

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Diplomová práce

**Implementace all-flash diskového pole v prostředí
VMware**

Bc. Martin Kolman

© 2020 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Provozně ekonomická fakulta

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Martin Kolman

Systémové inženýrství a informatika
Informatika

Název práce

Implementace all-flash diskového pole v prostředí VMware

Název anglicky

Implementation of all-flash disk array in VMware environment

Cíle práce

Hlavní cíle práce:

Hlavním cílem práce je implementace all-flash diskového pole do funkční infrastruktury s využitím virtualizace v prostředí VMware. Implementace by měla přinést významné odlehčení standardního diskového pole přesunutím klíčových systémů na nové řešení.

Dílčí cíle:

zpracování přehledu řešené problematiky s důrazem na prostředí VMware

analýza a příprava prostředí pro implementaci

implementace diskového pole

konfigurace FC Switche

konfigurace VMware vCenter

Testování a následná migrace kritických systémů

hodnocení implementovaného řešení a formulace závěrů práce

Metodika

Teoretická část je založena na analýze a studiu dostupných informačních zdrojů a literatury. V první části bude provedeno vyhodnocení současného stavu a jeho kritických problémů. Následně budou charakterizovány jednotlivé navazující technologie.

V praktické části bude provedena analýza a následná příprava prostředí. Dále bude provedena fyzická instalace, která bude následována základní instalací diskového pole. Poté bude provedena konfigurace FC switche s konfigurací jednotlivých portů a nastavení jednotlivých zón přístupů. Na straně diskového pole budou přidány jednotlivé ESXi hosty. Dalším krokem bude konfigurace VMware vCenter. Dále bude provedeno ladění a optimalizace s cílem zvýšení stability a výkonu celého systému.

Po dokončení implementace a optimalizace diskového pole, budou přesunuty, po jeho řádném otestování testovacími servery, stěžejní systémy pro odlehčení diskového pole, se standardními rotačními disky, a snížení diskové latence, která dosahuje kritických hodnot.

Syntézou poznatků z teoretické části a výsledků praktické části, budou vyvozeny závěry celé práce a predikce budoucího vývoje této implementace.

Doporučený rozsah práce

50 – 60 stran

Klíčová slova

All-flash, storage, virtualizace, VMware, vCenter, ESXi

Doporučené zdroje informací

Dell EMC Unity: Best Practices Guide [online], [cit. 2020-2-25], dostupné z:

<<https://www.dell.com/en-us/collaterals/unauth/white-papers/products/storage/h15093-dell-emc-unity-best-practices-guide.pdf>>

Dell EMC Unity:VMware vSphereBest Practices [online], [cit. 2020-2-25], dostupné z:

<<https://www.dell.com/resources/en-us/asset/white-papers/products/storage/h16391-dell-emc-unity-storage-vmware-vsphere.pdf>>

Storage Networking 101: Understanding the Fibre Channel Protocol [online], [cit. 2020-2-25], dostupné z:

<<http://www.enterprisenetworkingplanet.com/netsp/article.php/3690921/Storage-Networking-101-Understanding-the-Fibre-Channel-Protocol.htm>>

Předběžný termín obhajoby

2020/21 LS – PEF

Vedoucí práce

Ing. Jiří Vaněk, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 27. 8. 2020

Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 21. 10. 2020

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 04. 02. 2021

Čestné prohlášení

Prohlašuji, že svou diplomovou práci "Implementace all-flash diskového pole v prostředí VMware" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 22.03.2021

Poděkování

Rád bych touto cestou poděkoval Ing. Jířímu Vaňkovi Ph.D., za veškerou jeho pomoc, trpělivost a odborné vedení této diplomové práce. Dále bych rád poděkoval své manželce, Mgr. Kamile Kolmanové za veškerou pomoc, jakožto i hlídání našich dvou dětí, abych mohl studovat a psát tuto práci.

Implementace all-flash diskového pole v prostředí VMware

Abstrakt

Tato práce se zabývá implementací all-flash diskového pole v prostředí VMware. V první části bude provedeno vyhodnocení současného stavu, jeho kritických problémů a jaká jsou očekávání od implementovaného řešení. Následně budou charakterizovány jednotlivé navazující technologie.

V praktické části bude provedena analýza a následná příprava prostředí. Dalším krokem bude uskutečnění fyzické instalace, následované základní instalací diskového pole. Poté bude upravena konfigurace FC (Fibre Channel) switchu s konfigurací jednotlivých portů a nastavení jednotlivých zón přístupů. Na straně diskového pole budou přidány jednotlivé ESXi hosty. Dalším krokem bude konfigurace VMware vCenter. Dále bude optimalizací a laděním jednotlivých prvků zvýšena stabilita a výkon celého systému.

Po dokončení implementace a optimalizace diskového pole budou přesunuty, po jeho řádném otestování testovacími servery, stěžejní systémy pro odlehčení diskového pole se standardními rotačními disky. Bude tím dosaženo snížení diskové latence, která dosahuje kritických hodnot.

Klíčová slova: All-flash, virtualizace, VMware, vCenter, ESXi

Implementation of all-flash disk array in VMware environment

Abstract

This work deals with the implementation of an all-flash disk array in a VMware environment. In the first part, an evaluation of the current state, its critical problems and what is expected from the implemented solution will be performed. Subsequently, individual related technologies will be characterized.

In the practical part, the analysis and subsequent preparation of the environment will be performed. Next, a physical installation will be performed, followed by a basic disk array installation. Then the configuration of the FC switch with the configuration of individual ports and the setting of individual access zones will be performed. Individual ESXi guests will be added on the disk array side. The next step will be to configure VMware vCenter. Furthermore, debugging and optimization will be performed in order to increase the stability and performance of the entire system.

Once disk array implementation and optimization is complete, core disk lightening systems, with standard rotating disks, and disk latency reductions that reach critical values will be moved, after being properly tested by test servers.

Keywords: All-flash, virtualization, VMware, vCenter, ESXi,

Obsah

1 Úvod	11
2 Cíl práce a metodika	12
3 Teoretická východiska	13
3.1 Související technologie	13
3.2 Disková pole	14
3.3 SAN a NAS	14
3.4 RAID	15
3.4.1 Druhy RAID	Chyba! Záložka není definována.
3.5 Optická vlákna	21
3.6 Optické konektory	24
3.7 Fibre Channel (FC)	29
3.8 SFP moduly	32
3.9 Technologie SSD	34
3.10 VMware ESXi	36
3.11 VMware vCenter	37
4 Vlastní práce	39
4.1 Charakteristika aktuálního stavu a jeho zhodnocení	39
4.1.1 Blade chassis HP C3000 a servery	39
4.1.2 Brocade 300 - HP StorageWorks 8/8 and 8/24 SAN Switches ^[39]	41
4.1.3 HUAWEI Core Switch S7706 ^[40] (Obr. 22)	42
4.1.4 DELL – EMC Unity 300 – aktuální stav ^[43]	43
4.1.5 Kritické problémy a zhodnocení současného řešení	44
4.1.6 DELL – EMC Unity 450f – nové diskové pole ^[44] (Obr. 23)	44
4.1.7 Alternativy a výběr	45
4.2 Instalace a konfigurace pole	46
4.2.1 Fyzická instalace	46
4.2.2 Konfigurace síťových prvků	49
4.2.3 Základní část konfigurace diskového pole	49
4.2.4 Vytvoření disk pool	53
4.2.5 Dokončení základní konfigurace	54
4.3 Unisphere	57
4.4 Vytvoření datastoru a připojení jednotlivých ESXi hostů	58
4.5 Konfigurace SAN Switchů	62
4.6 Připojení pole k vCenter	65
4.7 Testování a migrace systémů	69
5 Výsledky implementace a doporučení	71
5.1 Výsledky implementace a doporučení	71

5.2	Problémy během implementace	73
5.3	Zlepšení současného řešení	73
5.4	Serverovna.....	74
5.5	Diskové pole Unity 450f – budoucnost.....	74
6	Závěr	76
7	Seznam použitých zdrojů	78

Seznam obrázků

Obr. 1	RAID 0 ^[6]	17
Obr. 2	RAID 1 ^[6]	18
Obr. 3	RAID 4 ^[6]	19
Obr. 4	RAID 5 ^[6]	19
Obr. 5	RAID 6 ^[6]	20
Obr. 6	RAID 10 ^[6]	21
Obr. 7	skladba optického vlákna ^[8]	22
Obr. 8	Deutsch 1000 ^[13]	24
Obr. 9	BIONIC ^[13]	25
Obr. 10	ST konektor ^[13]	25
Obr. 11	SC konektor ^[13]	26
Obr. 12	LC konektor [13]	27
Obr. 13	Duplexní LC a SC konektory	27
Obr. 14	MPO (MTP) kabely ^[25]	28
Obr. 15	FC Frame ^[19]	30
Obr. 16	SFP modul ^[24]	32
Obr. 17	QSFP-DD 8X ^[31]	33
Obr. 18	Bitový zápis NAND Flash čipů ^[53]	35
Obr. 19	ESXi host ^[32]	36
Obr. 20	svMotion mirroring ^[34]	37
Obr. 21	HPE ProLiant BL460c Gen9 Server Blade ^[35]	40
Obr. 22	HPE BladeSystem C3000 Enclosure ^[38]	41
Obr. 23	Huawei Core Switch S7706 ^[40]	43
Obr. 24	Dell EMC Unity 450F All-Flash Storage ^[45]	45
Obr. 25	Postup instalace ^[46]	46
Obr. 26	fyzické zapojení	48
Obr. 27	Průvodce konfigurací.....	50
Obr. 28	Admin a Service password	51
Obr. 29	Storage Tier	53
Obr. 30	konfigurace ESRS.....	55
Obr. 31	ESRS ověření emailu.....	56
Obr. 32	Unisphere Dashboard	57
Obr. 33	Performance	58
Obr. 34	VMware Datastore Type	59
Obr. 35	Konfigurace úložiště.....	60
Obr. 36	Datastore Summary	61
Obr. 37	SAN konfigurace	62
Obr. 38	SAN druhy portů	63
Obr. 39	konfigurace FC Portu.....	64

Obr. 40 SAN Zóny	65
Obr. 41 Storage adapters.....	66
Obr. 42 New Datastore	67
Obr. 43 Round Robin.....	68
Obr. 44 VMFS Devices	68
Obr. 45 ESXi RoundRobin	69
Obr. 46 UNITY 300 Latency	71
Obr. 47 UNITY 450f Latency.....	72
Obr. 48 Migrace SQL	72
Obr. 49 Reponse extream.....	73

Seznam tabulek

Tabulka 1 NTP servery ^[50]	51
Tabulka 2 Serverová konfigurace	70

1 Úvod

Technologie flash disků zažívá v současné době obrovský boom. Není tedy divu, že se tato technologie dostala i do velkých řešení, kde vytlačuje standardní konvenční disková pole, postavená na klasických rotačních discích. All-flash disková pole dokáží dodat výkon IOPS, kterého standardní disková pole s 15k disky nejsou schopna. Toto je dáno samotnou koncepcí flash disků, kde je signifikantní rozdíl v rychlosti zápisu a čtení. Jednou z výhod je velikost. Zatímco u kapacitních disků, pokud je potřeba vyšší kapacita, je nutná velikost 3,5“, v případě SSD disků je standardní velikost 2,5“. Další nespornou výhodou je spotřeba el. energie. Ta je díky použitým technologiím a absenci motorku nižší, což u velkých diskových polí má za následek významné snížení spotřeby, a tedy další snížení nákladů.

All-flash disková pole se tedy stávají nedílnou součástí řešení, kde je potřebný velký výkon. Tato pole se obecně používají kupř. pro SQL, exchange a webové servery. Tento typ diskových polí není vhodný jako úložiště dat pro uchovávání dokumentů. Pro ně jsou vhodná standardní disková pole s kapacitními Nearline (NL) disky, kde není potřeba výkon v IOPS, ale je potřeba velká kapacita.

2 Cíl práce a metodika

Hlavním cílem práce je implementace all-flash diskového pole do funkční infrastruktury s využitím virtualizace v prostředí VMware. Implementace by měla přinést významné odlehčení standardního diskového pole přesunutím klíčových systémů na nové řešení.

Teoretická část je založena na analýze a studiu dostupných informačních zdrojů a literatury. Následně budou charakterizovány jednotlivé navazující technologie, které jsou potřebné pro implementaci.

V praktické části bude provedeno vyhodnocení současného stavu a jeho kritických problémů. Poté bude vykonána samotná instalace a konfigurace celého řešení, následována fyzickou instalací, která bude pokračovat základní instalací diskového pole. Následně bude provedena konfigurace FC (Fibre Channel) switche s konfigurací jednotlivých portů a nastavení jednotlivých zón přístupů. Na straně diskového pole budou přidány jednotlivé ESXi hosty. Dalším krokem bude konfigurace VMware vCenter. Dále bude optimalizací a laděním jednotlivých prvků, zvýšena stabilita a výkon celého systému.

Po dokončení implementace a optimalizace diskového pole budou přesunuty, po jeho řádném otestování testovacími servery, stěžejní systémy pro odlehčení diskového pole se standardními rotačními disky. Bude tím dosaženo snížení diskové latence, která dosahuje kritických hodnot.

Syntézou poznatků z teoretické části a výsledků praktické části budou vyvozeny závěry celé práce a predikce budoucího vývoje této implementace.

V práci jsou využity znalosti a zkušenosti, které autor nasbíral během své dosavadní praxe. Problematika, kterou se tato diplomová práce zabývá, obsahuje řadu anglických termínů, pro které se nevyužívají české ekvivalenty, nebo nejsou odpovídající ekvivalenty k dispozici.

3 Teoretická východiska

3.1 Související technologie

V současné době společnosti i uživatelé řeší, kam a jak bezpečně ukládat jejich data. Existuje celá řada možných řešení. Mezi nejčastěji zastoupená jsou malá řešení v podobě NAS (Network Attached Storage) serverů, dedikovaných serverů, malých či větších diskových polí a cloudová řešení. Disková pole se dají rozdělit na tři základní NAS, SAN (Storage area Network), DAS (Direct attached storage). Všechna tato řešení nabízejí určitý způsob ochrany v podobě RAID (Redundant Array of Independent Disk). Pro zajištění konektivity k diskovým polím je možné využít metalickou nebo optickou síť. Pro SAN optickou konektivitu jsou nejčastěji využity SAN switche.

3.2 Disková pole

Diskové pole nebo také diskové úložiště je vyhrazené HW úložiště, které obsahuje buď klasické rotační disky (HDD – Hard Disk Drive)^[1] nebo SSD^[1] (solid-state drive). Pole kombinují jednotlivé disky. V současné době se jedná až o petabajty (PB) dat^[2]. Důvod vzniku bylo oddělení diskového úložiště od serverů a tím mohlo dojít ke vzniku velkých monolitických systémů pro blokové nebo souborové úložiště. Jako první diskové pole byl již v roce 1956 model IBM 350 disk storage unit^[3]. Současná disková pole poskytují pokročilé funkce, jako je deduplikace, šifrování, mezipaměť pro rychlejší přístup k datům, RAID a diskové sloty pro rychlou výměnu nebo doplnění jednotek. Větší disková pole obvykle poskytují kompletní redundanci u všech prvků. Celý systém se tak stává odolnějším proti výpadkům jednotlivých částí a nabízí i pohodlnou údržbu v případě SW update nebo HW upgrade či servisu. Většinu dílů je tedy možné měnit za provozu bez dopadu na běžící systém.

V současné době by se dala disková pole rozdělit na tři základní typy. Standardní disková pole osazená rotačními disky (HDD), hybridní pole – tedy kombinace SSD a HDD a nejmodernější all-flash disková pole osazená pouze SSD disky. SSD disky nabízejí mnohem vyšší rychlost zápisu a čtení, než jsou HDD schopna^[4] a díky tomu jsou schopny podat více IOPS s menším počtem disků. Přes nesporné výhody all-flash diskových polí si společnosti stále zachovávají i standardní HDD nebo hybridní disková pole pro méně náročné systémy nebo pro zálohování. Hlavním důvodem je nesporně cena řešení HDD jsou podstatně levnější než SSD, ale je otázkou času, kdy dojde k vyrovnání a konvenční HDD budou plně nahrazeny SSD disky. Všechna řešení používají identickou architekturu zapojení, a to včetně připojení k serverům pomocí adaptéru a switchů.

3.3 SAN a NAS

Tyto dva typy diskových polí, SAN a NAS, jsou nejrozšířenější. Dalším diskovým polem je DAS (Direct-attached storage)^[5], který se připojuje přímo k serveru, avšak postrádá sofistikovaný management dostupný právě u NAS a SAN.

SAN úložiště je primárně navrženo pro blokové úložiště a využívá se pro virtualizaci. NAS je pravým opakem. Je primárně navržen pro nestructurovaná data a souborový systém. Obě

tato řešení mohou pro svou konektivitu využívat jak metalické, tak optické spojení. Pro NAS je nejběžnější použití metalického spojení, naopak pro SAN je typické FC připojení, které pro to bylo speciálně vyvinuto.

NAS je koncipován pro sdílení souborů přes síť. Umožňuje klientům a serverům sdílet data přes počítačovou síť a tím eliminovat potřebu klasických souborových serverů. NAS obvykle využívá vlastní operační systém a integrované HW a SW komponenty k pokrytí použitých souborových služeb. Pro své funkce využívá více protokolů a funkcí, např. TCP/IP pro datový přenos, SMB, NFS a FTP pro sdílení souborů. NAS obecně přináší výhody v centralizaci datového úložiště, čímž zvyšuje zabezpečení dat a zjednodušuje správu a kontrolu nad daty a kapacitou úložiště. Dále lze pomocí centrálního systému zabezpečit přístupy k jednotlivým datům, autentizaci uživatelů a správu oprávnění^[16]. NAS řešení má zastoupení ve všech myslitelných velikostech. Od nejlevnějších pro domácí využití až po velká enterprise řešení. Typickým zástupcem pro domácí nebo small business jsou NAS od společností Synology, QNAP a mnoha dalších. V případě enterprise řešení jsou, stejně jako v případě SAN, aktuálními lídry na trhu velcí hráči, kupř. Dell-EMC, IBM, HP, NetApp a další.

Naopak SAN řešení je vysokorychlostní síťové úložiště pro více serverů s přístupem k diskovému fondům na blokové úrovni. Toto řešení poskytuje vysokou spolehlivost a škálovatelnost. Míra využitelnosti úložiště je podstatně vyšší než u konvenčního řešení, což často snižuje náklady na pořízená úložiště^[17]. SAN představuje hostiteli úložiště tak, že se zdá, že je připojeno lokálně. Pro připojení je nejčastěji využita technologie FC přes vysokorychlostní infrastrukturu. Lze využít i jiné technologie připojení, jako kupř. iSCSI, které je běžnější pro menší a středně velké organizace využívající toto řešení jako levnější alternativu k FC.

3.4 RAID

RAID^[6] – (redundantní pole nezávislých disků) jedná se o způsob zápisu dat na vícero rotačních nebo SSD disků. Tento způsob zápisu dat vytváří jejich ochranu před případným selháním disku. Další výhodou je překrývání vstupně /výstupních (I/O) operací a tím zvýšení výkonu úložiště. Použití vícero disků také znamená snížení zatížení, a tedy i zvýšení průměrné doby mezi případnými poruchami. RAID data nechrání před neúmyslným či

úmyslným smazáním dat a v žádném případě nenahrazuje zálohování. Využití RAID řešení má široké využití od domácích NAS až po enterprise řešení v serverech a diskových polích.

RAID lze implementovat buď pomocí fyzického řadiče (HW RAID), anebo pomocí ovladače v operačním systému (SW RAID). Obě varianty mají své výhody a nevýhody.

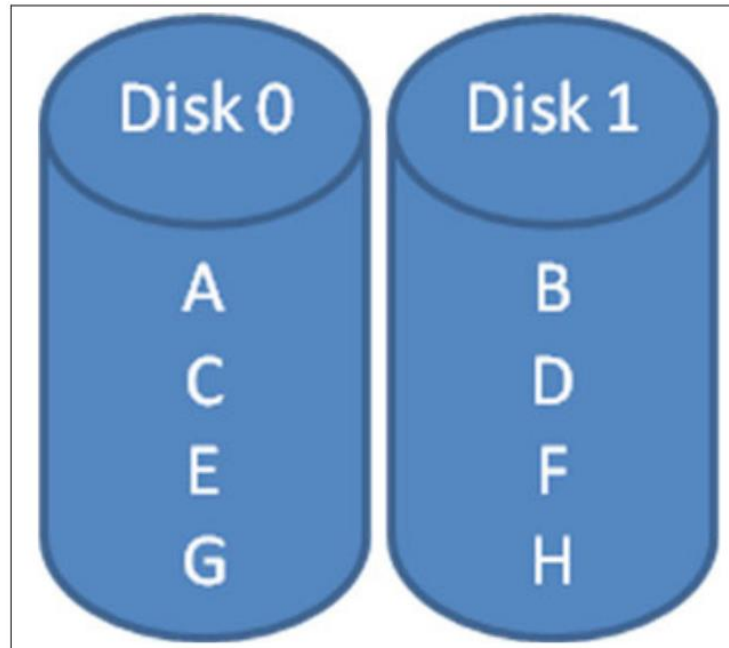
Hardware RAID, jakožto samostatný dedikovaný systém, nezatěžuje výpočetní výkon systému, ale veškeré výpočty probíhají na řadiči či kartě. Nicméně toto řešení je dražší a dále v případě poruchy řadiče je potřeba výměna za identický model, aby nedošlo ke ztrátě dat.

Software RAID používá systémové prostředky operačního systému. Nevyžaduje tedy žádné další investice. Nicméně toto řešení je náročnější na výpočetní výkon, a tak může v případě náročnějších úkonů dojít ke zpomalení serveru.

Rozhodnutí, zda použít SW nebo HW RAID, tedy primárně záleží na tom, k čemu je daný server určen a jaký je na něj rozpočet. Pokud je nízký rozpočet a jsou nízké požadavky na server, je poté SW RAID jasnou volbou. Pokud požadujeme špičkový výkon, nebo komplikovanější RAID, jako je RAID 10, je nutné použít HW RAID.

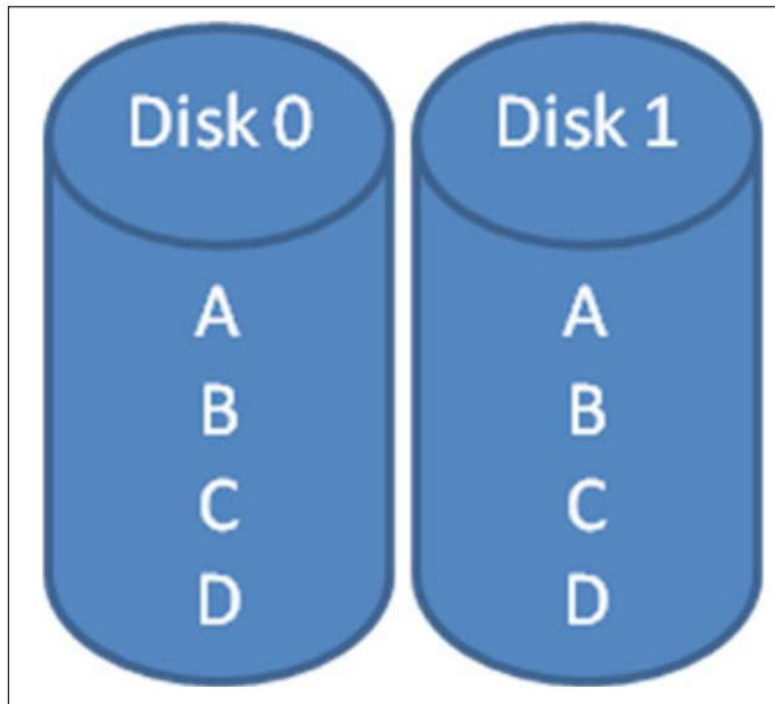
Existuje několik druhů RAID^[6] polí, avšak ne všechna chrání data před případným selháním. Níže budou vypsány pouze běžně používané RAID^[6] konfigurace.

RAID0^[6] – Nejedná se o skutečný RAID. Neobsahuje žádné redundantní informace a tak neposkytuje datům žádnou redundanci. Tedy při ztrátě elementu má za následek ztrátu dat. Jednotlivá data jsou spojena v logickém celku a tím vytváří kapacitu součtem všech členů. Je zapotřebí minimálně dvou disků pro toto řešení. Na tomto druhu RAIDu jsou data ukládána střídavě na jednotlivé disky. Prostor je rozdělen na pevné části a zápis nebo čtení většího úseku je tím pádem rychlejší, protože probíhá z více disků. Nicméně při poruše jednoho z disků je nepravděpodobné, že by zůstala některá data neporušená. Ukládání dat prokládaně zrychluje čtení i zápis, protože rychlost se zvyšuje s počtem použitých disků^(Obr. 1).



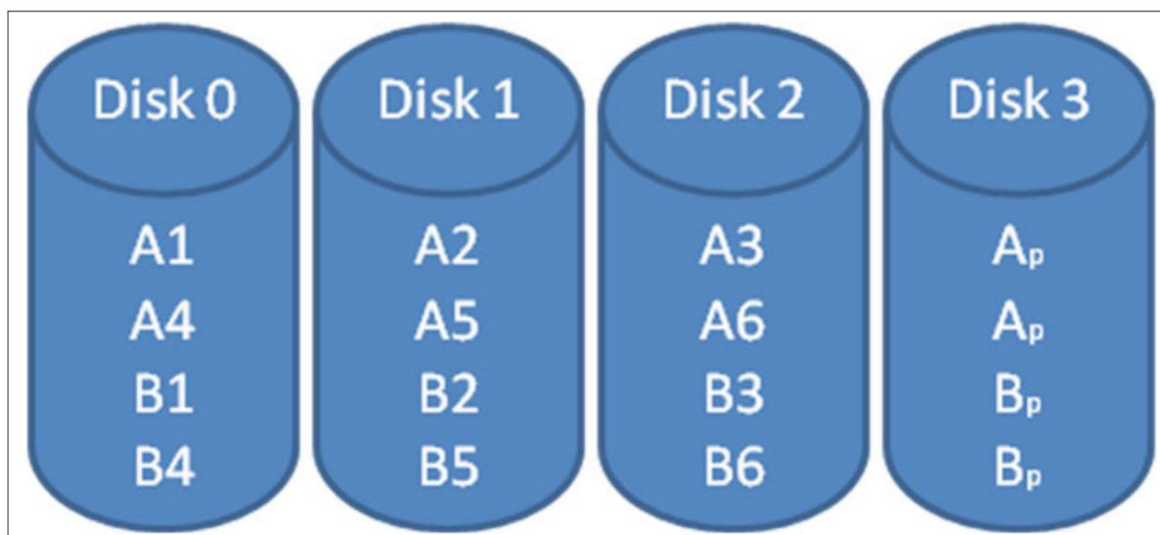
Obr. 1 RAID 0^[6]

RAID1^[6] – Mirroring. Jedná se o nejjednodušší, ale efektivní ochranu dat. Data se automaticky zapisují na dva disky najednou. Podobná technika může být použita i pro řadiče. Tato metoda se poté jmenuje duplexing a je odolná proti výpadku samotného řadiče. Zrcadlení zvyšuje rychlost čtení tam, kde se čte ze dvou disků zároveň, ale současně, díky zápisu na dva disky najednou, se zvyšuje jeho doba zápisu. Rychlost zápisu poté záleží na kvalitě řadiče a rychlosti disků. Tato technika zvyšuje bezpečnost dat proti ztrátě, nicméně je náročná na kapacitu, čímž zvyšuje cenu řešení. Je zapotřebí minimálně dvou disků pro tento RAID^(Obr. 2).



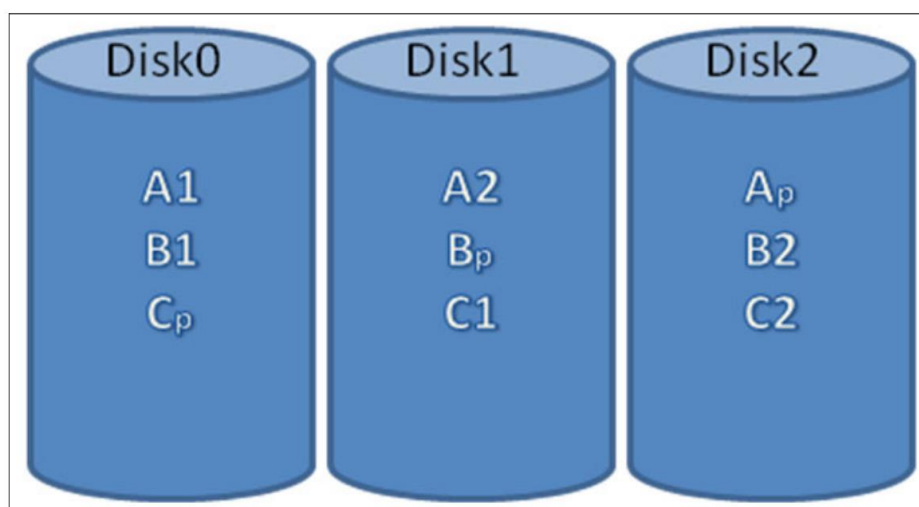
Obr. 2 RAID 1^[6]

RAID4^[6] – zde se dá již mluvit o plnohodnotném RAIDu. V tomto typu dochází k prokládání dat na vícero disků. Kromě samotného prokládání ukládá informace o paritě na samostatném disku. Toto řešení přináší nižší režijní kapacitu a zvyšuje rychlost paralelních I/O požadavků. Z důvodu uložení paritních bitů na jednu jednotku je rychlost řešení závislá na rychlosti paritního disku. To může, v určitých případech, vést ke zpomalení systému. Je zapotřebí minimálně tři disků pro tento RAID. Výrobce diskových polí NetApp využívá a rozšiřuje RAID4^[7] o další paritní disk a tím docílil stejné charakteristiky jako má RAID6. Toto řešení je dostupné pouze na ONTAP operačním systému určenému pouze pro NetApp disková pole (Obr. 3).



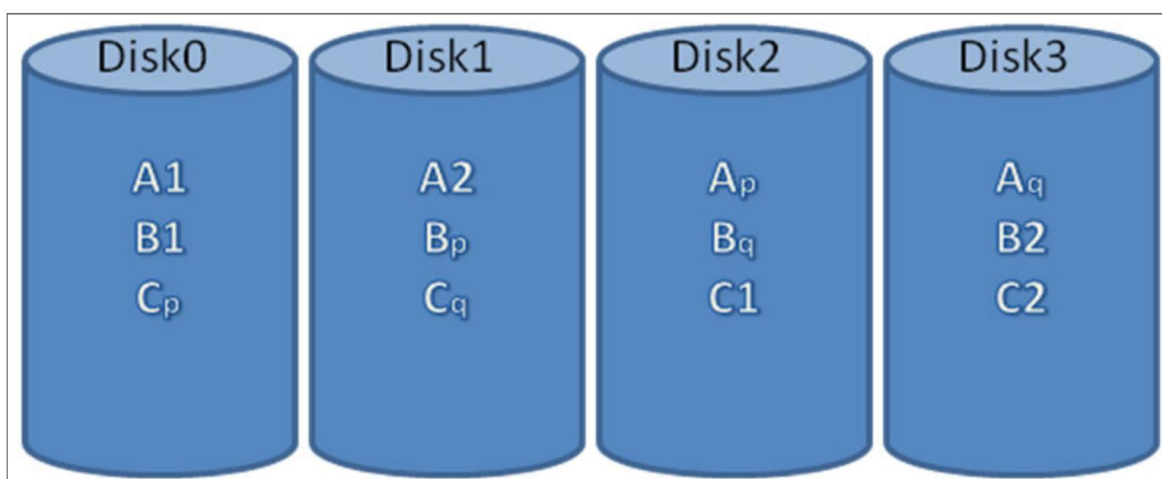
Obr. 3 RAID 4^[6]

RAID5^[6] – zde probíhá blokové prokládání dat s distribuovanou paritou. Tedy informace o paritě jsou distribuovány na všechny disky rovnoměrně, a ne na jeden disk, jako je tomu v případě RAID4. V případě selhání jednoho disku lze dopočítat informace bez ztráty dat. Je zapotřebí minimálně tři disků pro tento RAID. Všechny jednoparitní RAIDy, tedy i RAID5, mají u velkých implementací nevýhodu dlouhé doby sestavení při výpadku jedné jednotky a zvýšeného rizika výpadku druhé jednotky. To je způsobeno nutností číst data ze všech jednotek, což vede ke zvýšenému riziku dalšího výpadku, tedy k celkovému selhání pole. Tento druh RAIDu podporuje vícenásobné čtení a zápis najednou. I přes vyšší rizika, než u RAID6, je zde nejlepší poměr cena / výkon^(Obr. 4).



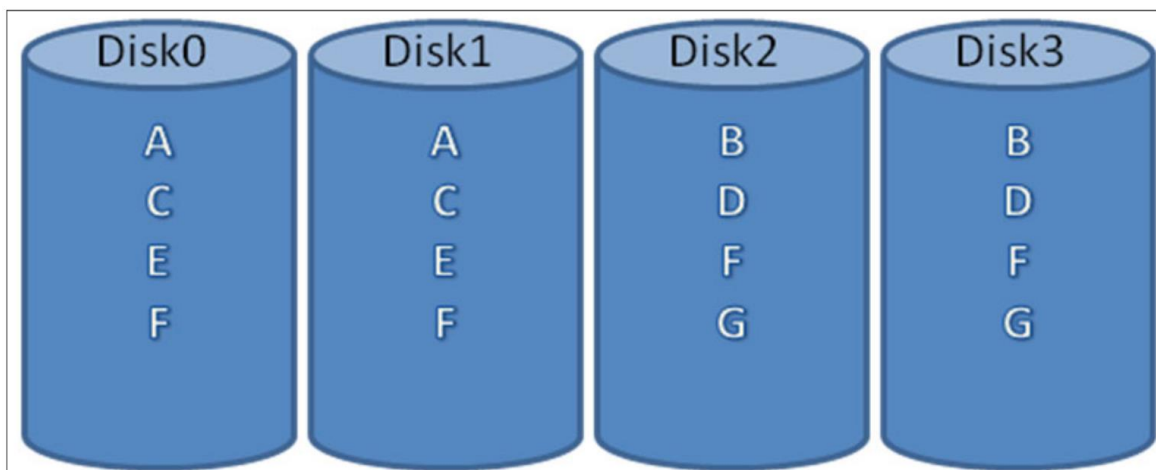
Obr. 4 RAID 5^[6]

RAID6 ^[6] – stejně jako u RAID5 i zde probíhá blokové prokládání dat, ale s dvojitou distribuovanou paritou. Dvojitá parita zajišťuje odolnost řešení proti výpadku až dvou jednotek. Výpadek i jedné jednotky má, stejně jako u RAID5, dopad na výkon celého pole. RAID6 bývá použit tam, kde jsou velká datová pole a požadavek na vysokou dostupnost dat. Výkon je v tomto případě o něco nižší než u RAID5, a to u zápisu, což je dáno zdvojenou paritou. Je zapotřebí minimálně pěti disků pro tento RAID, tři pro data a dva pro paritu ^(Obr. 5).



Obr. 5 RAID 6 ^[6]

RAID10 ^[6] – jedná se o kombinaci RAID1 a RAID0, někdy se také používá označení hybridní RAID. Tento druh RAID pole nabízí to nejlepší z obou RAIDů. Nabízí rychlost RAID0 a současně redundanci RAID1. Je zapotřebí minimálně čtyř disků pro tento RAID a v případě rozšíření je potřeba přidávat vždy po dvou discích. V případě tohoto RAIDu se nejedná o neekonomičtější řešení z důvodu zrcadlení, tedy poloviční kapacity řešení, to ale ani není účel tohoto RAIDu. Ten je zaměřený na vysoký výkon a využívá se převážně u aplikačních, poštovních a databázových serverů ^(Obr. 6).



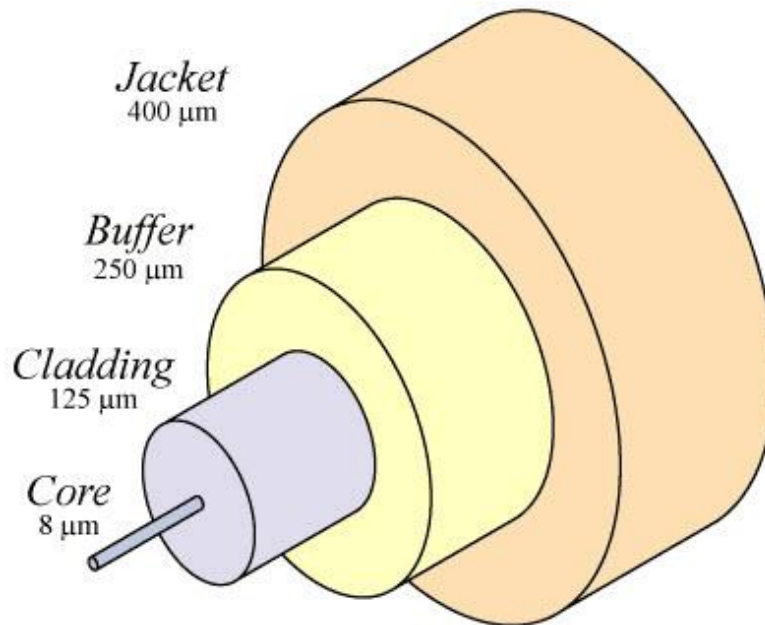
Obr. 6 RAID 10^[6]

JBOD^[6] – Just a Bunch Of Disks (je hromada disků). Nejedná se o RAID v pravém slova smyslu, protože funguje na principu řetězení disků, kdy je možné připojovat jakékoliv kapacity disků nebo svazků za sebou. Výhodou je použití libovolné velikosti disků a zvětšení kapacity úložiště. Ztráta jednoho elementu má za následek ztrátu pouze těch dat, která se nacházejí na daném disku. Ostatní data uložená na dalších discích zůstávají neporušena.

Dalšími typy standardních RAID jsou RAID2 a RAID3 a hybridních RAID řešení jsou RAID50, RAID60 a RAID100, ale jejich využití není běžné, a proto v této práci nebudou vypsána.

3.5 Optická vlákna

Jedná se o pružné průsvitné vlákno, jehož tloušťka je jenom o něco málo silnější než lidský vlas. Kompletní složení optického vlákna je jádro (plastové nebo skleněné), plášť (slouží k odražení a vedení světla uvnitř jádra), vnější plášť a obal^(Obr. 7).



Obr. 7 skladba optického vlákna ^[8]

Optická vlákna jako taková mají širokou škálu využití. Využívají se i jako dekorativní prvky s jednoduchým přenosem barevného světla. Speciálně upravená vlákna mohou být využita jako optické senzory nebo vláknové lasery. Optická vlákna mají oproti metalickým několik zásadních výhod. Díky menší ztrátě se dají využít pro přenos dat na větší vzdálenosti. Další nespornou výhodou je imunita proti elektromagnetickému rušení, což je jeden z hlavních problémů u metalického přenosu. V neposlední řadě je to také rychlost datového přenosu. V roce 2016 tým vědců z Nokia Bell Labs, Deutsche Telekom T-Labs a Technické univerzity v Mnichově v Německu úspěšně použil novou technologii k posunu 1Tbps přenosu dat v reálném světě přes kabel z čistě optických vláken, aby dosáhl téměř svého teoretického maxima a dosáhl přenosové rychlosti 1Tbit/s^[9].

Pro výrobu optických vláken se využívá buď plast, nebo častěji sklo. Plastové optické vlákno (POF) je jádro vyrobené z plastu nebo polymeru. Obvyklý materiál pro výrobu jádra je PMMA, tedy ^[10] akrylátová pryskyřice. Stejně jako u skleněných vláken i zde je pro přenos světla využito jádro vlákna. POF vlákna jsou obvykle multimode s velkým jádrem dosahujícím v průměru od 0,15 až do 2 mm. Hlavními výhodami mimo ceny jsou kupříkladu jednoduchost instalace, kterou zvládne každý uživatel i v domácím prostředí. Použití plastových (akrylových) jader má ale i některé nevýhody. Mezi hlavními je útlum signálu, který je signifikantní, a tak je využití omezeno na velmi krátké vzdálenosti. Další nevýhodou

je použití v extrémních podmínkách. Akrylové jádro neodolá extrémním teplotám tak, jako skleněné.

Jádro skleněného vlákna je vyráběno tažením skla a obkladu. Ten obaluje, chrání a zároveň odráží světlo zpět do jádra. Jedná se tedy o nejdražší, ale také o nejefektivnější typ kabelu. Standardně mají kabely použitelnost od -40°C až do cca 260°C , nicméně dají se vyrobit optická vlákna s použitelností až do cca 482°C . I přes své nesporné výhody není použití skleněných vláken pro každého. Pro manipulaci s ním musí být vyškolený personál a neodborná manipulace s vlákny může způsobit jeho zlomení a znehodnocení vlákna. Dále instalace nebo oprava poškozeného vlákna je určena pouze odborníkům. Nutností je naprosto precizní spoj, jinak kabel nebude funkční.

Jednotlivé kabely se poté dělí dle použité technologie a vzdálenosti, pro kterou jsou určeny, na single mode a multimode.

Single mode - jádro tohoto vlákna je tenčí a světelný signál může být poslán přímo středem jádra. Obecně se single mode vlákna používají na větší vzdálenosti, nad $550\text{m}^{[12]}$. Nicméně se využívají i pro krátké vzdálenosti. Průměr jádra je $9\mu\text{m}$ při tloušťce jádra $125\mu\text{m}^{[12]}$.

Multimode – optické signály jsou přenášeny více kanály. V tomto případě se využívají i kabely POF (s plastovým jádrem) pro přenos dat do 100m pro Gigabitový přenos^[11]. Dražší vlákna se skleněným jádrem jsou používána maximálně do vzdálenosti 550m . Nad tuto vzdálenost se již používají single mode kabely kvůli vysokému útlumu v multimode vodičích. Průměr jádra u multimode vodičů je $50\mu\text{m}$, respektive $62,5\mu\text{m}$ při tloušťce jádra $125\mu\text{m}^{[12]}$.

3.6 Optické konektory

Díky dynamickým začátkům s optickými vlákny a jejich konektory existuje přes 100 různých druhů optických konektorů. Avšak prvním komerčně úspěšným konektorem byl Deutsch 1000 (Obr. 8). Toto řešení bylo na konci 70 let to nejmodernější se ztrátou 3db [13]. Nicméně v současné době je tento konektor již zastaralý, a tedy se již nepoužívá.



Obr. 8 Deutsch 1000 [13]

Dalším pokrokovým konektorem byl BIONIC (Obr. 9). Tento konektor byl vyvinutý týmem vedeným Jackem Cookem v Bell Labs [13]. Konektor byl využíván jak pro single mode, tak pro multimode, rozlišených dle barev. Žlutý pro single mode a černý pro multimode.

Konektor byl formován z plastu naplněného sklem, který byl tvrdý téměř jako keramika. Při spojení konektorů zůstala nepatrná mezera, aby mohlo dojít k otáčení konektoru. To mělo za následek, že ztráta nebyla menší než 0,3dB, což bylo zapříčiněno fresnelovým odrazem [14]. Obvyklé ztráty se pohybovaly pro multimode mezi 0,5 – 1 dB a u single mode 0,7 dB a výše [13].



Obr. 9 BIONIC ^[13]

Pravděpodobně nejoblíbenějším konektorem pro multimode síť byl až do roku 2005 konektor ST ^[13] (Obr. 10). Tento konektor je válcového tvaru 2,5 mm dlouhý a má bajonetový držák, který drží vlákno. Dále obsahuje kroužek, který drží vlákno. Tento kroužek je většinou keramický, ale vyrábí se i kovový nebo plastový. Protože jsou konektory vybaveny pružinovým systémem, je nutné se při každém spojení přesvědčit o kvalitě spojení a případně je rozpojit a opětovně zapojit, aby nebyly vysoké ztráty ^[13].



Obr. 10 ST konektor ^[13]

Dalšími konektory, vycházejícími z předchozího modelu jsou konektory SC a FC. Tyto tři konektory ST, SC a FC vzájemně sdílejí stejnou velikost objímky 2,5 mm. Při použití hybridní redukce je možné tyto tři konektory vzájemně propojit.

FC konektor byl, stejně jako ST, po mnoho let jedním z nejpoužívanějších konektorů pro změnu pro single mode. Není zde použit bajonetový způsob zapojení, ale šroubovací. Nicméně tento model byl postupně nahrazen konektory SC a LC

SC konektor ^(Obr. 11) je také vybaven 2,5 mm objímkou, která se široce využívá pro svůj výkon. Tento konektor je však proti předchozím modelům zásuvný. Z počátku bránila jeho rychlému rozšíření vyšší cena konektoru. Nyní je sice stále o něco dražší než ST konektor, ale již je mnohem běžnější ^[13].



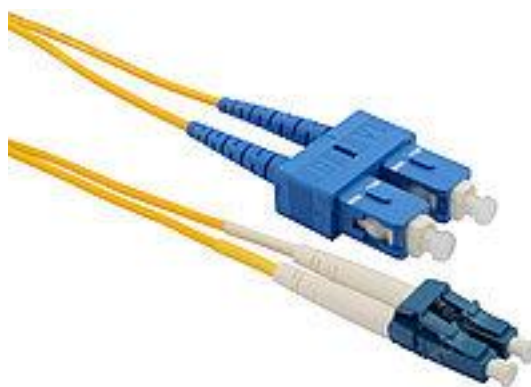
Obr. 11 SC konektor ^[13]

LC konektor ^(Obr. 12) je menší a využívá poloviční velikost objímky 1,25 mm oproti SC. Jedná se o standardní keramický spojovací kroužek snadno ukončitelný jakýmkoliv lepidlem. V posledních letech prošel konektor bouřlivým vývojem a v současné době patří mezi nejlepší varianty pro datové spoje.



Obr. 12 LC konektor [13]

Jak LC tak SC konektory ^(Obr. 13) jsou aktuálně nejčastěji využívány v jejich duplexních variantách. Za rozšířením duplexních variant obou konektorů a masivním rozšířením LC konektoru může i fakt, kdy do jednoho panelu je výrobce schopen dát 12 SC konektorů a 18 LC duplexních konektorů. Zde je již patrný nárůst 12 vláken SC a 36 vláken LC ^[13].



Obr. 13 Duplexní LC a SC konektory

Nejmladším konektorem je MPO (Multi-fiber push on) s nejdominantnější variantou MTP konektor. MTP konektory jsou optické konektory, které obsahují 8,12,16,24,48 a 72 optických vláken. Všechny tyto konektory mají velikost SC konektoru, a přitom pojmu až 72 vláken, což umožňuje připojení s velmi vysokou hustotou. Pro připojení vícero zařízení k jednomu MPO/MTP konektoru existují i hybridní řešení nebo rozdělená vlákna. Tedy na jedné straně může být MTP24 a na druhé straně může být kupříkladu 3x MTP8, anebo v kombinaci MTP – LC duplex ^[25]. Toto řešení je koncipováno pro rychlosti 40 Gbps, 100 Gbps a 400 Gbps. Respektive v případě MTP-LC duplex se jedná o 40 G to 10 G ^(Obr. 14).

MPO Patch Cables for 40G and 100G



MPO to LC Breakout Cables for 40G to 10G (x4)



Obr. 14 MPO (MTP) kably [25]

3.7 Fibre Channel (FC)

Vývoj Fibre Channel (FC) začal v roce 1988 jako součást projektu IPI a byl navržen k překonání rozhraní iSCSI a HIPPI ^[20]. Protokol prošel ANSI schválením 1. ledna 1994 ^[21].

V současné době stanovuje normy pro FC INCIS T11 ^[23]. Původně byl navržen pouze pro optická vlákna a později přibyla možnost využít ke spojení metalické kabely. Již v době schvalování bylo pracováno na dalším vývoji, a tak z původních 133 Mbit/s v roce 1993 byla již v roce 1997 rychlost 1 Gbps (1GFC) ^[22]. Od doby uvedení na trh se vždy, každých několik let, zdvojnásobuje rychlost na aktuální hodnotu 256 GFC. Každý FC je vždy zpětně kompatibilní s nejméně dvěma předchozími. Tedy 16 GFC je zpětně kompatibilní s 8 GFC a 4 GFC.

FC definuje komunikační vrstvy stejně, jako je tomu u modelů ISO/OSI a TCP/IP, avšak rozdělení je odlišné. FC, stejně jako v případě TCP/IP, rozděluje proces síťové komunikace do vrstev nebo do souvisejících skupin. Na rozdíl od ISO/OSI, který má vrstev sedm a TCP/IP, který má čtyři vrstvy, používá FC vrstev pět ^(Obr. 15):

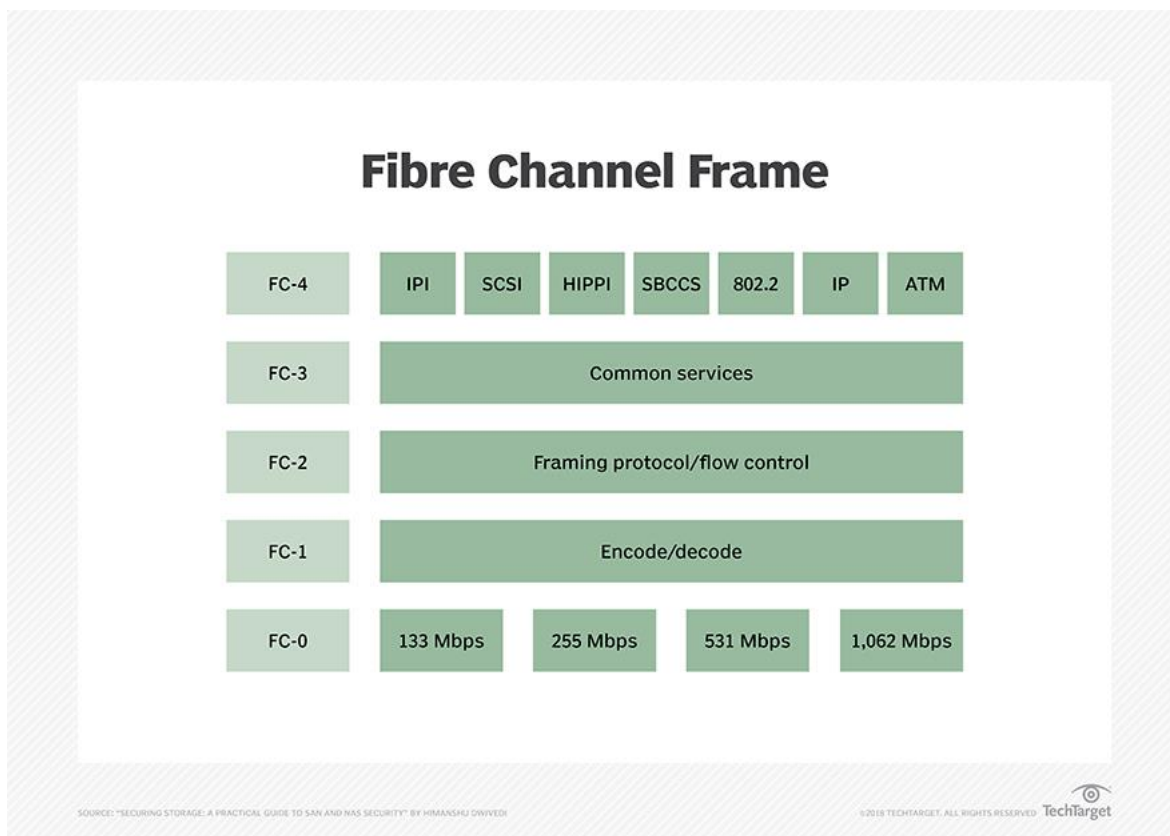
FC-4: vrchní vrstva služeb (iSCSI, HIPPI, IPI, ...)

FC-3: vrstva běžných služeb

FC-2: rámcový protokol, kontrola toku, zónování

FC-1: datová vrstva - kódování / dekodování

FC-0: fyzická vrstva – kabely, konektory, datový rámeček



Obr. 15 FC Frame ^[19]

Existuje několik druhů topologií FC. Nejpoužívanější jsou aktuálně dva. Bude zde vypsán i třetí, ale ten se již v současnosti běžně nevyskytuje.

PTP tedy Point-To-Point - jedná se o zapojení, které je nejjednodušší, tedy Host – Storage. Tento druh zapojení se většinou používá u serverů dedikovaných pro nějakou speciální činnost, kupříkladu u velkých SQL serverů, kde je kladen důraz na rychlost odezvy disků, vysoký počet IOPS a dále na možnost rychlého rozšíření diskového prostoru.

Kruh – toto zapojení se již běžně nevyskytuje. Jedná se o známé zapojení do kruhu, které sebou přináší nevýhody v případě výpadku celé sítě, v případě výpadku jediného prvku.

Hvězdice – jedná se, stejně jako v případě počítačové sítě, o standardní zapojení s centrálním FC switchem uprostřed. Toto řešení má také určité nevýhody v podobě závislosti na centrálním prvku, který v případě výpadku vyřadí celou síť. Pro eliminování tohoto problému se většinou přidává druhý FC switch. Tím je poté rozdělena zátěž, čímž nedochází

k přetěžování jednoho switche a v případě SW upgrade nebo HW výměny je vše provedeno bezvýpadkově.

Fyzická vrstva je založena na komunikaci přes sériové zapojení připojené pomocí optické, nebo metalické verze. Nicméně metalické se využívají tak málo, že se na ně nebude v této práci brát ohled. Pro připojení FC k jednotlivým zařízením existuje celá řada portů.

Nejpoužívanější porty jsou tyto:

L_Port – „Loop Port“ využívá se pro připojení právě do kruhu

E_Port – jedná se o trunk port propojení mezi jednotlivými FC switchi

F_Port – Fabrický port

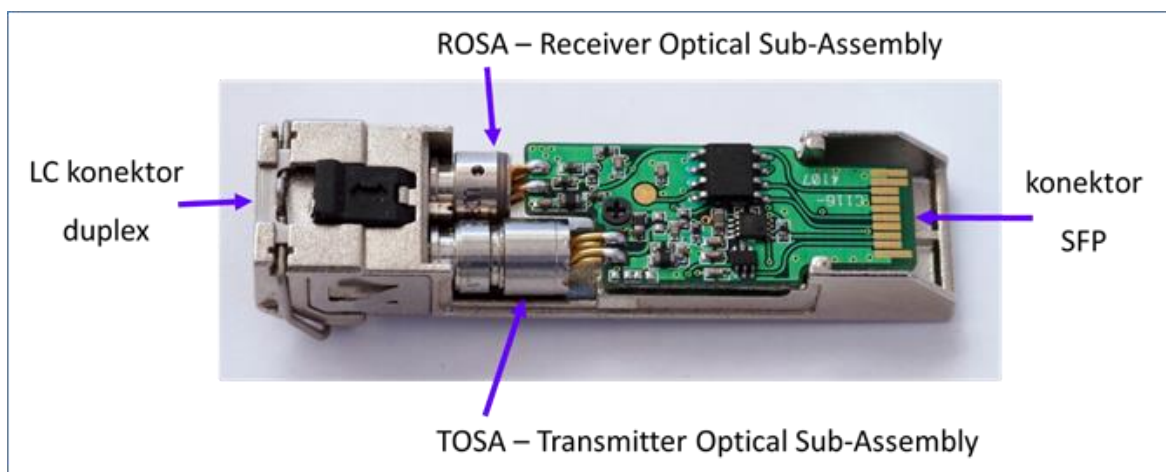
N-Port – používá se k hvězdicovému zapojení, případně může být připojen na F port

3.8 SFP moduly

Jedná se o malý zásuvný modul připojitelný za provozu (plug-and-play). Výhodou, která měla za následek i jeho masivní rozšíření a podporu mezi výrobci je, že port rozhraní může být dle potřeby vybaven jakýmkoliv vhodným typem transceiveru. Díky modulárnímu slotu lze využít specifické transceivery pro optické vlákno nebo pro metalický kabel.

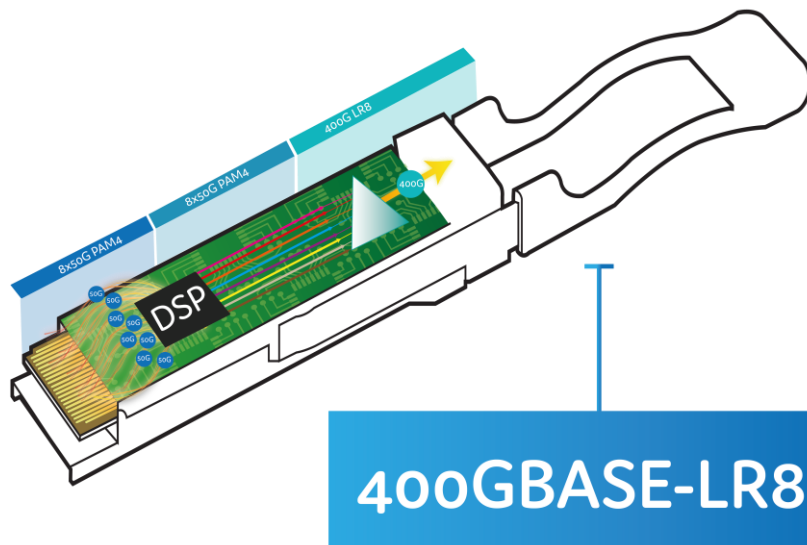
SFP moduly s metalickým připojením jsou vybaveny standardní zdířkou pro připojení konektoru RJ45. V případě optického SFP modulu je nejčastěji využit LC konektor.

SFP moduly se používají jak pro Gigabit Ethernet (GE) tak pro FC. V roce 2006 byla publikována specifikace SFP+. SFP+ podporují rychlosti 8 G pro FC a 10 G pro GE. Přesto, že standard SFP+ neobsahuje zmínku o 16 G FC, tak lze pro tuto rychlost využít. To je způsobeno metodou kódování, která je v případě 16 G - 64b/66b. Tato metoda umožňuje oproti 8 G – 8b/10b zdvojnásobení rychlosti přenosu dat bez nutnosti zdvojnásobení rychlosti linky (Obr. 16).



Obr. 16 SFP modul ^[24]

Další novinkou uvedenou v roce 2006 byly QSFP transceivery ^[27] (Obr. 17). Tento typ transceiveru umožňuje připojení 4 optických vláken a tedy další zhuštění, jelikož využívá stejnou velikost modulu. Pro připojení do QSFP se využívá optické vlákno MPO (MTP) nebo DAC (Direct-Attach Cable), který se využívá pouze na krátké vzdálenosti. Následně zrychlování QSFP přineslo v 2010 QSFP+ s rychlostí 40 Gbps ^[28] a dále QSFP28 s rychlostí 100 G GE a 32 G FC ^[29], QSFP56 s rychlostí 200 G GE a 64 G FC ^[30] až na QSFP-DD 400 Gbps.



Obr. 17 QSFP-DD 8X ^[31]

3.9 Technologie SSD

SSD (Solid State Disk) - technologie využívá k ukládání dat NAND Flash paměť. Oproti konvenčním rotačním diskům SSD neobsahuje žádné pohyblivé části, což činí toto úložiště odolnější proti nárazům. Při použití SSD technologie je důležité, jakými buňkami NAND chip disponuje. V současné době existuje několik typů NAND Flash pamětí, které jsou v technologii SSD využívány ^[51].

SLC – (Single Level Cell) oproti ostatním druhům tento chip přijme pouze jeden bit na paměťovou buňku (1,0). To má výhody v rychlosti a stabilitě, chip je méně náchylný k chybám. SSD osazeno těmito disky není určeno pro domácí prostředí, ale pro enterprise prostředí, kde je ztráta dat méně tolerována. Jejich nevýhodou je vysoká cena ^[51] (Obr. 18).

MLC – (Multi-Level Cell) proti SLC se do jedné buňky zapisují dva bity. To umožňuje vyšší hustotu než v případě SLC, ale také nižší stabilitu a rychlost. MLC má dále verzi eMLC, která má vyšší TBW a tím cílí na levnější enterprise řešení ^{[52][51]} (Obr. 18).

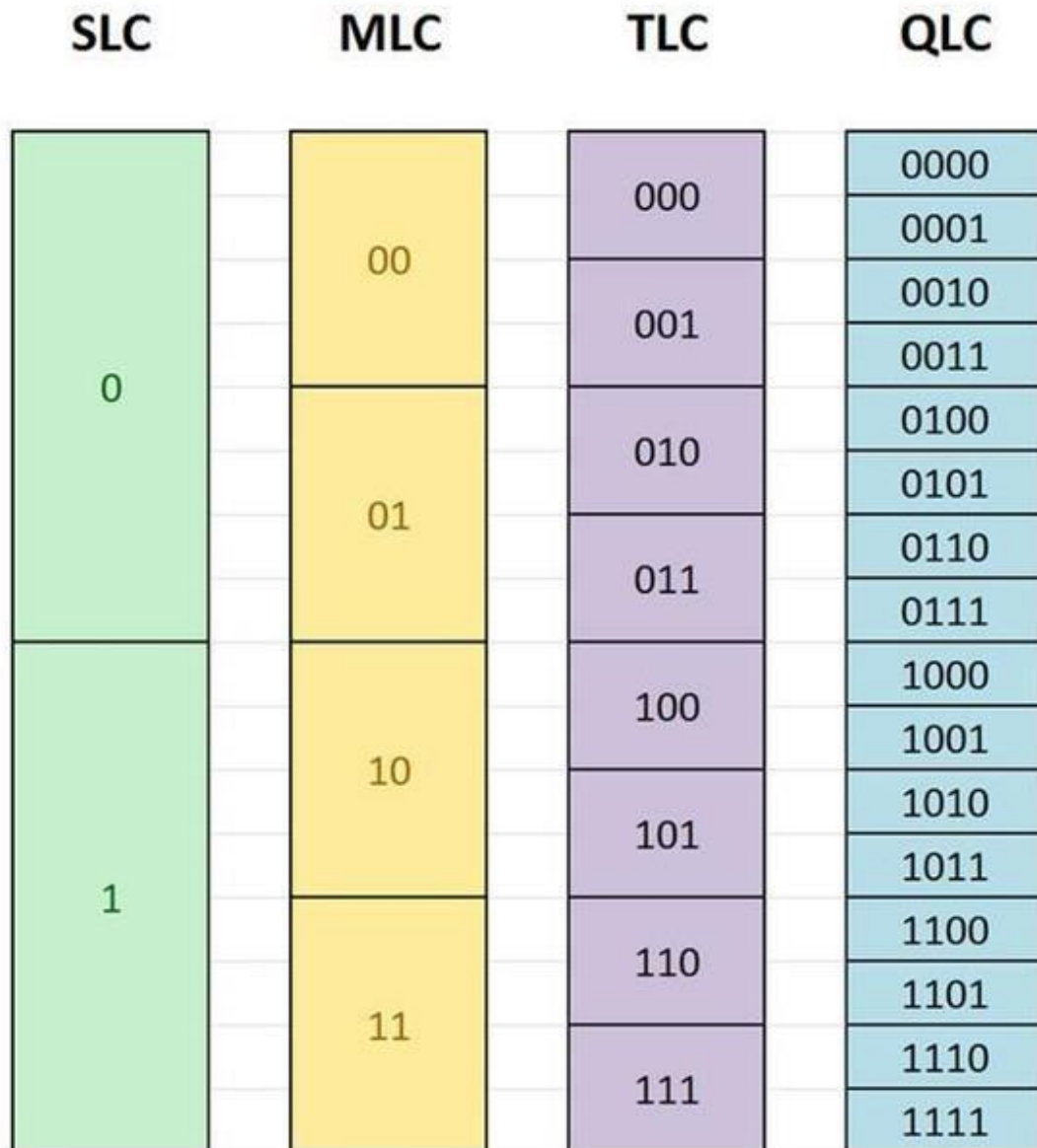
TLC – (Triple-Level Cell) jak je z názvu patrné, zde se již ukládají tři bity do každé buňky. Tento druh čipu je nejběžnější a jeho nejčastější užití je v koncových stanicích. Ať již pro firemní nebo domácí použití ^[51] (Obr. 18).

QLC – (Quad-Level cell) i zde je na základě názvu patrné, že se již ukládají čtyři bity. Tyto čipy jsou nejpomalejší a vyznačují se i menší stabilitou. Rychlost a stabilita je stále vyšší než u rotačních disků a dostatečná pro nenáročného domácího uživatele ^[51] (Obr. 18).

Životnost SSD se určuje dle toho, kolik terabyte je možné na disk zapsat (Terabytes Written – TBW). Tyto hodnoty jsou hodnoty bezpečné úrovně a jednotky SSD je běžně mnohonásobně překračují, hlavně v případě starších disků. U nejlevnějších SSD osazených QLC, pro domácí využití, jsou často hodnoty okolo 70 TBW a pro firemní koncová zařízení s TLC jsou hodnoty mezi 300-1000 TBW u SSD. U MLC jsou udávány hodnoty již 3-35 tisíc TBW, u SLC dosahují hodnoty 50-100 tisíc TBW. Z výše uvedených hodnot je patrné, že v případě průměrného SSD osazeného NAND pamětí TLC s životností 600TBW, je možné každý den po dobu pěti let zapsat na disk 329 GB dat.

$$(600000 \text{ GB} / (5 * 365) = 328,767 \text{ GB})$$

Tato hodnota je dostatečná pro firemní prostředí koncových uživatelů, nicméně v případě použití velkými SQL servery s náročným databázovým provozem nebo poštovními servery nemusí být tyto hodnoty dostatečné.



Obr. 18 Bitový zápis NAND Flash čipů ^[53]

3.10 VMware ESXi

VMware ESXi, dříve ESX, je hypervizor enterprise třídy-1 určený pro nasazení a správu virtuálních serverů. Do verze 4.1, vydané v roce 2010, kdy došlo i k přejmenování z ESX na ESXi, bylo při zavádění systému nejdříve spuštěno linuxové jádro. Až poté byly načteny jednotlivé specializované virtualizační komponenty včetně ESX, známé jako komponenta vmkernel. Od nové verze ESXi nahrazuje svůj základní operační systém Service Console pevně integrovaným operačním systémem. ESXi tedy funguje na „čistém železe“, tedy serveru bez operačního systému. ESXi obsahuje OS s vlastním jádrem.

ESXi poskytuje virtualizační vrstvu, která abstrahuje fyzické prostředky serveru pro více virtuálních serverů. Virtuální servery tedy využívají všechny prostředky bez přímého přístupu k fyzickému hardware podkladového serveru. VMware využívá hypervisoru vmkernel, který přijme požadavky od virtuálních strojů na zdroje a poté předkládá tyto požadavky fyzickému HW. Instalace ESXi lze provádět jak na interní HDD/SSD, usb nebo SD kartu, jelikož samotná instalace je miniaturní a spokojí se s cca 150 MB datového prostoru ^[32] (Obr. 199).



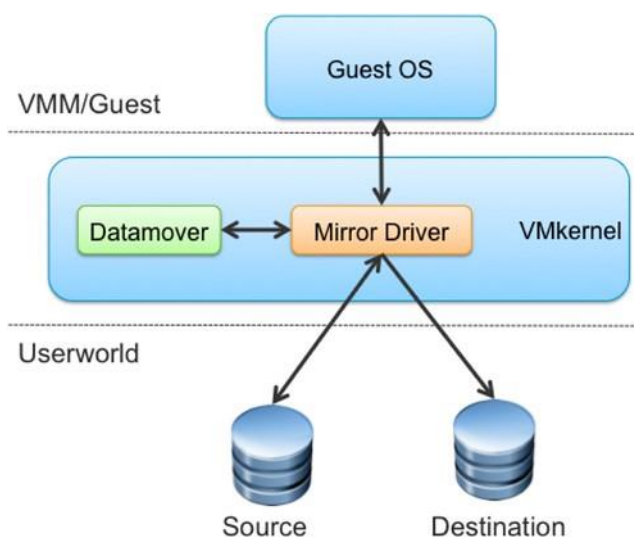
Obr. 19 ESXi host ^[32]

3.11 VMware vCenter

VMware vCenter je nástroj pro centralizovanou správu v prostředí VMware. Využívá se ke správě virtuálních strojů, ESXi hostitelů a všech závislých komponent. Pro správu ESXi hostů vCenter využívá dvou komponent vMotion a svMotion.

Komponenta vMotion (živá migrace), zajišťuje přesun virtuálního stroje (VM) na jiného ESXi hosta. Během migrace běžícího VM je obsah paměti odeslán z běžícího VM do nové instance VM na cílovém hostu ESXi. Jelikož se obsah paměti běžícího hosta neustále mění, využívá ESXi systém kontroly, který kontroluje stav paměti a případně odešle pouze změněná data. Před dokončením migrace dojde ke krátkému zmrazení stávající VM a přenesení do nového prostředí. Tento způsob přenesení minimalizuje dobu, po kterou je VM pozastaven [33].

Komponenta svMotion (Obr. 20) umožňuje přesun virtuálních disků včetně jejich adresářové struktury bez prostojů. svMotion využívá mirroring pro kopírování požadovaných dat ze zdrojového do cílového adresáře. Po celou dobu přesunu udržuje data synchronizovaná na obou místech, dokud není operace svMotion dokončena. Po dokončení migrace do cílové destinace jsou zdrojová data odstraněna. Spuštění svMotion má vliv na výkon diskových prostředků [34].



Obr. 20 svMotion mirroring [34]

Do verze vCenter 6.0 bylo možné k webovému klientu stále používat i tlustého klienta. Od verze 6.5 je již pouze webový klient, který právě touto verzí doznal několika změn. Asi nejvýraznější změnou je použití HTML5, která proti verzi FLEX, přinesla zrychlení odezvy celého systému.

4 Vlastní práce

4.1 Charakteristika aktuálního stavu a jeho zhodnocení

Současná infrastruktura, do které bude implementováno nové diskové pole, se skládá z menšího blade chassis C3000 od společnosti HP osazeného sedmi blade servery, dvou SAN switchů Brocade 300 a diskového pole Dell Unity 300. Infrastruktura dále využívá SW VMware. Pro provoz samotných serverů je využíván VMware ESXi Enterprise 6.0 a pro ucelenou zprávu serverů VMware Vcenter 6.5.

4.1.1 Blade chassis HP C3000 a servery

Toto chassis patří svojí velikostí do kategorie menších až středních, je vybaveno osmi sloty pro blade servery. Celý systém je navrhnutý jako dostatečně robustní a je osazen na současnou dobu výkonnými prvky. Chassis se dá rozdělit na tři části. Na čelní je device bay, kde se nacházejí blade servery a v zadní části na interconnect bay a power & thermal (Obr. 22²).

V device bay se nachází sedm z osmi možných serverů HP blade. Jak se servery postupně dokupovaly dle rostoucí poptávky po zdrojích pro virtualizaci, jsou zde různé modely jednotlivých blade serverů.

Nejstarší částí jsou dva servery BL 460 G6. Ty mají stejnou konfiguraci a jsou osazeny dvěma čtyřjádrovými CPU, Intel^(R) Xeon^(R) CPU E5540 @ 2.53 GHz, 196 GB RAM (osazeno 12x16 GB DDR3-1333Mhz). Na těchto dvou blade serverech již není možné další rozšíření a jsou tedy na maximu svých konfigurací. Dalšími prvky tohoto serveru jsou síťové a FC karty. Ty jsou na těchto serverech slabou stránkou celého řešení a jejich rozšíření není již možné. Síťové karty mají kapacitu pouze 1 Gbps a FC karty pouze 4 Gbps.

Další tři blade servery jsou BL 460 Gen8. Všechny mají opět stejnou konfiguraci a jsou osazeny dvěma šestijádrovými CPU, Intel^(R) Xeon^(R) CPU E5-2640 0 @ 2.50 GHz, 256 GB RAM (osazeno 8x32 GB DDR3-1333Mhz). Na těchto serverech je ještě možnost rozšířit paměť serverů na dvojnásobek, tedy dalších 8x32 GB až na maximální velikost 512 GB RAM. Dalšími prvky tohoto serveru jsou síťové a FC karty. Na těchto serverech je již

prostupnost síťové i FC karty dostatečnou. Jedná se o 10 GBps v případě síťové karty a 8 Gbps u FC karty.

Poslední dva servery jsou nejmladší a opět se jedná o stejné konfigurace a modelovou řadu BL 460 v generaci G9 (Obr. 21). Tyto modely jsou osazeny dvěma desetijádrovými CPU, Intel^(R) Xeon^(R) CPU E5-2640 v4 @ 2.40 GHz, 256 GB RAM (osazeno 8x32 GB DDR4-2133). Tyto servery se dají osadit až 16x64 GB moduly, takže výsledná kapacita by mohla být až 1 TB paměti. Dalšími prvky tohoto serveru jsou síťové a FC karty. I na těchto serverech je již prostupnost síťové i FC karty dostatečná. Jedná se o 10 GBps v případě síťové karty a 8 Gbps u FC karty.

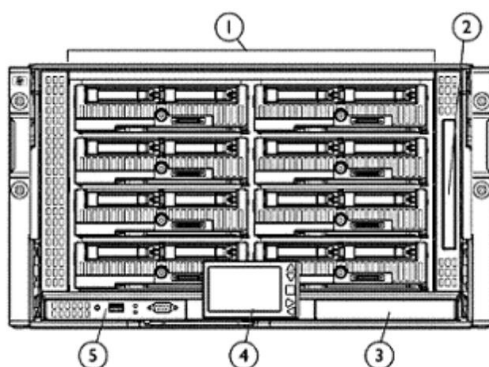


Obr. 21 HPE ProLiant BL460c Gen9 Server Blade^[35]

Interconnect bay je osazen dvěma HP VC Flex-10/10D^[36] moduly, které zajišťují síťovou komunikaci jak uvnitř blade chassis, tak i směrem ven. Do obou těchto modulů vede stejný počet optických vláken pro zajištění síťového provozu. Celkem se jedná o dvanáct vláken, tedy pro každý modul 6 single mode vláken s 10 Gbps SFP mouly. Dále je osazen dvěma HP VC 8 Gb 24-Port FC Moduly^[37]. Ty mají na starost komunikaci směrem k diskovým polím.

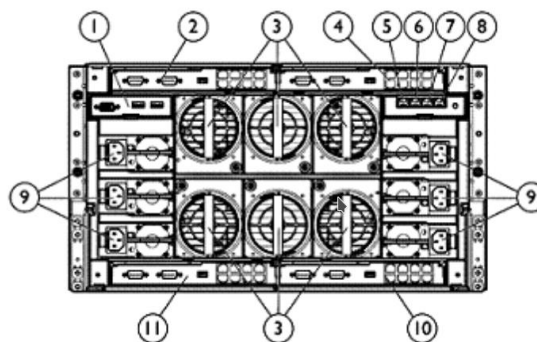
Zde je opět zapojeno 12 portů a všechny směřují do dvou SAN switchů Brocade 300. Použité SFP moduly jsou 8 Gbps. I zde je použito křížové zapojení pro minimalizaci nebezpečí výpadku.

Power & thermal část je plně obsazena. Je zde použito šest zdrojů, každý o jmenovitém výkonu 1200 wattů. Dva zdroje jsou aktuálně nastavené jako zálohy a jejich spuštění si řídí systém automaticky sám dle aktuálních požadavků, nebo v případě výpadku jednoho ze zdrojů. Řešení aktuálně spotřebovává maximálně cca 3000 wattů. A je zde tedy dostatek zdrojů pro případné rozšíření. Ventilačních jednotek je zde použito šest Active Cool FAN 100. Průměrné vytížení těchto ventilačních jednotek je cca 45 %. Systém pracuje s dostatečnou kapacitou ventilačních jednotek a je připraven pro výpadek jedné až dvou ventilačních jednotek bez hrozby přehřívání systému.



HP BladeSystem c3000 Enclosure – Rack – Front View

1. Device Bays 1 thru 8
2. Enclosure DVD Drive
3. Standby Onboard Administrator (reserved for Future)
4. Insight Display
5. Active Onboard Administrator



HP BladeSystem c3000 Enclosure – Rack – Rear View

1. Local KVM interface
2. Interconnect Bay 1
3. Active Cool Fans
4. Interconnect Bay 2
5. Enclosure Up-link and Service Port
6. Enclosure Down-link
- 7.
- 8.
- 9.
- 10.
- 11.

Obr. 22 HPE BladeSystem C3000 Enclosure [38]

4.1.2 Brocade 300 - HP StorageWorks 8/8 and 8/24 SAN Switches^[39]

Toto řešení je vybaveno dvěma stejnými jednotkami, aby byla zajištěna vysoká dostupnost celého řešení. Oba tyto switche mají identickou konfiguraci. Jedná se o switche, které jsou určené pro malé a střední enterprise řešení. Tento druh switche se základním typem pro SAN síť. Maximální rychlost, kterou je switch schopen přenést, je 8 Gbps na FC síti. Použitím FC switche dochází ke zjednodušení správy, zvýšení výkonu a snižuje náklady na použité úložiště. Tento typ switche lze vybavit FC moduly SFP a SFP+ a dosahovat rychlosti přenosu

1, 2, 4 a 8 Gbps v plně duplexním módu. Celé řešení umožňuje využít 24portu na každém zařízení. Tedy až 48 portů s maximální rychlostí 8 Gbps. Zde jsou použity výhradně 8 Gbps FC SFP+ moduly a jednotlivé prvky propojené k tomuto FC řešení využívají multi-mode optická vlákna. Dalším prvkem, který je zde využíván, jsou možnosti nastavení zónování. Pomocí zónování jsou zde nastaveny prvky, které na sebe vzájemně nevidí a obráceně jsou zde prvky, například zálohování, které vidí na všechna připojená zařízení. Jednou z hlavních nevýhod tohoto řešení je konfigurace. Pro konfiguraci tohoto řešení je sice využít webový přístup, nicméně je k němu vyžadována aplikace Java v určité verzi a s tím jsou spojené komplikace.

4.1.3 HUAWEI Core Switch S7706 ^[40] (Obr. 23)

Jedná se o duplexní řešení zajišťující síťovou konektivitu v celém řešení, včetně výchozí brány pro celou společnost. Oba switche mají identickou konfiguraci, jsou vzájemně propojené a veškerá zařízení jsou vždy zapojena do každého switche. V případě výpadku jednoho zařízení bude infrastruktura bez výpadku fungovat dále. Celá infrastruktura je navržena pro maximální dostupnost. Každý switch ^[41] je vybaven třemi 48 portovými kartami umožňujícími rychlost připojení 10/100/1000Mbps pro metalické spoje s konektory RJ45. Jako výchozí je zde nastavena 1 Gbps rychlost spojení. Veškeré metalické patch kabely, které jsou zde použity pro případné spojení se serverovou infrastrukturou, jsou v CAT6. Jedná se o případy, kde nelze použít optické kabely nebo je jejich využití zbytečné. Dále je v těchto zařízeních použita jedna karta ^[42] pouze pro SFP+ moduly a její rychlost je 10 Gbps. Do těchto modulů jsou svedeny všechny ostatní switche. Jedná se tedy o páteřní infrastrukturu společnosti nejen pro virtualizaci, ale i pro běžný provoz. Switche mají redundantní zapojení napájení, které je možné vyměnit za provozu.



Obr. 23 Huawei Core Switch S7706 [40]

4.1.4 DELL – EMC Unity 300 – aktuální stav [43]

Aktuální diskové pole DELL-EMC Unity 300 je vybaveno třemi lavicemi pro disky a je ve velikosti 6U. Diskové pole je rozděleno na tři části. Dvě části jsou určeny pro 2,5“ disky po 25 slotech pro disky a třetí část, která je určena pro disky velikosti 3,5“, je možné osadit až 15 disky. Diskové pole je kompletně redundantní pro zajištění maximální dostupnosti. Obsahuje dva kontrolery, přes které je rozložena zátěž a duální napájení. Toto diskové pole je vybaveno i NAS částí pro sdílení souborů. Disky mají celkovou kapacitu 40,8TB rozdělenou na tři části:

- ExtremePerformanceTier, pět 2,5“ FLASH disků zapojených do RAID 5 s celkovou kapacitou 1,4 TB
- PerformanceTier, patnáct 2,5“ SAS disků 15k, zapojených do RAID 5, s celkovou kapacitou 10,7TB
- CapacityTier, deset 3,5“ NL disků, zapojených v RAID 6 s celkovou kapacitou 28,6TB

Všechny disky jsou spojeny do jednoho diskového poolu, nad kterým je zapnutý automatický tiering. V pravidelném intervalu jednou za hodinu probíhá přesouvání dat mezi jednotlivými Tiery. Toto přesouvání dat probíhá na základě vnitřního algoritmu vyhodnocujícímu použití

dat. Tedy méně využívaná data jsou odkládána na kapacitní disky a nejčastěji využívaná data jsou přesouvána na PerformanceTier, případně do ExtremePerformanceTier.

4.1.5 Kritické problémy a zhodnocení současného řešení

Aktuální řešení jako takové je dobře navržené a je i dostatečně robustní. Nicméně řešení má i několik neduhů, které je potřeba odstranit. Nejpalčivějším problémem tohoto řešení je poddimenzované diskové pole. To se projevuje vysokou odezvou od diskového pole, která by optimálně měla být do 5ms. Pokud je odezva od diskového pole již kolem 20ms, je nutná investice do posílení a zrychlení. V tomto případě se ve špičkách, tedy při běžném denním provozu a spuštěném zálohování, objevují špičky v řádech několika vteřin. Takto obří výkyvy se již projevují na běžném provozu, kdy virtualizace směrem k diskovému poli má pause frames na FC. Pause frames se projevují jak zpomalením systému, ale i výpadky měřené pomocí příkazu ping, a tedy problémem s dostupností jednotlivých služeb. Dalším problémem je i celková disková kapacita stávajícího pole, která je rozdělena na čtyři části. SAN je rozdělen na tři datastore po 8TB a NAS 7TB. Toto rozdělení kapacity není při zpětném ohodnocení nejlepším řešením, jelikož zátěž na celé pole je enormní a kapacitní i flash část je naprosto nedostačující. Po dokončení implementace a odlehčení diskovému poli, bude provedena úprava konfigurace. Jeden datastore bude zrušen a ušetřené místo bude přerozděleno mezi zbývající dva datastorey s tím, že část zůstane prozatím volná, aby byla zachována rezerva. Na základě konzultace se specialistou ze společnosti DELL-EMC bylo rozhodnuto, že se do aktuálního diskového pole již nebude investovat, ale zvolí se nové řešení. Částka, která by byla nutná do tohoto diskového pole investovat, by převyšovala cca 1 500 000,- Kč bez DPH. Tato částka by situaci sice zlepšila, nicméně problém by setrval. Společnost se rozhodla zainvestovat do nového all-flash diskového pole DELL-EMC UNITY 450f^[44].

4.1.6 DELL – EMC Unity 450f – nové diskové pole^[44] (Obr. 24)

Toto diskové pole je svojí koncepcí již osazeno pouze 2,5“ SSD disky. Součástí instalace není NAS část, i když jí tento model 450f umožňuje. Zaměření diskového pole bude pouze na virtuální servery a požadavek na toto pole bude tedy čistě jen na hrubou sílu dodaného výkonu v IOPS a rychlosti odezvy. Samotné diskové pole je malé pouze 2U s možností až 25x2,5“ SSD disků s kapacitou od 400 GB až 15,36 TB a maximální kapacitou 4 PB. Jelikož

ne vždy je 25 x2,5“ SSD disků dostatečné, je zde možné připojit diskovou lavici DAE (Disk Array Enclosure). Ty se dodávají ve dvou provedeních, a to buď ve velikosti 2U s 25x2,5“ nebo 3U s 80x2,5“. Možností pro vytvoření RAID jsou 1/0, 5 a 6, což je pro určení diskového pole dostačující. Pro celé řešení jsou zde dva desetijádrové procesory Intel a 128 GB RAM. O připojení se starají dva kontrolery. Každý obsahuje vlastní management port, dva síťové a dva FC SFP+ moduly.

Pořízené řešení má základní konfiguraci a je vybaveno dvanácti 1,92 TB SSD disky v celkové využitelné kapacitě 14,5 TB.



Obr. 24 Dell EMC Unity 450F All-Flash Storage^[45]

4.1.7 Alternativy a výběr

Pro výběr diskového řešení existuje nepřehledné množství výrobců. Nicméně na základě osobních zkušeností autora a vedení společnosti byl výběr zredukován na čtyři výrobce DELL-EMC, HP, HUAWEI a NETAPP. Ve všech případech došlo ke kontaktování zástupců společností s poptávkou na disková pole. Na základě rozsáhlé vstupní analýzy, kterou ve spolupráci s autorem práce prováděli zástupci jednotlivých společností, byla následně vybrána tato řešení, DELL-EMC 450F, HP 3par 8200, Huawei Dorado 3000AFA a NETAPP EF280. Posledně jmenované bylo nabízeno jako Lenovo DE 4000F z důvodu nižší ceny. Jedná o stejné diskové pole, pouze nabízené pod značkou LENOVO. V rámci výběrového řízení všech čtyř polí byly prezentovány téměř identické možnosti výkonu. Každé pole nabízí navíc několik unikátních funkcí, avšak pro jeho určení jsou tyto funkce téměř zanedbatelné. Následně byla provedena další redukce výběru dle důvěryhodnosti, zkušenosti a vize společnosti. Jako první bylo vyřazeno diskové pole od spol. Huawei. Vedlo k tomu několik poznatků. Aktuálně spol. Huawei nemá z globálního hlediska úplně dobré resumé, a jelikož se společnost podílí na státních zakázkách, bylo od tohoto výrobce odstoupeno. Následně bylo vyřazeno Lenovo z velmi podobného důvodu. Sice se jedná o značku, o které se v souvislosti s Čínou vládou nemluví, avšak i zde má velké zastoupení. Po vyřazení těchto dvou výrobců

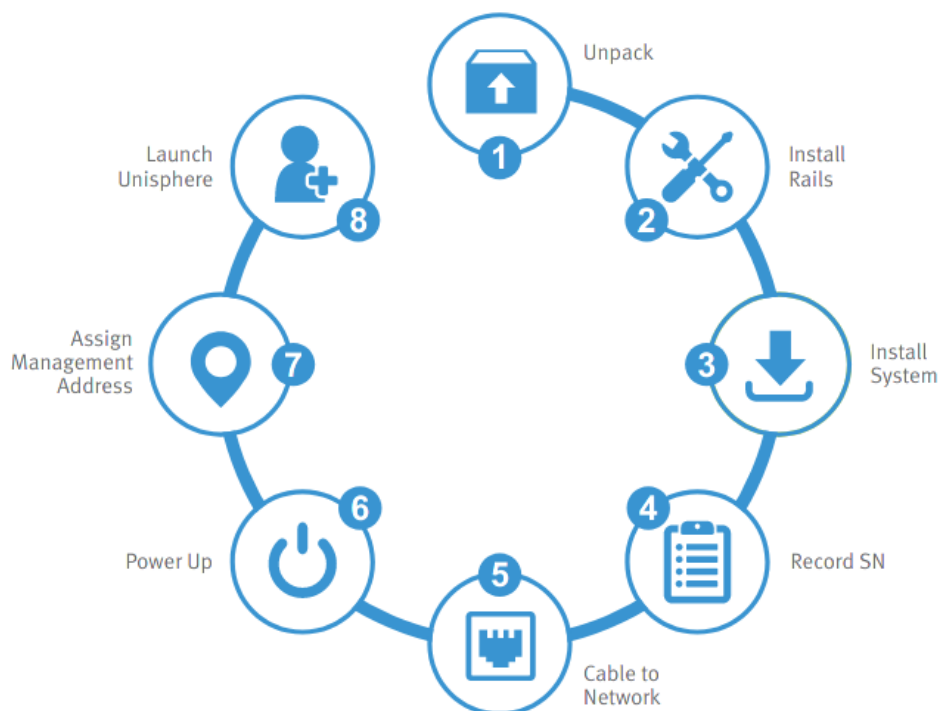
bylo kontaktováno zastoupení spol. NETAPP k dodání cenové nabídky. Jejich cena však byla téměř o třetinu vyšší než u zbývajících nabídek DELL-EMC a HP. K finálnímu rozhodnutí, zda jít cestou DELL-EMC nebo HP, které se nakonec přiklonilo na stranu DELL-EMC, byla v první řadě cena, která byla nepatrně nižší, cena disků pro pozdější rozšíření a poté vedení společnosti, které tomu bylo více nakloněno.

4.2 Instalace a konfigurace pole

Samotná instalace diskového pole bude provedena v několika krocích. Prvním krokem bude fyzická instalace pole včetně propojení FC a síťových optických vláken a metalických kabelů pro připojení k managementu pole. Dalším krokem bude konfigurace síťové části, poté samotná instalace diskového pole a konfigurace SAN switche. Na závěr bude provedena implementace nového řešení do prostředí VMWare.

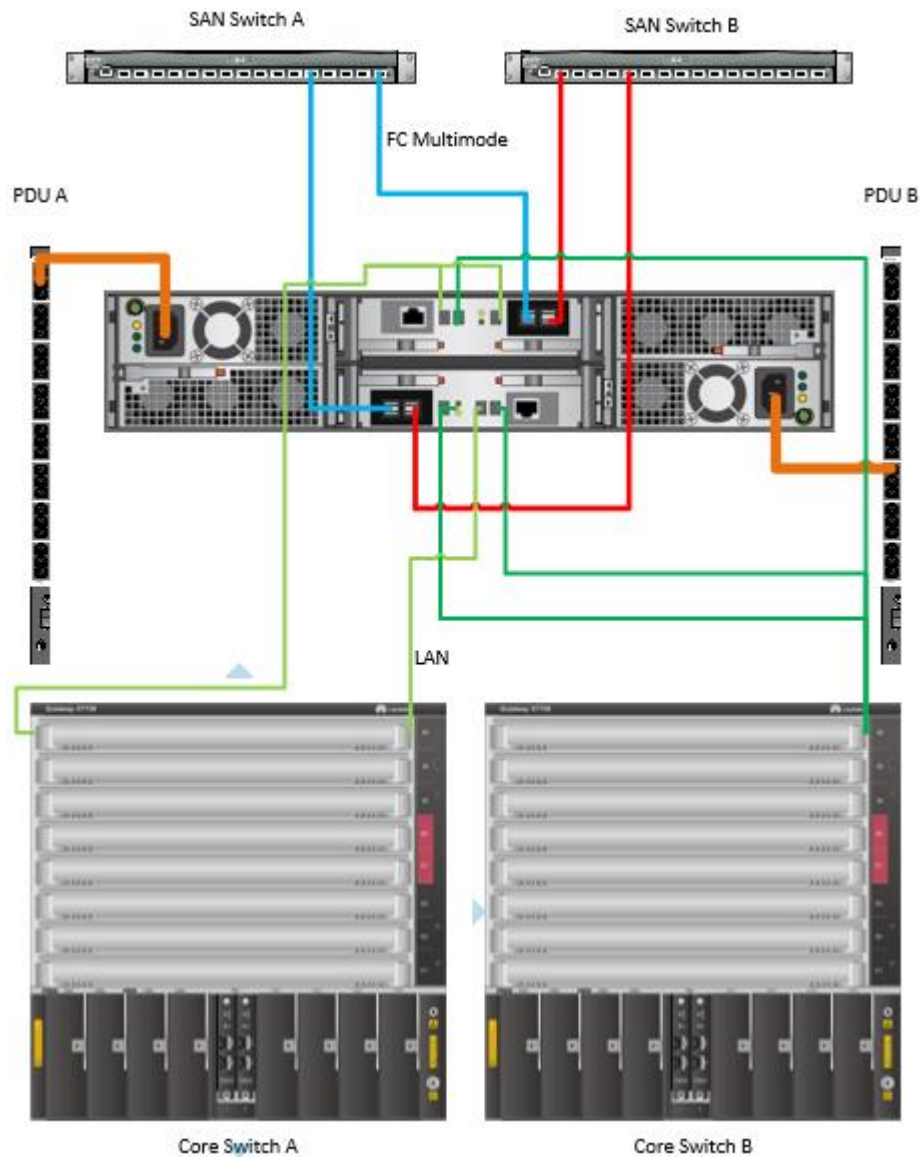
4.2.1 Fyzická instalace

Diskové pole je rozměry určeno jako „rackové“ a ne samostatně stojící. Je tedy nutné uskutečnit jeho fyzickou instalaci do racku. Celá instalace se provádí dle doporučení výrobce v několika krocích (Obr. 25).



Obr. 25 Postup instalace ^[46]

Aby bylo možné diskové pole uchytit v racku, je potřeba pro něj připravit prostor o velikosti 2U a nainstalovat kolejnice pro uchycení. Po namontování do racku je nutné diskové pole připojit k jednotlivým prvkům. Jelikož je celý systém navržen jako vysoce dostupný, tak i připojení všech prvků diskového pole budou zapojena tak, aby toto kritérium splňovala. V tomto případě se tedy jedná o následující zapojení, které je patrné z nákresu níže (Obr. 25). Každý zdroj pole bude zapojen do separátních PDU, které je každé zapojeno na separátní větvi napájení a zálohování. Optické spojení s jednotlivými SAN switchi bude provedeno z každého kontroleru do obou SAN switchů. Jedná se o propojení dohromady čtyřmi 8 Gbps SFP FC a multimode s duplexními LC/LC konektory. Dohromady tedy bude diskové pole propojeno osmi optickými vlákny. Podobný způsob zapojení bude proveden i v případě LAN spojení, tedy každý kontroler do obou Core switchů. Jelikož toto pole není koncipováno jako NAS nebo kombinace NAS/SAN, bude jeho síťová vrstva připojena pouze pomocí metalických kabelů s rychlostí 1 Gbps.



Obr. 26 fyzické zapojení

Na vyobrazeném zapojení (Obr. 26) je modrou a červenou barvou oddělené optické spojení. Světle zelenou a zelenou barvou jsou znázorněna zapojení datové části a managementu jednotlivých kontrolerů pomocí metalických kabelů UTP v kategorii Cat6. Dále je zde zobrazené připojení napájení do jednotlivých PDU. Každé PDU je napojeno na samostatnou napájecí větev. Celý systém je tedy zapojen s ohledem na maximální možné zabezpečení proti selhání. Pro toto zapojení je zde pouze jeden záložní zdroj UPS, který je s aktuálním řešením schopen zajistit 45minut chodu celé serverové infrastruktury na baterie při výpadku el. energie.

4.2.2 Konfigurace síťových prvků

Aby bylo možné se k samotnému diskovému poli připojit, je nutné změnit konfiguraci síťových prvků a jednotlivých portů. Úprava bude vykonána na obou síťových prvcích CORE switch. Konfigurace se provádí pomocí několika příkazů z terminálu. Pro připojení k požadovanému CORE switchi je využit protokol SSH.

```
ssh <ipadresa_switche> -l uzivatel
```

Zadáním hesla se přihlásíme na konfigurovaný switch a začneme provádět úpravy konfigurace. Jednotlivé příkazy se po zapsání potvrzují enterem. Na závěr je nutné pomocí příkazu save zapsat konfiguraci, aby nedošlo k její ztrátě po restartu zařízení.

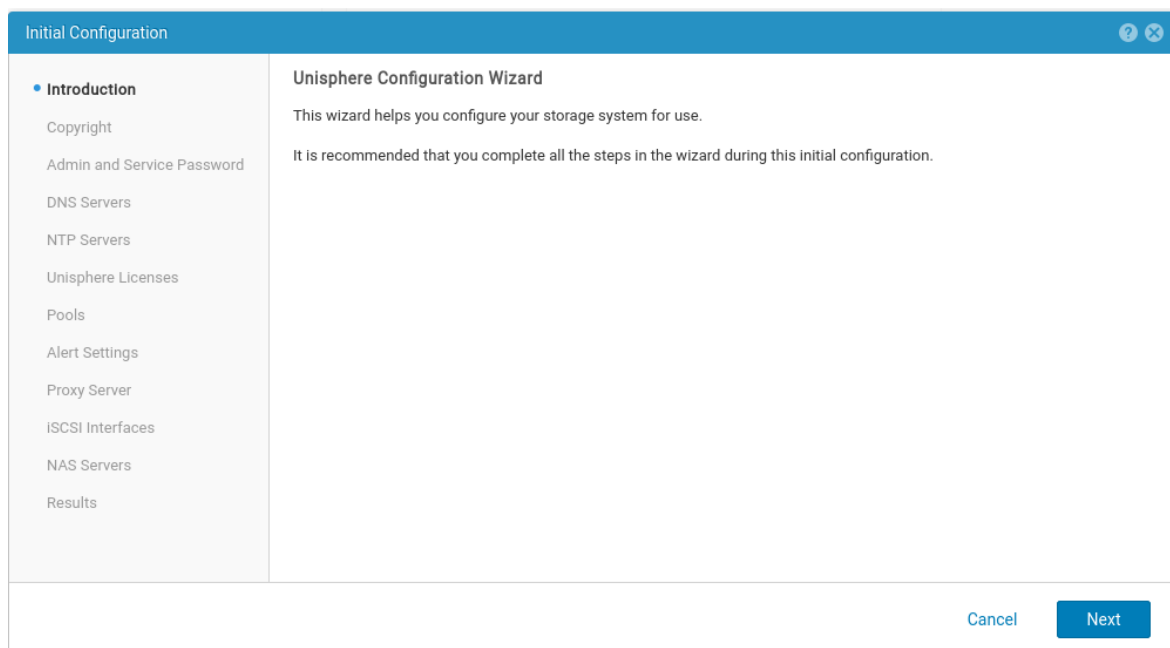
```
system-view // tento příkaz umožní přístup ke konfiguraci („conf t“ v prostředí cisco)  
interface XGigabiteEthernet 2/1/0/30 // tímto příkazem vybereme požadovaný port pro konfiguraci  
port link-type access // nastavení typu portu  
description U450f_CA // nastavení popisu portu pro větší přehlednost jednotlivých portů  
port default vlan 245 // definování VLAN  
undo shutdown // povolení portu  
quit // ukončení konfigurace interface  
quit // ukončení system-view  
save // uložení konfigurace, která se stane permanentní a nesmaže se po restartu zařízení
```

Tímto způsobem bude provedena konfigurace na všech požadovaných portech.

4.2.3 Základní část konfigurace diskového pole

Po kontrole veškerého zapojení bude provedeno první spuštění systému a následná konfigurace. Prvotní spuštění systému může trvat 15-60 minut ^[46].

Samotná konfigurace diskového pole je intuitivní. Po prvním přihlášení do systému se automaticky zobrazí průvodce konfigurací ^(Obr. 277).



Obr. 27 Průvodce konfigurací

Prvním krokem v průvodci je potvrzení autorských práv a následné definování administrátorského a servisního hesla. Tyto dva účty jsou základními účty, které tento systém používá. Administrátorský účet je využíván pro vytváření a řízení dalších uživatelských účtů a komplexní nastavování systému. Servisní účet je jedinečný účet v systému a nelze ho smazat nebo vytvořit nový ^[47]. Z tohoto důvodu je dobré vygenerovat dostatečně silné heslo jak pro servisní účet, tak i pro lokálního administrátora. Dle bezpečnostních doporučení je optimální délka hesla 12 znaků při použití malých a velkých písmen abecedy, čísel a speciálních znaků (*, @, !, #, &, +, /, atd.). Překonání takto komplexního hesla je dle webových analyzátorů ^{[48][49]}, cca 34tisíc let. Nicméně s vývojem výpočetní techniky se i toto heslo stane zranitelnější a doba prolomení se jistě zkrátí.

Při vyplňování hesla administrátorského účtu lze zaškrtnout použití stejného hesla i pro servisní účet. Z bezpečnostního hlediska je autorem práce doporučováno vygenerování nového samostatného hesla ^(Obr. 288). Heslo k servisnímu účtu je potřebné kupř. pro upgrade FW a v některých případech je nutný zásah od společnosti DELL-EMC, se kterou je poté heslo sdíleno. Tedy v tomto případě není žádoucí, aby heslo, které se může sdílet s cizí společností, bylo identické jak pro servisní, tak pro administrátorský účet.

Initial Configuration

- Introduction
- Copyright
- Admin and Service Password**
- DNS Servers
- NTP Servers
- Unisphere Licenses
- Pools
- Alert Settings
- Proxy Server
- iSCSI Interfaces
- NAS Servers
- Results

Admin and Service Password

Admin Password

Username: admin

Password: ●●●●●●●●

Confirm Password: * ●●●●●●●●

Set service password the same as admin password

Service Password

Username: service

Password: _____

Confirm Password: _____

Cancel Back Next

Obr. 28 Admin a Service password

Dalším krokem průvodce je nastavení DNS serverů. Na základě osobních zkušeností autora je doporučeno mít nastaven jeden z DNS serverů fyzický, aby nebyl závislý na virtualizaci.

Obecně je doporučováno mít alespoň jeden z AD serverů, který bývá i nositelem DHCP a DNS serveru, takto oddělen. Tento server by měl být po výpadku el. energie, nebo při startování systémů po nějaké nehodě, nastartován vždy jako první. Toto řešení nám zajistí funkční služby AD, DNS a DHCP nezávisle na virtualizaci a dále správnou konfiguraci jednotlivých prvků spouštěných posléze.

Při zadávání DNS serverů byly zadány tři servery, z nichž jeden je fyzický a zbylé dva virtuální. Tato konfigurace dává větší volnost administrátorům při aktualizacích a následných restartech jednotlivých serverů a tím i větší pohodlí při správě infrastruktury.

Následuje konfigurace NTP serverů, kde budou přidány čtyři servery. Jedná se o české servery (Tabulka 1).

0.cz.pool.ntp.org
1.cz.pool.ntp.org
2.cz.pool.ntp.org
3.cz.pool.ntp.org

Tabulka 1 NTP servery^[50]

Nastavení čtyř serverů má své opodstatnění v zabezpečení přesného času. V případě jednoho serveru není čas, který je získán ze serveru, ničím ověřen. Při zadání dvou serverů již dochází k ověřování času, avšak pokud jeden ze serverů bude mít špatný čas, bude vybrán server s rychlejší odpovědí, a tedy to nemusí být server se správným časem. Proto musí být minimálně tři servery. Ty jsou již schopny podstatně lépe určit přesný čas. Ani tři servery ale nejsou optimálním řešením. Pokud by jeden ze serverů vypadl a jeden měl chybný čas, vrátilo by nás to zpět do dvou serverového paradoxu, který je nežádoucí. Optimálním řešením jsou servery čtyři. Při výpadku jednoho serveru a chybovém čase druhého jsou stále dva servery, které zajistí správný čas a tím se zabrání co možná nejvíce chybám spojeným s nastavením času ^[50].

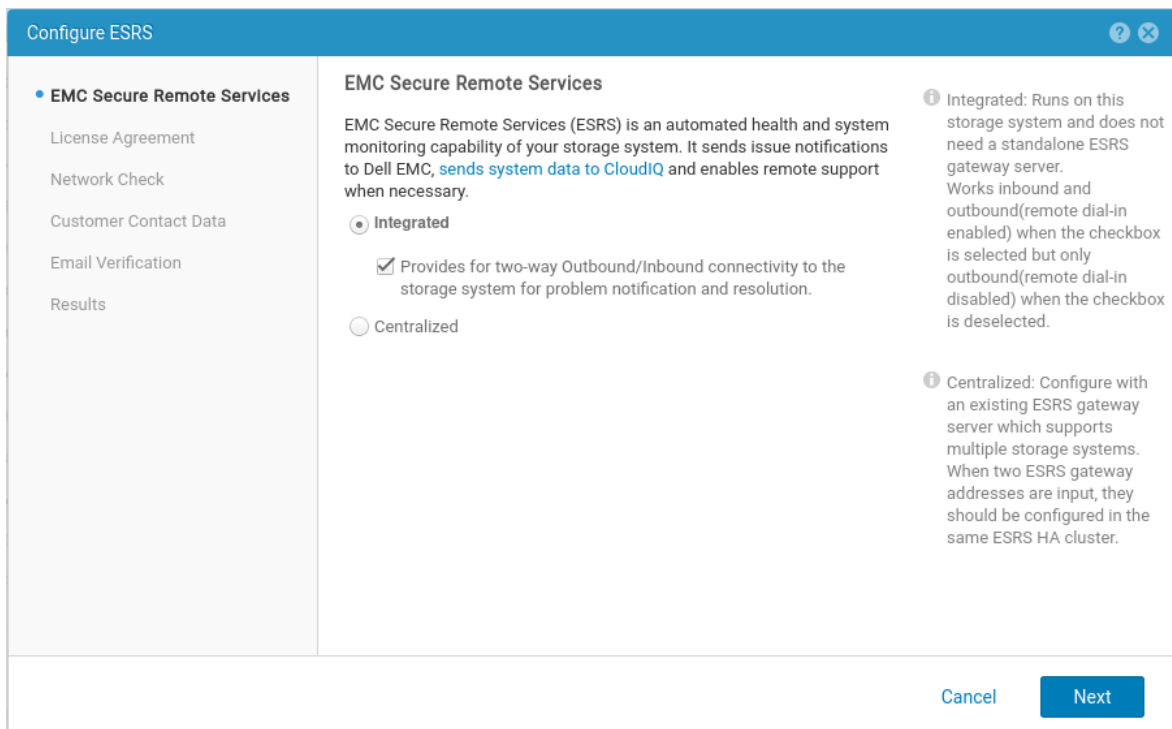
Následuje nahrání souboru, jenž obsahuje licenci pro jednotlivé funkcionality diskového pole. Veškeré nahrané licence jsou permanentní, a tedy není u nich čas vypršení.

počet SSD disků, které budou součástí aktuálně konfigurovaného poolu. Jelikož se jedná o diskové pole s menším počtem disků, bude do současného poolu přidáno všech 12 jednotek. V této části konfigurace se již nic dalšího nenastavuje a průvodcem bude dokončeno vytvoření poolu.

4.2.5 Dokončení základní konfigurace

Pro kompletní dokončení základní konfigurace je nutné provést několik dalších kroků. Nastavení výstrah bude v tomto kroku přeskočeno a bude případně konfigurováno později. Připojení k internetu pomocí proxy serveru bude ponecháno nevyplněné. Společnost sice proxy servery k připojení na internet využívá, ale pouze pro koncová zařízení. Serverová infrastruktura proxy server nevyužívá. V rámci zvýšení zabezpečení nemá každý server povolený přístup k internetu a na firewallu společnosti bude přidán záznam o tomto diskovém poli, aby byl přístup umožněn. Následuje vyplnění uživatelského jména a hesla pro komunikaci s podporou Dell-EMC. Po jejich zadání je nutné vyplnit kontaktní informace pro hlavní technický kontakt za společnost pro komunikaci s Dell-EMC.

V této části konfigurace je stěžejní nastavení ESRS (EMC Secure Remote Services), které slouží pro monitorování systému a odesílání servisních dat výrobci. Na základě těchto dat jsou zpět získávány informace o aktualizacích a bezpečnostních záplatách celého systému nebo nového Firmware pro celý systém i jednotky SSD ^(Obr. 30).



Obr. 30 konfigurace ESRS

V prvním kroku jsou na výběr dvě možnosti ESRS systému, Integrovaný nebo Centralizovaný. Integrovaný ESRS je součástí diskového pole a nic dalšího není vyžadováno, vyjma nastavených síťových prístupů. Na FW je nutné povolit obousměrnou komunikaci na portech 443 a 8443 s omezením na dostupnosti z adres **.emc.com* ^[54].

V případě centralizovaného řešení je předpoklad jiného již běžícího diskového pole, ke kterému by se nové pole připojilo. Vytvořil by se ESRS Cluster a komunikace by probíhala přes již jednou vytvořený síťový prístup ^[54].

V dalších krocích bude potvrzeno licenční ujednání a provedena kontrola síťové dostupnosti. I zde je možné vyplnit připojení pomocí proxy serveru. Dále jsou opětovně vyplněny kontaktní informace. Pro dokončení je nutná emailová verifikace (Obr. 31).

Configure ESRS

We've sent you an access code to: k*****@eltodo.cz.

- ✓ EMC Secure Remote Services
- ✓ License Agreement
- ✓ Network Check
- ✓ Customer Contact Data
- **Email Verification**

Results

Email Verification

1. Send an access code to the email address kolmanm@eltodo.cz which you registered with support.emc.com.
[Send access code](#)
2. Enter the access code you received. The code expires 30 minutes after it is sent.

Access code: *

[Alternative for Support personnel only](#)

If you did not receive an email from an EMC address, contact EMC Support for assistance. <https://support.emc.com>. Alternative authentication is available for EMC Support personnel only.

Cancel Back Next

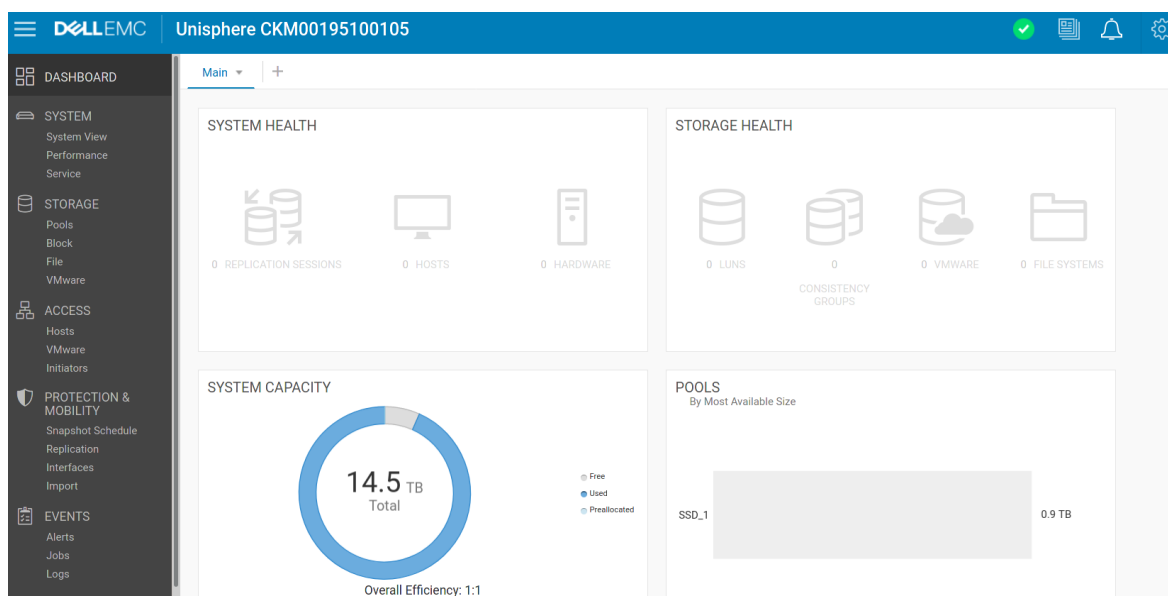
Obr. 31 ESRS ověření emailu

Po přijetí přístupového kódu je nutné ho zadat do pole pro jeho ověření. Po zadání kódu dojde k jeho ověření a je možné dokončit konfiguraci ESRS.

Zbývají poslední dva kroky průvodce, iSCSI Interface a NAS server. Toto diskové pole nebude připojeno pomocí iSCSI a neobsahuje NASovou část. Tyto dvě části tedy nebudou konfigurovány. Potvrzením poslední volby dokončíme prvotní konfiguraci a ukončíme průvodce.

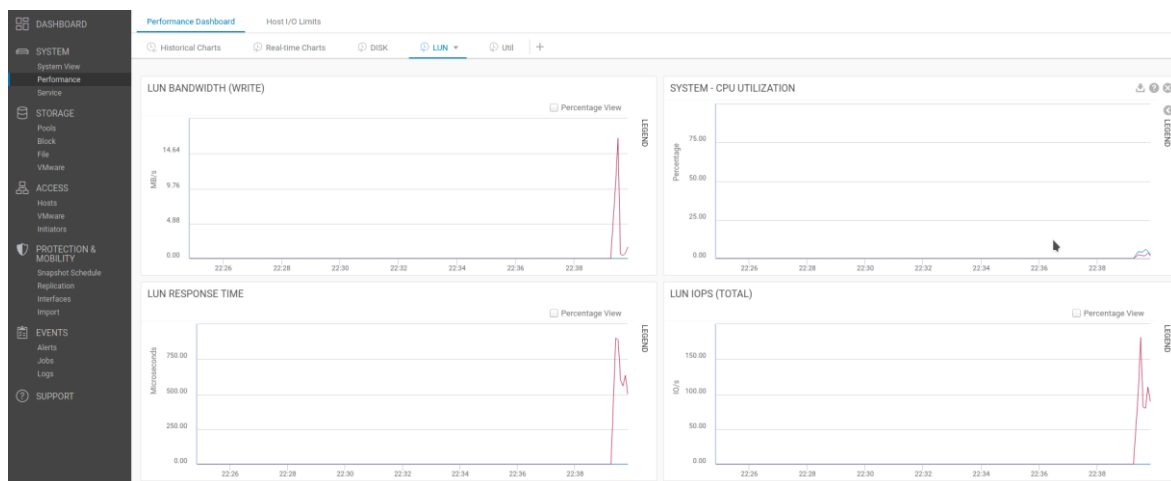
4.3 Unisphere

Po přihlášení k webovému managementu unisphere je zobrazený *DASHBOARD* s předkonfigurovanými dlaždicemi zobrazujícími základní informace o diskovém poli. Jednou z hlavních informací je *System Health*, *Storage Health*, *System Capacity* a zobrazení jednotlivých poolů (Obr. 32). *System* a *Storage Health* zobrazuje status jednotlivých částí, což umožňuje rychlou identifikaci případných problémů. *System capacity* zobrazuje celkovou kapacitu diskového pole a zbývající volné místo. *Pools* zobrazuje volné místo na jednotlivých poolích. V tomto případě zobrazuje pouze jeden, jelikož zde není více poolů vytvořeno.



Obr. 32 Unisphere Dashboard

Celý webový management je rozdělen na několik hlavních částí s jednotlivými pododkazy. *SYSTEM* slouží pro zobrazení systémových informací a v části „Service“ jsou možnosti nastavení jednotlivých služeb s možností jejich restartování nebo zastavení. Nachází se zde možnosti restartování jednotlivých kontrolerů v případě problémů nebo ručního upgrade FW. Další důležitou a hojně využívanou částí je pododkaz „Performance“. Tato část slouží pro kontrolu výkonu pole, rychlost odezvy pole a jednotlivých datastorů a disků. Rychlost odezvy diskového pole by měla dosahovat do 5ms. Pokud je rychlost nad 5ms, tak již může docházet ke zpomalení provozu (Obr. 33).



Obr. 33 Performance

Následuje položka *STORAGE*, která slouží pro vytvoření diskových poolů, blokových a souborových přístupů (NAS, SMB, NFS Share) a VMware. V poslední jmenované části lze vytvářet *datastore* nebo *Virtual Volumes*. Další položkou je *ACCESS*. Zde se provádí nastavení přístupů k diskovému poli. V položce *PROTECTION & MOBILITY* je možné nastavit kupř. snapshoty a replikace na jiné diskové pole. Předposlední možností jsou *EVENTS*, kde se nacházejí výstrahy na jednotlivé problémy, naplánované úlohy a logy. Poslední možností je položka *SUPPORT*, ve které se dá na základě jednoduchých odkazů kontaktovat podpora, diskuzní fórum, video návody, nebo objednáni nových, či vadných disků.

4.4 Vytvoření datastoreu a připojení jednotlivých ESXi hostů

Nyní bude vytvořen datastore, který bude sloužit jako úložiště pro virtuální stroje, VMware VMFS (Virtual Machine File System). Pro založení nového datastore bude v části *STORAGE* vybrán *VMware* a v záložce *datastore* zvoleno jeho vytvoření. Tímto krokem dojde ke spuštění průvodce pro vytvoření požadovaného úložiště (Obr. 34).

Create VMware Datastore

Type

Name

Storage

Access

Snapshot

Replication

Summary

Results

VMware Datastore Type

File
Access the VMware NFS datastore using the NFS protocol

Block
Access the VMware VMFS datastore using the block protocols

VMware VMFS version:
VMFS5

VVol (File)
Provision datastore to be used for VMware Virtual Volumes (File)

VVol (Block)
Provision datastore to be used for VMware Virtual Volumes (Block)

At least one NAS server is required to create File Datastore

At least one capability profile is required to create VVol Datastore

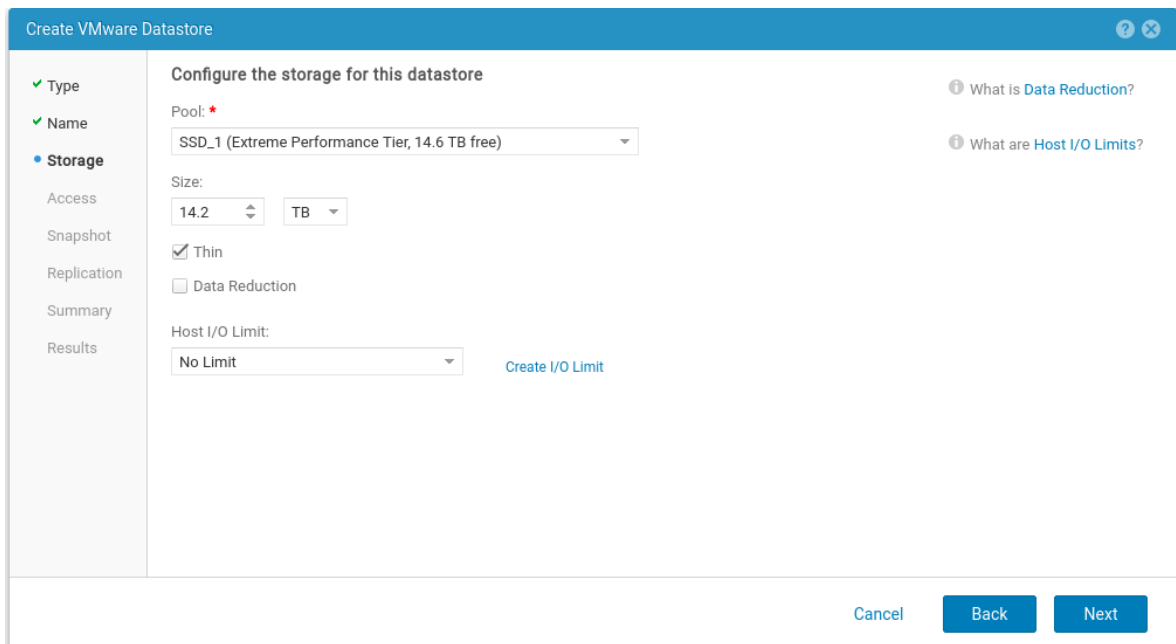
Cancel Next

Obr. 34 VMware Datastore Type

Jelikož společnost využívá i některé starší stroje, na které již nelze instalovat novější verzi ESXi a je nutné zachovat verzi VMware ESXi 6.0, bude zvoleno VMFS5 [82].

Hlavní výhodou, která použitím starší verze VMFS 5 nebude využita, je nová funkce „*Automatic Space Reclamation*“. Tato funkce umožňuje diskovému poli si vyžádat již nevyužívané místo po smazaných nebo přesunutých diskových blocích a využít jejich kapacitu jinde [82].

Potvrzením volby a následným přiřazením jména datastoru, které bude viditelné ve vCenter, bude průvodce pokračovat ke konfiguraci úložiště (Obr. 35).



Obr. 35 Kofigurace úložiště

Zde je možné vybrat z jednotlivých poolů. Toto řešení má pouze jeden pool, a tedy je vybrán vytvořený SSD_1. Dále bude alokováno maximální množství dostupného místa, což je 14,2TB. Velice důležitou volbou, která nesmí být opomenuta, je zaškrtnutí volby „Thin“.

Touto volbou je umožněno efektivnější nakládání s volným místem, které je alokováno pouze dle skutečných nároků na velikost úložiště. Toto řešení umožňuje lépe spravovat a předvídat růst dat na úložišti a dále snižuje počáteční náklady.

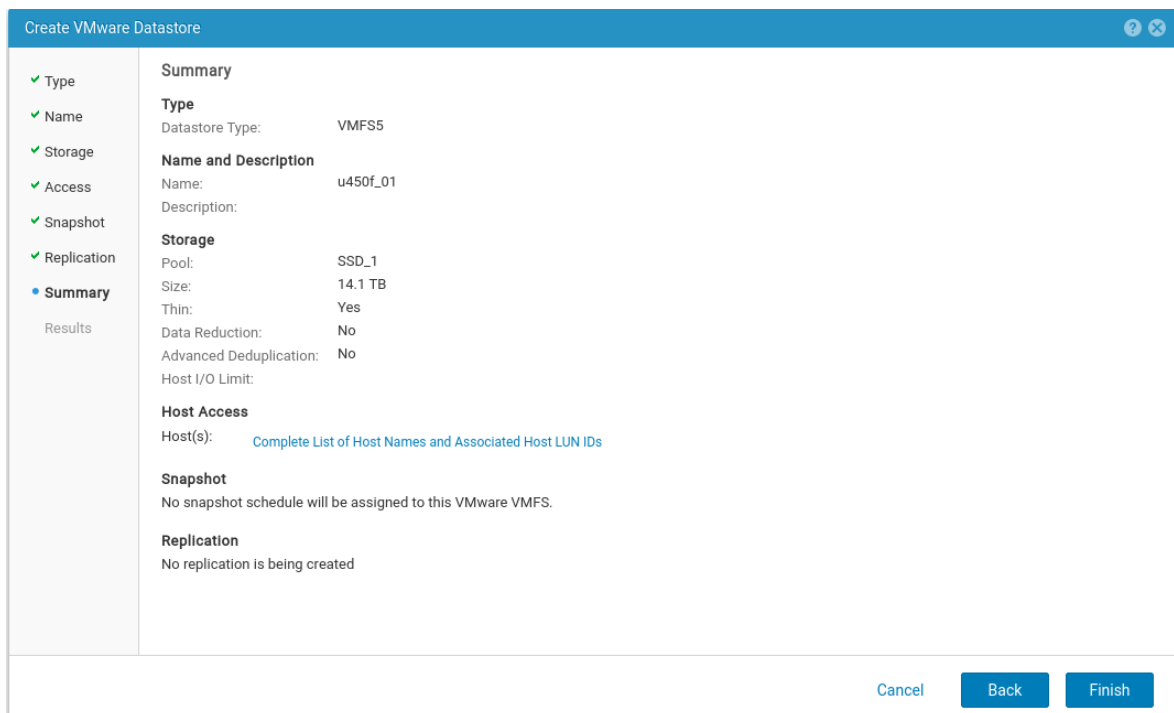
Další volbou je „Data reduction“. Ta sice umožňuje snížení nároků na místo pomocí deduplikace dat, ale jelikož na tomto poli budou hlavně náročné aplikace typu SQL a exchange server, tak tato volba zůstane nevybrána.

Poslední možností je nastavení volby Host I/O Limit. Výchozí volbou je možnost „No Limit“. Veškeré další volby je nutné vytvořit. Toto nastavení umožňuje definování maximální zátěže akceptovatelné pro tento datastore. Tato funkce se využívá v případech, kdy je na jednom diskovém poli více poolů nebo datastorů. Nastavováním limitů lze některé prioritizovat. Tato volba zůstane ve výchozím nastavení bez limitu.

V další části se nastavují možnosti připojení jednotlivých hostů, kterým je umožněn přístup na tento datastore. Autorem diplomové práce je doporučeno tento krok konfigurace nyní vynechat a nastavit přístupy až po dokončení vytvoření datastore.

Následujícím krokem konfigurace je možnost nastavení automatických snapshots. Jelikož tento datastore bude sloužit čistě pro virtualizaci a společnost vlastní sofistikovaný zálohovací software, není vytváření snapshotů touto cestou žádoucí.

Závěrečnou konfigurační položkou je možnost nastavení replikace a RPO (Recovery Point Objective). Tato funkcionality je využívána pro replikaci dat na jiné diskové pole. V případě využití jednoho pole je možná replikace na jiný diskový pool. Replikace na jiné diskové pole má nesporné výhody. Pokud synchronizujeme disková pole mimo budovu, zvyšujeme zabezpečení vysoké dostupnosti. Tedy pokud by se s prvním polem cokoliv stalo (chyba HW, požár nebo jiné), jsou data dostupná na jiné lokalitě. Potvrzením volby se zobrazí shrnutí jednotlivá nastavení tohoto datastore (Obr. 36).



Obr. 36 Datastore Summary

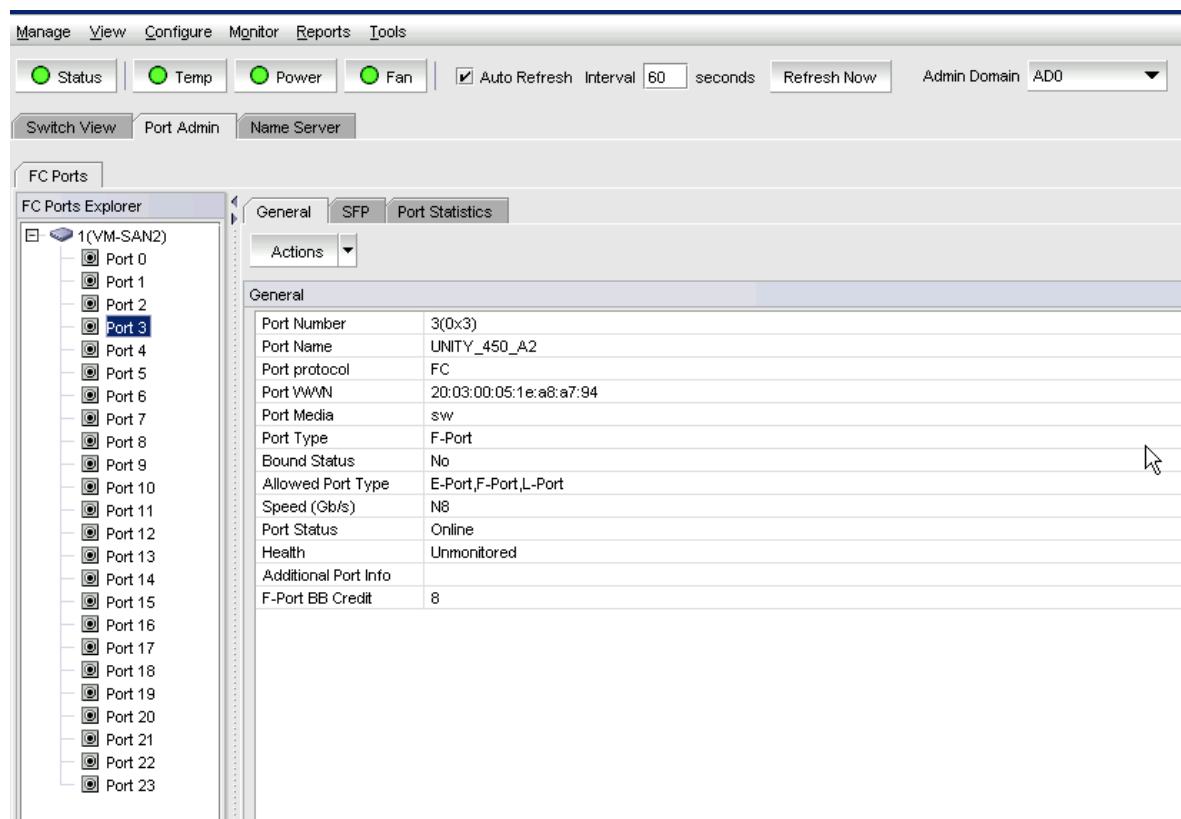
Nyní bude nakonfigurováno připojení jednotlivých ESXi hostů, tedy krok, který byl vynechán v předchozí části. Zde se nabízejí dvě možnosti připojení hostů. Buď postupným připojováním jednotlivých hostů, stejně jako v předchozím průvodci, nebo připojením vCenter serveru a tím připojení všech hostů, které jsou k vCenter připojeny.

Připojení bude provedeno v části *ACCESS*, v položce *VMware*. Definováním vCenter serveru a přihlašovacími údaji dojde k zobrazení hostů. V této části je možné vybrat všechny hosty

nebo případně provést individuální výběr. Po provedení výběru bude dokončena konfigurace a ukončen průvodce.

4.5 Konfigurace SAN Switchů

Nyní je nutné provést konfiguraci SAN switchů a jejich zónování, aby bylo diskové pole dostupné pro vCenter server. V prvním kroku konfigurace je nutné provést nastavení jednotlivých portů (Obr. 37). Nejprve provedeme u jednotlivých portů pojmenování, abychom jim stanovili jednoznačnou identifikaci.



Obr. 37 SAN konfigurace

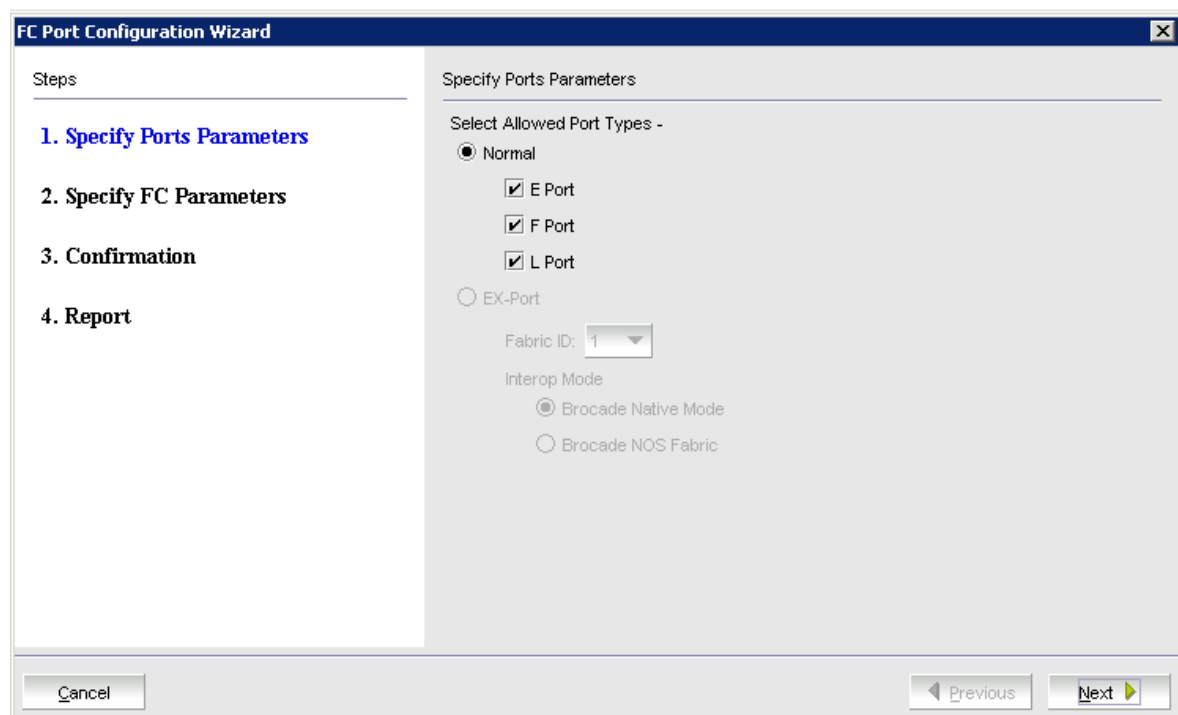
Dále bude nastaven typ portu, respektive druhy portů, které jsou pro daný SFP modul povoleny (Obr. 38). SAN switch umožňuje vybrat tři typy portů:

E port: Expansion port, slouží ke kaskádovému spojení jednotlivých SAN switchů [56].

L port: Loop port, slouží k připojení k FC smyčce (v nové generaci switchů, po roce 2010, se již nepoužívá) [56].

F port: Fabrický port, slouží pro připojení jednotlivých zařízení k SAN switchi [56].

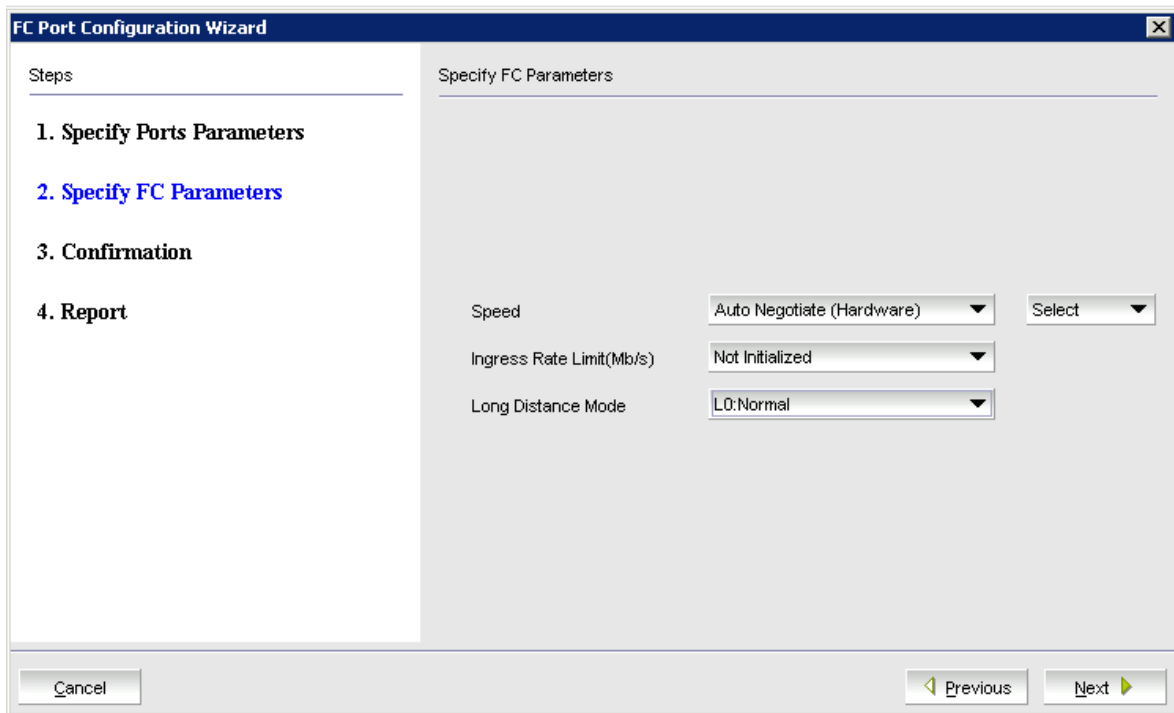
V tomto případě bude ponechán zaškrtnutý F port, jelikož zbytek portů není pro tuto konfiguraci potřebný.



Obr. 38 SAN druhy portů

Následuje specifikace FC parametrů, kde je umožněno specifikovat rychlost portu dle zvoleného FC SFP/SFP+ modulu. Na základě autorových osobních zkušeností je lepší ponechat nastavení v automatickém módu HW (Auto Negotiate Hardware). Switch sám provede detekci SFP/SFP + modul a dle toho určí maximální rychlost přenosu. Toto nastavení je nejbezpečnější v případě chyby na portu. Dále je možné nastavit maximální přenosovou

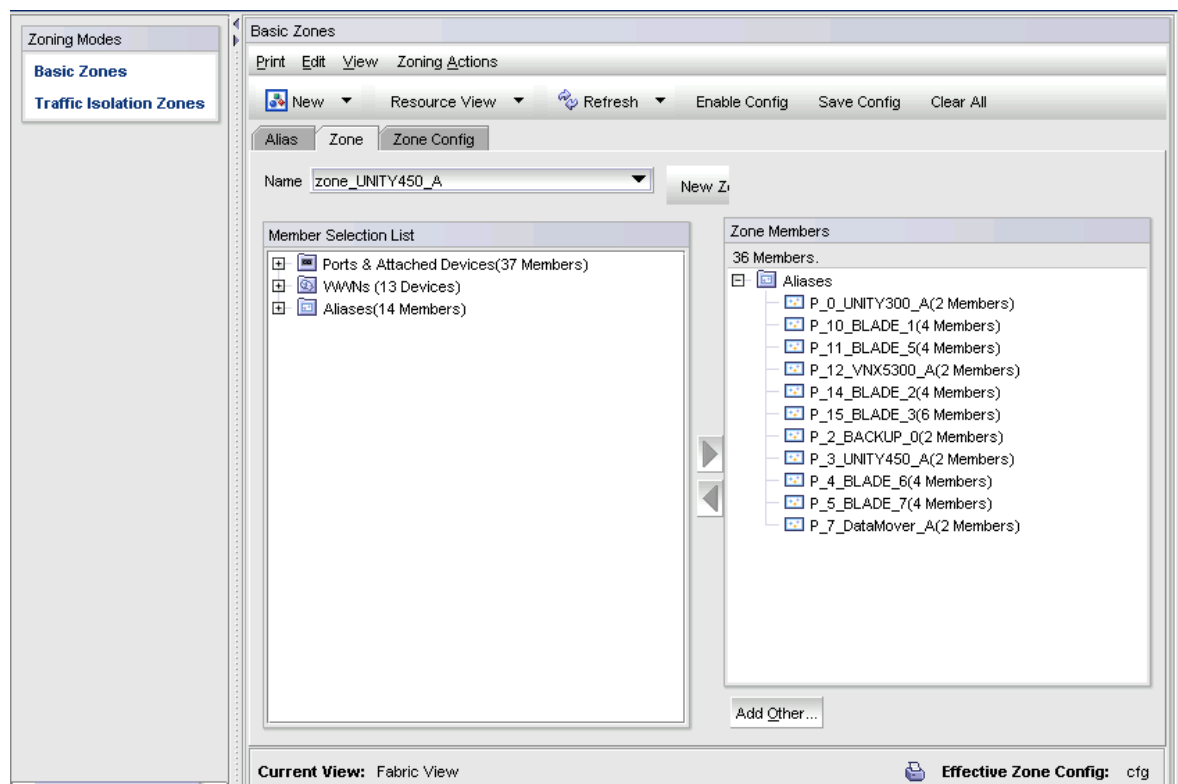
rychlost. Tyto switche umožňují dosahovat rychlosti až 8 Gbps, diskové pole umožňuje až 16 Gbps a pro zachování nejvyšší možné rychlosti je vyžadováno 8 Gbps. Omezení rychlosti tedy není žádoucí a bude ponecháno výchozí nastavení. Poslední možností je nastavení *Long distance mode*. Ten slouží k propojení na velmi velké vzdálenosti nad 10 km. Tato volba tedy zůstane ve výchozím, prázdném, nastavení a bude dokončen konfigurační průvodce (Obr. 39).



Obr. 39 konfigurace FC Portu

Nyní je nutné provést konfiguraci Zón SAN switchů. Každý switch musí mít nakonfigurované zóny, aby bylo možné nastavit oprávnění jednotlivým zařízením a přiřazení k jejich zdrojům. V prvním kroku bude provedeno vytvoření nového aliasu pro jednotlivé porty. Po vytvoření aliasu (P_3_UNITY450_A a P6_UNITY450_B) budou oběma aliasům přiřazené konkrétní fyzické porty. Po dokončení vytvoření aliasu je možné pokračovat dále a vytvořit jednotlivé zóny. U vytváření zón bude zachována jmenná konvence, použitá u aliasů pro přehlednost celého systému. Na každém switchi budou vytvořeny dvě zóny (zone_UNITY450_A, zone_UNITY450_B), do každé zóny budou připojeny již dříve definované aliasy. Přidáním, respektive nepřidáním jednotlivých aliasů pro danou zónu, jsou nastavovány přístupy k portům, a tedy k jednotlivým zařízením připojeným k SAN switchům. Do jednotlivých zón jsou přidány porty pro ostatní disková pole. Tato konfigurace umožňuje přímý přístup mezi diskovými poli. Dalšími porty, které budou k těmto zónám připojeny, jsou veškeré porty

určené ke spojení s Blade servery (Obr. 40). Následně budou přidány porty pro připojení zálohování. Na závěr je nutné provést uložení konfigurace.

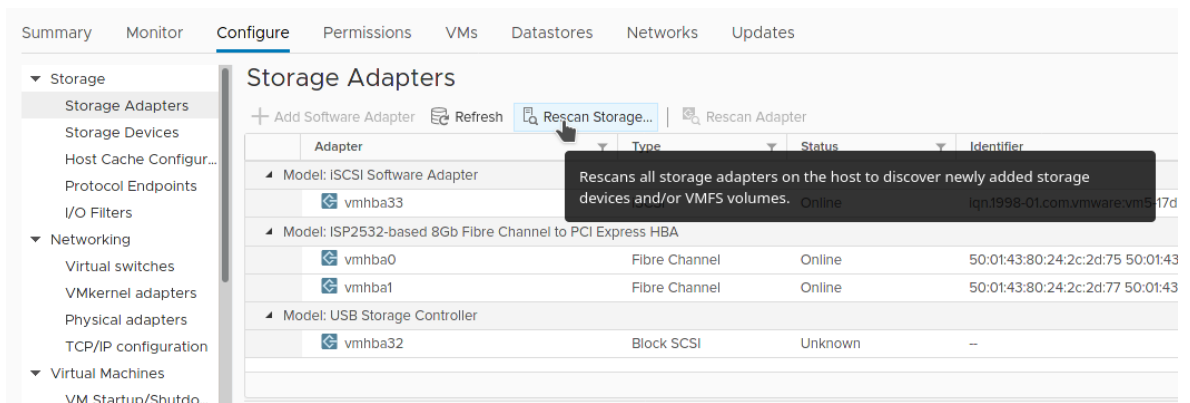


Obr. 40 SAN Zóny

4.6 Připojení pole k vCenter

Aby bylo umožněno využívat vytvořený datastore, je nutné provést jeho přidání v prostředí VMware vCenter. VMware vCenter je ve verzi 6.5 s připojenými hosty verze VMware ESXi verze 6.0. Toto heterogenní prostředí je zapříčiněno na jedné straně stářím některých hostů, na které již není možné instalovat vyšší verzi ESXi a na druhé straně potřebou aktualizovat verzi vCentra z verze 6 na verzi 6.5. Verze 6 využívá pro svojí funkcionalitu zastaralé flashové prostředí a verze 6.5 již využívá HTML5.

Implementace nového datastore do prostředí VMware vCenter se provádí přes grafické prostředí. Aby bylo možné připojit nový datastore, je nutné na všech ESXi hostech provést scanování portů. Toto scanování se provádí na každém hostu v konfiguraci, čímž dojde k otestování vybraných portů a detekci diskového pole (Obr. 41).



Obr. 41 Storage adapters

Po dokončení detekce na všech hostech bude možnáno přejít k další části vytvoření nového datastore. Vybráním položky „New Datastore“ bude spuštěn průvodce, kde je nutné vybrat nejdříve typ datastore. Zde bude vybrána možnost VMFS ^(Obr. 42), která kopíruje nastavení použitá při vytváření datacentra na diskovém poli. V následujícím kroku bude provedeno pojmenování datacentra a definování hostů, kteří jej budou moci využít. V následném kroku bude vybrána verze VMFS, kterou bude, dle předchozí konfigurace, zvolena možnost VMFS 5.

New Datastore

✓ 1 Type
2 Name and device selection
3 VMFS version
4 Partition configuration
5 Ready to complete

Type
Specify datastore type.

VMFS
Create a VMFS datastore on a disk/LUN.

NFS
Create an NFS datastore on an NFS share over the network.

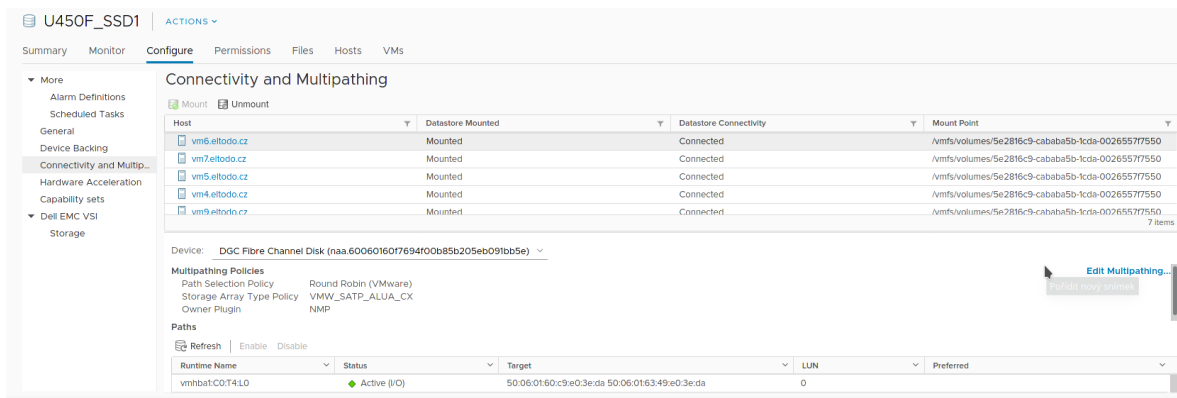
VVol
Create a Virtual Volumes datastore on a storage container connected to a storage provider.

CANCEL BACK NEXT

Obr. 42 New Datastore

Poté bude potvrzena konfigurace přiřazeného oddílu a dokončeno vytvoření diskového pole. Po ukončení průvodce je nutné provést dokončení konfigurace.

Aby bylo možné diskové pole začít využívat, je nutné provést post instalační optimalizaci pro dané diskové pole. Jelikož veškeré ESXi hosty využívají více než jednu FC cestu, tedy používají ke svému propojení s diskovými poli vícenásobné připojení, bude provedena optimalizace FC spojení. První nastavení bude provedeno v konfiguraci datastore pomocí webového prostředí vCenter. Pro nejrychlejší připojení k diskovému poli bude využita funkcionality *ESXi Round Robin*, která využívá algoritmus pro vyvážení zátěže jednotlivých spojení FC. Vybraná cesta je využívána, dokud není naplněna a vyčerpána její kapacita. Poté bude využita jiná cesta, která je uvedena v seznamu. Nastavením limitu na jednotlivé cesty bude určeno, jak rychle se mají měnit (Obr. 43).



Obr. 43 Round Robin

Po zvolení možnosti Round Robin je nutné provést úpravu cesty dle doporučení DELL Best practice^[58]. Tato část se provádí na jednotlivých hostech pomocí příkazů v terminálu. Pomocí ssh (Putty, terminál) bude provedeno spojení k ESXi hostům dle níže uvedeného příkazu.

```
ssh <adresa_esxi_hosta> -l root
```

Po zadání přihlašovacího hesla budou pomocí příkazu níže vypsána veškerá vmfs zařízení s jejich kapacitou (Obr. 44).

```
esxcfg-scsidevs -c
```

Device UID	Device Type	Console Device	Size	Multipath Plugin	Display Name
mpx.vmhba32:C0:T0:L0	Direct-Access	/vmfs/devices/disks/mpx.vmhba32:C0:T0:L0	2615MB	NMP	Local USB Direct-Access (mpx.vmhba32:C0:T0:L0)
naa.6006016071602d00246d5973d7dde011	Direct-Access	/vmfs/devices/disks/naa.6006016071602d00246d5973d7dde011	838868MB	NMP	DGC Fibre Channel Disk (naa.6006016071602d00246d5973d7dde011)
naa.60060160bd7042009928235874f271f4	Direct-Access	/vmfs/devices/disks/naa.60060160bd7042009928235874f271f4	10366976MB	NMP	DGC Fibre Channel Disk (naa.60060160bd7042009928235874f271f4)
naa.60060160bd704200ff2823580600e385	Direct-Access	/vmfs/devices/disks/naa.60060160bd704200ff2823580600e385	10366976MB	NMP	DGC Fibre Channel Disk (naa.60060160bd704200ff2823580600e385)
naa.60060160f7694f00b85b205eb091bb5e	Direct-Access	/vmfs/devices/disks/naa.60060160f7694f00b85b205eb091bb5e	14889779MB	NMP	DGC Fibre Channel Disk (naa.60060160f7694f00b85b205eb091bb5e)

Obr. 44 VMFS Devices

Na základě výsledku předchozího příkladu bude určeno, pro které zařízení bude proveden níže uvedený příkaz, kterým bude nastaven Round Robin na hodnotu 1 dle Dell Knowledge base^[59].

```
for i in `esxcfg-scsidevs -c |awk '{print $1}' | grep
naa.60060160f7694f00b85b205eb091bb5e`; do esxcli storage nmp psp roundrobin
deviceconfig set --type=iops --iops=1 --device=$i; done
```

Toto nastavení upraví výchozí počet vstupně výstupních (I/O) operací z 1000 na 1. Tedy pro každou I/O operaci bude zvolena vlastní cesta. Takto rychlé přepínání zajistí vyšší rychlost vyřizování požadavků^[59]. Pro kontrolu bude proveden následující příkaz.

esxcli storage nmp device list

Tímto příkazem dojde k vypsání podrobného seznamu jednotlivých zařízení s jejich nastavením. Z toho výpisu je patrná hodnota „*iops=1*“, která byla nastavena (Obr. 45).

```
naa.60860168f7694f08b85b285eb91bb5e
Device Display Name: DDC Fibre Channel Disk (naa.60860168f7694f08b85b285eb91bb5e)
Storage Array Type: VMW_SATP_ALUA_CK
Storage Array Type Device Config: (navireg=on, ipfilter=on){implicit_support=on; explicit_allow=on; alua_followover=on; action_onRetryErrors=off; {TPG_id=2,TPG_state=AND}{TPG_id=1,TPG_state=AD}}
Path Selection Policy: VMW_PSP_RR
Path Selection Policy Device Config: (policy-iops.iops=1,bytes=10485760_useAND=0; lastPathIndex=1; NumIOsPending=0,numBytesPending=0)
Path Selection Policy Device custom Config:
Working Paths: vmhba0:C0:T4:L0, vmhba1:C0:T4:L0
Is USB: false
```

Obr. 45 ESXi RoundRobin

Tato nastavení bude provedeno na všech ESXi hostech, které jsou připojeny ke konfigurovanému diskovému poli. Na rozdíl od spousty jiných úprav, tato nevyžaduje restart hostů. Po dokončení konfigurace bude provedena kontrola jednotlivých systémů, zda nevykazují chyby zapojení nebo konfigurace.

4.7 Testování a migrace systémů

Testování diskového pole bude probíhat v několika krocích. Vytvořením nového virtuálního stroje přímo na diskovém poli a následnou migrací mezi diskovými poli. V dalším kroku bude tento server přidán do zálohování, aby byla otestována odpověď pole při zálohách. Následným krokem bude migrace lehčích systémů a na závěr dojde k migraci stěžejního systému MS Exchange a MS SQL server řešení MS Dynamix NAV.

Prvním krokem je tedy vytvoření nového virtuálního stroje. Jelikož společnost aktuálně prochází i migrací MS Exchange verze 2010 na verzi 2016 je nutné pro tuto migraci vytvořit i nový server. Bude tedy vytvořen takový server, který bude odpovídat budoucím požadavkům na nový MS Exchange server 2016. Server bude mít následující konfiguraci, která bude nadhodnocovat standardní požadavky MS Exchange serveru 2016, a zároveň bude odpovídat aktuálním požadavkům na velikost diskového úložiště dle současného řešení^[60](Tabulka 2).

Serverová konfigurace	
CPU	<i>2 x 4 core</i>
RAM	<i>64 GB</i>
HDD System	<i>450 GB – Thick provision</i>
HDD DB	<i>5,2 TB – Thick provision</i>
HDD Logs	<i>650 GB – Thick provision</i>
Net adapter 1	<i>Standardní síť</i>
Net adapter 2	<i>DAG Network</i>
OS	<i>Windows 2016 Standard</i>

Tabulka 2 Serverová konfigurace

Pro dostatečné zatížení pole bude při vytváření virtuálního stroje vybráno pro jednotlivé disky nastavení, které alokuje celou kapacitu disků na diskovém poli. Po vytvoření virtuálního stroje a jeho nainstalování bude otestována migrace virtuálního stroje. Nejprve bude provedena migrace z nového diskového pole na druhé pomocí vCenter vMotion, následně bude provedena migrace zpět na diskové pole s migrací na jiného ESXi hosta. Během těchto migrací bude provedena kontrola ve vCenter, zda nedochází k chybám a na diskovém poli kontrola jeho performance a LUN Reponse Time. Zde by hodnoty neměly přesáhnout pět milisekund ^[61]. Až bude dokončena migrace na všechny hosty, aby bylo dostatečně otestováno spojení mezi hosty a novým diskovým polem, bude provedena migrace některého z náročnějších serverů. Optimálním se, z produkčního prostředí, jeví jeden z aplikačních serverů MS Dynamics NAV. Tento server, označený jako NAVAPL4, slouží pro testování úprav v prostředí MS Dynamics NAV před plným nasazením a není na něm takový provoz, jako na plně produkčních serverech. Provoz je ale dostatečný pro otestování diskového pole. Pokud by byla chyba v konfiguraci, tak by se v tomto bodu již projevila chybovost. Následně se na nové diskové pole přesunou servery, které nejvíce vytěžují současné diskové pole. Jedná se hlavně o dva servery, MS SQL server určený pro MS Dynamics NAV a MS Exchange server.

5 Výsledky implementace a doporučení

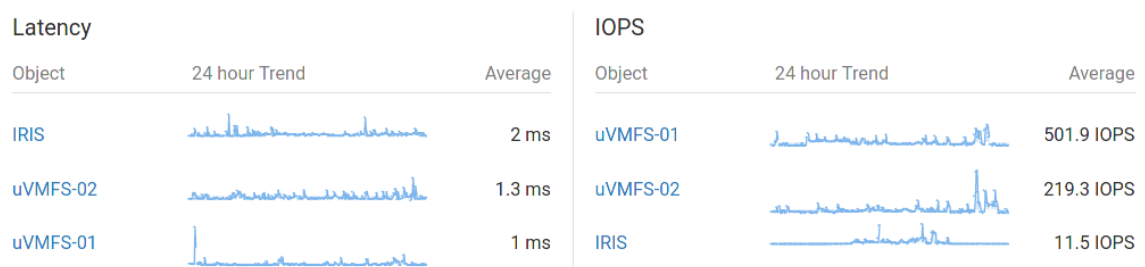
V této části jsou prezentovány výsledky implementace diskového pole, včetně vybraných problémů během implementace a možnosti, kam by se společnost měla, nebo mohla ubírat dále se současným řešením.

5.1 Výsledky implementace a doporučení

Implementace diskového pole do současného systému splnila veškerá očekávání i přes prvotní problémy, které se ze začátku provozu vyskytly. Původní diskové pole, na kterém byla umístěna většina virtuálních strojů, mělo průměrnou latenci nad 75 milisekund s občasnými peaky k několik vteřinám. Nyní po dokončení migrace došlo k narovnání jednotlivých částí a srovnání latence na únosnou míru. Aktuální zatížení původního pole je na SANové části mezi 1 – 1,3 milisekund. Takováto doba odezvy je oproti přechozímu extrémnímu zatížení obrovským posunem v dostupnosti služeb, které přestaly vypadávat.

Na výkon staršího diskového pole měla dále velmi pozitivní vliv redukce datastore. Z původních tří uVMFS-01 – 03 došlo k redukci na dva a volné místo na discích bylo rozděleno z větší části pro zbývající datastore a zbylá část pro NAS. Na základě této úpravy došlo i k rozdělení zátěže, kdy každý datastore využívá vlastní kontroler, což má také, i když menší vliv na výkon.

Object Activity

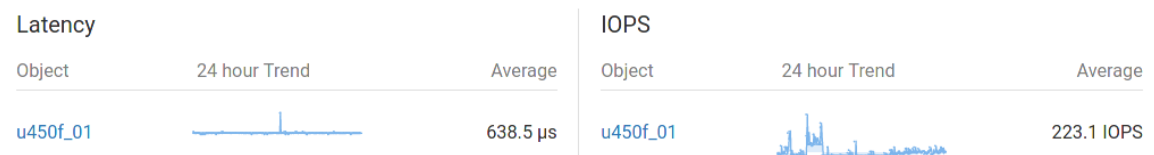


Obr. 46 UNITY 300 Latency

Na implementované pole byly přesunuty nejnáročnější servery, a to MS SQL server pro MS Dynamics NAV a MS Exchange server. Oba servery jsou náchylné na rychlost odezvy diskového pole. MS Dynamics NAV využívá cca 300 aktivních uživatelů a má alokovanou kapacitu 1,5 TB. Server MS Exchange využívá cca 1500 uživatelů s téměř dvojnásobnými

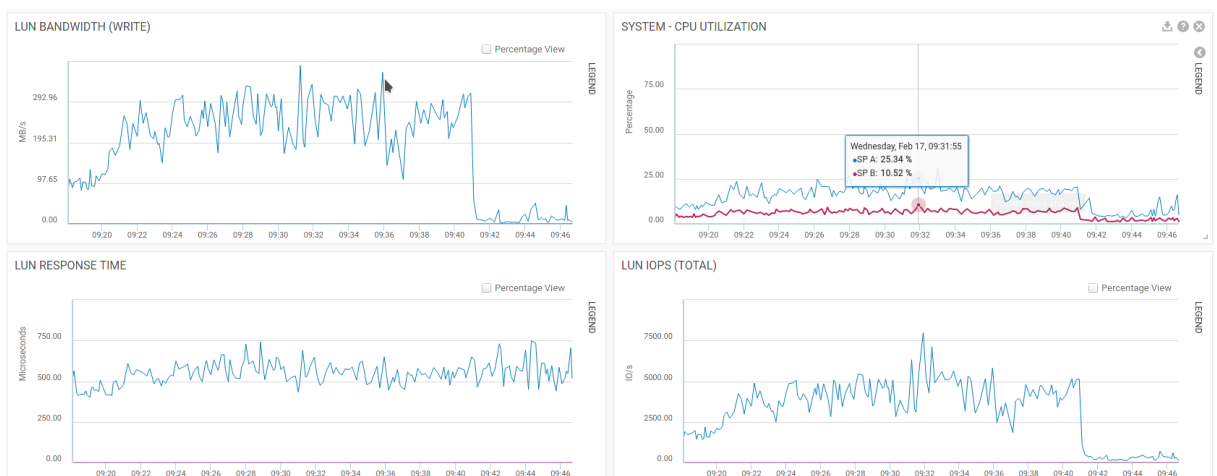
požadavky na počet spojení, způsobenými připojením mobilních zařízení, a alokovanou kapacitou 6,3 TB. Společně s těmito zásadními virtuálními stroji je zde dále pět virtuálních strojů a společně zabírají asi 8,6 TB diskového prostoru. I přes tuto zátěž nedosahuje latence na diskovém poli vysokých hodnot. Průměrná latence je cca 0,638 milisekundy a průměrným požadavkem 223,1 IOPS (Obr. 47).

Object Activity



Obr. 47 UNITY 450f Latency

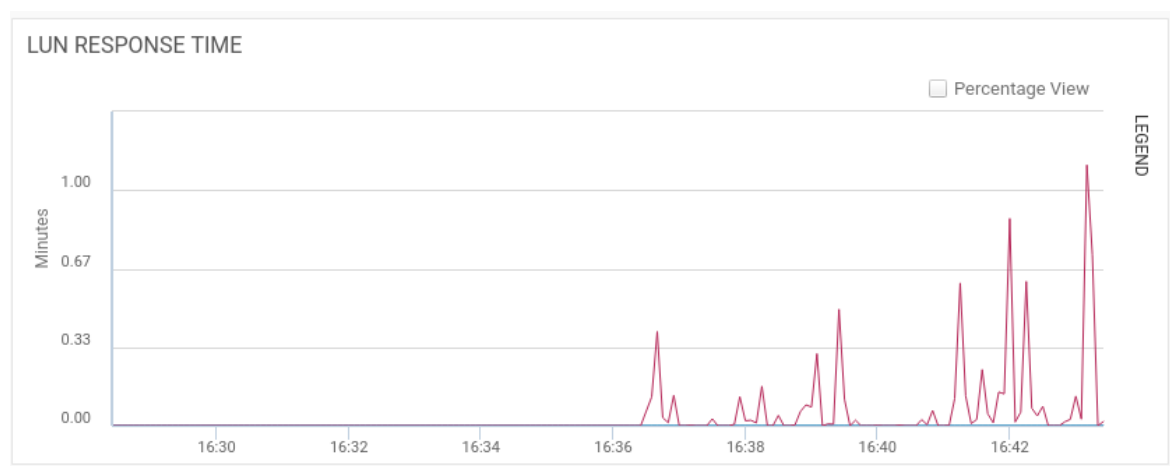
V době zálohování je možné pozorovat výkyvy, které dosahují v nárazech okolo 1,47k IOPS. Jak je z grafů patrné, vytížení původního diskového pole je nyní na únosné mezi a na implementovaném poli jsou hodnoty hluboko pod hodnotou 1ms. A tedy latence diskového pole je v pořádku. Pro ověření byla provedena migrace dalšího virtuálního stroje s MS SQL server určeného pro MS Sharepoint. Tento SQL server je využíván také nadstavbou MS Sharepoint, NINTEX^[63]. Tato nadstavba slouží společnosti pro vytváření, správu a automatizaci procesů a jsou nad ním vyvíjeny aplikace, které využívají všechna oddělení ve společnosti. Během migrace na nové diskové pole byly sledovány hodnoty latence diskového pole, šířka pásma pro zápis a počet IOPS. Jak je vidět z grafů na obrázků níže, tak se zatížení neprojeví na latenci diskového pole, ale je patrné na požadavcích na šířku pásma a IOPS (Obr. 48).



Obr. 48 Migrace SQL

5.2 Problémy během implementace

Během implementace se vyskytlo několik problémů, které byly postupně eliminovány. Zásadním problémem, který vznikl hned po implementaci, bylo opomenutí několika drobností, včetně konfiguračních nastavení, která měla zásadní vliv na výkon diskového pole v prostředí VMware. Při vytváření a migraci virtuálního stroje na diskové pole končily požadavky chybou, oznamující dlouhý časový požadavek. Po kontrole na straně diskového pole ukazatel výkonu byly výkyvy latence v mezích, které naprosto neodpovídaly rychlosti SSD disků a FC připojení. Při pokusu o nakopírování několika souborů na implementované diskové pole byla pozorována hodnota, LUN Response time, přesahující 1 minutu (Obr. 49).



Obr. 49 Reponse extream

Jako hlavním problémem bylo opomenutí správné konfigurace multipathingu v prostředí VMware. Dalším problémem bylo nedokonalé domáčknutí optických kabelů a FC SFP+ modulů. To mělo za následek vysoký útlum v signálech a velkou chybovost na portech.

5.3 Zlepšení současného řešení

Jednotlivé části celého systému je možné vylepšit. Záleží pouze na rozpočtu a zhodnocení obchodního modelu, co se společnosti vyplatí a co je již zbytečný luxus. Toto rozhodnutí je na vedení společnosti, které musí určit cenu výpadku systému společnosti na jeden den, určit rizika a dle toho rozvrhnout investice do infrastruktury.

5.3.1 Serverovna

Pro serverovnu existuje několik možností jak vylepšit vysokou dostupnost. První z možností je pořízení druhé jednotky UPS, která by zajistila jak redundanci samotného systému v případě výpadku el. energie, tak i prodloužení doby chodu na baterie. Toto řešení ale není nejlevnější, pořizovací cena současného řešení byla téměř 451.000,-Kč. Další možností, jak zlepšit zajištění vysoké dostupnosti během výpadku el. energie, by bylo pořízení diesel agregátu, který by v případě výpadku el. energie mohl napájet celou infrastrukturu po téměř neomezenou dobu. Nejlepším, ale také nejdražším řešením by byla kombinace obou možností. Pořízení druhé UPS a diesel agregátu je nejdražší řešení z důvodu pořízení nejvíce prvků, ale také z důvodu největších stavebních zásahů v prostorách budovy a úpravy vedení el. energie pro serverovnu. Cena druhé UPS byla stanovena dle pořizovací ceny první UPS. Je předpoklad, že cena nové UPS by byla podobná, ale s lepší konfigurací.

Orientační cena nového dieselagregátu s výkonem 40KVA činí cca 380.000,- Kč bez DH. Výkon toho agregátu by měl být dostatečný pro provoz takto malé serverovny. Další investicí by byla cena za elektroinstalační materiál a práci. V tomto případě je odhadována cena na 100.000,- Kč. Výsledná suma celého řešení je odhadována na cca 930.000,- Kč bez DPH. I přes vyšší pořizovací náklady se jedná o nejlepší řešení v rámci ochrany prostředí před výpadky el. energie.

5.3.2 Diskové pole Unity 450f – budoucnost

Po dokončení implementace a ostrém provozu s kritickými systémy se společnost rozhodla dále zainvestovat do diskového pole a rozšířit jeho kapacitu. V této době již probíhá poptávka u společnosti DELL na dodání dalších disků pro diskové pole. Diskové pole je možné rozšířit vždy minimálně po blocích, které obsahují šest disků. Společnost aktuálně poptává jeden blok s kapacitou disků 1,92TB. Cena těchto disků se na internetu pohybuje okolo 6650\$^[62] za 1 SSD disk. Aktuální cenová nabídka spol. DELL – EMC je okolo 3300\$ za jeden SSD disk. Cenová úspora oproti internetu je zde patrná. Po instalaci dalších disků bude provedeno navýšení aktuálního diskového poolu a zároveň bude provedena úprava konfigurace datastore. Jelikož poptávané množství není dostatečné pro vytvoření samotného datastore, bude současný datastore nepatrně ponížen, aby bylo možné vytvořit nový s využitelnou kapacitou a tím vytvořit dva stejně veliké datastore. Tato úprava bude provedena za účelem dalšího zvyšování performance celého diskového pole. Každému datastore bude přidělen samostatný kontroler, (SPA/SPB). Tím bude docíleno rozdělení zátěže mezi oba kontrolery a jejich

optimálnější využití. Dále bude umožněna funkcionalita HA, jelikož v případě chyby umožňuje diskové pole plynulé převzetí datastore druhým kontrolerem.

6 Závěr

Cílem práce byla implementace all-flash diskového pole do prostředí VMware pro zrychlení infrastruktury a odlehčení původnímu diskovému poli s rotačními disky a tím zrychlení krizových aplikací společnosti. Problematika, kterou se tato diplomová práce zabývá, obsahuje řadu anglických termínů, pro které se nevyužívají české ekvivalenty, nebo nejsou odpovídající ekvivalenty k dispozici.

Teoretická část se v první části zaměřuje na disková pole obecně a na jejich dělení. Dále jsou rozepsány jednotlivé druhy RAIDů a jejich určení. Následující část je zaměřena na optické spojení, vlákna, konektory, FC a SFP moduly. V další části je provedena charakteristika SSD a použité technologie. Závěr teoretické části je věnován VMware ESXi a vCenter, u kterých byla provedena stručná charakteristika.

V praktické části je nejprve provedena charakteristika současného stavu celého systému s kritickými chybami. Dále se nachází specifikace nového diskového pole a vybíraných alternativ.

Ve druhé části je provedena kompletní instalace diskového pole po jednotlivých krocích. Poté byla uskutečněna fyzická instalace a úprava konfigurace síťových prvků. Následně byla vykonána základní konfigurace diskového pole s vytvořením datastoru, určeného k připojení do prostředí VMware. Dále byla upravena konfigurace SANové části společně s nastavením zón pro přístup k diskovému poli. Následně bylo nastaveno připojení k prostředí VMware vCenter a otestování propojení. Poté byl vytvořen testovací virtuální server a byla provedena jeho migrace mezi diskovými poli. Po otestování byly přesunuty stěžejní virtuální servery na implementované pole.

V poslední části byly interpretovány výsledky implementace s výsledným zatížením obou diskových polí. Došlo ke snížení latence původního diskového pole z neúnosných hodnot, které v průměru dosahovaly nad 75ms na optimální zatížení kolem 1,3 ms. Implementované pole dosahuje nízké latence menší než 1ms. Dále byly vypsány jednotlivé zásadní implementační problémy s výsledky jejich řešení. Hlavním problémem bylo opomenutí jednoho konfiguračního příkazu, což mělo za následek extrémní latenci diskového pole v řádu minut. Následně byly shrnuty návrhy zlepšení serverovny pro dosažení vyšší dostupnosti,

kupříkladu pořízení druhé UPS jednotky a vybavení serverovny diesel agregátem. Tento návrh byl ekonomicky zhodnocen. Navrženo budoucí směřování all-flash diskového pole s cenovou kalkulací na jeho rozšíření a budoucí rozdělení datastore.

Výstupem této práce je implementace diskového pole ve společnosti. Použité postupy a jednotlivá řešení je možné využít, s úpravou dle prostředí, na podobné implementace. Společnosti přineslo toto řešení výrazné snížení zatížení konvenčního diskového pole. Dalším přínosem zrychlení celé infrastruktury jsou i ekonomické výhody v podobě zrychlení práce s firemními aplikacemi. Implementované řešení má velké možnosti rozšíření do budoucna s předpokladem pokračující migrace infrastruktury na nové diskové pole.

7 Seznam použitých zdrojů

- [1] Joel Santo Domingo, SSD vs. HDD: What's the Difference? [online] PCMAG, 2.9.2020 [cit. 2020-08-04]. Dostupné z: <https://uk.pcmag.com/ssd/8061/ssd-vs-hdd-whats-the-difference>
- [2] Dave Raffo, Disk array [online] TechTarget 2018-11 [cit. 2020-08-04]. Dostupné z: <https://searchstorage.techtarget.com/definition/array>
- [3] IBM, IBM 350 disk storage unit [online] IBM [cit 2020-08-04] Dostupné z: https://www.ibm.com/ibm/history/exhibits/storage/storage_350.html
- [4] Robert Cook, SSD Throughput, Latency and IOPS Explained – Learning To Run With Flash [online] TheSSDReview [cit 2020-08-04] Dostupné z: <https://www.thessdreview.com/featured/ssd-throughput-latency-iopsexplained/>
- [5] Caesar Wu and Rajkumar Buyya, *Cloud Data Centers and Cost Modeling : A Complete Guide to Planning, Designing and Building a Cloud Data Center.* [online] Elsevier Science & Technology 2015-03-09 [cit. 2020-08-04] ISBN 9780128016886
- [6] Scott Vetter, Swarna Narendra Babu, Harihara Balakrishnan, *IBM Power Systems RAID Solutions Introduction and Technical Overview.* [online] IBM Redbooks 2018-12 [cit. 2020-08-04] ISBN 9780738454344
- [7] What RAID-DP protection is [online] NETAPP 2015-03 [cit. 2020-08-04] Dostupné z: <https://library.netapp.com/ecmdocs/ECMP1636022/html/GUID-DAEB7BDC-4E6E-4CEE-9B3D-F40962C3D428.html>
- [8] ELPROCUS, What are the Basic Elements of a Fibre Optic Communication System? [online] ELPROCUS [cit. 2020-08-05] Dostupné z: <https://www.elprocus.com/basic-elements-of-fiber-optic-communication-system-and-its-working/>
- [9] Mark Jackson, Scientists Successfully Push Fibre Optic Transmissions Close to the Limit [online] ISPPreview 19.09.2016 [cit. 2020-08-05] Dostupné z: <https://www.isppreview.co.uk/index.php/2016/09/scientists-successfully-push-fibre-optic-data-speeds-close-limit.html>
- [10] Polymethylmethacrylate (PMMA, Acrylic) [online] MakeItFrom 30.5.2020 [cit. 2020-08-05] Dostupné z: <https://www.makeitfrom.com/material-properties/Polymethylmethacrylate-PMMA-Acrylic>
- [11] Graded-Index Polymer Optical Fiber (GI-POF) [online] THORLABS [cit. 2020-08-05] Dostupné z: <https://www.thorlabs.com/catalogPages/1100.pdf>
- [12] Tone Casazza, Fiber Type Vs. Speed and Distance [online] lanshack [cit. 2020-08-07] Dostupné z: <https://www.lanshack.com/Fiber-Type-vs-Speed-and-Distance.aspx>
- [13] THEFOA, The Fiber Optic Association - Tech Topics, *Connector Identifier* [online] THEFOA 2019 [cit. 2020-08-07] Dostupné z: <https://www.thefoa.org/tech/connID.htm>

- [14] Ronald G. Driggers, Craig Hoffman *Encyclopedia of Optical Engineering* [online] 2003 [cit. 2020-08-10] ISBN 9780824709402
- [15] Jon Tate, Pall Beck, Hector Hugo Ibarra, Shanmuganathan Kumaravel, Libor Miklas, *Introduction to Storage Area Networks* [online] IBM Redbooks 2017-12-19 [cit. 2020-08-11] ISBN 9780738442884
- [16] DELL Technologies, Network-Attached Storage [online] DELL Technologies [cit. 2020-08-12] Dostupné z: <https://www.delltechnologies.com/sr-me/glossary/network-attached-storage.htm>
- [17] DELL Technologies, SAN Storage [online] DELL Technologies [cit. 2020-08-12] Dostupné z: <https://www.delltechnologies.com/en-us/glossary/san-storage.htm>
- [18] WORDISK, Network-attached storage [online] WORDISK [cit. 2020-08-12] Dostupné z: https://worddisk.com/wiki/Network_attached_storage/
- [19] James Alan Miller, Fibre Channel [online] TechTarget 2018-08 [cit. 2020-08-13] Dostupné z: <https://searchstorage.techtarget.com/definition/Fibre-Channel>
- [20] Jim Bell, The HIPPI Protocol [online] WUSTL 1995 [Cit. 2020-08-13] Dostupné z: <https://www.cse.wustl.edu/~jain/cis788-95/ftp/hippi/index.html>
- [21] ANSI, FC-PH ANSI X3.230 [online] ANSI 1994-01-01 [cit. 2020-08-13] Dostupné z: <https://standards.globalspec.com/std/453244/ANSI%20X3.230>
- [22] JENS, Generations of Fibre Channel and their Differences [online] Fibre-Blog 2020-04-02 [cit. 2020-08-13] Dostupné z: <http://www.fiber-blog.de/en/280-generations-of-fibre-channel-and-their-differences.html>
- [23] JUNIPER, Overview of Fibre Channel [online] Juniper 2018-02-02 [cit. 2020-08-13] Dostupné z: https://www.juniper.net/documentation/en_US/junos/topics/concept/fibre-channel-overview.html
- [24] RLC, Optical transmitters and network technologies [online] RLC [cit. 2020-08-14] Dostupné z: <https://rlc.cz/en/produkty/transceivers-and-technologies/>
- [25] CompleteConnect, MTP (MPO) Fibre Networking FAQ [online] CompleteConnect [cit. 2020-08-14] Dostupné z: <https://www.completeconnect.co.uk/mpo-mtp-fibre-networking-faq/>
- [26] fibrain. SFP+ [online] fibrain [cit. 2020-08-14] Dostupné z: <https://fibrain.com/produkty/sfp-plus,142/>
- [27] SNIA, QSFP (Quad Small Formfactor Pluggable) Transceiver [online] SNIA 2016-11-01 [cit 2020-08-04] Dostupné z: <https://www.snia.org/node/4792>
- [28] SNIA, QSFP+ 4X 10 Gb/s Pluggable Transceiver [online] SNIA 2018-11-31 [cit 2020-08-04] Dostupné z: <https://www.snia.org/node/4888>

- [29] SNIA, QSFP+ 28 Gb/s 4X Pluggable Transceiver Solution (QSFP28)[online] SNIA 2015-06-29 [cit. 2020-08-14] Dostupné z: <https://www.snia.org/node/4662>
- [30] QSFP-DD MSA, QSFP-DD Hardware Specification [online] QSFP-DD MSA 2020-08-07 [cit. 2020-08-14] Dostupné z: <http://www.qsfp-dd.com/wp-content/uploads/2020/08/QSFP-DD-Hardware-rev5.1.pdf>
- [31] ProLabs, What's the difference between 400 GBASE-LR8 and 400 GBASE-LR4? [online] ProLabs [cit. 2020-08-14] Dostupné z: <https://www.prolabs.com/pulse/400g-from-50gbps-pam4-to-100gbps-pam4>
- [32] Geek-university, What is VMware ESXi? ? [online] Geek-university [cit. 2020-08-14] Dostupné z: <https://geek-university.com/vmware-esxi/what-is-vmware-esxi/>
- [33] Kyle Gleed, vMotion – what's going on under the covers? [online] VMware 2011-02-25 [cit. 2020-08-15] Dostupné z: <https://blogs.vmware.com/vsphere/2011/02/vmotion-whats-going-on-under-the-covers.html>
- [34] Duncan Epping, vSphere 5.0: Storage vMotion and the Mirror Driver [online] YELLOW BRICKS 2011-07-14 [cit. 2020-08-15] Dostupné z: <http://www.yellow-bricks.com/2011/07/14/vsphere-5-0-storage-vmotion-and-the-mirror-driver/>
- [35] HPE, ProLiant BL460c Gen9 Server Blade - Overview [online] HPE [cit. 2020-10-25] Dostupné z: https://support.hpe.com/hpesc/public/docDisplay?docId=emr_na-c04434541
- [36] HPE, QuickSpec, Virtual Connect Flex-10/10D Module [online] HPE [cit. 2020-10-25] Dostupné z: <https://h20195.www2.hpe.com/v2/GetDocument.aspx?docname=c04154336>
- [37] HPE, QuickSpec, Virtual Connect 8 Gb 24-Port Fibre Channel Module [online] HPE [cit. 2020-10-25] Dostupné z: <https://h20195.www2.hpe.com/v2/GetPDF.aspx/c04154358.pdf>
- [38] HPE, QuickSpec, BladeSystem C3000 Enclosure [online] HPE [cit. 2020-10-25] Dostupné z: http://www.hp.com/canada/promotions/midmarket/hpww_lmmc/resources/HP_BladeSystem_c3000_Enclosure.pdf
- [39] HPE, QuickSpec, HPE8/8 Base SAN Switch and 8/8 SAN Switch Family [online] HPE [cit. 2020-10-22] Dostupné z: <https://h20195.www2.hpe.com/v2/GetPDF.aspx/c04123371.pdf>
- [40] Huawei, Campus switch S7700 Series, S7706 [online] Huawei [cit. 2020-10-22] Dostupné z: <https://support.huawei.com/enterprise/en/switches/s7706-pid-22346878>

- [41] Huawei, S7700 Series Switches Hardware Description, ES0DG48TFA00-48-Port [online] Huawei [cit. 2020-10-22] Dostupné z: <https://support.huawei.com/enterprise/en/doc/EDOC1000013755/d45ab409/es0dg48tfa00-48-port-10-100-1000base-t-interface-card-fa-rj45>
- [42] Huawei, S7700 Series Switches Hardware Description, ES1D2X48SX2S-48-Port [online] Huawei [cit. 2020-10-22] Dostupné z: <https://support.huawei.com/enterprise/en/doc/EDOC1000013755/2a835157/es1d2x48sx2s-48-port-10gbase-x-interface-card-x2s-sfp>
- [43] DELL Technologies, DELL EMC UNITY HYBRID STORAGE [online] DELL Technologies [cit. 2020-10-22] Dostupné z: <https://www.dell.com/en-us/collaterals/unauth/data-sheets/products/storage/h14958-unity-hybrid-family-ss.pdf>
- [44] DELL EMC Specification, DELL EMC UNITY ALL-FLASH STORAGE [online] DELL Technologies [cit. 2020-10-22] Dostupné z: <https://www.dell.com/en-us/collaterals/unauth/data-sheets/products/storage/h16018-unity-all-flash-family-ss.pdf>
- [45] Dell EMC Unity 450F All-Flash Storage [online] DELL Technologies [cit. 2020-10-22] Dostupné z: <https://www.delltechnologies.com/en-us/storage/unity/unity-450f-all-flash-storage.htm#>
- [46] EMC Unity™ Quick Start [online] EMC Corporation 2015 [cit. 2021-01-18] Dostupné z: <https://www.delltechnologies.com/zh-cn/collaterals/unauth/technical-guides-support-information/products/storage/docu69356.pdf>
- [47] Unity Family technical documentation, Access Control [online] DELL Technologies [cit. 2021-01-18] Dostupné z: <https://www.delltechnologies.com/cs-cz/documentation/unity-family/unity-p-security-config-guide/02-vxe-c-sec-config-chap-access-control.htm>
- [48] Mike Burnet, How Secure Is My Password [online] [cit. 2021-01-18] Dostupné z: <https://howsecureismypassword.net/>
- [49] Security.Org, How Secure Is My Password [online] Security.Org [cit. 2021-01-18] Dostupné z: <https://www.security.org/how-secure-is-my-password/>
- [50] NTP Pool Project [online] NTP Pool Project [cit. 2021-01-20] Dostupné z: <https://www.ntppool.org>
- [51] Ian Paul, Multi-Layer SSDs: What Are SLC, MLC, TLC, QLC, and PLC? [online] How to geek 2019 [cit. 2021-01-20] Dostupné z: <https://www.howtogeek.com/444787/multi-layer-ssds-what-are-slc-mlc-tlc-qlc-and-mlc/>
- [52] Speedguide, SLC, MLC or TLC NAND for Solid State Drives? [online] Speedguide [cit. 2021-01-21] Dostupné z: <https://www.speedguide.net/faq/slc-mlc-or-tlc-nand-for-solid-state-drives-406>

- [53] ELINFOR, We can trust the quality of QLC SSD? [online] ELINFOR 2019 [cit. 2021-01-21] Dostupné z: <https://www.elinfor.com/knowledge/we-can-trust-the-quality-of-qlc-ssd-p-11102>
- [54] EMC, Secure Remote Services Requirements and Configuration [online] EMC 2015 [cit. 2021-01-21] Dostupné z: <https://www.delltechnologies.com/el-gr/collaterals/unauth/technical-guides-support-information/products/storage/docu69327.pdf>
- [55] Mohammed Raffic, Difference between VMFS 5 & VMFS 6 and Configuration [online] VMware Arena 27.9.2017 [cit. 2021-01-26] Dostupné z: <http://www.vmwarearena.com/difference-between-vmfs-5-vmfs-6/>
- [56] NetApp, There are different types of Fibre Channel ports, what are they and what type of port does the NetApp filer have? [online] NetApp knowledge base 15.4.2020 [cit. 2021-01-30] Dostupné z: https://kb.netapp.com/Advice_and_Troubleshooting/Data_Storage_Software/ONTAP_OS/There_are_different_types_of_Fibre_Channel_ports%2C_what_are_they_and_what_type_of_port_does_the_NetApp_filer_have%3F
- [57] VMware, Create a VMFS Datastore [online] VMware Docs 31.5.2019 [cit. 2021-01-31] Dostupné z: <https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.storage.doc/GUID-5AC611E0-7CEB-4604-A03C-F600B1BA2D23.html>
- [58] DELL Technologies, Dell EMC Unity: Best Practices Guide, [online] DELL Technologies 2020 [cit. 2021-02-02] Dostupné z: <https://www.delltechnologies.com/en-vn/collaterals/unauth/white-papers/products/storage/h15093-dell-emc-unity-best-practices-guide.pdf>
- [59] VMware, Adjusting Round Robin IOPS limit from default 1000 to 1 [online] VMware kb 2069356 11.6.2020 [cit. 2021-02-02] Dostupné z: <https://kb.vmware.com/s/article/2069356>
- [60] Todd Nelson, Exchange 2016 Requirements and Prerequisites [online] ODDYTEE 5.10.2015 [cit. 2021-02-03] Dostupné z: <https://oddytee.wordpress.com/2015/10/05/exchange-2016-requirements-and-prerequisites/>
- [61] Enterprise Storage Guide, Latency: The King of Storage Performance Metrics [online] ActualTech media 17.4.2016 [cit. 2021-02-03] Dostupné z: <https://www.actualtechmedia.com/io/latency-king-storage-performance-metrics/>
- [62] Itprice, EMC Price list 2021 [online] itprice [cit. 2021-02-15] Dostupné z: <https://itprice.com/emc-price-list/d3f-2sfxl2-1920u.html>
- [63] Nintex, [online] nintex [cit. 2021-02-15] Dostupné z: <https://www.nintex.com/>