stejný typ (model) procesoru, stejnou konfiguraci paměti RAM a budou pracovat ve shodném provozním nastavení (frekvence, časování, nastavení vlastností, atp.).
Každý Výpočetní server s akcelerací obsahuje dva akcelerátory shodného typu Intel Xeon Phi 7120P.
Výpočetní cluster obsahuje 576 Výpočetních serverů bez akcelerace.
Výpočetní cluster obsahuje 432 Výpočetních serverů s akcelerací, tj. 864 akcelerátorů.

Všechny MIC akcelerátory Intel Xeon Phi 7120P Výpočetních serverů s MIC akcelerací splňují požadavky zadávací dokumentace následujícími parametry:

- 61 jader architektury x86
- výpočetní výkon Rpeak 1.2Tflop/s v režimu dvojité přesnosti (double precision)
- paměť 16GiB v režimu ECC

Všechny Výpočetní servery mají jednotný 64-bitový operační systém s jádrem Linux CentOS 6.5.

Výpočetní servery jsou určeny výhradně pro výpočty, žádný Výpočetní nebude využit pro zajištění jiné funkcionality.

Výpočetní servery bez akcelerace, detainí popis:

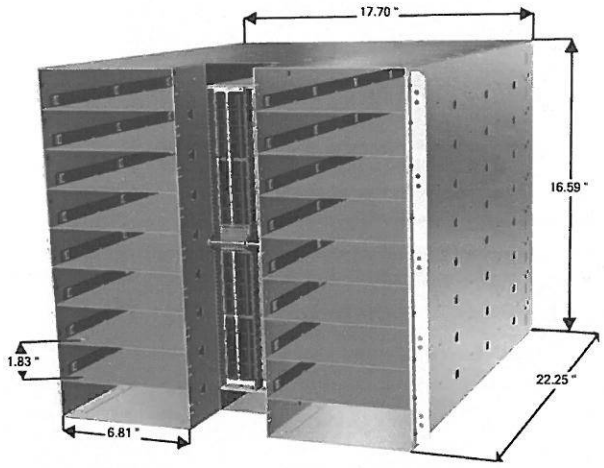
Výpočetní servery bez akcelerace budou řešeny prostřednictvím specializovaného clusterového systému ICE-X v provedení M-Cell.

ICE X - IRU

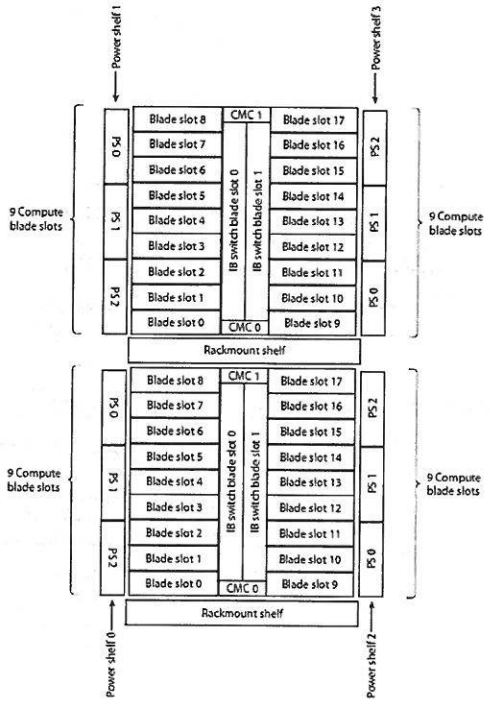
Základním stavebním kamenem systému SGI ICE X je jedna nezávislá racková jednotka (IRU - Independent rack unit) o velikosti 9.5U, která bude osazena do M-Racku (vnější šíře skříně 28"). Tato IRU má v sobě zaintegrovány veškeré potřebné komponenty tak, aby bylo možno postavit cluster do 864 CPU jader (s 12-jádrovými procesory, pouze M-Rack) bez potřeby jakékoli externí kabeláže. Každá IRU má 18 pozic pro osazení hot-plug blade serverů a při využití duálních (twin) blade serverů může obsahovat až 36 výpočetních uzlů. Součástí IRU je dále IB FDR (56Gbit/s) infrastruktura, redundantní násobná infrastruktura managementu a správy postavená na sítích Ethernet a hiearchickém principu systémových kontrolerů a redundatní soustava napájení.

IRU bude umístěna v SGI M-Racku, a ty pak tvoří základ M-Cellu a umožňují chlazení horkou vodou.

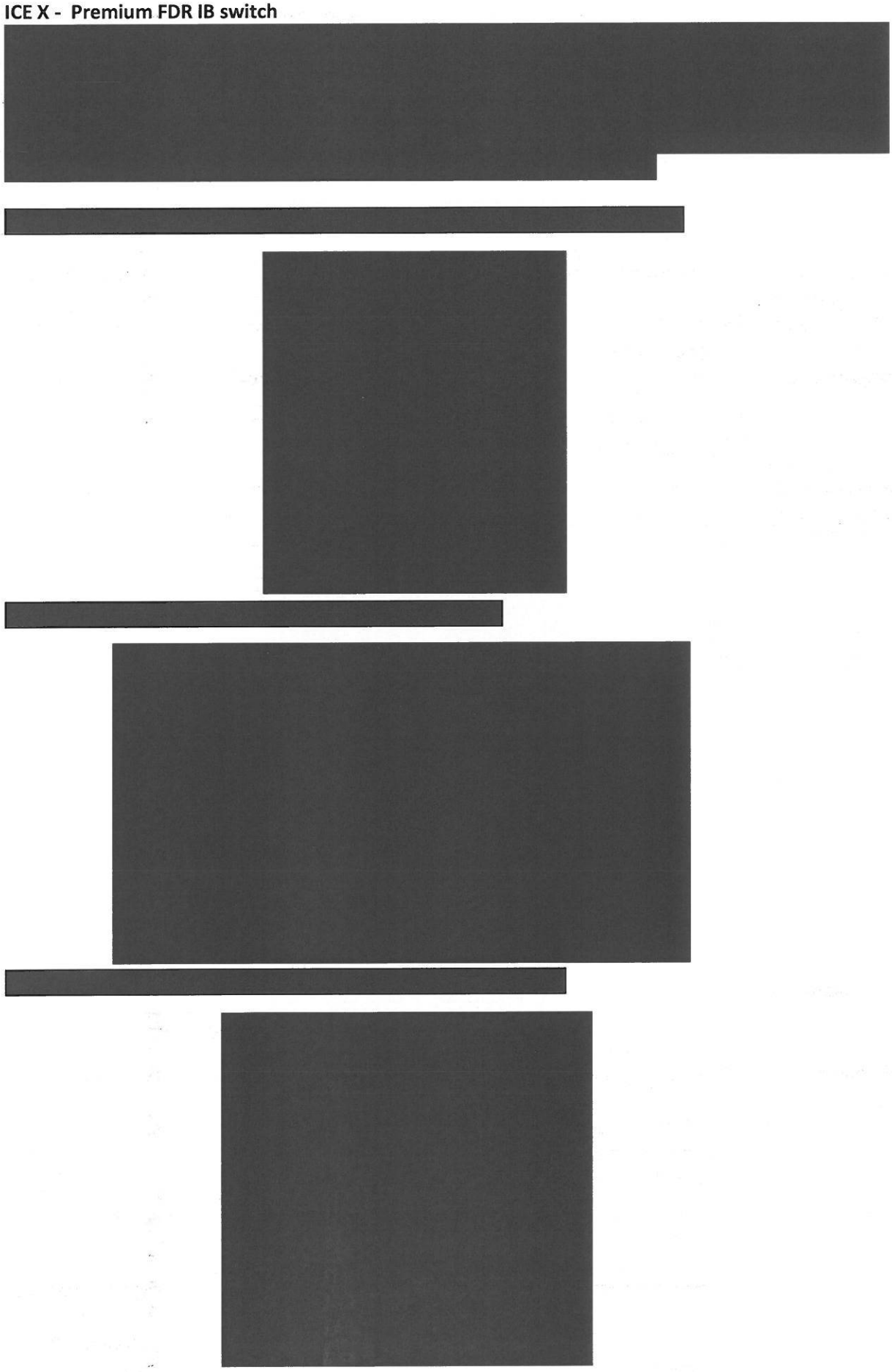
Následující obrázek ukazuje čelní pohled na jednu IRU s rozměrem 9.5U:



Následující obrázek ukazuje čelní pohled na stavební blok SGI ICE X skládající se ze dvou IRU a určený pro osazení v M-Racku:



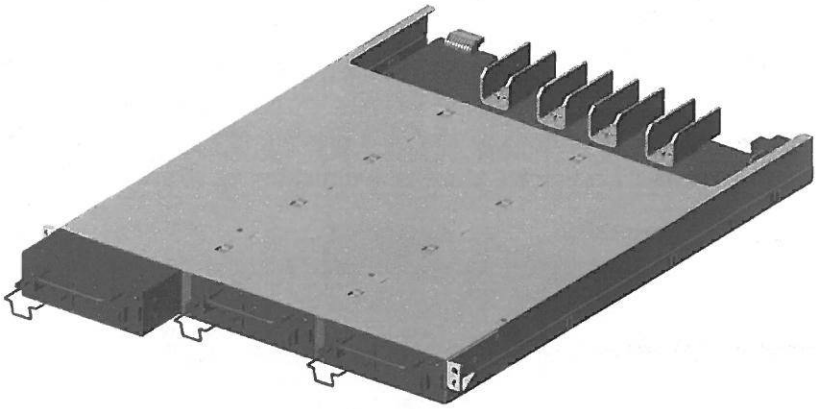
ICE X - Premium FDR IB switch



ICE X - napájení

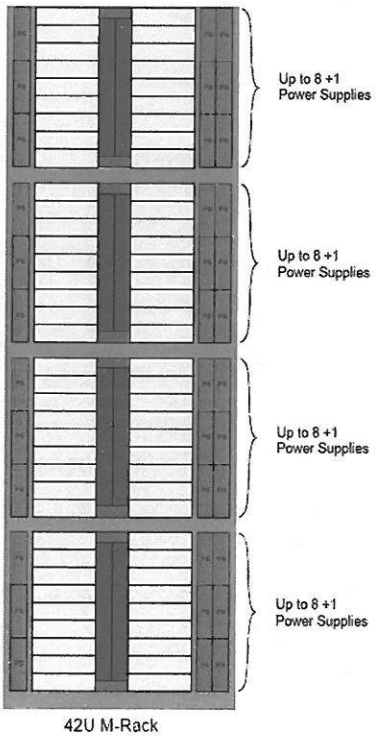
Nejmodernější generace SGI ICE X má vylepšený design napájení, který umožňuje škálovat počet zdrojů a jejich redundanci nezávisle na IRU, dle potřeb konkrétní konfigurace či zákazníkových požadavků (400-1440W na výpočetní uzel). Tím, že je power shelf nezávislý na IRU umožňuje snadný přechod na novější technologii napájení v rámci stejné IRU pouze změnou power shelfu. V případě M-Racku jsou power shelfy umístěny z boku M-Racku.

Následující obrázek zobrazuje jeden power shelf, který obsahuje 3 x napájecí zdroj, jeden M-Cell bude obsahovat celkem 48 takovýchto power shelfů, celkem tedy 144 napájecích zdrojů využívajících redundanci zapojení N+1 (8+1 pro každou IRU):



ICE X – M-Rack

Jednotlivá IRU s blade servery budou umístěna v M-Racku. M- Rack je speciálně navržený rack, který může být provozován pouze v rámci M-Cellu, umožňuje maximální výpočetní hustotu a maximálně efektivní chlazení. Každý z celkem 4 M-Racků v M-Cellu bude osazen čtyřmi IRU s maximálním počtem 144 dvouprocesorových výpočetních nodů v provedení twin-blade. M-Rack podporuje přímé chlazení horkou vodou samotného výpočetního čipu – procesoru. V případě M-Racku jsou jednotlivá IRU instalována bez jakýchkoliv chladících ventilátorů, chlazení je zajištěno chladícími věžemi, které jsou umístěny v M-Cellu.

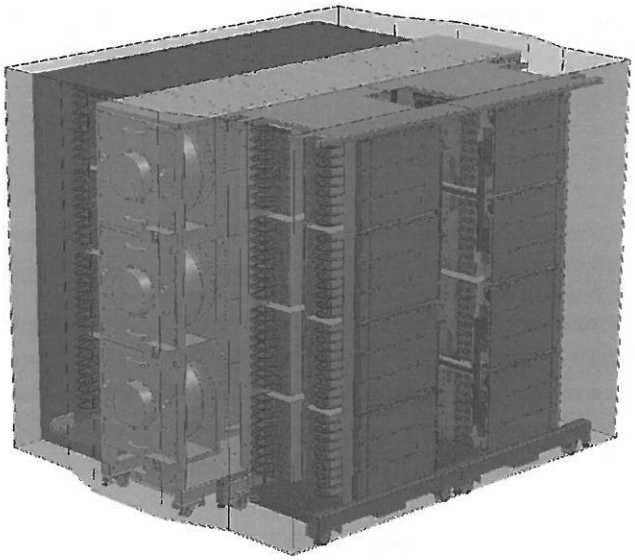


42U M-Rack

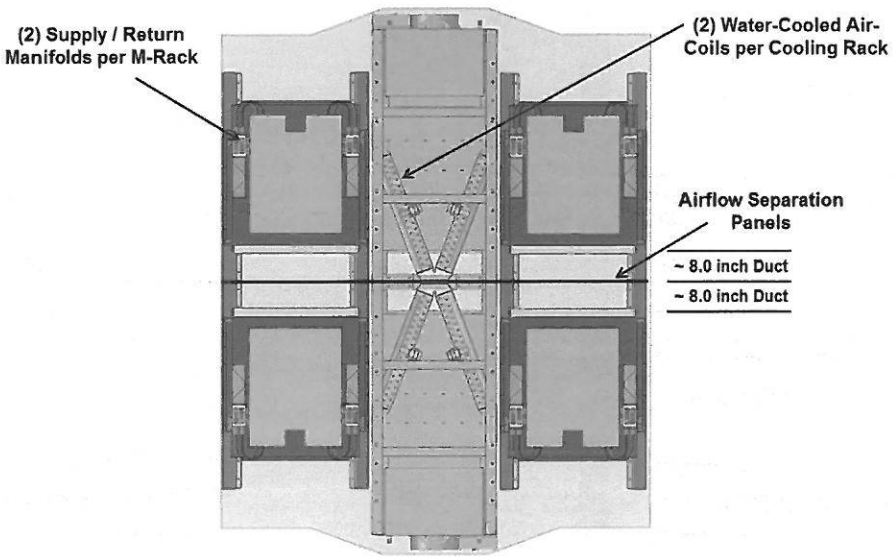
ICE X – M-Cell

M-Racky budou umístěny v M-Cellu, což je jakési zapouzdření M-Racků. M-Cell je založen na hybridním chlazení a využívá přímého chlazení kapalným médiem pro výpočetní procesory a chlazení pomocí proudícího vzduchu mezi výpočetním rackem (M-Rack) a chladícím rackem pro chlazení ostatních komponent jako jsou paměti, síťové adaptéry, switche, napájecí zdroje atd... Z tohoto pohledu se jedná o jediné a unikátní technologické řešení, které uchladí 100% generovaného tepla výpočetním systémem pomocí chlazení horkou vodou. Chladící rack je osazen velkým hliníkovým chladičem, který využívá ke chlazení stejný zdroj horké vody, který je využíván distribuční jednotkou (CDU) pro přímé chlazení procesorů kapalným médiem. Proudění vzduchu uvnitř M-Cellu je zajištěno velkými ventilátory (turbínami), které směrují studený vzduch do výpočetních uzlů a nasávají horký vzduch přes výměník tepla. M-Cell je plně zakrytován čímž je zajištěno, že vzduch uvnitř M-Cellu není za standardního provozu míchán s okolním vzduchem. Jeden plný M-Cell je osazen čtyřmi výpočetními M-Racky a dvěma chladícími racky (Cooling rack) a připojeno je jedno CDU.

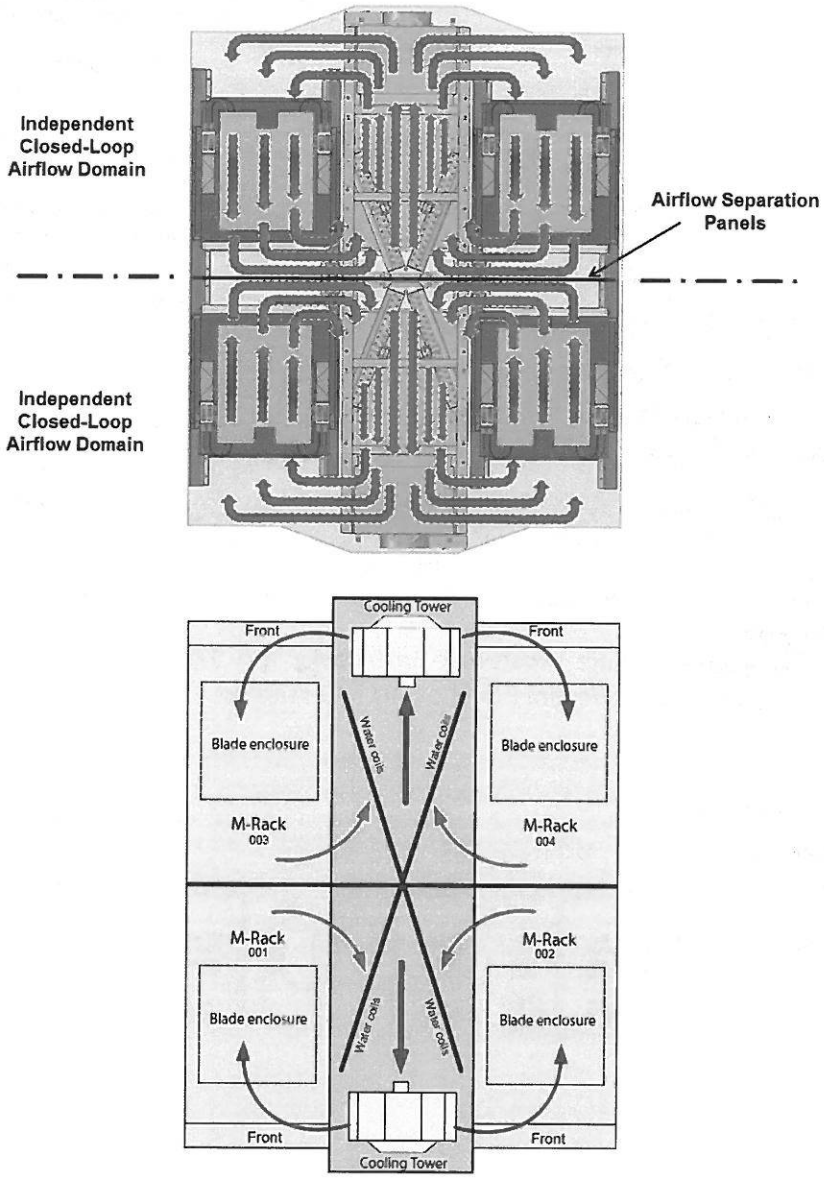
Následující obrázek zobrazuje zakrytovaný M-Cell v provedení Cube (stěny znázorněny průhledně), M-Rack PDU a kontroler chladícího racku, které jsou umístěny na střeše M-Cellu nejsou zobrazeny:



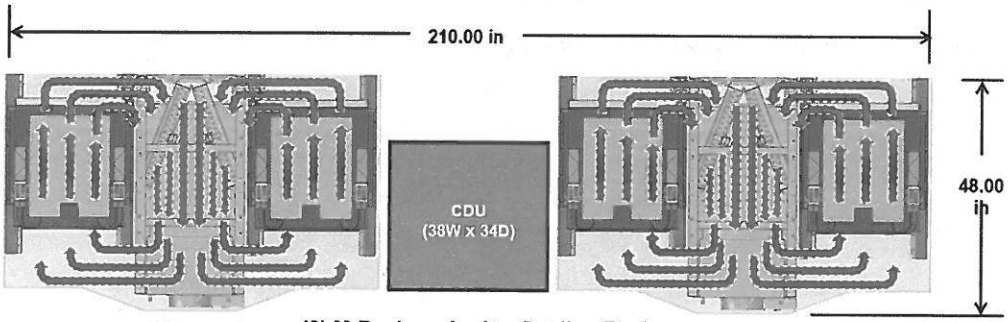
Následující obrázek zobrazuje pohled na M-Cell shora:



Následující obrázky zobrazují proudění vzduchu v rámci jednoho M-Cellu v provedení Cube:

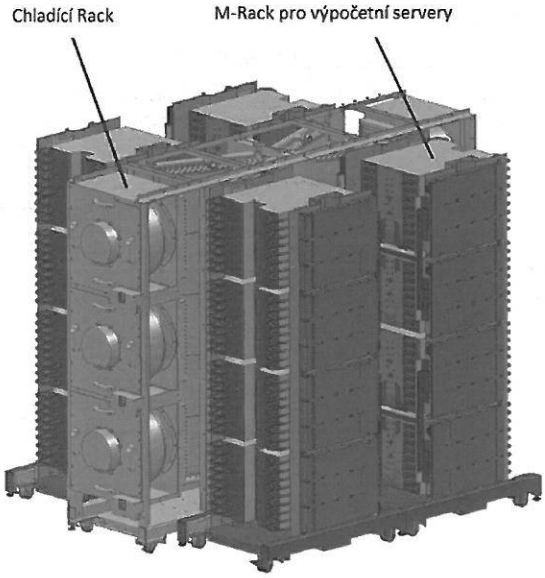


Následující obrázek zobrazuje proudění vzduchu v rámci jednoho M-Cellu v provedení Isle, toto rozložení M-Celu je předmětem této nabídky:

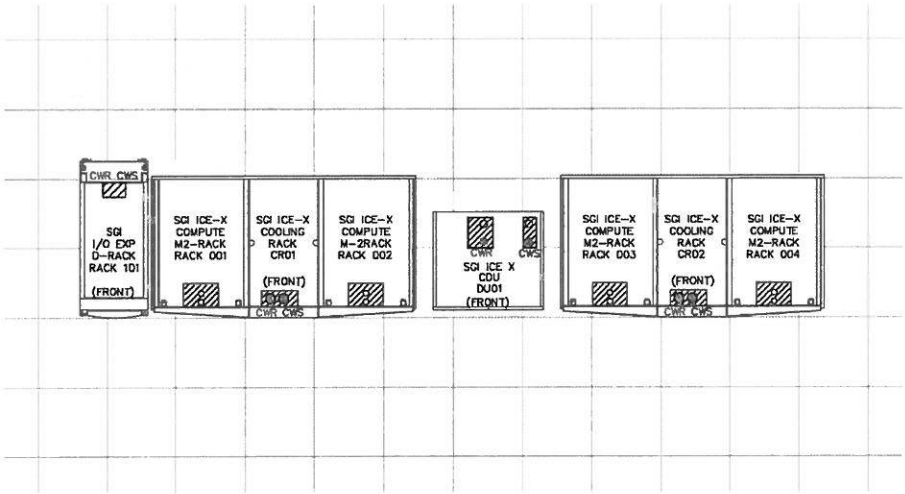


(2) M-Racks paired to Cooling Rack
(Top View)

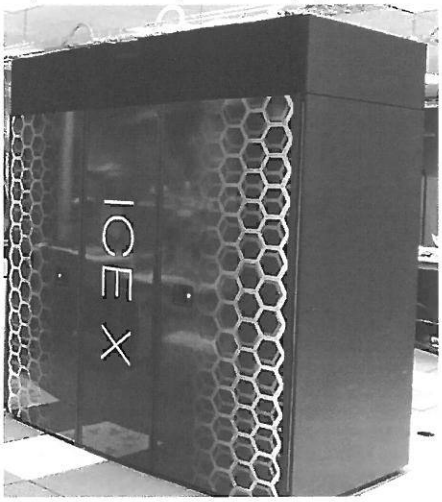
Následující obrázek zobrazuje odkrytovaný M-Cell v provedení Cube, zobrazující chladící racky (Cooling Rack) a výpočetní racky (M-Rack):



Následující obrázek znázorňuje prostorové požadavky pro 1xM-Cell (576 výpočetních nodů, 288 dualních blade serverů) v provedení Island s CDU a IO rackem:



Následující obrázek znázorňuje prostorové požadavky pro 1/2xM-Cell (zapouzdřený 2xM-Rack a Chladící rack).

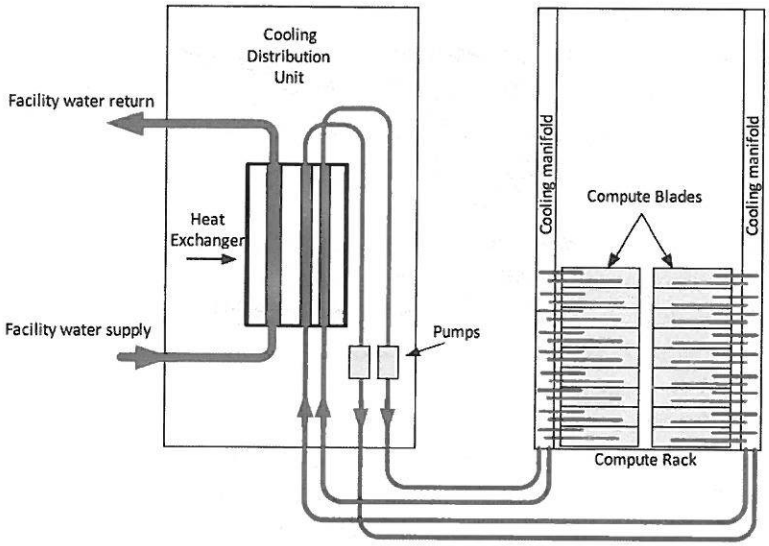


ICE X - CDU

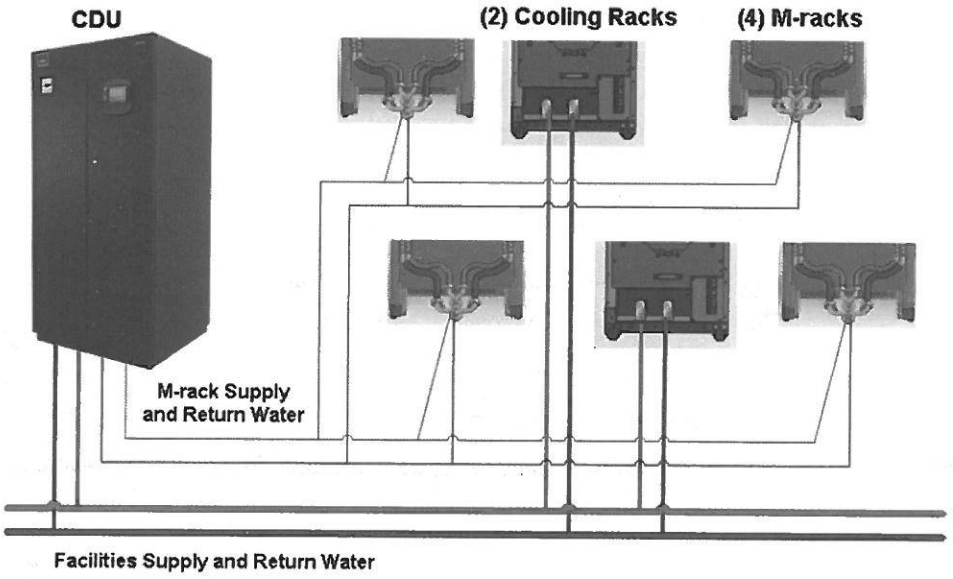
Nezbytnou součástí systému s přímým chlazením procesorů kapalným médiem je CDU (Cooling distribution unit) jednotka. Jedna CDU jednotka distribuje chladící médium až do čtyř M-Racků neboli jednoho kompletního M-Cellu. Ochlazená voda přichází z CDU dvěma potrubími do každého M-Racku, kde je distribuována až k samotným procesorům, tam absorbuje teplo a vrací se do CDU. CDU dále přenáší teplo do chladícího okruhu samotného datového centra. CDU kontroluje teplotu vody, která je dodávána do M-Racků a zajišťuje , aby byla vyšší než je rosný bod okolí, tak aby nedocházelo ke kondenzaci.

CDU obsahuje dvě čerpadla, využíváno je pouze jedno, druhé plní redundantní funkci.

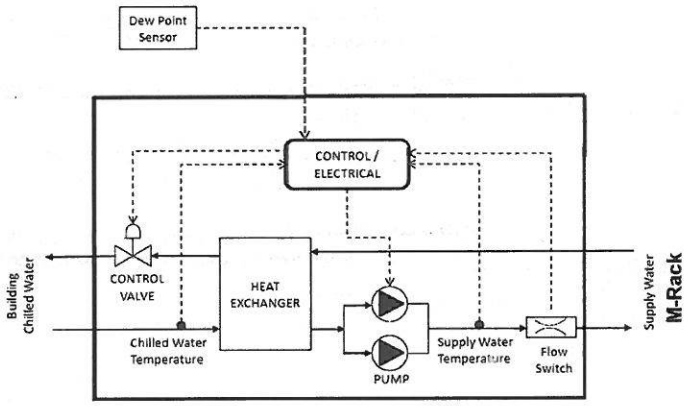
Následující obrázek zobrazuje připojení M-Racku k CDU:



Následující obrázek zobrazuje obecné schematické připojení čtyř M-Racků (1 M-Cell) k CDU a připojení CDU a chladících racků k jednomu chladícímu okruhu datového centra:



Následující obrázek zobrazuje detailně funkční schéma CDU:



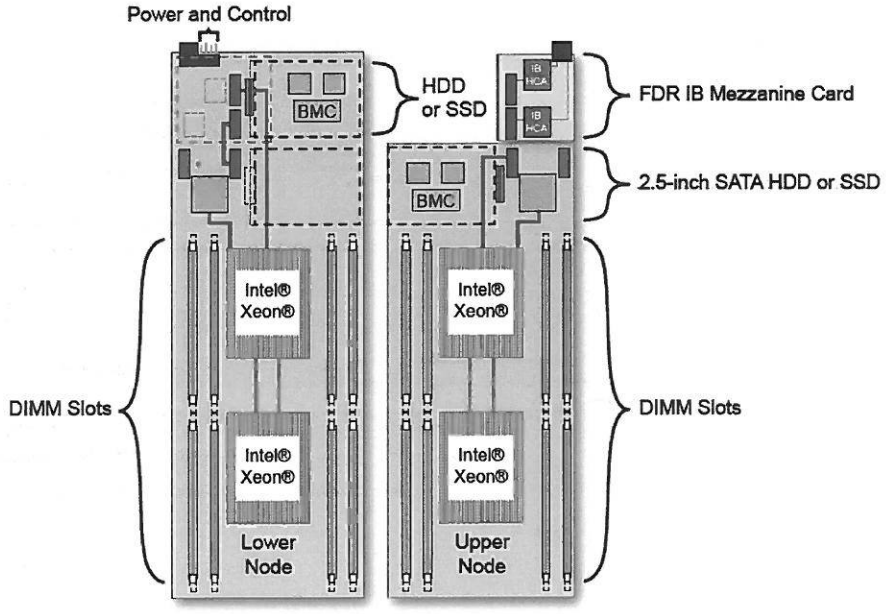
M-Cell může být chlazen studenou nebo teplou vodou v širokém rozsahu teplot s tím, že se zvyšující se teplotou vstupního kapalného média pro přímé chlazení samozřejmě klesá i celkové PUE. Náš návrh počítá s využitím vstupního média pro „teplovodní" přímé chlazení o teplotě minimálně 32C.

ICE X – Twin Blade servery

Jako výpočetní servery bez akcelerace budou použity blady pro systém ICE X s označením IP133, které mohou být osazeny procesory řady Intel Xeon E5-2600v3 „Haswell" a které se pyšní speciální HPC blade základní deskou – obsahující skutečně jen ty komponenty, které je třeba mít osazeny v HPC výpočetním uzlu. Zásluhou speciální základní desky tento blade umožňuje vysokou hustotu výp.výkonu – jedná se o speciální blade – tzv. twin-blade, který umožňuje do jednoho blade slotu v IRU osadit dvojnásobnou výpočetní kapacitu oproti standardnímu výpočetnímu bladu. Každý výpočetní uzel – twin blade s označením IP133 - bude disponovat dvěma základními deskami z nichž každá bude osazena dvěma procesory. Twin Blade využívá chlazení procesoru pomocí kapalného média a proto může být osazen nejvýkonnějšími procesory s vysokými hodnotami TDP. Twin blade disponuje šestnácti paměťovými pozicemi DDR4 (4 na procesorovou patici) s podporou frekvence 2133MHz a jedním mezzanine slotem. Twin blade je hot-plug což umožňuje servisovat jeden každý twin blade bez potřeby odstavit z provozu jakýkoli jiný prvek systému.

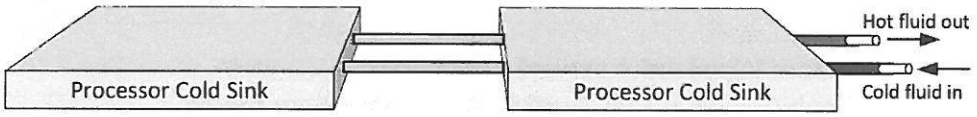
Twin Blade IP133 je možné osadit dvěma pevnými disky (1 disk na výpočetní server), nicméně námi navrhovaná konfigurace této možnosti nevyužívá, blady nebudou osazeny lokálními disky.

Následující obrázek zobrazuje vzhled twin bladu IP133 - dvě základní dvouprocesorové desky které zabírají v IRU společně jeden blade slot:



Následující obrázek zobrazuje Twin-blade IP133 - dvě základní desky vložené do sebe včetně chladícího prvku umožňujícího přímé chlazení kapalným médiem:

Následující obrázky zorazují podobu chladícího prvku pro Twin blade IP133:



Detailní konfigurace Výpočetních serverů bez akcelerace:

- 1x M-Cell v provedení Isle (ostrov), 1x CDU, 2x chladící rack, 4x M-Rack s výpočetními servery, 16x ICE X IRU
- Každé IRU bude osazeno 18 x dualním bladem IP133, celkem bude tedy osazeno v jednom M-Cellu 288x IP133 bladů – každý blade obsahuje dva výpočetní dvouprocesorové uzly, celkem tedy bude mít tato konfigurace 576 výpočetních uzlů
- Všechny Výpočetní servery bez akcelerace budou osazeny dvěma procesory Intel Xeon E5-2680v3, 12-jader 2.5GHz
- Každý dvouprocesorový výpočetní uzel bude osazen 8x16GiB DDR4 RAM moduly, celkem tedy 128GiB RAM, tj. 5.3GiB RAM na jedno fyzické výpočetní jádro CPU
- Výpočetní uzly nebudou osazeny lokálními disky, bootovaní bude zajištěno z bootovacích nodů

Výpočetní servery s akcelerací, detailní popis:

Výpočetní servery s akcelerací budou řešeny prostřednictvím specializovaného clusterového akcelerovaného dvouprocesorového 1U uzlu SGI Rackable C1104-GP1. Jedná se o speciální clusterový uzel vyvinutý jako ideální serverový uzel pro osazení akcelerátory.

- Celkem bude dodáno 432 x 1U fyzický dvouprocesorový uzel architektury x86-64 s označením SGI Rackable C1104-GP1
- Každý akcelerovaný výpočetní uzel bude osazen dvěma procesory Intel Xeon E5-2680v3, 12-jader 2.5GHz, celkem tedy bude uzel osazen 24 fyzickými CPU jádry na server
- Každý akcelerovaný výpočetní uzel bude osazen 8x16GiB DDR4 RAM moduly, celkem tedy 128GiB RAM, tj. 5.3GiB RAM na jedno fyzické výpočetní jádro CPU
- Každý akcelerovaný výpočetní uzel bude osazen dvěma shodnými akcelerátory Intel Xeon Phi 7120P, celkem tedy bude dodáno 864 akcelerátorů tohoto typu
- Každý akcelerovaný výpočetní uzel bude osazen 1xFDR InfiniBand portem 56Gb/s
- Každý akcelerovaný výpočetní uzel bude osazen 2x1Gb/s Ethernet portem
- Výpočetní servery s akcelerací nebudou osazeny lokálními disky, bootovaní bude zajištěno z bootovacích nodů, akcelerované výpočetní uzly podporují bootovaní ze sítě

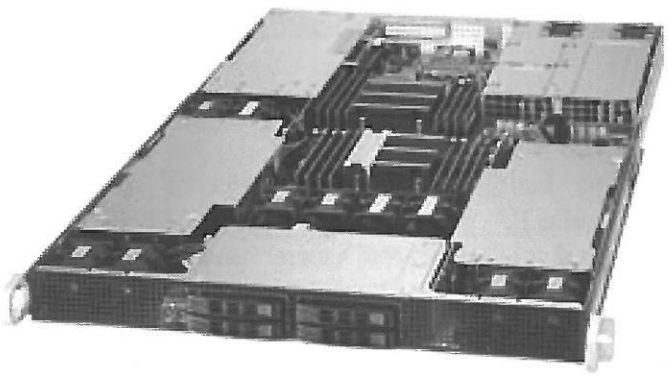
Vlastnosti SGI® Rackable™ C1104-GP1:

Provedení:	1U
Chipset:	Intel® C610
Procesor:	2xIntel® Xeon® E5-2600 v3
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	16
Počet pozic pevných disků:	4x3,5"
Rozšiřující sloty:	3x PCIe Gen 3.0 x16
	1x PCIe Gen 3.0 x8
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1660W redundantní zdroj

Detailní konfigurace Výpočetních serverů s akcelerací:

- Provedení: 1U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2680v3, 12 jader, 2.5GHz
- Operační paměť RAM: 128GiB DDR4
- Lokální disky: nejsou osazeny
- Diskový řadič RAID: onboard diskový řadič
- Konektivita Výpočetní síť: 1x FDR 56Gb/s
- Konektivita Ethernetová síť: 2x 1Gb/s
- Boot: bootovaní operačního systému ze sítě
- Zdroj: redundantní, za provozu vyměnitelné napájecí zdroje 1660W
- Operační systém: 64-bitový operační systém s jádrem Linux CentOS 6.5

Výpočetních server s akcelerací, obrázek:

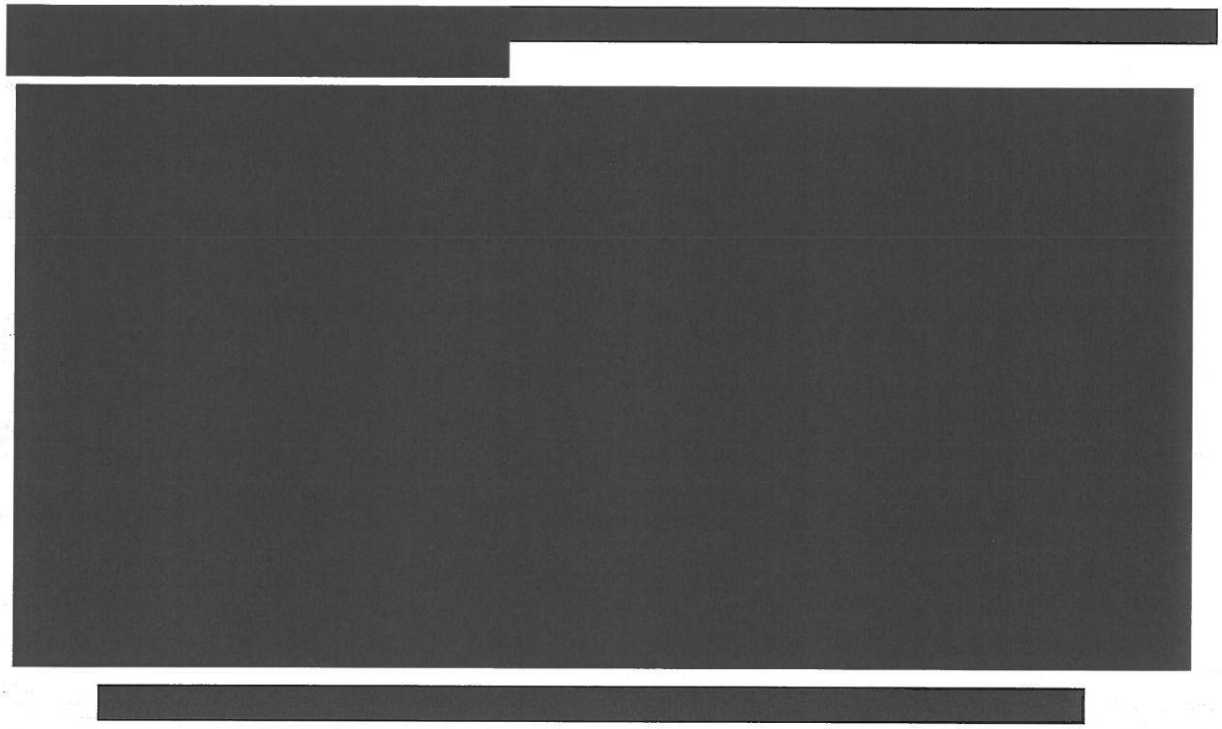


Na všech Výpočetnich serverech s akcelerací a na všech Výpočetních serverech bez akcelerace bude provozován operační systém 64-bitový operační systém s jádrem Linux CentOS 6.5

1.5 Výpočetní síť

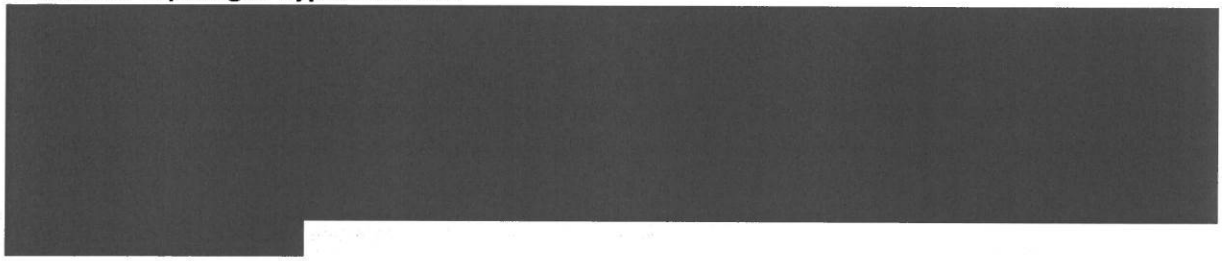
Rychlost komunikace libovolného páru Výpočetních serverů ve Výpočetní síti bude minimálně 5GB/s. Pro měření rychlosti bude použit PinPong benchmark (HPC Challenge Benchmark - Bandwidth-Latency-Benchmark - Min Ping Pong Bandwidth nebo Intel MPI Benchmarks – PingPong) s velikostí bloku 4MiB.

Celá Výpočetní síť bude zkonfigurována jako jedna homogenní síť v topologii 7D Enhanced hypercube.



Výpočetní síť Enhanced hypercube dosahuje vysoké efektivity komunikace. Efektivita komunikace Výpočetní sítě bude měřena testem PTRANS ze sady testů HPC Challenge Benchmark, verze 1.4.2. Test bude spouštěn jako 1 proces na jádro, v celkovém počtu procesů X, které je stanoveno jako celkový počet fyzických CPU jader vybraných 256 Výpočetních serverů. Nejlepší čas, dosažený paralelním během testu přes Výpočetní servery v počtu X, nebude více než 1.6x vyšší než nejlepší čas (ze standardních pěti běhů testu) pro paralelní běh testu přes 16 náhodně vybraných Výpočetních serverů. Řádek 34 souboru hpccinf.txt (označen „values of N") bude pro 16 Výpočetních serverů obsahovat hodnotu N1, vyšší nebo rovnou 262144 a pro X Výpočetních serverů minimálně hodnotu N2, vyšší nebo rovnou N1*sqrt(X/16), kde sqrt() označuje druhou odmocninu.

Nabízená topologie Výpočetní sítě



1.6 Přístupové servery

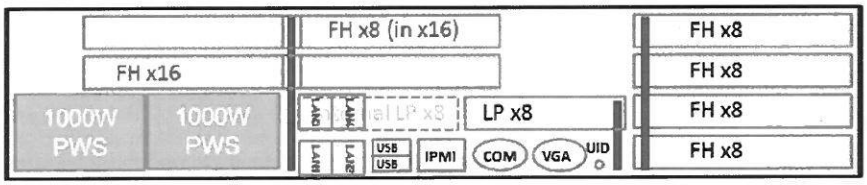
Řešení Velkého clusteru obsahuje dodávku **čtyř** Přístupových serverů v konfiguraci detailně uvedené dále v této kapitole. Každý Přístupový server splňuje veškeré požadavky zadávací dokumentace.
Přístupové servery budou určeny výhradně pro zajištění přístupu uživatelů a pro práci uživatelů, žádný Přístupový server nebude využit pro zajištění jiné než popsané funkcionality.
Přístupové servery budou poskytovat přístup uživatelům protokolem SSH2 a poskytovat služby pro přenos souborů SCP a SFTP.
Všechny přístupové servery budou mít stejnou hardwarovou konfiguraci.
Všechny přístupové servery používají stejnou x86 technologii procesorů a stejný operační systém CentOS 6.5 jako Výpočetní servery.

Přístupové servery budou řešeny prostřednictvím serveru **SGI® Rackable™ C2112-GP2**.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL) 4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Přístupového serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2695v3, 14 jader, 2.3GHz
  Celkový teoretický výpočetní výkon (Rpeak) serveru je 1030Gflop/s a to bez využívání dočasného přetaktování procesorů či jiných podobných vlastností
- Operační paměť RAM: 128GiB DDR4 (8x16GiB DIMM)
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS diskovy řadič s funkcionalitou HW RAID
- Konektivita Výpočetní síť: 1x FDR port 56Gbit
- Konektivita Ethernetová síť: 2x1Gb/s, 1x 10Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W
- Operční systém: 64-bitový operační systém s jádrem Linux CentOS 6.5

1.7 Vizualizační servery

Řešení Velkého clusteru obsahuje **dva** Vizualizační servery v konfiguraci detailně uvedené dále v této kapitole. Každý Vizualizační server splňuje veškeré požadavky zadávací dokumentace.
Vizualizační servery budou poskytovat vzdálenou hardwarově akcelerovanou vizualizaci s podporou OpenGL.

Řešení Vizualizačních serverů splňuje pro každý jednotlivý Vizualizační server následující požadavky:

1. Výkon při plném 4K rozlišení (4096 x 2560 bodů), při plné barevné hloubce (30 bit) alespoň 30 fps (snímků za vteřinu) při interaktivní práci jednoho uživatele v OpenGL akcelerované aplikaci.
2. Možnost současně probíhající interaktivní práce alespoň dvou uživatelů na dvou zcela odlišných úlohách s výkonem 30 fps při plném HD rozlišení (1920 x 1080) a plné barevné hloubce (30 bit).
3. Možnost současné kolaborativní práce dvou uživatelů nad jednou úlohou (sdílení obrazovky).

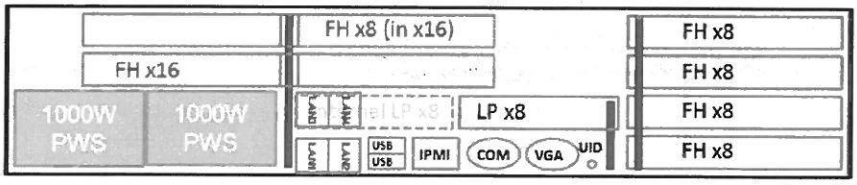
Vizualizační servery budou plně integrovány do prostředí Velkého clusteru ve stejném rozsahu jako Výpočetní servery, jde zejména o integraci uživatelských účtu, autentizace, integrace a zpřístupnění služeb Vizualizačních serverů pomocí Plánovače a dostupnost Souborových datových úložišť.
Vizualizační servery jsou zamýšleny výhradně pro vizualizaci dat uživateli, žádný Vizualizační server nebude použit pro zajištění jiné funkcionality.
Vizualizační servery mají stejnou hardwarovou konfiguraci.
Vizualizační servery používají stejnou technologii x86 procesorů a stejný operační systém CentOS 6.5 jako Výpočetní servery.
Součástí dodávky obou Vizualizačních serverů je dodávka vizualizačního software SGI VizServer [redacted] pokrývající v plném rozsahu požadavky zadávací dokumentace na funkcionalitu vzdálené vizualizace a podpory kolaborativní práce.

Vizualizační servery budou řešeny prostřednictvím serveru **SGI® Rackable™ C2112-GP2**.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL) 4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x GigE onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Vizualizačního serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2695v3, 14 jader, 2.3GHz
  Celkový teoretický výpočetní výkon (Rpeak) serveru je 1030Gflop/s a to bez využívání dočasného přetaktování procesorů či jiných podobných vlastností
- Operační paměť RAM: 512GiB DDR4 (16x 32GiB DIMM)
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS diskovy řadič s funkcionalitou HW RAID
- GPU karty: [redacted]
- Konektivita Výpočetní síť: 1x FDR port 56Gbit
- Konektivita Ethernetová síť: 2x 1Gb/s, 1x 10Gb/s
- Zdroj: redundantní, za provozu vyměnitelné napájecí zdroje 1000W
- Operační systém: 64-bitový operační systém s jádrem Linux CentOS 6.5

1.8 Datová úložiště

1.8.1 Datová úložiště - společné

Tato kapitola obsahuje popis architektury a řešení Datových úložišť pro Velký cluster. Řešení datových úložišť je redundantní, tedy neobsahuje komponentu, jejíž výpadek by způsobil nefunkčnost služeb datového úložiště. Komponenty řešení datového úložiště – zejména disky, napájecí zdroje, řadiče diskových polí, switche, servery jsou redundantní a vyměnitelné za provozu bez výpadku služeb datového úložiště. Veškerá disková pole řešení datového úložiště zajišťují takové zabezpečení dat, že selhání libovolných dvou disků nezpůsobí ztrátu dat. Služby a provoz datových úložišť se vzájemně negativně neovlivňují. Deklarované parametry agregované rychlosti a výkonu náhodných I/O operací jsou dosažitelné i při současném, paralelním zatížení všech datových úložišť.

Souborová datová úložiště poskytují služby síťového souborového systému. Souborové systémy, soubory realizované v souborovém datovém úložišti jsou zpřístupněné - sdílené přes síť na všech Výpočetních serverech Výpočetního clusteru, na všech Přístupových serverech, na všech Vizualizačních serverech a na Management serverech určených pro management uživatelů a dat. Na všech klientech souborových datových úložišť je poskytována obvyklá funkcionalita souborového systému.

Souborová datová úložiště jsou na straně klientů transparentně integrována do operačního systému, umožňují obvyklé souborové operace a realizaci obvyklé sémantiky nativních souborových systémů, podporují nativní rozhraní (API) souborového systému operačního systému a umožňují integrovat uživatele operačního systému jako uživatele souborového systému.

Souborová datová úložiště podporují Unicode ve jménech souborů a dlouhá jména souborů, umožňují řízení přístupu, přístupová práva na úrovni standardních Unixových práv (čtení, zápis, spuštění; uživatel, skupina, ostatní) a rozšířená ACL. Dále podporují uživatelské a skupinové kvóty, tedy omezení využité kapacity a počtu souborů nastavitelné individuálně na každého uživatele a na každou skupinu. Nabízené typy souborových datových úložišť podporují soubory o velikosti větší než 1TB, počet souborů 10 miliard a vytváření symbolických linků.

Souborové datové úložiště se z pohledu uživatele chová jako jediná, souvislá datová oblast s jednotným prostorem jmen, kde uživatel Souborového datového úložiště pro přístup k souborům Souborového datového úložiště používá jednotný prostor jmen a v rámci tohoto jednotného prostoru jmen je bez omezení dostupná veškerá kapacita úložiště a vlastnosti úložiště.

Data všech Souborových datových úložišť jsou dostupná z externích lokalit (z Internetu) protokoly SFTP, SCP a nástrojem sshfs (vše verze 2, 128bit). Řešení poskytuje reálnou agregovanou propustnost (rychlost sekvenčních operací nad Souborovými datovými úložišti) protokoly SFTP a SCP minimálně 5GB/s. Řešení je redundantní dle bodu SPEC_71 zadávací dokumentace - služba je poskytována i v případě výpadku či odstávky libovolného jednoho serveru zajišťujícího službu, avšak výkonové parametry pak budou při takovém režimu provozu nižší. Pro zajištění této funkcionality nebudou použity Přístupové, Vizualizační nebo Výpočetní servery či Další serverové systémy, pro tuto službu budou použity vyhrazené servery.

Data všech Souborových datových úložišť jsou dostupná interně ve vnitřní Ethernetové síti protokoly SFTP, SCP, nástrojem sshfs (vše verze 2, 128bit), dále protokoly NFS v4 a SMB/CIFS. Export protokolem SMB/CIFS umožňuje použití exportovaných dat v prostředí MS Active Directory. Řešení je redundantní dle bodu SPEC_71 Zadávací dokumentace - služba je poskytována i v případě výpadku či odstávky libovolného jednoho serveru zajišťujícího tuto službu.

Data všech Souborových datových úložišť jsou dostupná protokolem SMB/CIFS pro vizualizační jeskyni. (Vizualizační jeskyně není předmětem této veřejné zakázky.) Řešení poskytuje reálnou agregovanou propustnost (rychlost sekvenčních operací nad Souborovými datovými úložišti) protokolem SMB/CIFS klientům s 64-bitovým operačním systémem MS Windows 7, MS Windows 8 nebo MS Windows 8.1 infrastruktury vizualizační jeskyně minimálně 2.5GB/s na vyhrazeném optickém rozhraní QSFP IB QDR určeném pro připojení infrastruktury vizualizační jeskyně. Optické rozhraní nebude zapojeno do Ethernetové sítě. V případě výpadku či odstávky zařízení/serveru zajišťujícího tuto službu služba nebude poskytována.

Údaje o čisté kapacitě a výkonových parametrech datových úložišť jsou uvedeny ve formě vyplněné Excelovské tabulky (dle závazného vzoru Přílohy č.4 zadávací dokumentace) a jsou nedílnou součástí tohoto dokumentu v kapitole 6.

1.8.2 Souborové datové úložiště HOME

Souborové datové úložiště HOME je určeno výhradně pro data uživatelů Velkého clusteru a z pohledu uživatele se chová jako jediná, souvislá datová oblast s jednotným prostorem jmen. Datové úložiště HOME splňuje požadavek na čistou dostupnou (nekomprimovanou) kapacitu minimálně 0.5PB ($0.5x10^{15}$ byte) a splňuje požadavek na dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB minimálně 6 GB/s ($6x10^{9}$ byte/s). Tato požadovaná rychlost je reálně dosažitelná z Výpočetních serverů Výpočetního clusteru. Dále datové úložiště HOME splňuje požadavek na dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 75%/25% minimálně 5 kIOPs.

Souborové datové úložiště HOME je navrženo jako dvou-vrstvý (vrstva = tier) disk-na-disk/páska HSM systém, kde první tier je diskové pole a druhý tier reprezentuje NL-SAS diskové pole společně s jednou částí (partition) páskové knihovny T950B [redacted]. HSM software zajišťující automatické migrace dat mezi oběma tiery je SGI DMF (Data Migration Facility) [redacted]

Navržené souborové datové úložiště HOME poskytuje služby síťového souborového systému. Souborové systémy, soubory realizované v souborovém datovém úložišti jsou zpřístupněné - sdílené přes Výpočetní síť FDR InfiniBand na všech Výpočetních serverech Výpočetního clusteru, na všech Přístupových serverech, na všech Vizualizačních serverech a přes síť Ethernet na Management serverech určených pro management uživatelů a dat. Na všech klientech souborových datových úložišť je poskytována obvyklá funkcionalita souborového systému.

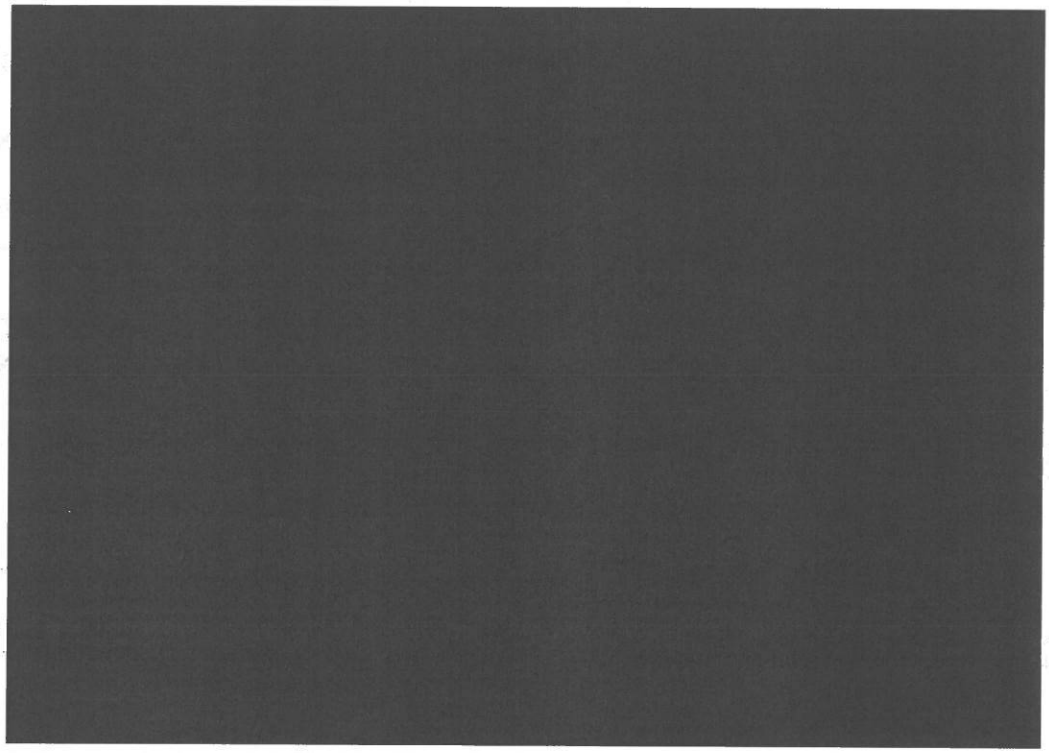
Souborové datové úložiště HOME je realizováno výkonným a spolehlivým clusterovým souborovým systémem CXFS (Clustered XFS). [redacted] Toto datové úložiště je na straně klientů transparentně integrováno do operačního systému, umožňuje obvyklé souborové operace. Datové úložiště HOME se z pohledu uživatele může chovat jako jediná, souvislá datová oblast s jednotným prostorem jmen – jeden velký filesystém.

Pro efektivní a rychlé přenosy dat mezi souborovým datovým úložištěm HOME a jinými úložišti dat vně výpočetního systému jsou CXFS Edge Servery Souborového datového úložiště HOME připojeny do sítě LAN dvěma 10GE síťovými spoji na server. Jejich prostřednictvím je úložiště zpřístupněno z externích lokalit (z Internetu) protokoly SFTP, SCP a nástrojem SSHFS a dosahuje agregovanou propustnost (rychlost sekvenčních operací nad Souborovými datovými úložišti) protokoly SFTP a SCP minimálně 5GB/s. Interně jsou data Soubového datového úložiště HOME dostupná prostřednictvím FDR InfiniBand a Ethernet sítě protokoly SFTP, SCP, NFS, CIFS a nástrojem SSHFS.

[redacted]

1. RAID Diskové pole pro Tier1 (2ks)
   [redacted]
2. NL-SAS diskové pole pro Tier2 (1ks)
   [redacted]
3. CXFS/DMF servery (2ks)
   [redacted]
4. DMF pDMO server (parallel data mover=server pro paralelní přesun dat) (1ks)
   [redacted]
5. CXFS Edge Servery (3ks)
   [redacted]
6. Pásková knihovna pro Tier2/Zálohovací systém (1ks)
   [redacted]
7. Prvky síťové infrastruktury pro zajištění komunikace (sdílené v rámci celého řešení)
8. licence pro CXFS/DMF řídící software a operační systémy RedHat 6.5 na všech výše uvedených serverech
9. Další potřebné příslušenství nezbytné k řádnému provozu sestavy datového úložiště (napájecí kabely, adaptéry, propojovací kabely mezi servery a switchi, případně další nezbytné síťové i jiné komponenty).

Schéma zapojení a integrace souborového datového úložiště HOME do prostředí řešení Velkého clusteru:



RAID disková pole pro Tier1 oblast jsou reprezentována dvěma systémy SGI Infinite Storage 5600, každý se dvěma redundantními kontroléry, každý vybaven 12GiB zrcadlenou cache chráněnou baterií po dobu případného přesunu obsahu cache na flash paměť a pamětí s ECC. Systém disponuje osmi 16Gbit FibreChannel host porty, čtyřmi 1GE management porty, redundantními zdroji s ventilátory, 60-bay diskovými policemi s 10K RPM SAS disky. VolumeGroup oblasti jsou chráněné proti poruše disků RAID6 zabezpečením.

NL-SAS diskové pole pro Tier2 oblast tvoří jeden systém SGI Infinite Storage 5100 se dvěma redundantními kontroléry, každý vybaven 2GiB zrcadlenou cache chráněnou baterií po dobu případného přesunu obsahu cache na flash paměť a pamětí s ECC. Systém disponuje čtyřmi 16Gbit FibreChannel host porty, čtyřmi 1GE management porty, redundantními zdroji s ventilátory, 60-bay a 12-bay diskovými policemi s 7200 RPM NL-SAS disky.

Pásková knihovna (=TapeLibrary), první partition pro kopii (=zálohu) Tier2 datové oblasti je reprezentována moderní páskovou knihovnou SpectraLogic T950B [redacted]

1.8.3 Souborové datové úložiště SCRATCH

Souborové datové úložiště SCRATCH je určeno výhradně pro data uživatelů Velkého clusteru. Bude používáno pro krátkodobá data úloh (část označená jako TEMP) a střednědobá data úloh a projektů (část označená jako WORK). Nabízené datové úložiště SCRATCH splňuje požadavek na čistou dostupnou (nekomprimovanou) kapacitu minimálně 1.5PB ($1.5x10^{15}$ byte) a splňuje i požadavek na dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1MiB minimálně 30 GB/s ($30x10^{9}$ byte/s). Tato rychlost je reálně dosažitelná z Výpočetních serverů Výpočetního clusteru. Dále SCRATCH splňuje požadavek na dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 75%/25% minimálně 20 kIOPs.

Dle bodu SPEC_95 Zadávací dokumentace je nabízenou variantou řešení datového úložiště SCRATCH varianta 3: Kde souborové datové úložiště SCRATCH se z pohledu uživatele chová jako jediná, souvislá datová oblast s jednotným prostorem jmen, kde jsou umístěna data TEMP a WORK, a dosahuje požadovaných a deklarovaných výkonových parametrů Souborového datového úložiště SCRATCH. Data TEMP a WORK jsou vzájemně oddělena (jako adresáře souborového systému). Řešení umožňuje umístění dat TEMP a WORK libovolné velikosti, omezením je pouze celková kapacita Souborového datového úložiště. Řešení poskytuje uživatelské a skupinové kvóty - nastavitelné nepřekročitelné omezení využité kapacity a počtu souborů nastavitelné individuálně na každého uživatele a na každou skupinu, které lze nastavit a uplatňovat pro celé úložiště SCRATCH, avšak nelze je nastavit a uplatňovat zvlášť pro data TEMP a zvlášť pro data WORK. Řešení poskytuje uživatelské a skupinové soft kvóty – upozornění generované v případě překročení hranice využité kapacity a počtu souborů nastavitelné individuálně na každého uživatele a na každou skupinu, které lze nastavit zvlášť pro data TEMP a zvlášť pro data WORK. Řešení soft kvót je rychlé a efektivní a datové úložiště nadměrně nezatěžuje.

Pro clustery o velikosti nabízené tímto řešením je námi doporučovaným datovým úložištěm, které by splňovalo náročné výkonové požadavky pro SCRATCH oblast výpočetního systému, paralelní filesystém Lustre. Lustre je masivní paralelní souborový systém, který je schopen nabízet petabajty úložné kapacity výpočetním clusterům o tisících uzlech. Vysokou úroveň technické podpory SGI pro Lustre filesystému podtrhuje i nedávné uvedení a vývoj řešení SGI® DMF™ pro Lustre, které umožňuje rozšířit Lustre o nativní archivační vrstvu v podobě DMF HSM systému. Otevřením prostředí archivačních vrstev a implementací HSM funkcionalit ve filesystému Lustre 2.5 může být nyní DMF použito jako archivační prostředí pro tento vysokorychlostní souborový systém. Lustre je častou volbou filesystému mnoha našich zákazníků, kteří nasazují do provozu mnoha uzlové scale-out clustery, jako například nabízený cluster SGI ICE-X.

Navržené souborové datové úložiště SCRATCH poskytuje služby síťového souborového systému. Souborové systémy, soubory realizované v souborovém datovém úložišti jsou zpřístupněné - sdílené přes Výpočetní síť FDR InfiniBand na všech Výpočetních serverech Výpočetního clusteru, na všech Přístupových serverech, na všech Vizualizačních serverech a přes síť Ethernet na Management serverech určených pro management uživatelů a dat.

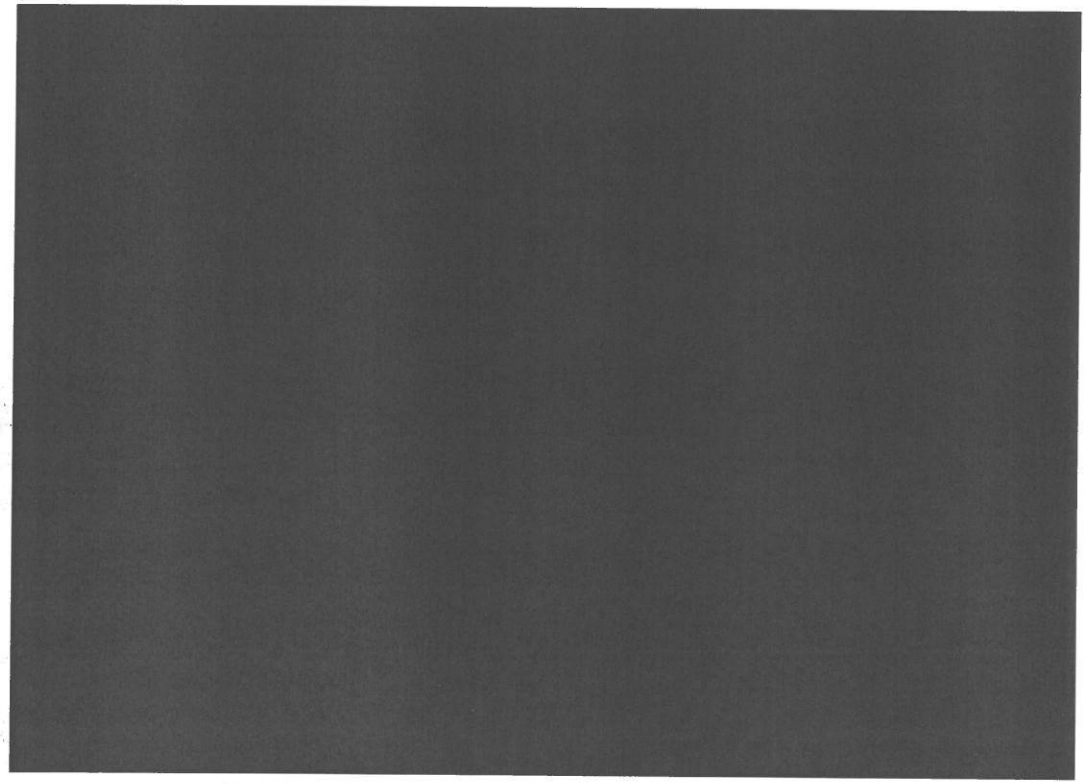
Díky dostupnosti obou datových úložišť HOME i SCRATCH přes rychlou FDR InfiniBand síť na dedikovaných serverech (Lustre a CXFS klientech), umožňuje uživatelům a úlohám prováděným na

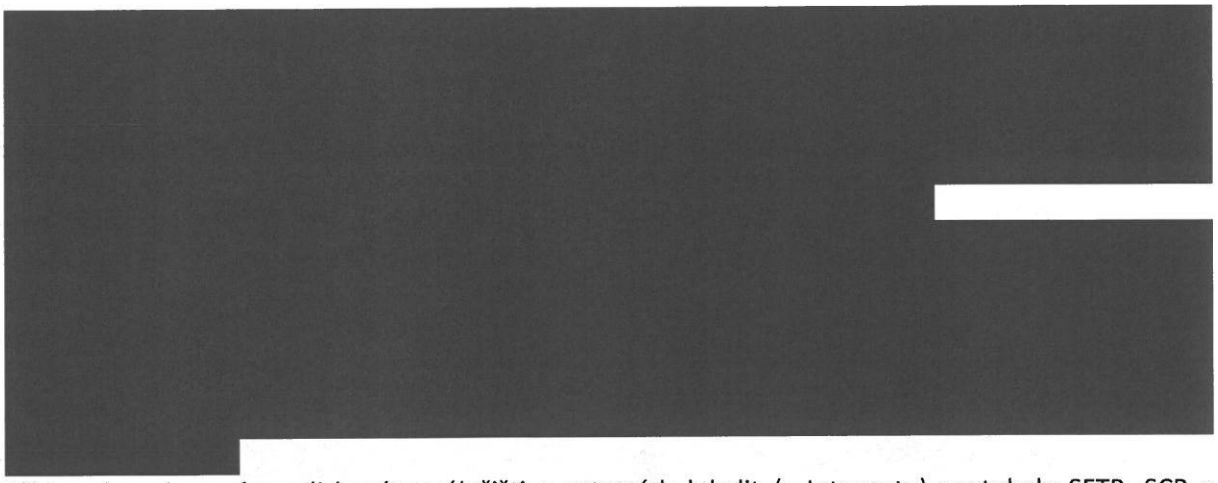
Výpočetních serverech clusteru provádět efektivní a rychlé přenosy dat mezi souborovými datovými úložišti SCRATCH a HOME.

[redacted]

1. Disková pole pro (OST) object storage targets [redacted]
2. Diskové pole pro metadata (MDT) filesystém [redacted]
3. Metadata (MDS) servery [redacted]
4. Object storage (OSS) servery [redacted]
5. LFSaccess servery (Lustre/CXFS klienti) [redacted]
6. Prvky síťové infrastruktury pro zajištění komunikace
7. Řídicí software Lustre a instalace všech potřebných operačních systémů a jejich HA rozšíření, včetně licencí těchto operačních systémů, na všech výše uvedených serverech.
8. Další potřebné příslušenství nezbytné k řádnému provozu sestavy datového úložiště (napájecí kabely, adaptéry, propojovací kabely mezi servery a LAN switchi, případně další nezbytné síťové i jiné komponenty).

Schéma zapojení a integrace souborového datového úložiště SCRATCH do prostředí řešení Velkého clusteru:





Přístup k souborovému diskovému úložišti z externích lokalit (z Internetu) protokoly SFTP, SCP a nástrojem sshfs zajišťuje pár LFSaccess serverů v active/active HA konfiguraci, pro zajištění vysoké dostupnosti dat, na platformě x86-64. Servery jsou připojeny do sítě LAN čtyřmi 10GE síťovými spoji na server. Jejich prostřednictvím je datové úložiště SCRATCH zpřístupněno z externích lokalit (z Internetu) protokoly SFTP, SCP a nástrojem SSHFS a splňuje požadavek na agregovanou propustnost (rychlost sekvenčních operací nad Souborovými datovými úložišti) protokoly SFTP a SCP minimálně 5GB/s. Interně jsou data Souborového datového úložiště SCRATCH dostupná prostřednictvím FDR InfiniBand a Ethernet sítě protokoly SFTP, SCP, NFS, CIFS a nástrojem SSHFS. Zároveň budou tyto servery využity pro efektivní a rychlé přenosy dat mezi souborovými datovými úložišti SCRATCH a HOME a to tak, že budou klienty obou filesystémů (Lustre a CXFS).

LFSaccess a Lustre MDS i OSS servery pro SCRATCH jsou reprezentovány [redacted] SGI Rackable ISS3104-RP10 1U servery s Intel Xeon E5 procesorem, DDR3 pamětí s ECC, dvěma interními disky, dvěma 1GE porty, čtyřmi 6Gbit SAS porty, dvěma FDR InfiniBand porty, redundantními zdroji a ventilátory. Servery mají vzdálený síťový management formou BMC kontroléru, který poskytuje mimo jiné i grafickou konzoli a připojení virtuálních médií.

Lustre OST disková pole pro SCRATCH jsou reprezentována [redacted] SGI Infinite Storage 5600 storage systémy se dvěma redundantními kontroléry, každý vybaven 12GiB zrcadlenou cache chráněnou baterií po dobu připadného přesunu obsahu cache na flash paměť a pamětí s ECC. Systémy disponují osmi 6Gbit SAS host porty, čtyřmi 1GE management porty, redundantními zdroji s ventilátory, 60-bay a 12-bay diskovými policemi se SAS disky.

Lustre MDT diskové pole pro SCRATCH tvoří SGI Infinite Storage 5100 storage systém se dvěma redundantními kontroléry, každý vybaven 2GiB zrcadlenou cache chráněnou baterií po dobu případného přesunu obsahu cache na flash paměť a pamětí s ECC. Systém disponuje čtyřmi 12Gbit SAS host porty, čtyřmi 1GE management porty, redundantními zdroji s ventilátory, 24-bay diskovou policí se SAS disky.

Nabízené řešení je flexibilní jak ke zvyšování kapacit na úrovni diskových polí připojených k OSS serverům, tak i ke zvyšování propustnosti celé Lustre sestavy. Řešení je tedy velmi dobře připravené na případné tváření vlastností podle reálného provozu, podle vývoje workflow uživatelů datového úložiště a jejich potřeb.

1.8.4 Datové úložiště infrastruktury

Datové úložiště infrastruktury je určeno pro potřeby zadavatele. Datové úložiště infrastruktury poskytuje diskový prostor blokovým protokolem a umožňuje rozdělení datové kapacity na logické části požadované velikosti a zpřístupnění těchto logických částí pouze na vybrané servery dle budoucích požadavků zadavatele. Datové úložiště infrastruktury je duálními cestami zpřístupněno virtualizačním serverům Virtualizační infrastruktury, k některým serverům řešení dodavatele, k

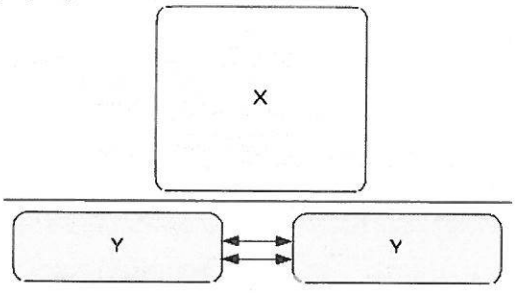
fyzickým serverům určených pro běh specifických služeb zadavatele (tzv. Dalším serverovým systémům) a je využitelné virtuálními servery ve Virtualizační infrastruktuře. Datové úložiště infrastruktury poskytuje funkcionalitu snapshotů a kopií.

Datové úložiště infrastruktury poskytuje dvě nezávislé diskové kapacity:

1) Diskovou datovou kapacitu X o čisté velikosti splňující požadavek na min. kapacitu 270TB
   a) která splňuje požadavek na dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1 MiB minimálně 5 GB/s ($5x10^9$ byte/s) a
   b) která splňuje požadavek na dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 75%/25% minimálně 20 kIOPs
2) Diskovou datovou kapacitu Y o čisté velikosti splňující požadavek na min. kapacitu 50TB
   a) která splňuje požadavek na dlouhodobě udržitelnou agregovanou rychlost sekvenčních operací pro velikost bloku 1 MiB minimálně 1 GB/s ($1x10^9$ byte/s) a
   b) která splňuje požadavek na dlouhodobě udržitelný výkon I/O operací náhodného charakteru o velikosti bloku 4KiB v režimu čtení/zápis 75%/25% minimálně 3 kIOPs
   c) která je synchronně replikovaná na dvě fyzicky oddělená umístění (v různých racích) a zajišťuje dostupnost dat i v případě výpadku zařízení v jednom umístění.

Disková datová kapacita X je realizována na jiném úložném zařízení než Disková datová kapacita Y. Disková datová kapacita X a Disková datová kapacita Y nesdílí žádné diskové pole. Datové úložiště infrastruktury poskytuje rovněž souborové služby. Souborové služby Datového úložiště infrastruktury jsou dostupné všem serverům řešení Velkého clusteru (fyzickým i virtuálním). Souborové služby Datového úložiště infrastruktury budou poskytovat souborový systém s uživatelským aplikačním vybavením (s aplikacemi) Přístupovým, Výpočetním, Vizualizačním a dalším serverům, poskytují obrazy centrálního úložiště instalačních obrazů pro instalaci serverů, poskytují sdílená data, sdílený souborový systém dalším službám běžícím v clusteru serverů. Řešení je redundantní dle požadavků bodu SPEC_71 Zadávací dokumentace – souborové služby jsou poskytovány i v případě výpadku či odstávky libovolného jednoho serveru zajišťujícího službu.

Datové úložiště infrastruktury je pro dosažení požadovaného výkonu a možnosti replikace navrženo jako dvě symetrická disková pole pro datovou oblast Y s funkcionalitou RVM (Remote Volume Mirroring přes FC porty), VolumeCopy a SnapShot, plus jedno diskové pole pro datovou oblast X s funkcionalitou VolumeCopy a SnapShot. Tato bloková disková pole poskytují diskový prostor prostřednictvím FibreChannel protokolu a redundantními cestami v síti SAN mohou být zpřístupněna všem serverům do této sítě zapojeným.



Součástí sestavy Datového úložiště infrastruktury, schematicky znázorněné na obrázku níže jsou následující komponenty:

1. Blokové Diskové pole pro kapacitu X (1ks)

[redacted]

2. Blokové Diskové pole pro kapacitu Y (2ks)

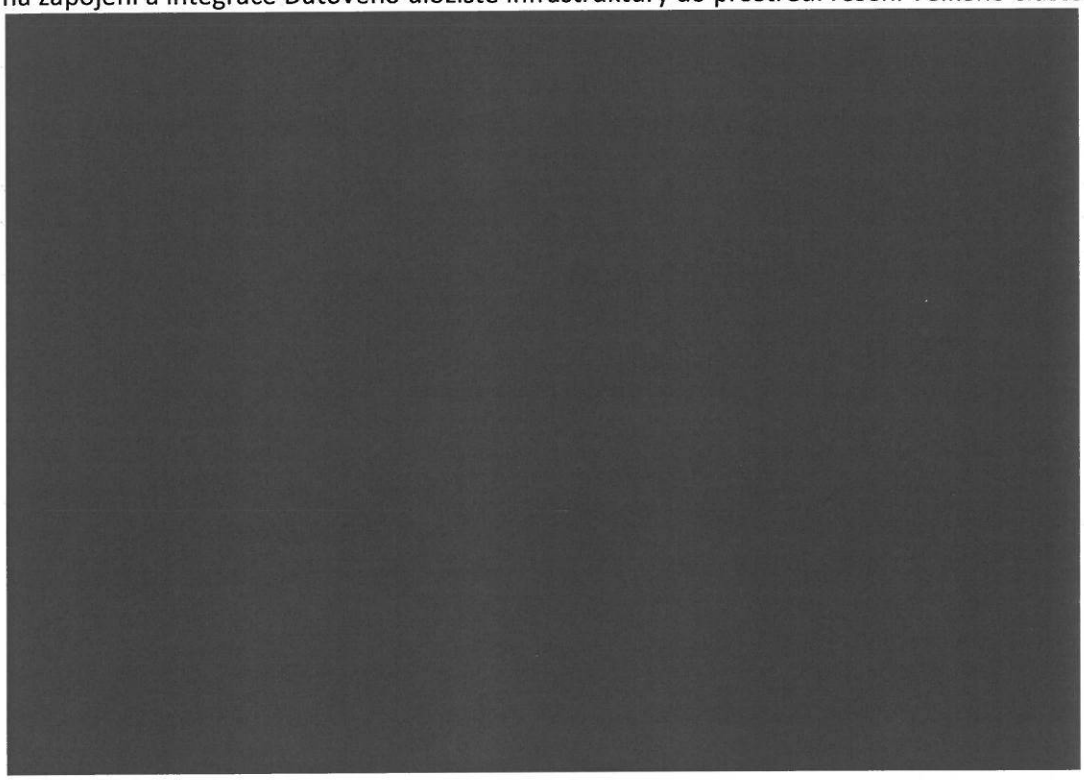
[redacted]

3. FILE access servery (2ks)

[redacted]

4. Prvky síťové infrastruktury pro zajištění komunikace a RVM replikace
5. Další potřebné příslušenství nezbytné k řádnému provozu sestavy datového úložiště (napájecí kabely, adaptéry, propojovací kabely, případně další nezbytné síťové i jiné komponenty).

Schéma zapojení a integrace Datového úložiště infrastruktury do prostředí řešení Velkého clusteru:



Obě bloková Disková pole pro datovou oblast Y jsou připojena třemi FC kabely z jednoho a třemi FC kabely z druhého kontroléru do redundantní 16Gbit FibreChanel SAN sítě, do které jsou rovněž připojeny Virtualizační servery, Další serverové systémy, DMF/CXFS servery a Zálohovací servery. Blokové Diskové pole je reprezentováno dvěma systémy SGI Infinite Storage 5000 se dvěma redundantními kontroléry, každý je vybaven 2GiB zrcadlenou cache chráněnou baterií po dobu případného přesunu obsahu cache na flash paměť a pamětí s ECC. Každý systém IS5000 disponuje osmi 8Gbit FibreChannel host porty, dvěma 1GE management porty, redundantními zdroji s ventilátory, třemi 24-diskovými policemi, které jsou osazeny rychlými 10K 2.5" SAS disky.

Třetí blokové Diskové pole pro datovou oblast X je připojeno dvěma FC kabely z jednoho a dvěma FC kabely z druhého kontroléru do redundantní 16Gbit FibreChanel SAN sítě, do které jsou rovněž připojeny Virtualizační servery, Další serverové systémy, DMF/CXFS servery a Zálohovací servery. Blokové Diskové pole je reprezentováno systémem SGI Infinite Storage 5600 se dvěma redundantními kontroléry, každý je vybaven 12GiB zrcadlenou cache chráněnou baterií po dobu případného přesunu obsahu cache na flash paměť a pamětí s ECC. Celkem systém disponuje osmi 16Gbit FibreChannel host porty, dvěma 1GE management porty, redundantními zdroji s ventilátory, šesti 60-diskovými policemi, které jsou osazeny 7200 RPM NL-SAS disky.

Souborový přístup k datovému úložišti INFRASTRUKTURY je realizován pomocí běžných NAS (NFS,CIFS,...) protokolů, dostupných v redundantní síti LAN a Výpočetní síti InfiniBand, prostřednictvím páru x86-64 serverů v Active/Passive HA konfiguraci, pro zajištění vysoké dostupnosti této služby. Tyto servery jsou reprezentovány dvěma SGI Rackable ISS3104 1U servery s Intel Xeon E5 procesorem, DDR3 pamětí s ECC, dvěma interními disky v RAID 1, dvěma 1GE porty, dvěma 10GE porty, dvěma 16Gbit FibreChannel porty, dvěma FDR InfiniBand porty, redundantními zdroji a ventilátory. Servery mají vzdálený síťový management formou BMC kontroléru, který poskytuje mimo jiné i grafickou konzoli a připojení virtuálních médií.

1.8.5 Disková pole SGI® InfiniteStorage

SGI® InfiniteStorage 5000

SGI® InfiniteStorage 5000 je RAID storage systém, který poskytuje vyspělé vlastnosti a konfigurační flexibilitu, aby uspokojil širokou škálu požadavků na uložení dat s výkonem střední třídy a příznivou cenou. SGI InfiniteStorage 5000 poskytuje zákazníkům lepší výkon i škálovatelnost, více-protokolovou uživatelskou konektivitu, flexibilní diskovou podporu, ochranu dat a vyspělé technologie, které šetří elektrickou energii.

SGI® InfiniteStorage 5000 nabízí připojení hostujícího počítače, buď pomocí SAS 2.0, 8Gb Fibre Channel nebo iSCSI připojení, v závislosti na zvolené SAN infrastruktuře. Diskové pole může být osazeno jedním řadičem nebo dvěma řadiči. Každý řadič disponuje v základu dvěma fixními 6Gb SAS porty a slotem pro HIC kartu prostřednictvím které můžeme přidat další SAS, FC nebo iSCSI porty. K dispozici jsou následující konfigurace dual-kontroléru: 4x 6Gb SAS, 8x 6GB SAS, 4x 6Gb SAS + 8x 8Gb FC nebo 4x 6Gb SAS + 8x 1Gb iSCSI.

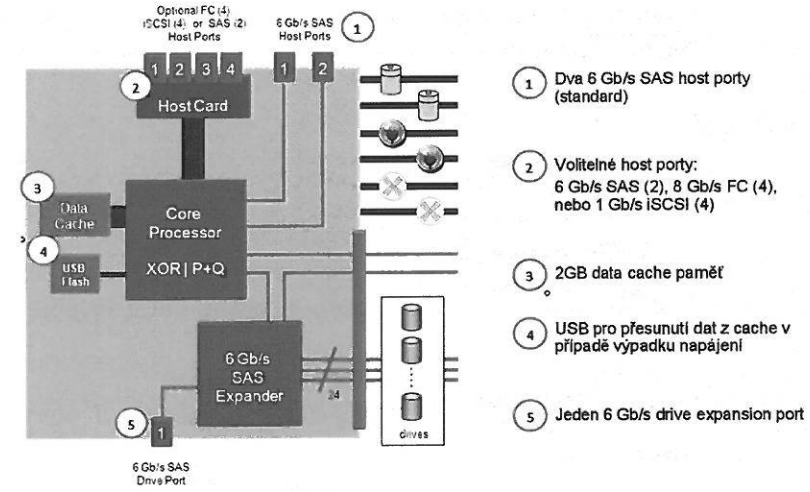
Typ pole	**SGI InfiniteStorage 5000**
Zapojení řadičů	2 redundantní řadiče, active-active
Cache velikost	4 GiB zrcadlená a zálohovaná baterií
Rozhraní pro připojení serverů	6Gb SAS (4/8 portů), 8Gb FC (8 portů), 10GE iSCSI (4 porty),
Partitions	128 max. (2 v základu)
Výška v racku	2 nebo 4 U - dle použité diskové police

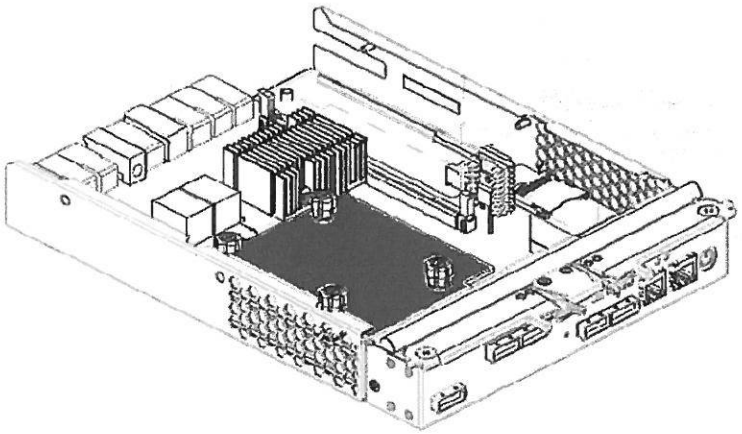
Vlastnosti řadičů	
RAID úrovně	0, 1, 3, 5, 6, 10 and Dynamic Disk Pools
Cache zajištění	Cache zálohovaná baterií je přesunuta na Flash v případě výpadku proudu
LUNy	256 LUN / partition, 512 celkem
Počet globálních hot-spare disků	Neomezený
Napájení a chlazení	Dvojí, redundantní, hot swap
Počet podporovaných disků celkem	192 ve 24-iř, 180 v 60-ti a 192 ve 12-ti diskových policích
Volitelné vlastnosti řadičů	
Max snapshots	1024
Volume Copy	podporuje
Remote Volume Mirroring	podporuje

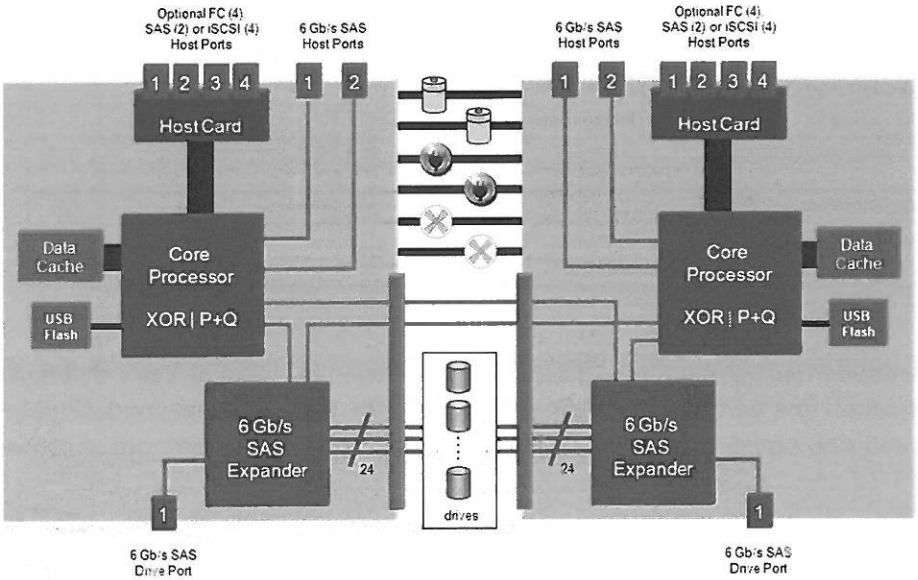
Plně redundantní komponenty, automatický fail-over (přepínání) cest, dynamické rekonfigurace a schopnost on-line údržby, jsou zde navrženy tak, aby bylo zajištěno, že vaše data jsou k dispozici 24 hodin 365 dní v roce. Veškeré redundantní komponenty, které jsou pro zajištění nepřetržité služby důležité (disky, řadiče, komunikační rozhraní, větráky, zdroje) jsou vyměnitelné za chodu (hot swapable). Přístup ke všem hot-plug diskům je díky zdvojení cest mezi oběma řadiči a disky také redundantní. Diskové pole umožňuje umístění do standardního racku. Cache diskového pole je zabezpečena proti ztrátě nebo poškození dat při výpadku napájení baterií, která zajistí napájení po dobu automatického překopírování obsahu cache na flash. Zrcadlení obsahu cache do druhého řadiče

pro případ poruchy jednoho z řadičů je možné vypnout, nebo jen nastavit, že v případě poruchy baterie se zrcadlení automaticky vypne. Read cache prefetch automaticky rozpoznává sekvenční nebo „nějak" sekvenční požadavky na čtení a čte dopředu i nepožadovaná data, u kterých předpokládá, že budou následně požadována.

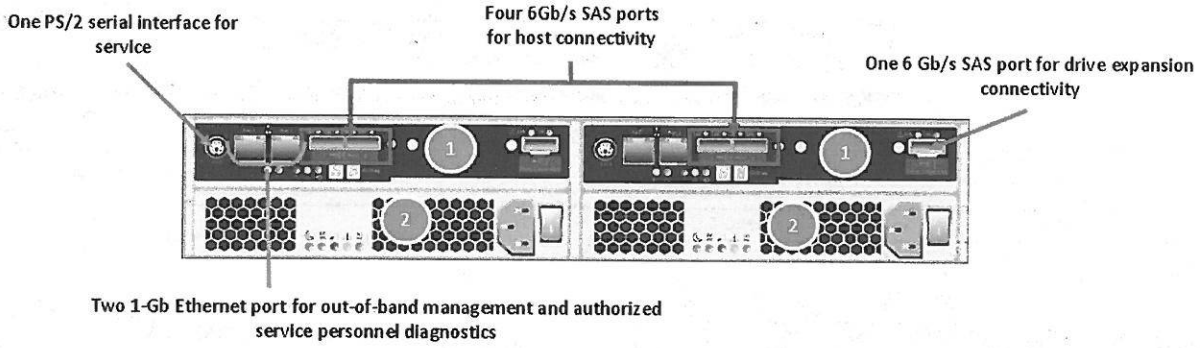
Následující tří schémata ukazují rozložení komponent v modulu IS5000 řadiče. Na zadní straně modulu je konektor pro připojení na midplane diskové police. Konektory pro připojeni dalších diskových polic, management ETH a SAN sítě jsou dostupné z přední části modulu.







Následující schéma čelní strany dual-controller modulů zobrazuje z leva dva Ethernet konektory s možností připojeni pole do dvou nezávislých sítí pomocí kterých je možné toto diskové pole konfigurovat a administrovat. Čtyři 6Gb SAS konektory tvoří základní host interface. Jeden a jeden 6Gb 4x SAS porty v pravo slouží k připojení dalších diskových polic.



SGI® InfiniteStorage 5100

SGI IS5100 je nejnovějším modelem blokových diskových polí SGI Infinite Storage řady 5X00. Je navrženo tak, aby vyhovovalo současným i budoucím náročným požadavkům, nabízelo trvalý výkon, umožnilo multi-dimenzionální škálovatelnost a prokázalo absolutní spolehlivost. Vyvážený výkon a schopnost vyniknout ve smíšené pracovní zátěži umožňuje jejich nasazení jak na podporu transakčních aplikací, jako jsou databáze a OLTP, tak i na propustnost náročné aplikace pro HPC a multimédia. SGI InfiniteStorage 5100 je prvním úložným systémem od SGI, který nabízí SAS 3.0 technologii o rychlosti 12 Gb/s. SGI InfiniteStorage 5100 poskytuje zákazníkům lepší výkon i škálovatelnost, více-protokolovou uživatelskou konektivitu, flexibilní diskovou podporu, ochranu dat a vyspělé technologie, které šetří elektrickou energii. Tento úložný systém nabízí bezkonkurenční přizpůsobivost měnícím se požadavkům zákazníků. Lineární škálovatelnost, online expanze a dynamické rekonfigurace jsou jen náznakem možností tohoto pole.

SGI® InfiniteStorage 5100 nabízí připojení hostujícího počítače, buď pomocí SAS 2.0/3.0, 16Gb Fibre Channel nebo připravovaného 10GE iSCSI rozhraní, v závislosti na zvolené SAN infrastruktuře. Diskové pole může být osazeno jedním nebo dvěma řadiči. Každý řadič disponuje v základu dvěma fixními 12Gb SAS porty a slotem pro HIC kartu prostřednictvím které můžeme přidat další dva/čtyři 12Gb SAS, dva/čtyři 16Gb FC nebo čtyři 10GE iSCSI porty. K dispozici jsou následující konfigurace dual-kontroléru: 4x 12Gb SAS, 8x 12GB SAS, 12x 12GB SAS, 4x 12Gb SAS + 4x 16Gb FC, 4x 12Gb SAS + 8x 16Gb FC nebo 4x 12Gb SAS + 4x 10Gb iSCSI.

Plně redundantní komponenty, automatický fail-over (přepínání) cest, dynamické rekonfigurace a schopnost on-line údržby, jsou zde navrženy tak, aby bylo zajištěno, že vaše data jsou k dispozici 24 hodin 365 dní v roce. Veškeré redundantní komponenty, které jsou pro zajištění nepřetržité služby důležité (disky, řadiče, komunikační rozhraní, větráky, zdroje) jsou vyměnitelné za chodu (hot swapable). Přístup ke všem hot-plug diskům je díky zdvojení cest mezi oběma řadiči a disky také redundantní. Diskové pole umožňuje umístění do standardního racku. Cache diskového pole je zabezpečena proti ztrátě nebo poškození dat při výpadku napájení baterií, která zajistí napájení po

dobu automatického překopírování obsahu cache na flash. Zrcadlení obsahu cache do druhého řadiče pro případ poruchy jednoho z řadičů je možné vypnout, nebo jen nastavit, že v případě poruchy baterie se zrcadlení automaticky vypne. Read cache prefetch automaticky rozpoznává sekvenční nebo „nějak" sekvenční požadavky na čtení a čte dopředu i nepožadovaná data, u kterých předpokládá, že budou následně požadována.

Typ pole	**SGI InfiniteStorage 5100**
Zapojení řadičů	2 redundantní řadiče, active-active
Cache velikost	4/8/16 GiB zrcadlená a zálohovaná baterií
Rozhraní pro připojení serverů	12Gb SAS (4/8/12 portů), 16Gb FC (4 nebo 8 portů), 10GE iSCSI (4 porty),
Partitions	128 v základu
Výška v racku	2 nebo 4 U - dle použité diskové police

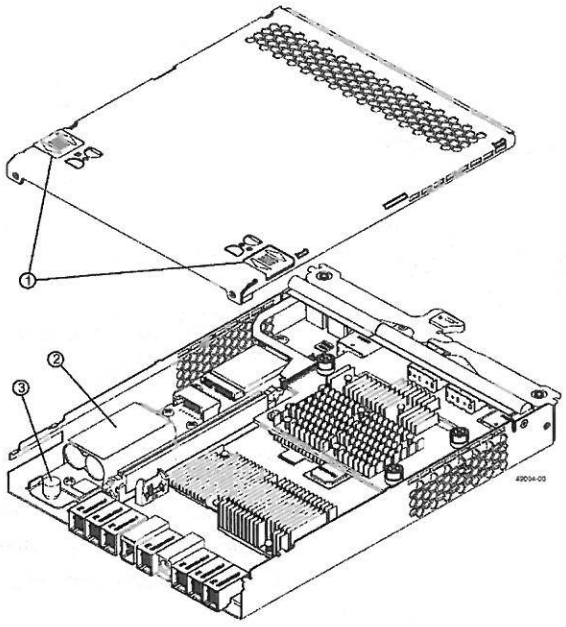
Vlastnosti řadičů	
RAID úrovně	0, 1, 3, 5, 6, 10 and Dynamic Disk Pools
Cache zajištění	Cache zálohovaná baterií je přesunuta na Flash v případě výpadku proudu
LUNy	256 LUN / partition, 512 celkem
Počet globálních hot-spare disků	Neomezený
Napájení a chlazení	Dvojí, redundantní, hot swap
Počet podporovaných disků celkem	192 ve 24-iř, 180 v 60-ti a 192 ve 12-ti diskových policích
Volitelné vlastnosti řadičů	
Max snapshots	1024
Volume Copy	podporuje
Remote Volume Mirroring	podporuje

IS5100 řadičový modul

je nejnovější generací RAID platforem s LSI SAS 3108 ("Invader") dual core ROC procesorem. Jeho moderní design umožňuje dodávat vysoký výkon pro širokou škálu IOPS a bandwidth intenzivních aplikací. Tento řadičový modul je možné používat ve všech třech typech diskových polic – 12-pozicové, 24-pozicové i 60-pozicové.

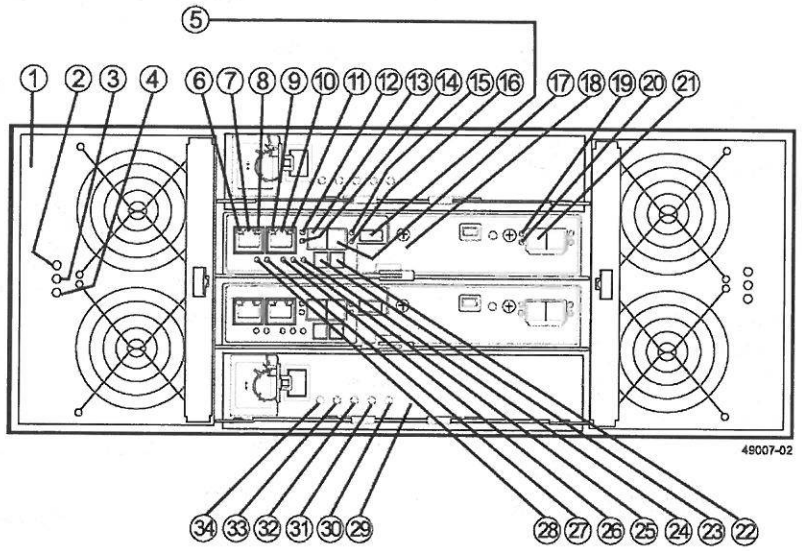
Následující obrázek ukazuje rozložení komponent v modulu IS5600 řadiče. Na zadní straně modulu je konektor pro připojení na midplane diskové police. Konektory pro připojeni dalších diskových polic, management ETH a SAN sítě jsou dostupné z přední části modulu.

1 – Lid
2 – Battery PCB
3 – Battery screw
4 – Memory DIMM
5 – IS5100 controller
6 – HIC Subplate
7 – Flash Drive
8 – HIC (SAS version)



Následující obrázek čelní strany řadičového modulu zobrazuje z leva dva Ethernet konektory s možností připojeni pole do dvou nezávislých sítí pomocí kterých je možné toto diskové pole konfigurovat a administrovat. Service display zobrazuje ID diskové police v rámci celého diskového pole a v případě poruchy číselný kód chyby. Čtyři 12Gb SAS konektory uprostřed tvoří základní + rozšiřující host interface. Čtyři 12Gb 4x SAS porty v pravo slouží k připojení dalších diskových polic.

1. Fan Canister
2. Fan Power LED
3. Fan Service Action Requir LED
4. Fan Service Action Allowed LED
5. USB Connector
6. Ethernet Link 1 Active LED
7. Ethernet Connector 1
8. Ethernet Link 1 Rate LED
9. Ethernet Link 2 Active LED
10. Ethernet Connector 2
11. Ethernet Link 2 Rate LED
12. Host Link 1 Fault LED
13. Base Host SFF-8644 Conn 1
14. Host Link 1 Active LED
15. Host Link 2 Fault LED
16. Host Link 2 Active LED
17. Base Host SFF-8644 Conn 2
18. Controller A Canister
19. Expansion Fault LED
20. Expansion Active LED
21. Expansion SFF-8644 Port Conn
22. Second Seven-Segment Display
23. First Seven-Segment Display
24. Cache Active LED
25. Ctrl A Service Action Req LED
26. Ctrl A Service Action Allowed LED
27. Battery Service Action Required LED
28. Battery Charging LED
29. Power Canister
30. Power AC Power LED
31. Power Service Action Required LED
32. Power Service Action Allowed LED
33. Power DC Power LED
34. Power Standby Power LED



SGI® InfiniteStorage 5600

SGI IS5600 je navrženo tak, aby vyhovovalo současným i budoucím náročným požadavkům, nabízelo trvalý výkon, umožnilo multi-dimenzionální škálovatelnost a prokázalo absolutní spolehlivost. Vyvážený výkon a schopnost vyniknout ve smíšené pracovní zátěži umožňuje jejich nasazení jak na podporu transakčních aplikací, jako jsou databáze a OLTP, tak i na propustnost náročné aplikace pro HPC a multimédia.

Tento úložný systém nabízí bezkonkurenční přizpůsobivost měnícím se požadavkům zákazníků. Lineární škálovatelnost, online expanze a dynamické rekonfigurace jsou jen náznakem možností tohoto pole. IS5600 lze také přizpůsobit vývoji síťové infrastruktury prostřednictvím výměnné karty hostitelského rozhraní, tak zvané HIC karty (Host Interface Card). Např. diskové pole IS5600 s nabízenou základní konfigurací osmi 16Gb Fibre Channel portů lze přidáním HIC karty rozšířit o osm 6Gb SAS portů, nebo je možné později dodat host karty v té době dostupné (např. iSCSI 10Gbit, atd...)

Plně redundantní komponenty, automatický fail-over (přepínání) cest, dynamické rekonfigurace a schopnost on-line údržby, jsou zde navrženy tak, aby bylo zajištěno, že vaše data jsou k dispozici 24 hodin 365 dní v roce. Veškeré redundantní komponenty, které jsou pro zajištění nepřetržité služby důležité (disky, řadiče, komunikační rozhraní, větráky, zdroje) jsou vyměnitelné za chodu (hot swapable). Přístup ke všem hot-plug diskům je díky zdvojení cest mezi oběma řadiči a disky také redundantní. Diskové pole umožňuje umístění do standardního racku. Cache diskového pole je zabezpečena proti ztrátě nebo poškození dat při výpadku napájení baterií, která zajistí napájení po dobu automatického překopírování obsahu cache na flash. Zrcadlení obsahu cache do druhého řadiče pro případ poruchy jednoho z řadičů je možné vypnout, nebo jen nastavit, že v případě poruchy

baterie se zrcadlení automaticky vypne. Read cache prefetch automaticky rozpoznává sekvenční nebo „nějak" sekvenční požadavky na čtení a čte dopředu i nepožadovaná data, u kterých předpokládá, že budou následně požadována.

Typ pole	**SGI InfiniteStorage 5600**
Zapojení řadičů	2 redundantní řadiče, active-active
Cache velikost	24GiB zrcadlená a zálohovaná baterií
Rozhraní pro připojení serverů	6Gb SAS (8 portů), 16Gb FC (8 portů)
Partitions	512 v základu
Výška v racku	2 nebo 4 U - dle použité diskové police

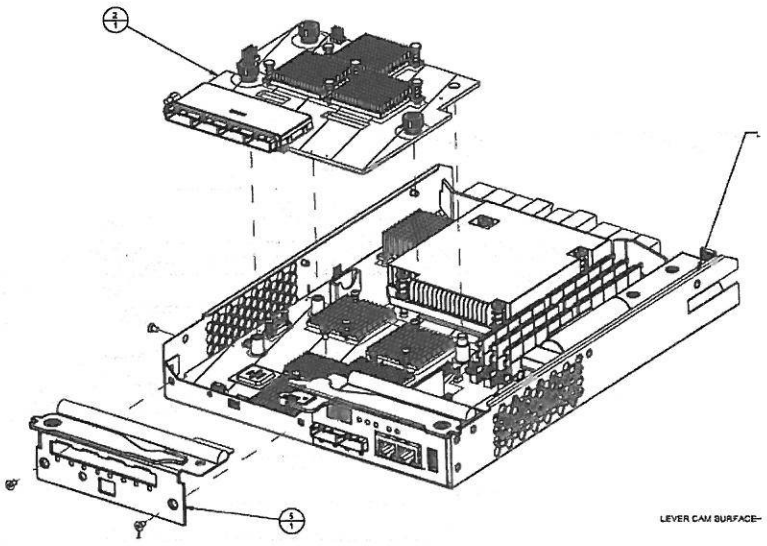
Vlastnosti řadičů	
RAID úrovně	0, 1, 3, 5, 6, 10
Cache zajištění	Cache zálohovaná baterií je přesunuta na Flash v případě výpadku proudu
LUNy	256 LUN / partition, 2048 celkem
Počet globálních hot-spare disků	Neomezený
Napájení a chlazení	Dvojí, redundantní, hot swap
Počet podporovaných disků celkem	348 ve 24-iř, 360 v 60-ti a 192 ve 12-ti diskových policích
Volitelné vlastnosti řadičů	
Max snapshots	1024
Volume Copy	podporuje
Remote Volume Mirroring	podporuje

IS5600 řadičový modul

je nejnovější generací RAID platforem se 4-core Intel Sandy Bridge procesorem. Jeho moderní design umožňuje dodávat vysoký výkon pro širokou škálu IOPS a bandwidth intenzivních aplikací. Tento řadičový modul je možné používat ve všech třech typech diskových polic – 12-pozicové, 24-pozicové i 60-pozicové.

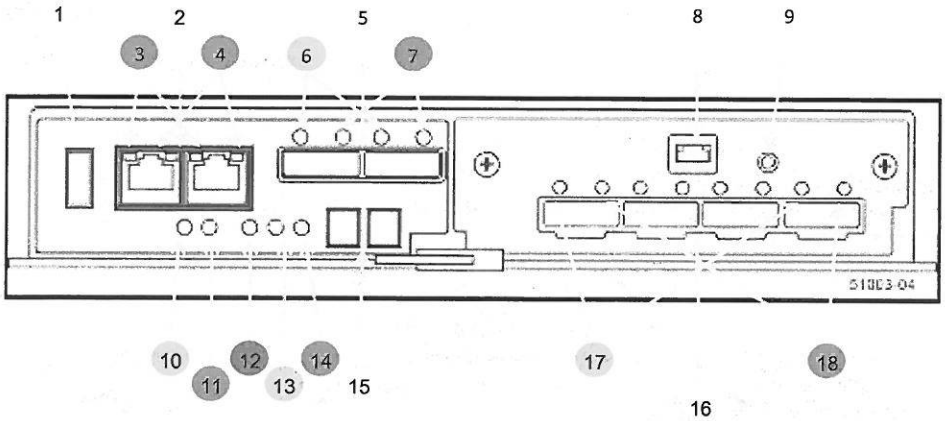
Následující obrázek ukazuje rozložení komponent v modulu IS5600 řadiče. Na zadní straně modulu je konektor pro připojení na midplane diskové police. Konektory pro připojeni dalších diskových polic, management ETH a SAN sítě jsou dostupné z přední části modulu.

1 – HIC (SAS version)
2 – IS5600 Controller
3 – Base Canister
4 – Memory DIMM (x3)
5 – Backup Battery
6 – HIC Subplate
7 – Flash Drive
8 – Lid (not shown)



Následující obrázek čelní strany řadičového modulu zobrazuje z leva dva Ethernet konektory s možností připojení pole do dvou nezávislých sítí pomocí kterých je možné toto diskové pole

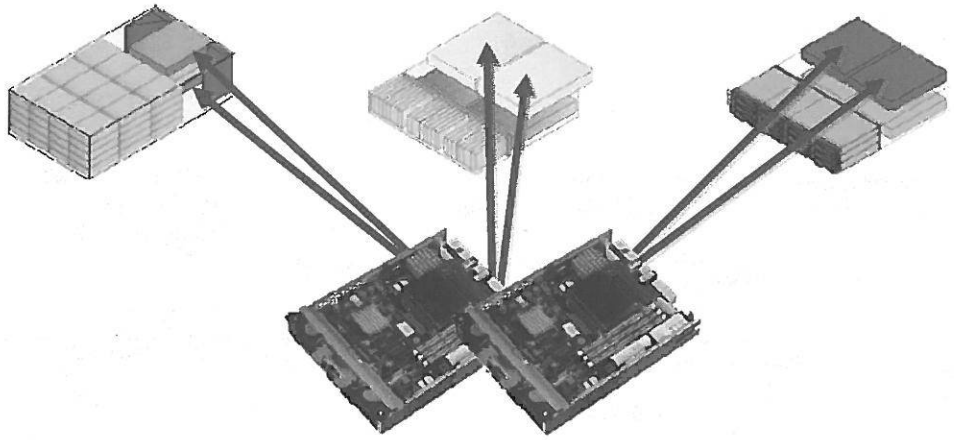
konfigurovat a administrovat. Service display zobrazuje ID diskové police v rámci celého diskového pole a v případě poruchy číselný kód chyby. Čtyři 16Gb FC konektory vpravo tvoří základní + rozšiřující host interface. Dva 6Gb 4x SAS porty uprostřed slouží k připojení dalších diskových polic.



1 – USB
2 – Ethernet (x2)
3 – Link Rate (green)
4 – Link Activity (green
5 – Expansion Ports 6Gb SAS (x2)
6 – Fault (amber)
7 – Activity (green)
8 – Mini USB RS232 Diagnostics Port
9 – Reset button (FW reset)
10 – Battery Fault (amber)
10 – Battery Fault (amber)
11 – Battery Charging (green)
12 – Service Action Allowed (blue)
13 – Controller Fault (amber)
14 – Cache Active (green)
15 – Seven Segment Status Display
16 – Host Ports 16Gb FC (x4)
17 – Fault (amber)
18 – Activity (green)

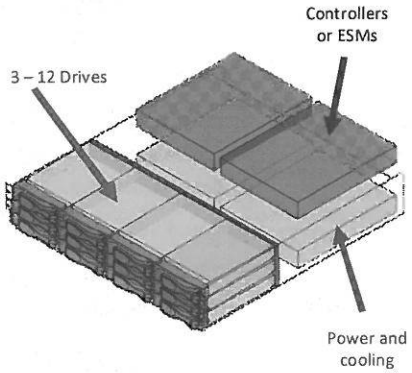
Diskové police SGI Infinite Storage 5x00 řady

Všechna disková pole IS5X00 řady umožňují kombinovat 12, 24 a 60-ti diskové police a provozovat v nich jak disky s vysokým výkonem typu SAS a SSD, tak i vysokokapacitní NL-SAS disky, což dělá z těchto polí ideální platformu pro široké spektrum poptávaných typů datových úložišť. **Disková pole využitá pro Home/Scratch/Infrastrukturní úložiště jsou tak založena na shodných komponentech, což má velmi pozitivní vliv na servisovatelnost a administraci.**



12-ti pozicová disková police

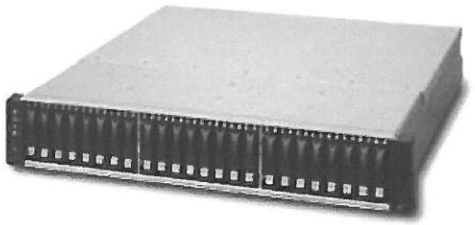
– umožňuje konfigurovat až dvanáct 3,5" disků typu SAS s 15000 ot/min nebo typu NL-SAS s 7200 ot/min. a kapacitou aktuálně až 4TB. Dva redundantní napájecí zdroje s chlazením a dva redundantní ESM (enviromental system monitoring) moduly zajišťují vysokou dostupnost a spolehlivost této diskové police.

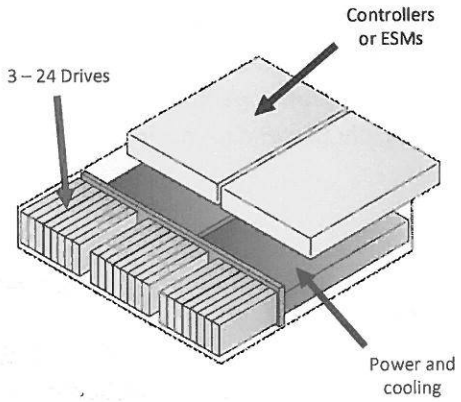


12-pozicová disková police		
Výška v racku	2 U	
Maximální váha	59.52 lbs (27 kg)	
Max. AC napájení	399W s kontroléry	276W bez kontrolerů
Max. tepelné vyzařování	1366 BTU/h s kontroléry	945 BTU/h bez kontrolérů

24-pozicová disková police

– nabízí dvacetčtyři pozic pro 2,5" disky typu SSD MLC, SAS 10000 ot/min a nižších kapacit (do 1.2TB). Dva redundantní napájecí zdroje s chlazením a dva redundantní ESM moduly zajišťují vysokou dostupnost a spolehlivost této diskové police.



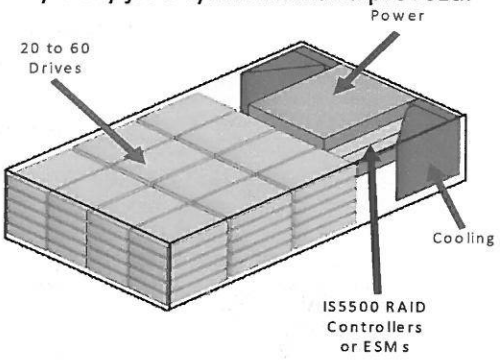


24-pozicová disková police		
Výška v racku	2 U	
Maximální váha	57.32 lbs (26 kg)	
Max. AC napájení	330W s kontroléry	240W bez kontrolerů
Max. tepelné vyzařování	1127 BTU/h s kontroléry	821 BTU/h bez kontrolérů

60-pozicová disková police

- je navžena tak aby dosahovala maximální hustoty dat na jednotku v racku. Může hostovat jak 3,5" tak i 2,5" disky jak typu SSD, SAS i NL-SAS, různých kapacit i otáček. Pět za provozu lehce vysunutelných šuplíků (drawers) umožňuje bezproblémový přístup ke všem diskům. Dva redundantní

napájecí zdroje, dva redundantní ventilátory a dva redundantní ESM moduly zajišťují vysokou dostupnost a spolehlivost této diskové police. Všechny disky jsou vyměnitelné za provozu.



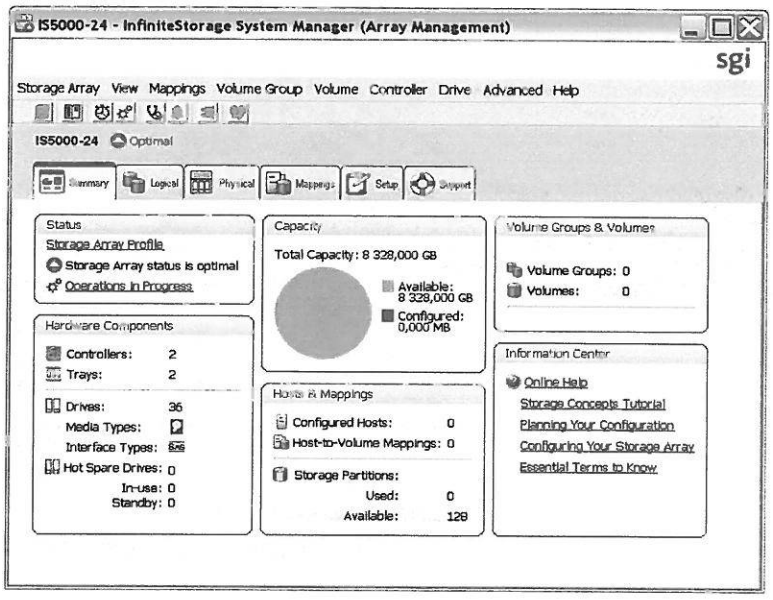
60-pozicová disková police		
Výška v racku	4 U	
Maximální váha	232.00 lbs (105.2 kg)	
Max. AC napájení	1222W s kontroléry	820W bez kontrolerů
Max. tepelné vyzařování	4180 BTU/h s kontroléry	2799 BTU/h bez kontrolérů

SW pro správu a monitoring SGI InfiniteStorage 5X00 řady

InfiniteStorage System Manager (ISSM)

ISSM je softwarový balík nástrojů pro, konfiguraci, správu, monitoring a diagnostiku diskových polí. Babízí jak command line interface (CLI) tak i intuitivní grafické uživatelské rozhranní (GUI), které zjednodušuje inicializaci diskového pole, umožňuje jednoduché rozšiřování kapacity, dynamické rozšiřování/změnu úrovní RAID group 0, 1, 3, 5, 6 a 10, on-line firmware update a další úkoly. Ty jsou mnohem rychlejší právě s využitím management softwaru a jeho grafickým rozhraním. Je možné jej nainstalovat jak na OS Linux tak Windows. Komunikace s kontroléry diskových polí je možná jak out-of-band metodou po ethernetu, tak i in-band metodou přes FibreChanel.

Ilustrativní obrázek níže zobrazuje Summary stavu a konfigurace jednoho z diskových polí.



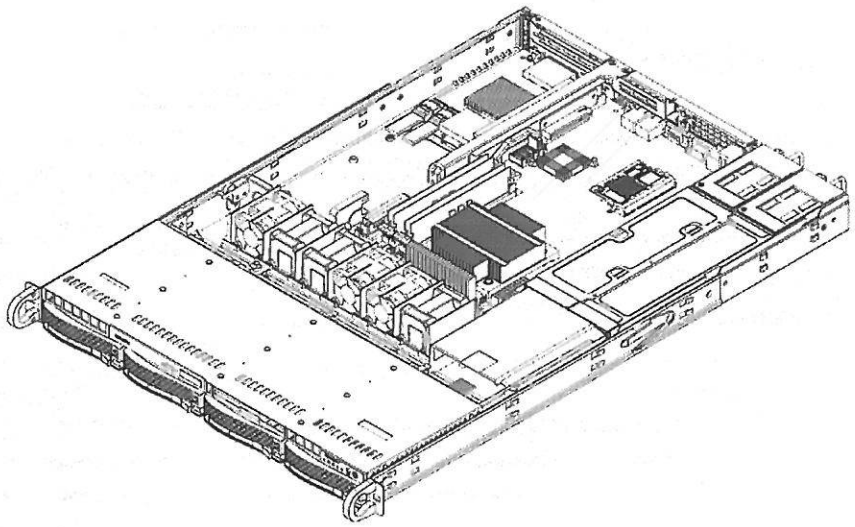
Tabulka podporovaných Premium features Keys v ISSM diskových polí IS5100 a IS5600

	IS5100	IS5600
Included Premium Feature Keys		Hyper Performance (Turbo) SANShare 512, SAS "SSD" Disk Drives SAS "FDE" Disk Drives
Optional Premium Feature Keys	Hyper Performance (Turbo) SANShare 512, SAS "SSD" Disk Drives SAS "FDE" Disk Drives SSD Cache Advanced software feature to unlock between 241-300 slots Advanced software feature to unlock between 301-384 slots Volume Copy Snapshot Consistency Group Thin Provisioning Checkpoint Asynchronous Mirroring (ARVM) Legacy Synchronous Mirroring (RVM)	SSD Cache Advanced software feature to unlock between 241-300 slots Advanced software feature to unlock between 301-384 slots Volume Copy Snapshot Consistency Group Thin Provisioning Checkpoint Asynchronous Mirroring (ARVM) Legacy Synchronous Mirroring (RVM)

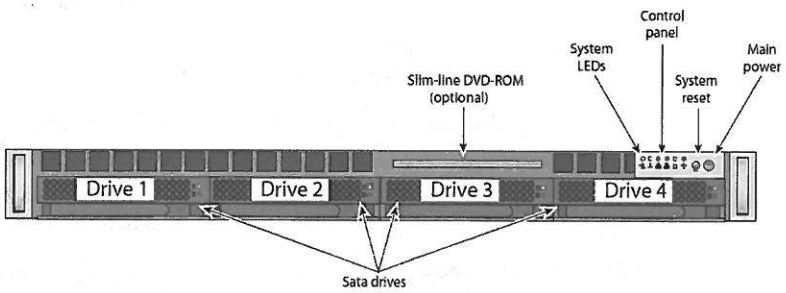
1.8.6 Server SGI Rackable ISS3104-RP10

Jedná se o standardní storage server, který může sloužit k různým účelům. Bývá ve storage řešeních často využíván jako Lustre OSS node, CXFS MDS node, CXFS Edge server, DMF pDMO node, nebo jako všeobecný NAS node. Jedná se o 1U server osaditelný do standardního 19" serverového stojanu. Tento 1U server doporučujeme osadit 3,5GHz 6-core (12-thread) CPU, 128GB paměti, 1-2 disků plus rozhraní pro I/O konektivitu potřebnou pro konkrétní nasazení. V takovéto konfiguraci dosahuje díky vysoké frekvenci CPU bez nutnosti sahat do paměti druhého socketu velice dobrých výkonových parametrů.

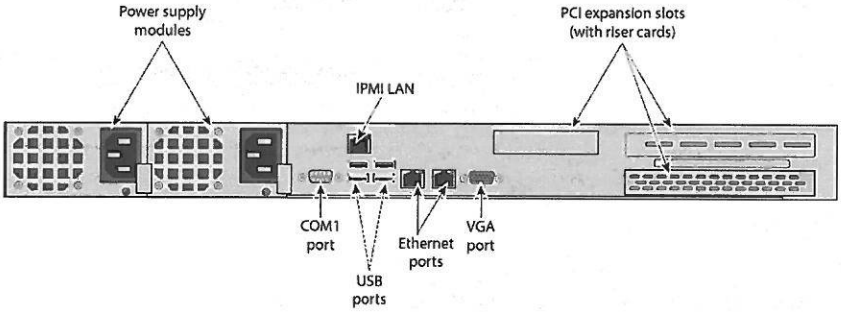
Schematické znázornění a rozložení jednotlivých komponent uvnitř storage serveru Rackable™ ISS3104-RP10



Pohled na uzel ISS3104-RP10 server ze předu:



Pohled na uzel ISS3104-RP10 server ze zadu:



Vlastnosti ISS3104-RP10 serveru:

Procesor
- Jeden Intel® Xeon® Processor(s) E5-1600v2 Ivy Bridge - EP
- 8GT/s Intel® QuickPath Interconnect (Intel® QPI)
- Thermal Design Power (TDP) up to 130 W

Čipset
- Intel® Patsburg C600-A/D čipset

Paměť
- 8 DIMM slotů
- 4 paměťové kanály na CPU
- 2 DIMM slot na paměťový kanál
- Podpora 1866/1600/1333/1066 and 800 MT/s ECC Registered DDR3 Memory

Integrated I/O
- DB-15 Video konektor
- 2x RJ-45 NIC konektor 10/100/1000 Mb LAN
- 4x USB 2.0
- 2x AHCI SATA 6 Gbps konektor
- 4x AHCI SATA 3 Gbps konektor
- 2x full-height x16 PCIE Gen3 slots
- 1x low-profile x8 Gen2 PCIe slot

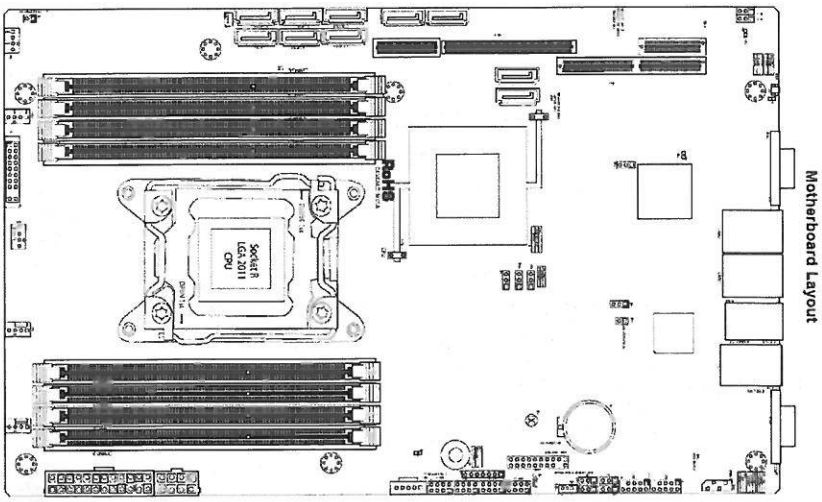
System Management
- Intel® Light-Guided Diagnostics on field replaceable units
- One IPMI 2.0 (RJ45 connector) Ethernet port

Integrovaný SAS kontroler
- Intel RSTe SW RAID 1

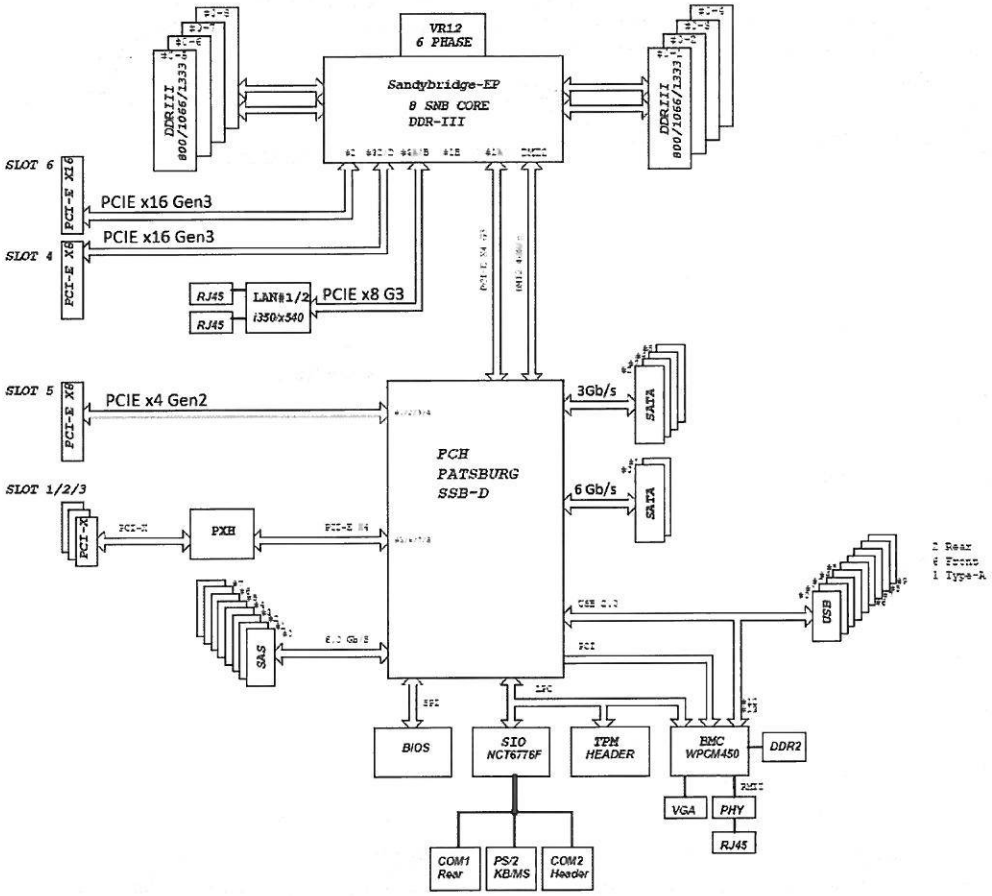
Integrated Serial ATA kontroler
- 2x AHCI SATA 6Gbps konektor
- 4x AHCI SATA 3Gbps konektor

Integrovaný LAN kontroler
- 2x RJ-45 Intel Ethernet Controller I350 1Gb porty

Základní deska serveru:



Funkční blokové schéma základní desky serveru ISS3104-RP10:



1.9 Infrastrukturní servery

Infrastrukturní servery a jejich infrastruktura jsou navrženy a dimenzovány tak, aby zajistily spolehlivý, bezpečný, rychlý a efektivní provoz Velkého clusteru.

Každý Infrastrukturní server realizovaný fyzickým serverem splňuje následující požadavky:

- Diskový řadič RAID, disky v RAID s redundancí dat
- Za provozu vyměnitelné (hot-swap) disky
- Redundantní, za provozu vyměnitelné (hot-swap) napájecí zdroje, redundantní napájení
- Konektivita Ethernetová síť

Všechny Infrastrukturní servery jsou řešeny pouze fyzickými servery.
Infrastrukturní servery jsou navrženy tak, že výpadek nebo odstávka libovolného jednoho Infrastrukturního serveru nezpůsobí přerušení probíhajících úloh a provozu Výpočetního clusteru, Přístupových serverů, Vizualizačních serverů, Souborových datových úložišť, Datového úložiště infrastruktury, Management serverů, Virtualizační infrastruktury, Dalších serverových systémů, Zálohování či Síťové infrastruktury.

Celkem bude dodáno 18 Infrastrukturních serverů v následujícím členění:

Infrastrukturní servery pro běh základních infrastrukturních služeb

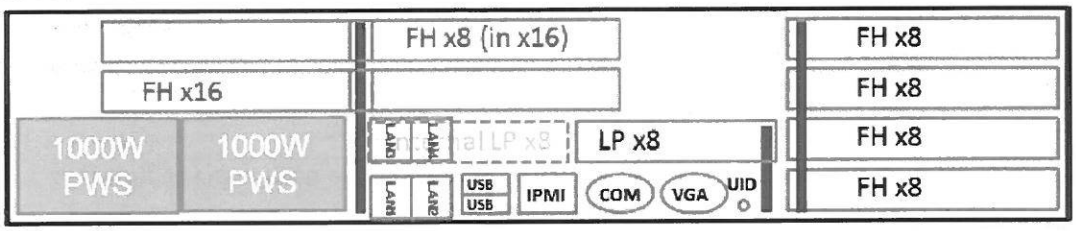
Návrh řešení zahrnuje **6 serverů** pro běh základních infrastrukturních služeb. Jedná se o servery zajišťující základní spolehlivou (aplikace HA funkcionality) serverovou platformu pro infrastrukturní služby výpočetního centra (řešení Velkého clusteru) tak, aby řešení bylo v maximální míře autonomní, nezávislé na externích systémech a službách a jeho chod byl spolehlivý bezpečný a efektivní. Mezi základní funkcionality provozované na těchto serverech patří zajištění DHCP, DNS, LDAP, služby licenčních serverů, běh plánovače úloh, monitorování, logování a ukládání vybraných dat o provozu, atd...).

Tyto servery budou řešeny prostřednictvím serveru SGI® Rackable™ C2112-GP2.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL) 4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2620v3, 8 jader, 2.4GHz
- Operační paměť RAM: 64GiB DDR4
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Ethernetová síť: 2x 1Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W
- Operční systém: 64-bitový operační systém s jádrem Linux RedHat 6.5

Infrastrukturní server - gateway IB-ETH

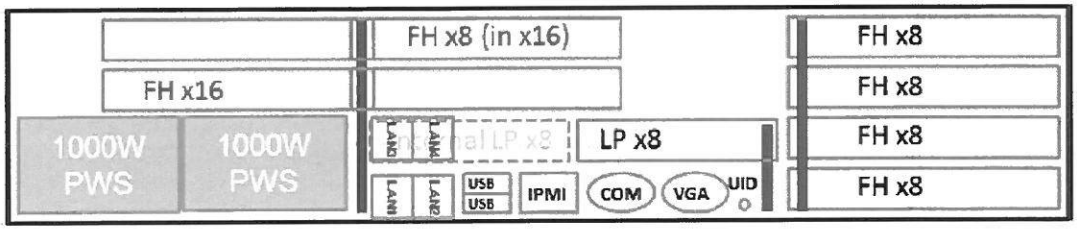
V návrhu řešení jsou **2 kusy** těchto serverů pro potřeby propojení interní FDR IB výpočetní sítě výpočetního clusteru a prostředí Ethernetové sítě.

Tyto infrastrukturní servery budou řešeny prostřednictvím serveru SGI® Rackable™ C2112-GP2.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL) 4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Gateway serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2620v3, 8 jader, 2.4GHz
- Operační paměť RAM: 64GiB DDR4
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Výpočetní síť: IB FDR 2x56Gb/s
- Konektivita Ethernetová síť: 2x 1Gb/s, 2x10Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W
- Operční systém: 64-bitový operační systém s jádrem Linux RedHat 6.5

Infrastrukturní server pro přístup a administraci výpočetního clusteru

Tyto infrastrukturní servery jsou v konfiguraci Velkého clusteru **dva** (pro zajištění dostupnosti) a primárně slouží pro administraci a provoz infrastruktury výpočetního clusteru.

Tnto server je zdroje SW instalací v clusteru a:

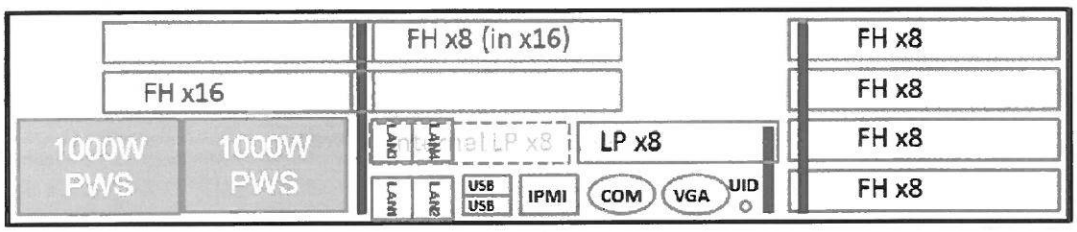
- poskytuje management SW nástroje
- poskytuje master boot image
- je primárním místem pro agregaci management dat z celého systému prostřednictvím komunikace s nižšími úrovněmi v hiearchické struktuře (rack leader nody, atd...)
- udržuje cluster databázi
- primární DNS pro interní infrastrukturu clusteru
- je primárním administrativním rozhraním

Tyto servery budou řešeny prostřednictvím serveru SGI® Rackable™ C2112-GP2.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL)
	4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Admin serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2620v3, 8 jader, 2.4GHz
- Operační paměť RAM: 64GiB DDR4
- Lokální disky: 2x 2TB SATA,RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Výpočetní síť: není
- Konektivita Ethernetová síť: 2x 1Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W
- Operční systém: 64-bitový operační systém s jádrem Linux RedHat 6.5

Rack Leader server výpočetního clusteru

Rack Leader serverů je v navrženém řešení **8 kusů**. Drží na sobě boot image (initrd a root fs) – z tohoto serveru bootují Výpočetní servery a mají na něm nesdílenou read-write systémovou oblast.

Rack Leader server:

- umožňuje vzájemnou zástupnost (redundanci) pro poskytované služby
- dostává a agreguje management data z hiearchicky podřízených komponent a poskytuje je dále do admin nodu
- sekundární DNS pro interní infrastrukturu clusteru
- zjednodušuje servisní model
- běží na něm IB fabric manager

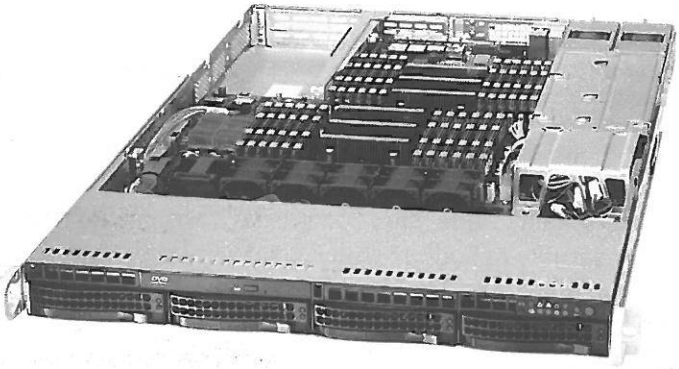
Rack Leader servery budou řešeny prostřednictvím serveru **6017R-N3RF4+.**

Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Může být v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server. Jedná se o 1U server osaditelný do standardního 19" serverového stojanu. Tento 1U server nabízí velkou flexibilitu osazení CPU, paměti, disků.

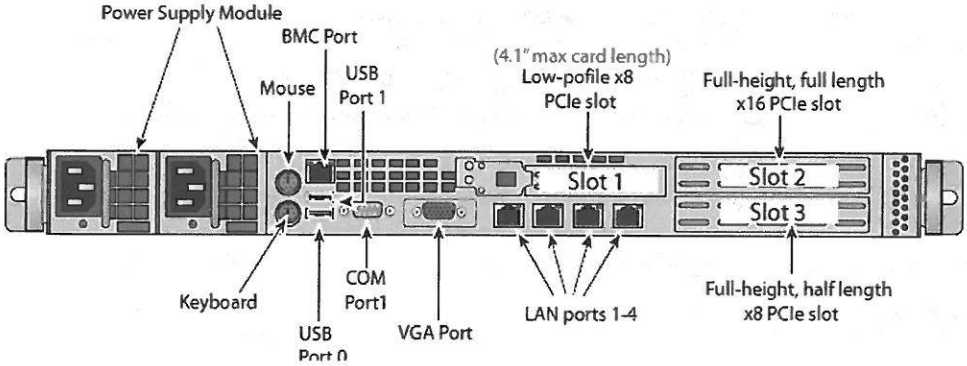
Vlastnosti serveru 6017R-N3RF4+:

Velikost:	1U
Chipset:	Intel® C606
Procesor:	2xIntel® Xeon® E5-2600 v2
Typ paměti:	1866 MHz DDR3 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	4x3,5"
Rozšiřující sloty:	1x PCIe 3.0 x8 (LP), 1x PCIe 3.0 x16, 1x PCIe 3.0 x8
Ethernet:	4x 1Gb/s onboard
Management:	IPMI
Zdroj:	2x750W redundantní zdroj

Pohled na server 6017R-N3RF4+ :



Pohled na server 6017R-N3RF4+ zezadu :



Detailní nabízená konfigurace jednoho Rack Leader serveru:

- Provedení: 1U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2603v2, 4 jádra, 1.8GHz
- Operační paměť RAM: 32GiB DDR4
- Lokální disky: 2x 1TB SATA, RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Výpočetní síť: 1x56Gb/s (pouze první HA dvojice těchto serverů v řešení)
- Konektivita Ethernetová síť: 4x 1Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 750W
- Operční systém: 64-bitový operační systém s jádrem Linux RedHat 6.5

1.10 Management servery

Velký cluster obsahuje **dva** vyhrazené Management servery určené pro správu Velkého clusteru.

Management servery budou poskytovat následující funkcionalitu:

a) Správu uživatelských účtů a skupin Velkého clusteru, zejména
   (a) Vytváření, modifikace a odstranění uživatele (uživatelského účtu)
   (b) Nastavení/změna hesla uživatele
   (c) Vytváření, modifikace a odstranění skupiny
   (d) Vkládání a odstraňování uživatele/uživatelů do/ze skupiny

b) Správu Souborových datových úložišť, zejména
   (a) Vytváření, modifikace a odstranění adresářů a souborů
   (b) Nastavení vlastníků a přístupových práv souborů
   (c) Nastavení uživatelských a skupinových kvót

c) Správu Výpočetních, Přístupových, Vizualizačních, Infrastrukturních a dalších serverů, zejména
   (a) Vzdálené vykonávání příkazů na serverech
   (b) Přenos souborů z a na servery

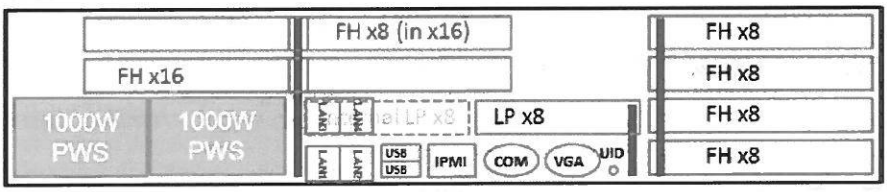
Uvedená funkcionalita bude plně, komplexně dostupná na dvou Management serverech současně.
Management servery nebudou řešeny virtuálními servery.
Řešení Management serverů bude poskytovat služby (zejména požadovanou funkcionalitu) i při výpadku či odstávce jednoho libovolného Management serveru.
Management servery budou poskytovat přístup administrátorům protokolem SSH2 a poskytovat služby pro přenos souborů SCP a SFTP.
Management servery budou vyhrazené, nebudou použity pro poskytování jiných služeb či zajištění další funkcionality než správa.

Management servery, jejich určení a detailní konfigurace:
Managament servery budou řešeny prostřednictvím serveru **SGI® Rackable™ C2112-GP2**.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL), 4x PCIe Gen 3.0 x8
Ethernet:	2x GigE onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Managament serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2620v3, 8 jader, 2.4GHz
- Operační paměť RAM: 64GiB DDR4
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Výpočetní síť: není
- Konektivita Ethernetová síť: 2x 1Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W
- Operační systém: 64-bitový operační systém s jádrem Linux RedHat 6.5

1.11 Virtualizační infrastruktura

Virtualizační infrastruktura bude poskytovat a zajišťovat běh virtuálních serverů.
Virtualizační infrastruktura bude určena pro potřeby zadavatele, virtualizační infrastruktura nebude využita k jiným účelům.
Virtualizační infrastruktura bude obsahovat 8 fyzických virtualizačních serverů a 1 fyzický server pro management Virtualizační infrastruktury.

Všechny Virtualizační servery splňují následující požadavky

- Fyzický server, architektura x86-64
- Minimálně 24 fyzických CPU jader na server
- Paměť RAM min. 256GiB
- CPU výpočetní výkon Rpeak serveru minimálně 1000Gflop/s bez využívání dočasného přetaktování procesorů či jiných podobných vlastností
- Redundantní konektivita Ethernetová síť min. 2x 10Gb/s
- Redundantní konektivita Datové úložiště infrastruktury min. 2x 8Gb/s
- Redundantní, hot-swap provedení

Všechny Virtualizační servery budou mít stejnou hardwarovou konfiguraci.
Virtualizační servery budou používat stejnou technologii procesorů jako Výpočetní servery.
Virtualizační servery budou fyzicky nezávislá zařízení, tj. nemají společnou komponentu.
Virtualizační servery budou rovnoměrně fyzicky rozděleny do 2 nezávislých racků.
Server pro management Virtualizační infrastruktury nebude sdílený s jiným serverem.
Virtualizační infrastruktura bude využívat jako datové úložiště Datové úložiště infrastruktury.
Virtualizační servery budou zapojeny do Ethernetové sítě.

Řešení virtualizace splňuje následující požadavky

- Podpora virtuálních 64 bitových (guest) serverů OS Linux a MS Windows
- Přenos virtuálních serverů mezi fyzickými servery za běhu (Live migration)
- Přenos obrazů virtuálních serverů mezi úložišti za běhu (Live Storage Migration)
- Automatický fail-over virtuálních serverů mezi fyzickými virtualizačními servery
- Automatické přesunutí virtuálních serverů mezi virtualizačními servery v případě nedostatku volných zdrojů virtualizačního serveru
- Automatické vypnutí/uspání virtualizačních serverů v případě velkého nadbytku volných výpočetních zdrojů, automatické zapnutí uspaného virtualizačního serveru v případě požadavku na zdroje
- Podpora přímého přístupu virtuálních serverů na disková bloková zařízení
- Grafické rozhraní pro správu Virtualizační infrastruktury
- Podpora automatické aktualizace virtualizačních serverů v clusteru
- Integrace do řešení Zálohování (viz ZD požadavky v kapitole 1.13 Zálohování)
- Rozdělení virtuální serverů do skupin s možností omezení dostupných zdrojů CPU a operační paměti.
- Ochrana běhu virtuálního serveru proti výpadku virtualizačního serveru bez nutnosti restartu virtuálního serveru
- Podpora IPv6 pro virtuální servery

Řešení virtualizace bude efektivně využívat zdroje Virtualizační infrastruktury, Virtualizační servery a Datového úložiště infrastruktury.

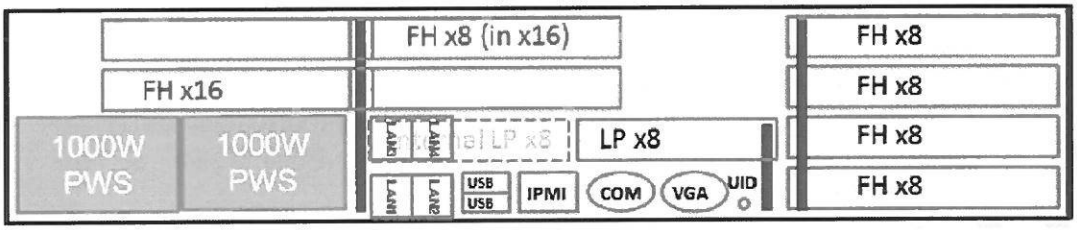
Řešení bude umožňovat přístup zadavatelem určených virtuálních serverů na Souborová datová úložiště HOME a SCRATCH, pro zajištění této funkcionality bude použit požadovaný interní export Souborových datových úložišť protokoly NFS a CIFS.
Virtualizační infrastruktura bude umožňovat současný běh minimálně 256 virtuálních serverů.

Počet serverů Virtualizační infrastruktury, jejich určení a detailní konfigurace:
Virtualizační servery a Management server Virtualizační infrastruktury budou řešeny prostřednictvím serveru **SGI® Rackable™ C2112-GP2**.
Jde o standardní a velmi univerzální serverový uzel, který může sloužit k různým účelům. Bývá v komplexních systémech využíván jako login node, head node, storage přístupový server, databázový server nebo jako server pro virtualizační infrastrukturu. Jedná se o 2U server osaditelný do standardního 19" serverového stojanu. Tento 2U server nabízí velkou flexibilitu osazení CPU, paměti, disků a zejména rozhraní pro I/O konektivitu.

Vlastnosti SGI® Rackable™ C2112-GP2:

Velikost:	2U
Chipset:	Intel® C612
Procesor:	2xIntel® Xeon® E5-2600 v3, maximálně 145W
Typ paměti:	2133 MHz DDR4 ECC reg.
Počet paměťových slotů:	24
Počet pozic pevných disků:	12x3,5"
Rozšiřující sloty:	2x PCIe Gen 3.0 x16 (FHFL) 4x PCIe Gen 3.0 x8 (2 FHHL, 2 low-profile)
Ethernet:	2x 1Gb/s onboard
Management:	IPMI 2.0
Zdroj:	2x1000W redundantní zdroj

Pohled na SGI® Rackable™ C2112-GP2 zezadu:



Detailní nabízená konfigurace jednoho Virtualizačních serveru:

- Provedení: 2U, Fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2695v3, 12 jader, 2.3GHz
- Operační paměť RAM: 256GiB DDR4
- Lokální disky: nejsou osazeny, pro boot použita USB flash paměť
- Diskový řadič RAID: pouze onboard diskový řadič
- Konektivita Výpočetní síť: 1x FDR port 56Gbit
- Konektivita Ethernetová síť: 8x 1Gb/s, 2x10Gb/s
- FC konektivita: 2x8Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W

Celkem bude dodáno 8 Virtualizačních serverů.

Detailní nabízená konfigurace Management serveru pro Virtualizační infrastrukturu:

- Provedení: 2U, fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2620v3, 6 jader, 2.4GHz
- Operační paměť RAM: 32GiB DDR4
- Lokální disky: 2x 300GB, 15krpm v RAID1, Hot-Swap
- Diskový řadič RAID: SAS HW RAID 1
- Konektivita Ethernetová síť: 4x 1Gb/s
- Zdroj: redundantní, za provozu vyměnitelný napájecí zdroj 1000W

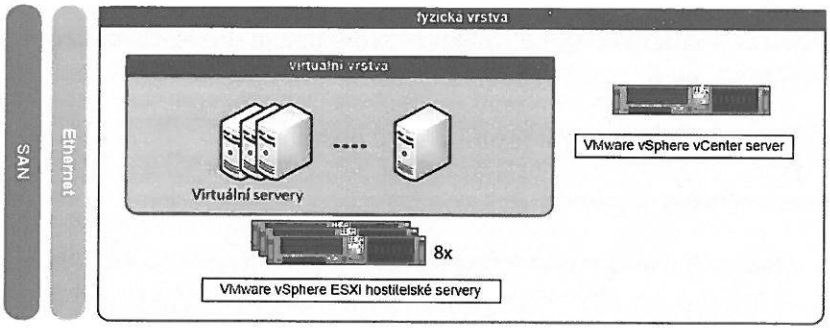
Celkem bude dodán 1 fyzický Management server Virtualizační infrastruktury.

Popis softwarového řešení Virtualizační infrastruktury:

Serverové virtualizační prostředí bude založeno na platformě VMware vSphere. Na 8 fyzických hostitelských serverech bude nainstalován VMware ESXi hypervizor. Pro správu virtuální infrastruktury bude sloužit dedikovaný VMware vCenter server běžící na samostatném fyzickém serveru.

Všechny virtualizační servery budou mít redundantní SAN FC konektivitu (2x 8 Gbit FC HBA) pro přístup k Datovému úložišti infrastruktury.

Hostitelské ESXi servery budou vybaveny 2x 10GE síťovým rozhraním ethernet pro redundantní připojení k síti a 8x 1GE rozhraním ethernet pro zajištění služeb samotné virtualizační infrastruktury.



VMware vSphere podporuje širokou paletu guest 64 bitových operačních systémů z rodiny Windows a řadu komerčních i volně šířených distribucí OS Linux.

Přenos virtuálních serverů mezi fyzickými servery za běhu (Live migration) bude zajištěn pomocí funkce vSphere vMotion s těmito charakteristikami:

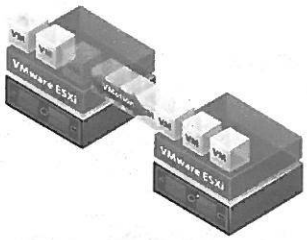
Umožňuje přesun běžících virtuálních strojů mezi fyzickými servery

Je nezávislé na operačním systému VM

Nedochází k výpadku při údržbě

Nabízí trvalou dostupnost služeb

Podporuje SAN, SCSI, NAS



Přenos obrazů virtuálních serverů mezi úložišti za běhu (Live Storage Migration) bude zajištěn pomocí funkce vSphere Storage vMotion s těmito charakteristikami:

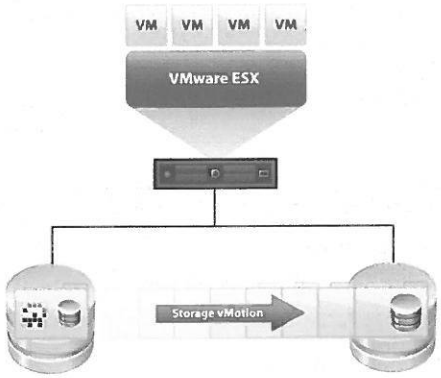
Umožňuje přesun běžících virtuálních strojů mezi fyzickými datovými úložišti

Nedochází k výpadku při údržbě

Podporuje iSCSI, FC a NFS úložiště

Podporuje přesun mezi různými typy datových VMFS storage

Umožňuje při migraci konvertovat diskový formát



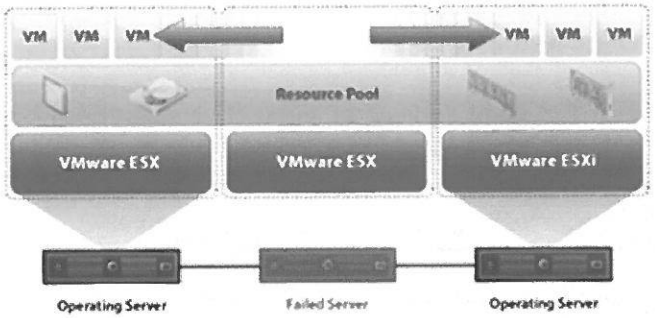
Automatický fail-over virtuálních serverů mezi fyzickými virtualizačními servery bude zajištěn pomocí funkce vSphere High Availability s těmito charakteristikami:

Umožňuje automatický restart virtuálních strojů pokud dojde k výpadku fyzického hostitelského serveru

Jedná se o nákladově efektivní řešení vysoké dostupnosti

Vhodné pro jakékoliv aplikace

Řízení vysoké dostupnosti (HA) je možné nastavit pro detekci výpadku, akcí pro obnovení. Sledovány mohou být i virtuální stroje a v případě jejich nedostupnosti je možné provést jejich restart.



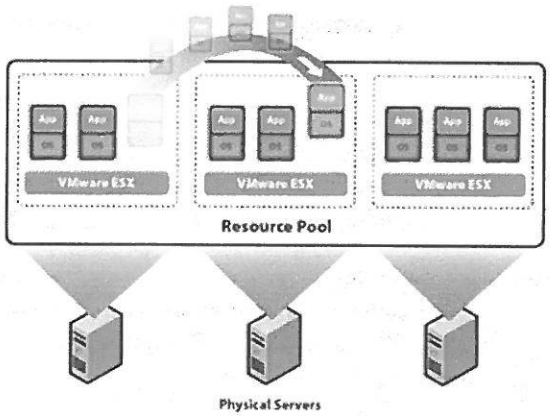
Automatické přesunutí virtuálních serverů mezi virtualizačními servery v případě nedostatku volných zdrojů virtualizačního serveru bude zjištěna funkcí vSphere DRS (Distributed Resource Scheduling) která:

Umožňuje dynamickou alokaci zdrojů a vyvažování zátěže hostitelských serverů

Vyrovnává dostupnost zdrojů s předdefinovanými prioritami

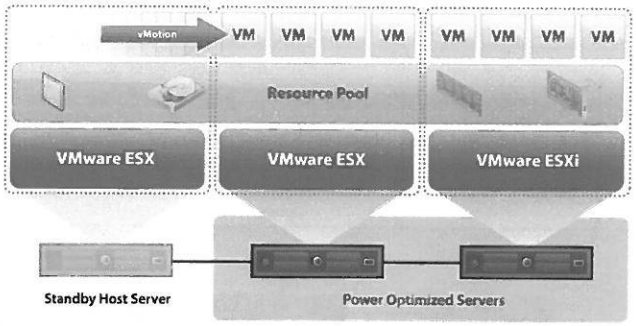
Dynamicky alokuje a vyrovnává výpočetní kapacity napříč hardwarovými zdroji

Trvale monitoruje utilizaci zdrojů a inteligentně alokuje dostupné zdroje



Automatické vypnutí/uspání virtualizačních serverů v případě velkého nadbytku volných výpočetních zdrojů, automatické zapnutí uspaného virtualizačního serveru v případě požadavku na zdroje bude zajištěno pomocí funkce VMware Distributed Power Management (DPM):

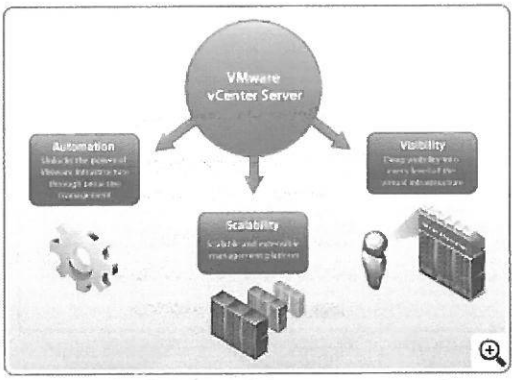
DPM monitoruje utilizaci CPU a paměti ve vSphere clusteru, a určuje, zda má být jeden nebo více ESXi serverů zapnuto či vypnuto, a to v závislosti na nastavení úrovně využití zdrojů. Když je využití zdrojů nízké, DPM vypne (resp. přepne do standby módu) jeden nebo více ESXi serverů, a opačně. Hardware, navrhovaný pro hostitelský server, funkci DPM podporuje.



Podpora přímého přístupu virtuálních serverů na disková bloková zařízení je zajištěno pomocí vlastnosti VMware vSphere tzv. "Raw Device Mapping" (RDM). Tato metoda zajišťuje přímý přístup virtuálních strojů na iSCSI nebo FC úložiště. Přístup je možný ve dvou módech: virtual compatibility mode a physical compatibility mode.

Grafické rozhraní pro správu Virtualizační infrastruktury bude zjištěno pomocí vSphere Web Client.vSphere Web Client nabízí možnost spravovat vSphere infrastrukturu pomocí podporovaného internetového prohlížeče.

Pro centrální správu hostitelských serverů a virtuálních strojů slouží VMware vCenter server, který umožňuje detailní správu všech prvků (serverů, clusterů, datastorů, virtuálních serverů, virtuální LAN struktury) z jednoho místa. Umožňuje automatizaci plánovaných kroků (přesun VM, vytváření nových VM) a zprostředkovává funkce HA, DRS, FT, DPM a další, popsané výše a níže.



Podpora automatické aktualizace virtualizačních serverů v clusteru bude zajištěna pomocí VMware vSphere Update Manager, který automatizuje správu aktualizací virtualizačních serverů a virtuálních strojů a dále:

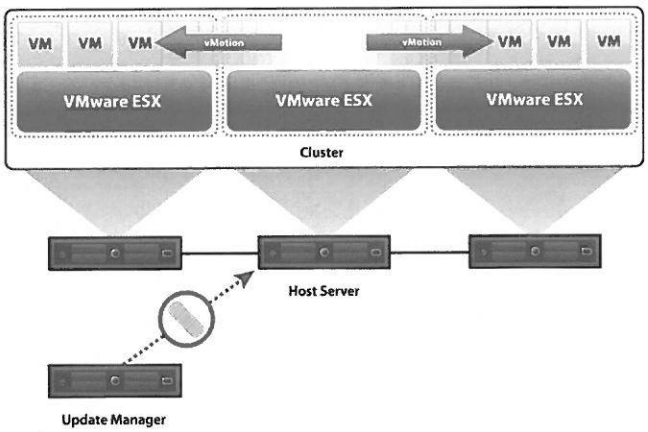
Porovnává aktuální stav ESXi hostitelských serverů s definovaným požadovaným stavem a aplikuje aktualizace a opravy, aby zajistil shodu.

Vizualizuje stav aktualizací v přehledném grafickém rozhraní

Zařazuje a plánuje aktualizaci jednotlivých hostitelských serverů

Pravidelně stahuje a ukládá aktualizace z webu

S využitím funkce DRS automaticky migruje virtuální stroje z ESXi serveru, na kterém probíhá aktualizace, a po provedení aktualizace migruje tyto virtuální stroje zpět.



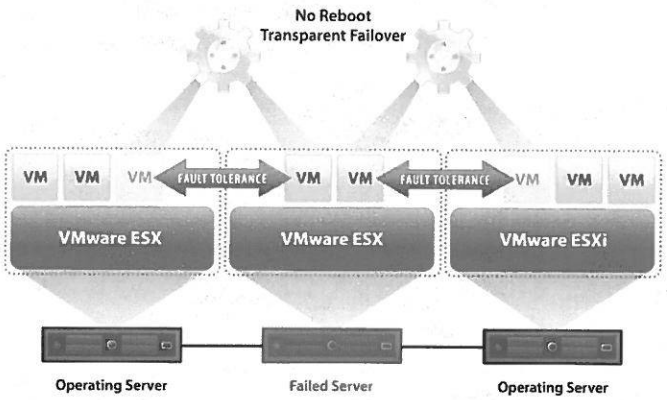
Integrace do řešení zálohování bude řešena pomocí softwarového produktu EMC NetWorker. Zálohování a obnova virtuálních serverů na úrovni virtualizační infastruktury na bázi produktu NetWorker Virtual Edition Client a splňuje všechny požadavky na zálohování virtualizované infrastruktury v Zadávací dokumentaci.

Rozdělení virtuální serverů do skupin s možností omezení dostupných zdrojů CPU a operační paměti je ve VMware vSphere zajištěno pomocí vytváření tzv. „Resource Pool": seskupení fyzických zdrojů (CPU, paměť a další) a jejich přiřazení skupinám virtuálních strojů. Resource pooly jsou hierarchické a umožňují tak další dělení zdrojů pro různé skupiny a různé účely.

Ochrana běhu virtuálního serveru proti výpadku virtualizačního serveru bez nutnosti restartu virtuálního serveru je zajištěna pomocí funkce vSphere Fault Tolerance (FT):

Jedná se o funkci odolnosti vůči výpadku, která zabezpečuje stálou dostupnost virtuálního počítače při poruše fyzického serveru. Povolením Fault Tolerance na vybraném virtuálním

počítači dochází k vytvoření kopie virtuálního počítače na jiném serveru pomocí technologie Lockstep. Všechny změny na primárním VM se přenášejí na kopi, a v případě výpadku je okamžitě k dispozici tato kopie, na kterou jsou přesměrovány klientské požadavky. Ihned po přepnutí na záložní kopii je vytvářena další záloha na dalším členu HA clusteru.



Virtualizační platforma VMware vSphere podporuje použití protokolu IPv6 ve virtuálních serverech. Podpora protokolu IPv6 je dále zahrnuta ve správě ESXi hostitelských serverů a vCenter serveru, a dále při přístupu ke zdrojům přes NFS a iSCSI, ve funkci vMotion atd.

Řešení virtualizace bude efektivně využívat zdroje Virtualizační infrastruktury, Virtualizační servery a Datového úložiště infrastruktury. VMware vSphere platforma je ze své podstaty určená k efektivnímu využití virtualizační infrastruktury a souvisejících komponent (Virtualizační servery, Datové úložiště). Tento požadavek je tedy naplněn volbou produktu VMware vSphere.

Přístup k Souborovým datovým úložištím HOME a SCRATCH pomocí protokolů NFS a CIFS je dán funkčností virtuálních serverů a je umožněn na úrovni Virtualizační infrastruktury připojením určeného virtuálního serveru k příslušné síti poskytující HOME a SCRATCH souborové úložiště.

Produkt VMware vSphere podporuje běh až 4000 virtuálních strojů v rámci clusteru, až 512 virtuálních strojů na jednom hostitelském ESXi serveru, a až 2048 běžících virtuáních strojů na datastore (při běhu v HA clusteru).

Pro zajištění všech zmíněných finkcinalit bude dodán pro všechny virtualizační servery potřebný počet licencí VMware vSphere 5 Enterprise a pro management server VMware vCenter Server 5 Standard for vSphere 5 a Microsoft® Windows® Server Standard 2012 R2.

1.12 Další serverové systémy

Součástí dodávky jsou rovněž další fyzické servery pro běh specifických služeb zadavatele (databázových, portálových a dalších) tzv. Další serverové systémy.

Velký cluster bude obsahovat 16 fyzických serverů kategorie Další serverové systémy. Každý server kategorie Další serverové systémy splňuje veškeré požadavky zadávací dokumentace.

Specifikace severů kategorie Další serverové systémy:

- 4x Server typ A - Operační paměť RAM minimálně 64GiB
- 4x Server typ B - Operační paměť RAM minimálně 128GiB
- 4x Server typ C - Operační paměť RAM minimálně 256GiB
- 4x Server typ D - Operační paměť RAM minimálně 512GiB

Polovina počtu serverů každého typu dle požadavků zadávací dokumentace bude vybavena redundantní konektivitou do Ethernetové sítě 2x10Gb/s, druhá polovina bude vybavena konektivitou do Ethernetové sítě 1x10Gb/s.

Další serverové systémy budou používat stejnou technologií procesorů a stejný operační systém jako Výpočetní servery.
Další serverové systémy jsou určeny výhradně pro využití zadavatelem.

Detailní nabízená konfigurace jednoho Dalšího serverového systému:

- Provedení: 2U, fyzický server architektury X86-64
- Procesor: 2x Intel Xeon E5-2695v3, 14 jader, 2.3GHz
  Celkový teoretický výpočetní výkon (Rpeak) serveru je 1030Gflop/s a to bez využívání dočasného přetaktování procesorů či jiných
- Operační paměť RAM: 64GiB DDR4 typ A, 128GiB DDR4 typ B, 256GiB DDR4 typ C, 512GiB DDR4 typ D
- Lokální disky: 2x 300GB, 15krpm disky v RAID1, Hot-Swap
- Diskový řadič RAID: SAS diskový řadič s funkcionalitou HW RAID
- Konektivita Výpočetní síť: 1x FDR port 56Gbit
- Konektivita Ethernetová síť: 2x 1Gb/s, 2x10Gb/s typ B a C, 2x 1Gb/s, 1x10Gb/s typ A a B
- FC konektivita: 2x8Gb/s
- Zdroj: redundantní, za provozu vyměnitelné napájecí zdroje 1000W
- Podpora operačních systémů: Linux a MS Windows
- Operační systém: 64-bitový operační systém s jádrem Linux CentOS 6.5

1.13 Zálohování

Součástí dodávky je také komplexní řešení zálohování dat – systém Zálohování. Účelem zálohování je disaster recovery.

Řešení Zálohování zajišťuje:

- zálohování všech dodávaných fyzických serverů včetně fyzických serverů určených pro běh specifických služeb zadavatele (tzv. Další serverové systémy) s výjimkou Výpočetních serverů; U Výpočetních serverů je instalace/reinstalace/obnova z připravených jednotných obrazů, zálohují se obrazy Výpočetních serverů.
- zálohování Virtualizační infrastruktury a virtuálních serverů; Zálohují se všechny virtuální servery a až 100 virtuálních serverů zadavatele.
- zálohování Souborového datového úložiště HOME; Souborové datové úložiště SCRATCH se nezálohuje.
- zálohování Datové úložiště infrastruktury; Datové úložiště infrastruktury se zálohuje na úrovni souborového systému realizovaném na logických částech Datového úložiště infrastruktury.

Řešení Zálohování umožňuje obnovu dat ze zálohy dat ne starší 1.5 dne. Zálohování dat bude probíhat s periodou maximálně 1 den a umožňuje/poskytuje obnovu dat z posledních 7 dnů, z posledních 7 denních záloh a dále umožňuje/poskytuje obnovu dat ze zálohy provedené před 2 až 4 týdny.

Priority zálohování a obnovy dat (pořadí dle priority od nejvyšší k nejnižší):

1. Nejvyšší priorita: Datové úložiště infrastruktury, replikovaná část Y
2. Vybrané virtuální servery zadavatele, vybrané fyzické servery určené pro běh specifických služeb zadavatele (tzn. vybrané servery kategorie Další serverové systémy)
3. Servery řešení nezbytné pro poskytování služeb systému, Virtualizační infrastruktura, Datové úložiště infrastruktury, nereplikovaná část X
4. Další servery
5. Nejnižší priorita: Souborové datové úložiště HOME

Doba obnovy Souborového datového úložiště HOME je dimenzována na maximálně 4 dny za předpokladu 90% obsazení celkové kapacity úložiště daty typickými pro HOME v prostředí HPC centra obdobné velikosti, a za předpokladu 5% změny kapacity úložiště denně a 20% změny kapacity úložiště za týden. Zálohování je řešeno tak, aby mělo minimální negativní dopad na provoz a výkon Velkého clusteru. Zálohovací systém efektivně využívá hardwarové prostředky.

Řešení pro zálohování a obnovu dat splňuje následující základní vlastnosti:

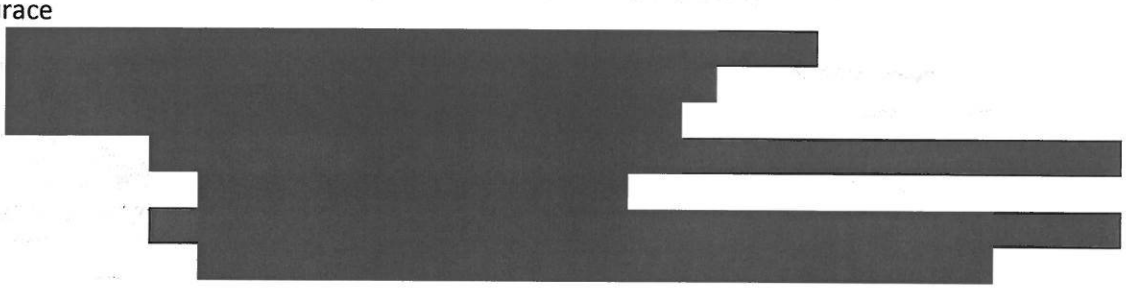
a) časový harmonogram záloh různého typu
b) granulární obnova individuálních souborů a složek
c) obnova vlastníků, práv a atributů souborů a složek
d) paralelní běh záloh a obnov

Řešení poskytuje zálohování virtuálních serverů na úrovni virtualizační infastruktury a možnost obnovy na úrovni souborového systému virtualizačních serverů/obrazů virtuálních serverů a rovněž na úrovni souborových systémů, souborů virtuálních serverů.

Řešení zálohování obsahuje páskovou knihovnu o celkové kapacitě 4PB (bez započítání komprese), součástí nabídky jsou datová média o nekomprimované kapacitě 3PB a čistící média v počtu dvojnásobku páskových mechanik.
Licence zálohovacího systému pokrývají všechny uvedené potřeby v maximální konfiguraci včetně variability provozovaného operačního systému Linux/Windows na serverech zadavatele (Další serverové systémy a virtuální servery zadavatele).

Paměti RAM všech serverů a řadičů diskových polí používají mechanismus detekce a opravy chyby - Error-correcting code memory (ECC). Běžný provoz a dostupnost deklarovaných kapacit zařízení nevyžaduje zásah obsluhy. Požadovaná datová kapacita páskové knihovny je dostupná bez jakéhokoliv manuální obsluhy zařízení - všechny pásky jsou umístěny ve slotech zařízení, čištění mechanik probíhá bez zásahu obsluhy. Řešení používá jednoznačnou identifikaci pásek (jedinečné čárové kódy). Všechna zařízení a systémy jsou spravovatelné vzdáleně. Servery mají vzdálený síťový management nezávislý na provozu operačního systému serveru poskytující grafickou konzoli a připojení virtuálních médií. Všechna zařízení jsou fyzicky označena jednoznačnou identifikací, která je na zařízeních snadno dostupná a čitelná, a vhodným způsobem evidována.

Konfigurace



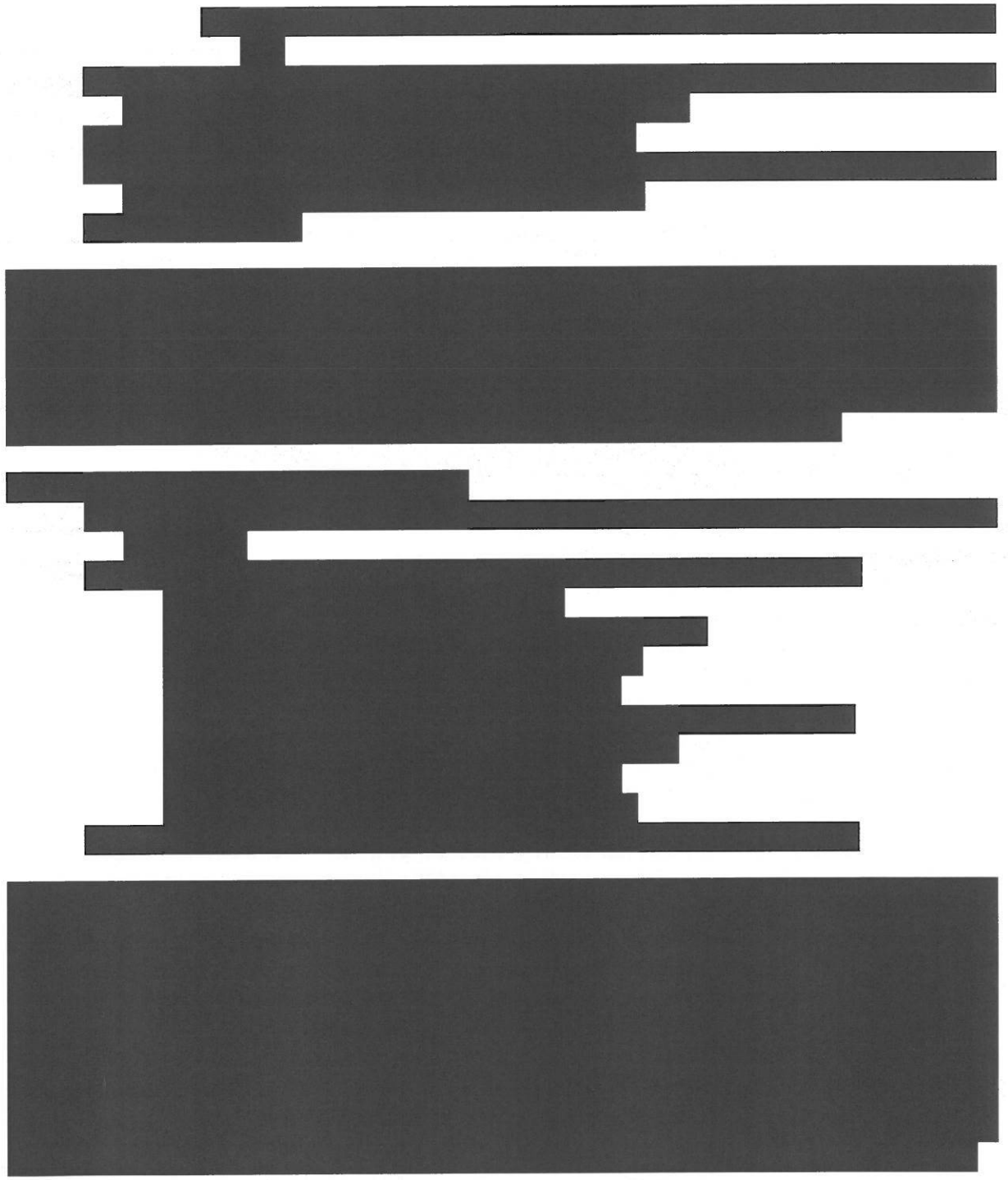
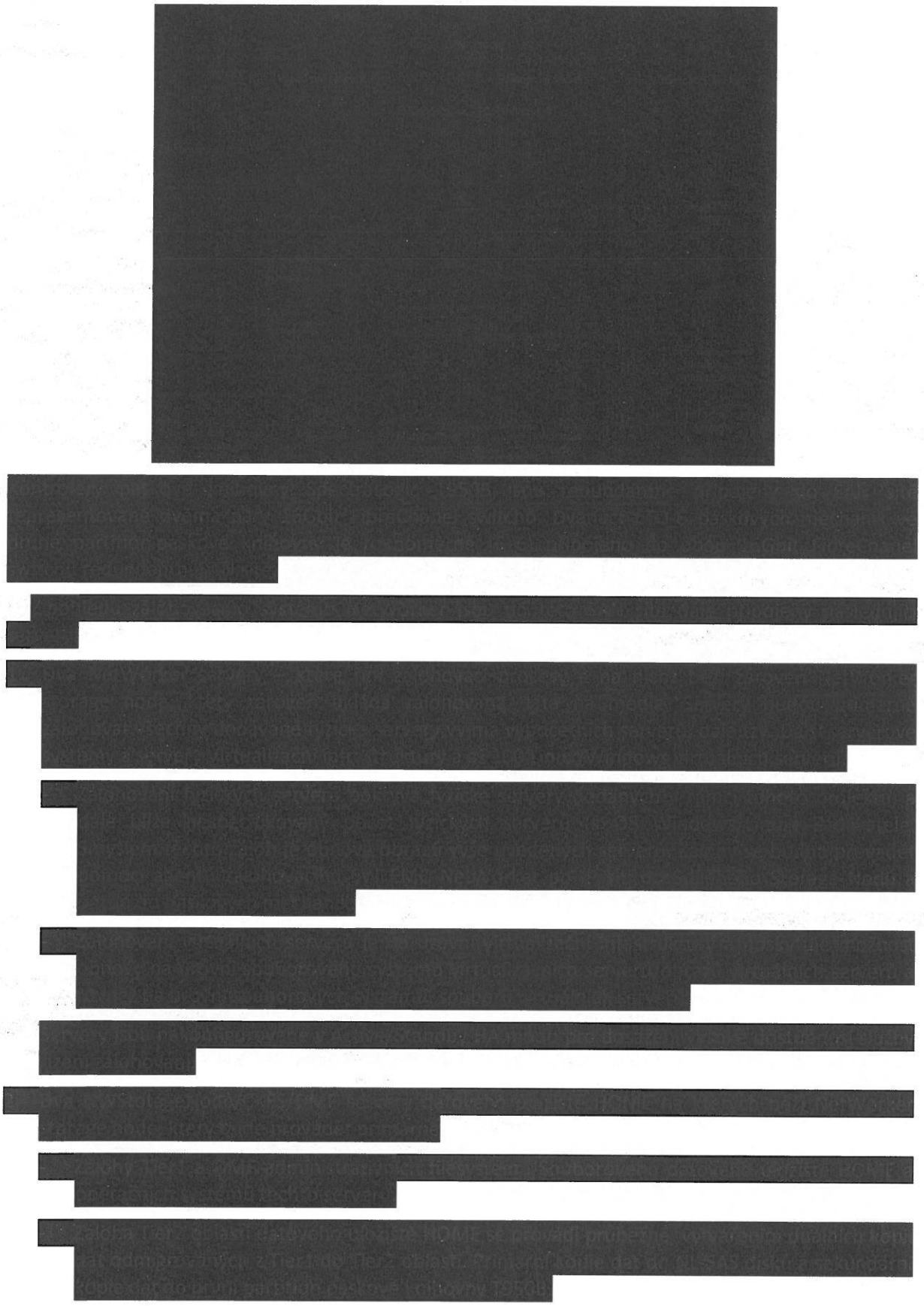


Schéma systému Zálohování:



[redacted]

Fyzické servery a tzv. Další serverové systémy

[redacted]

Virtualizační infrastruktura a virtuální servery

[redacted]

Souborového datového úložiště HOME

[redacted]

[redacted]

Datové úložiště INFRASTRUKTURY

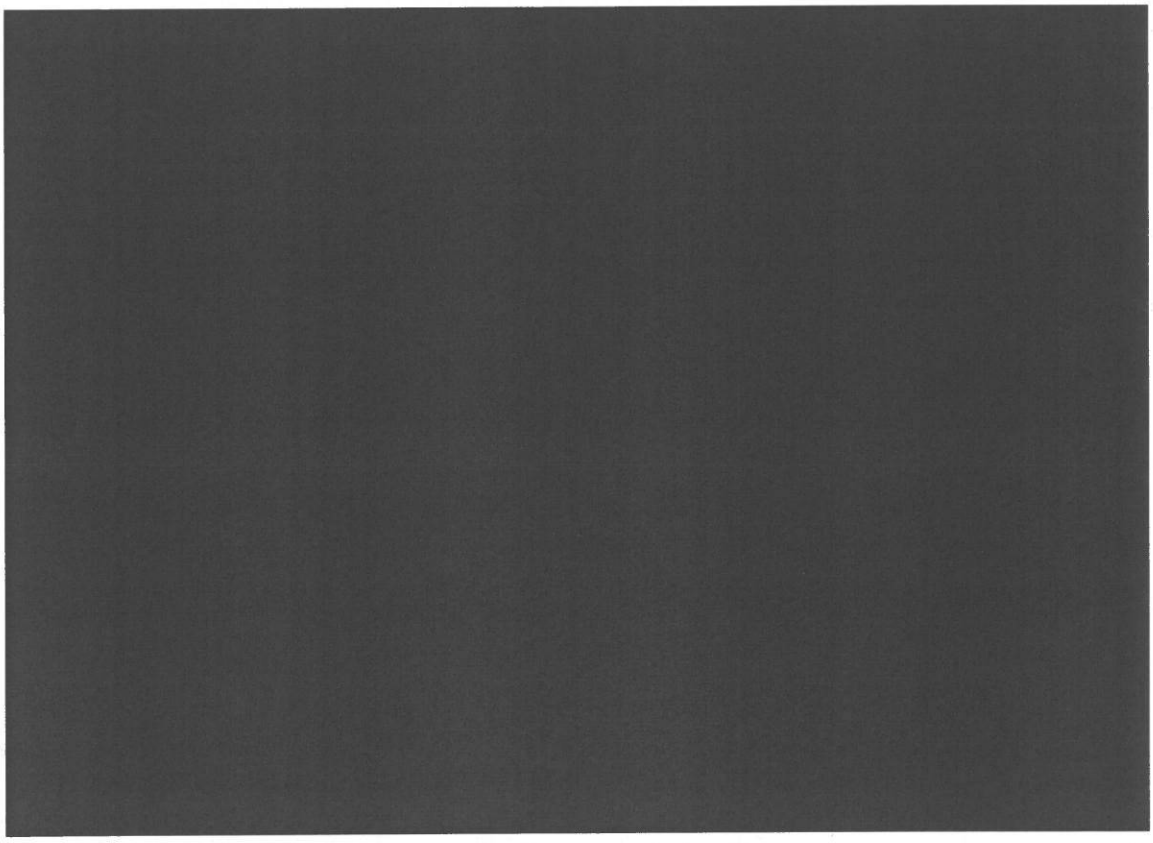
[redacted]

1.14 SAN infrastruktura

SAN, tedy Storage Area Netwotk navrhujeme postavit na technologii 4/8/16Gbit Fibre Channel a to tak aby zajistila dostatečnou propustnost, spolehlivost a bezpečnost provozu. Celou SAN tvoří dvě na sobě nezávislé Fibre Channel sítě, reprezentované dvěma dvojicemi 48-port + 24-port 16Gbit Fibre Channel switchů.
Součástí sítě SAN, schematicky znázorněné na obrázku níže jsou následující základní stavební prvky:

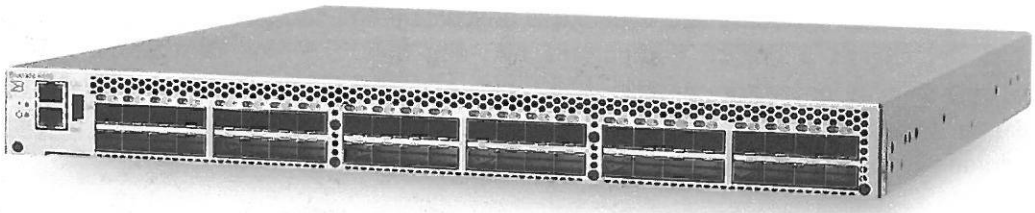
1. 48-portový Fibre Channel switch Brocade 5610 4/8/16 Gbps (2ks)
2. 24-portový Fibre Channel switch Brocade 5610 4/8/16 Gbps (2ks)
3. PoD licence a SFP+ moduly na 144 portů
4. Licence na Brocade ISL Trunking (4ks)
5. FC adaptéry serverů, diskových polí, páskové knihovny a mechanik jsou součástí konfigurací příslušných systémů
6. Další potřebné příslušenství nezbytné k řádnému provozu sestavy datového úložiště (napájecí kabely, adaptéry, propojovací kabely, případně další nezbytné síťové i jiné komponenty).

Schéma zapojení a integrace sítě SAN do prostředí Systému pro náročné výpočty:



SAN síť je postavena na čtyřech síťových prvcích Brocade 5610. Jedná se o 16Gbps Fibre Channel switch, který disponuje 24 až 48 4/8/16 Gbit SFP+ porty (v krocích po 12-ti portech), redundantními zdroji a ventilátory. Switche mají vzdálený síťový management prostřednictvím 1GE potu, který poskytuje mimo jiné přístup k Brocade Fabric Vision technology managementu. Všechny porty jsou vybaveny Brocade hot-pluggable SFP+ moduly s LC connectorem (16 Gbps SWL, LWL, ELWL). Latence

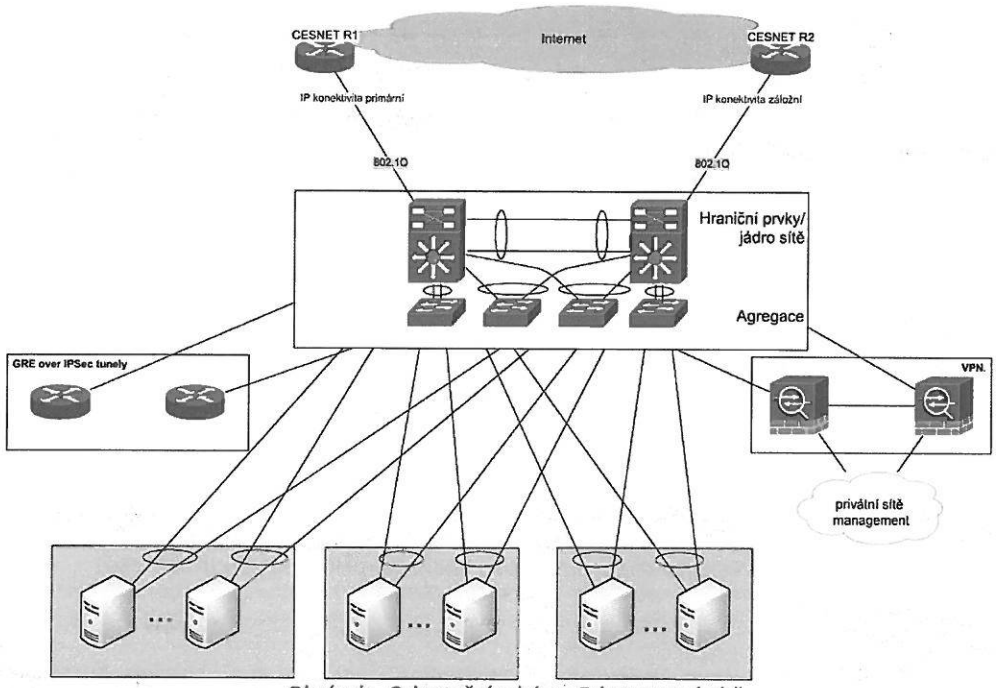
lokálně přepínaných portů nepřesahuje 700ns. Dva a dva síťové prvky Brocade 5610 jsou mezi sebou propojeny 8-mi ISL linkami, které za pomocí nainstalované Trunking licence spojují každé dva prvky v jedenu FibreChannel SAN síť (FC switched fabric).



Do obou sítí SAN jsou připojeny všechny požadované a pro funkčnost Systému pro náročné výpočty potřebné systémy, jako jsou Zálohovací systém, Souborové datové úložiště HOME, Datové úložiště INFRASTRUKTURY, Virtualizační infrastruktura a Další serverové systémy.

1.15 Ethernetová síť

Ethernetová síť bude zajišťovat komunikaci mezi zařízeními uvnitř Velkého clusteru, mezi Velkým clusterem, dalšími systémy provozovanými zadavatelem (Malý cluster, Specializovaný systém, atd.) a internetem. Součástí řešení je rovněž zabezpečení této komunikace.



Obrázek - Orientační schéma Ethernetové sítě

Ethernetová síť obsahuje veřejné a privátní části sítě.
Veřejné části sítě (dále jen veřejné sítě), ve kterých jsou použity veřejné IPv4 a IPv6 adresy, slouží pro poskytování služeb dostupných z internetu.
Privátní části sítě (dále jen privátní sítě), ve kterých jsou použity privátní IPv4 adresy, slouží pro vnitřní služby a pro management zařízení.

Součásti řešení Ethernetové sítě je remote VPN, site to site VPN a NAT.
Ethernetová síť bude obsahovat oddělenou OOB management síť.

Zařízením v privátních sítích bude umožněn přístup do internetu přes NAT.
Z internetu do privátních sítí bude možné přistupovat pouze přes remote VPN.
Privátní sítě budou směrovány na jiném L3 zařízení nebo v jiné routovací instanci než veřejné sítě.

Řešení Ethernetové sítě je odolné vůči výpadku či odstávce jedné komponenty – linky, transceiveru, modulu, zdroje, chassis a dalších komponent tak, aby byla zajištěna dostupnost připojení do internetu, byly splněny požadavky na připojení funkčních celků uvedené v zadávací dokumentaci a byla zajištěna dostupnost služby Remote VPN. Služby site to site VPN a OOB managementu sítě nejsou odolné vůči výpadku či odstávce jedné komponenty.

Jádro sítě bude postaveno jako multichassis se společným data plane a podporou multichassis etherchannel.

Propustnost spoje mezi prvky sítě bude 80 Gb/s v normálním stavu a minimálně 40 Gb/s v případě výpadku jedné komponenty sítě s výjimkou stavů, při nichž se stane nedostupný celý síťový prvek – porucha chassis nebo procesoru.

Prvky jádra sítě budou použity také jako hraniční prvky pro přístup do internetu (dále jen „hraniční prvky").

V případě výpadku procesoru prvku jádra sítě nebo hraničního prvku směrování kompletně zkonverguje do 60 sekund.

V případě výpadku poloviny napájecích zdrojů každého prvku jádra sítě nebo hraničního prvku, nebude provoz sítě žádným způsobem ovlivněn.

Navržené Prvky jádra sítě a hraniční prvky podporují export informací o tocích dat NetFlow v9 nebo vyšší.

Prvky jádra sítě a hraniční prvky budou podporovat minimálně 10 oddělených routovacích instancí bez použití MPLS.

Koncová zařízení je možno připojit do agregačních síťových prvků, požadavky na agregovanou propustnost a dostupnost uvedené v dalších kapitolách zadávací dokumentace budou splněny.

1.15.1 Připojení do internetu

Ethernetová síť obsahuje právě dva hraniční prvky Nexus 7009 pro přístup do internetu. Každý z obou hraničních prvků bude vybaven těmito rozhraními pro realizaci připojení do internetu:

- 5x 10GBASE-LR

Oba hraniční prvky budou v dodané konfiguraci rozšiřitelné (instalací rozšiřujícího modulu, karty) o optické rozhraní 100Gb/s (long range), tato rozhraní bude možno zapojit neblokujícím způsobem.

Primární připojení do internetu bude realizováno z obou hraničních prvků přes rozhraní 10GBASE-LR do infrastruktury CESNETu, na každém hraničním prvku bude vytvořen etherchannel 4x10Gb/s.
Další připojení do internetu bude realizováno z obou hraničních prvků přes rozhraní 10GBASE-LR do infrastruktury uzlu CESNETu v Ostravě.