

Datová úložiště pro korporátní sféru a jejich nasazení

Enterprise data storage solutions and their deployment

Ondřej Kulišťák



Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

akademický rok: 2011/2012

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Ondřej KULIŠŤÁK**
Osobní číslo: **A08234**
Studijní program: **B3902 Inženýrská informatika**
Studijní obor: **Informační a řídicí technologie**
Forma studia: **prezenční**

Téma práce: **Datová úložiště pro korporátní sféru a jejich nasazení**

Zásady pro vypracování:

1. Analyzujte stávající datové úložiště společnosti ON Semiconductor. Popište jeho slabé stránky a nevýhody, které vedly k rozhodnutí migrace na nový typ úložiště.
2. Analyzujte situaci na trhu datových úložišť, proveďte výběr nového datového úložiště a navrhnete způsob jeho zapojení (iSCSI, Fiber Channel, Infiniband, ...) do stávající síťové infrastruktury společnosti.
3. Analyzujte možnost migrace dat ze stávajícího úložiště na nové, pomocí různých síťových protokolů a softwarů (CIFS, NFS, NetApp SnapMirror, Data Motion..), včetně řešení deduplikací souborů.
4. Proveďte migraci dat a zhodnoťte její průběh.

Rozsah bakalářské práce:

Rozsah příloh:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

1. Cisco: optimalizace aplikací. 1. vyd. Praha: Grada, 2008. ISBN 80-247-1610-0.
2. CLARK, Tom. Designing storage area networks: a practical reference for implementing fibre channel and IP SANs. 2nd ed. Boston: Addison-Wesley, 2003, 572 s. ISBN 03-211-3650-0.
3. TROPPENS, Ulf. Storage networks explained: basics and application of Fibre Channel SAN, NAS, iSCSI, InfiniBand and FCoE. 2nd ed. Chichester, West Sussex, U.K.: Wiley, 2009, 564 s. ISBN 04-707-4143-0.
4. POELKER, Christopher. Storage area networks for dummies. 2nd ed. Indianapolis, IN: Wiley Pub., Inc., 2008. ISBN 04-703-8513-8.
5. VAUGHN, Stewart. NetApp and VMware VSphere – Storage Best Practices. 1st ed. Lulu Enterprises Inc, 2010, 124 s. ISBN 05-570-9451-8.

Vedoucí bakalářské práce:

Ing. Tomáš Dulík, Ph.D.

Ústav informatiky a umělé inteligence

Datum zadání bakalářské práce:

28. července 2012

Termín odevzdání bakalářské práce:

4. září 2012

Ve Zlíně dne 24. února 2012



prof. Ing. Vladimír Vašek, CSc.
děkan



prof. Ing. Vladimír Vašek, CSc.
ředitel ústavu

ABSTRAKT

Tato bakalářská práce se zabývá datovými úložišti pro korporátní sféru. Jejích architekturou, způsobech zapojení a implementací do stávající síťové infrastruktury firmy ON Semiconductor Czech Republic, s.r.o.

Klíčová slova:

Úložiště, Datová úložiště, NetApp, ON Semiconductor, Migrace dat

ABSTRACT

This Bachelor thesis is focused on the enterprise data storage solutions. Specifically on their architecture and deployment into the existing corporate network infrastructure of the ON Semiconductor Czech Republic, s. r. o. company.

Keywords:

Storage, Data Storage, NetApp, ON Semiconductor, Data Migration

Na tomto místě chci poděkovat Ing. Tomáši Dulíkovi Ph.D. za poskytnutí možnosti pracovat právě na mnou vybraném tématu. A také Ing. Jiřímu Laresovi a Jiřímu Zelenkovi z firmy ON Semiconductor, kteří mi umožnili podílet se na tomto projektu.

Prohlašuji, že

- beru na vědomí, že odevzdáním bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na bakalářské práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně 3.9.2012

.....
podpis diplomanta

OBSAH

ÚVOD.....	9
I TEORETICKÁ ČÁST.....	10
1 DATOVÁ ÚLOŽIŠTĚ	11
1.1 ARCHITEKTURY PŘIPOJENÍ DATOVÝCH ÚLOŽIŠŤ	12
1.1.1 DAS -Directly Attached Storage.....	12
1.1.2 NAS - Network Attached Storage.....	13
1.1.3 SAN – Storage area network.....	14
1.1.4 Srovnání popsanych řešení.....	15
1.2 DISKOVÁ ROZHRANÍ.....	16
1.2.1 PATA - Parallel Advance Technology Attachment.....	16
1.2.2 SATA - Serial ATA	16
1.2.3 SCSI - Small Computer Systems Interface	16
1.2.4 SAS - Serial Attached SCSI.....	16
1.2.5 Fiber Channel - FC.....	16
1.3 PROTOKOLY	17
1.3.1 CIFS	17
1.3.2 NFS	18
1.3.3 FC.....	18
1.3.4 FcoE	18
1.3.5 iSCSI	18
1.3.6 InfiniBand	19
1.3.7 NDMP	19
1.4 OCHRANA DAT	19
1.4.1 RAID	19
1.4.1.1 Raid 0.....	20
1.4.1.2 Raid 1	20
1.4.1.3 Raid 3	20
1.4.1.4 Raid 5.....	21
1.4.1.5 Raid 6.....	21
1.4.1.6 Raid 10.....	21
1.4.1.7 Hot-spare.....	21
1.4.2 SnapShot	22
1.4.3 SnapVault.....	22
1.4.4 SnapMirror (Replikace)	23
1.5 POJMY DATOVÝCH ÚLOŽIŠŤ.....	23
1.5.1 Filer	23
1.5.2 Qtree	24
1.5.3 HBA - Hot Bus Adapter	24
1.5.4 FileOver (TakeOver).....	25
1.5.5 Cluster active/active	25
1.6 SITUACE NA TRHU DATOVÝCH ÚLOŽIŠŤ.....	26
1.6.1 NetApp	26
2 ANALÝZA DATOVÉHO ÚLOŽIŠTĚ ON SEMICONDUCTOR.....	28

2.1	ON SEMICONDUCTOR	28
2.2	TYPY DATOVÝCH ÚLOŽIŠŤ V ON SEMICONDUCTOR	29
2.2.1	Win cluster + IBM FastT200	29
2.2.2	EMC CX500.....	29
2.2.3	EMC CX3-20	30
2.2.4	NetApp FAS270	30
2.2.5	NetApp FAS2050	30
2.3	ZAPOJENÍ V SÍŤOVÉ INFRASTRUKTUŘE SPOLEČNOSTI	31
2.4	HODNOCENÍ STAVU	31
3	VÝBĚR NOVÉHO DATOVÉHO ÚLOŽIŠTĚ	32
3.1	POŽADAVKY NA NOVÉ DATOVÉ ÚLOŽIŠTĚ	32
3.1.1	Kapacitní nároky	32
3.1.2	Nároky na aplikační prostředí	32
3.2	NOVÉ DATOVÉ ÚLOŽIŠTĚ	33
II	PRAKTICKÁ ČÁST	36
4	MIGRACE DAT	37
4.1	KONFIGURACE HARDWARU	37
4.1.1	Konfigurace oddílů SATA disků	37
4.1.2	Konfigurace oddílů SAS disků.....	40
4.2	MIGRACE POMOCÍ CIFS	43
4.3	MIGRACE SNAPVAULT	44
5	HODNOCENÍ PRŮBĚHŮ MIGRACÍ.....	46
	ZÁVĚR	47
	ZÁVĚR V ANGLIČTINĚ.....	48
	SEZNAM POUŽITÉ LITERATURY.....	49
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....	50
	SEZNAM OBRÁZKŮ	52
	SEZNAM TABULEK.....	53
	SEZNAM PŘÍLOH.....	54

ÚVOD

Informace a data hrají v našem životě stále významnější roli. Sběr/shromažďování dat, jejich uchovávání, vyhodnocování a následné využívání je spojeno prakticky se všemi oblastmi lidského života. Datová úložiště jsou tedy významným prvkem, zásadně ovlivňujícím kvalitu a úspěšnost tohoto konání. Objemy dat neustále rostou a je třeba s nimi pracovat. Zejména firmy musí dbát nejen na spolehlivé ukládání dat, ale i na jejich zabezpečení a vysokou dostupnost. Proto je toto odvětví, obzvláště v posledních letech s rostoucími nároky firem, středem pozornosti.

I. TEORETICKÁ ČÁST

1 DATOVÁ ÚLOŽIŠTĚ

Datová úložiště jsou významným prvkem, zásadně ovlivňujícím kvalitu a úspěšnost každé firmy. Na otázku jaká datová úložiště jsou pro danou oblast podnikání nejvhodnější není jednoduchá a jednoznačná odpověď. Přesto lze však nastínit několik typizovaných modelových řešení. Základním vodítkem mohou být požadavky na výkon, kapacitu, způsob využití, dostupnost, stabilitu a toleranci při ztrátě dat.

Dynamické požadavky dnešních výpočetních systémů nejlépe postihuje centralizovaný a konsolidovaný diskový prostor sdílený více servery s řadou významných atributů:

- lze jej lépe rozdělovat a přerozdělovat, podle momentálních kapacitních a výkonových potřeb, jednotlivým serverům a aplikacím
- jednoduchá rozšiřovatelnost kapacity zamezuje plýtvání zdroji a využitelnost instalované kapacity bývá podstatně lepší nežli v klasickém modelu lokálních disků
- celkově výkonnější, spolehlivější a otevřený systém s plnou redundancí všech komponent, s duálními přístupovými cestami k serverům, s možností on-line expanze a migrace RAIDů a s konektivitou k serverům s různou architekturou a s různými operačními systémy
- snazší centralizovaný management šetří personální náklady - jediný člověk zvládá zpravidla obsluhovat až čtyřnásobek svěřené kapacity

Z hlediska topologie připojení datového úložiště k serverům je běžné využití fibre channel SAN struktury.

Z hlediska rozhraní k diskům se nejčastěji využívá FC rozhraní pro aplikace vyžadující vysoký transakční výkon (databázové aplikace a elektronická pošta). Pro fileserverové a archivní aplikace nebo aplikace zálohování do disku se s úspěchem využívá SATA rozhraní na velkokapacitní ekonomické disky. Perspektivními datovými úložišti schopnými uspokojit rozmanité aplikační potřeby jsou systémy se SAS rozhraním na disky a možností do libovolné diskové pozice instalovat podle potřeby buď výkonný SAS disk, nebo kapacitní a ekonomický SATA disk.

Z hlediska stability a dostupnosti datového úložiště lze využívat systému snapshotů pro rychlou obnovu provozu narušených dat. Systém synchronního zrcadlení nebo asynchronní vzdálené replikace datových úložišť zajišťuje zachování provozu i při výpadku celého datového úložiště provozované lokality v důsledku požáru nebo jiné živelné pohromy.[1]

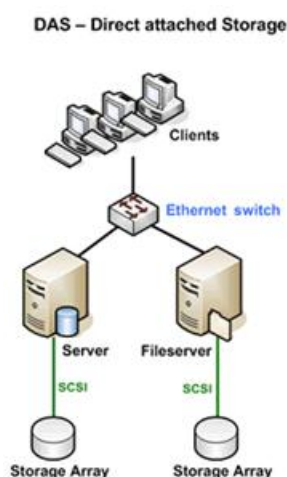
1.1 Architektury připojení datových úložišť

1.1.1 DAS -Directly Attached Storage

Původní a dnes již tradiční způsob připojení disků DAS - Directly Attached Storage (přímo připojená úložiště), čili ukládací zařízení přímo připojené k serveru. Tento na první pohled nejlogičtější způsob má řadu nevýhod. Při poruše serveru jsou data nepřístupná. Všechny informace směrem z i do úložní jednotky musí procházet serverem, takže často dochází k přetížení serveru a k malé výkonnosti tohoto řešení. I rozšiřitelnost diskové kapacity i spravovatelnost je značně omezená. Omezená je taktéž vzdálenost samotného serveru i jeho diskového subsystému, buď jsou disky přímo v serveru, nebo externím diskovém poli, v každém případě však délky kabelů (typicky SCSI) představují limitující faktor.

Mezi DAS zařízení můžeme zařadit např. interní disky samotného serveru nebo jednoduché diskové pole postavené zpravidla na technologii SATA nebo SCSI disků připojených k serveru buď technologií SCSI nebo Fiber Channel v konfiguraci Point-to-Point.

Celkové dlouhodobé náklady na správu tohoto řešení, jak již z výše uvedených záporů můžeme očekávat, budou vysoké. I přes veškeré tyto nevýhody je toto řešení hojně využíváno v menších firmách, kde hrají velkou roli počáteční náklady a náklady na správu nejsou příliš vysoké vzhledem k rozsahu ukládaných dat v těchto firmách.[2]



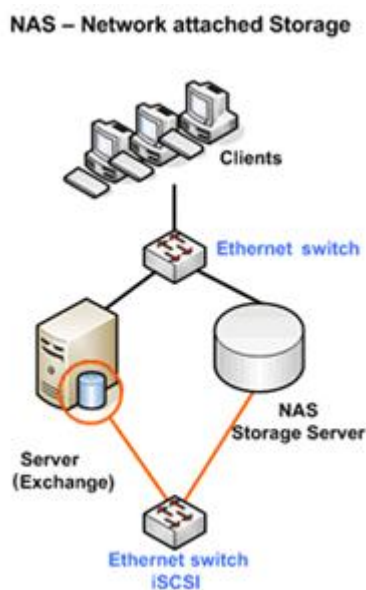
Obrázek 1 Řešení typu DAS [4]

Na konceptovém obrázku vidíme tradiční firemní síť s několika klientskými systémy a servery, každý se svým samostatným diskovým polem.

1.1.2 NAS - Network Attached Storage

Další velmi oblíbený mezistupeň ve vývoji datových úložišť představuje NAS. Hlavní odlišností od předchozího řešení je osamostatnění datového úložiště, které je nyní připojeno přímo do sítě. Přístup je realizován na souborové úrovni pomocí file protokolů NFS (Network File System) a CIFS (Common Internet File System), které jsou specializované na přenos a vyhledávání souborů – NAS tedy oproti DAS řešení neumožňuje přístup na blokové úrovni. Z toho důvodu není vhodný např. pro databáze a podobný SW vyžadující block-level přístup. NAS nabízí nákladově efektivní úložné řešení, které je mnoha firmami nasazováno s cílem snížit zatížení intenzivně využívaných serverů, aniž by bylo třeba jakkoliv modifikovat existující infrastrukturu. Je tedy vhodný pro subjekty, které chtějí pouze přidat úložnou kapacitu. Ideální je typicky pro rozrůstající se podniky – těm dovoluje snadnou migraci a přidávání storage systémů přímo za chodu

K nevýhodám patří velké zatížení klasické LAN, do které bývá NAS řešení obvykle připojováno. [3]



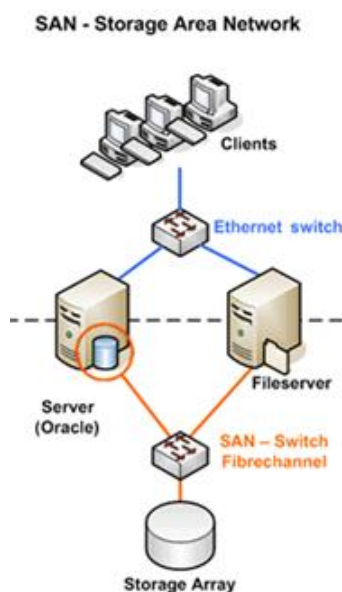
Obrázek 2 Řešení typu NAS [4]

1.1.3 SAN – Storage area network

San - Storage area network je dedikovaná (oddělená od LAN, WAN, atd) datová síť, která slouží pro připojení externích zařízení k serverům (disková pole, páskové knihovny a jiná zálohovací zařízení). SAN vznikla hlavně kvůli narůstajícím potřebám na zabezpečení a konsolidaci dat. Díky poměrně vysokým pořizovacím nákladům se SAN budují hlavně ve větších společnostech (bankovní sektor, automobilový průmysl, média), které vyžadují vysokou dostupnost svých služeb, rychlé odezvy v transakčně orientovaných aplikacích a co největší škálovatelnost (rozšiřitelnost). V poslední době se však díky klesající ceně začínají malé SAN rozšiřovat i do stále menších společností. SAN zařízení tak postupně nahrazují DAS (Direct Attach Storage) technologii.

SAN oproti DAS nabízí:

- Fyzické oddělení dat a serverů. Jednotlivé prvky mohou být od sebe vzdáleny až desítky km
- Sdílení zdrojů (diskových zařízení, zálohovacích zařízení) mezi jednotlivými servery (velmi omezeně umí toto i DAS)
- Vyšší propustnost
- Definici redundantních cest ke zdrojům (velmi omezeně umí toto i DAS)
- Podporu pro clusterová řešení a architekturu "no single point of failure"
- Podpora pro tzv. Disaster Recovery



Obrázek 3 Řešení typu SAN [4]

U SAN se používají protokoly, které pracují přímo s diskovými bloky. Základem je SCSI protokol, který je zabalen pro přenos přes Fibre Channel (FC) pomocí Fibre Channel Protocol (FCP) nebo pro přenos přes TCP/IP pomocí Internet Small Computer System Protocol (iSCSI). Případně některé další protokoly, které nejsou nyní příliš rozšířené jako Fibre Channel over Ethernet (FCoE), SCSI over Ethernet nebo další. Z výše uvedeného plyne, že disk (zařízení) připojený přes SAN v systému vypadá, jako by byl přímo připojený.

SAN diskové pole se skládá z kontroleru a diskových polic (Shelf). Na kontroleru se vytváří určité diskové prostory - virtuální disky (identifikují se pomocí Logical Unit Number - LUN), které se přiřadí určitým serverům [4]

1.1.4 Srovnání popsaných řešení

Vzhledem k detailnímu popisu všech řešení v předchozích kapitolách bych srovnání provedl formou srovnávací tabulky, abych neopakoval již jednou uvedená fakta.

Tabulka 1 Porovnání DAS, NAS, SAN

	DAS	NAS	SAN
Výhody	<ul style="list-style-type: none"> • nejjednodušší zřízení • nízké náklady • nevyžaduje dlouhé zaškolení 	<ul style="list-style-type: none"> • umožňuje sdílenou práci s uživatelskými daty • jednoduchá správa • nízké náklady 	<ul style="list-style-type: none"> • neomezuje provoz v podnikové síti (FC SAN) • jednoduchá správa • vysoká kapacita • snadnost rozšíření
Nevýhody	<ul style="list-style-type: none"> • nízká flexibilita • složité rozšiřování • nízký výkon • rychlé zastarávání 	<ul style="list-style-type: none"> • vysoká zátěž stávající sítě • nevhodné pro intenzivní práci s daty • jen souborová úroveň 	<ul style="list-style-type: none"> • vysoké pořizovací náklady (FC SAN) • složitost zřízení (FC SAN)

1.2 Disková rozhraní

1.2.1 PATA - Parallel Advance Technology Attachment

Standardní rozhraní pro připojení zařízení pro ukládání dat (jako pevné disky a CD-ROM mechaniky) uvnitř počítače. Využívá paralelní komunikaci. Dříve bylo označováno pouze jako ATA. Existují (skoro) synonyma jako IDE (Integrated Drive Electronics), EIDE (Enhanced IDE) a ATAPI (AT Attachment Packet Interface). [5]

1.2.2 SATA - Serial ATA

Sběrníková technologie pro počítače, která je primárně určena pro přenos dat na a z pevných disků a optických mechanik. Modernější sériová verze klasického ATA. SATA 2 teoreticky $3\text{Gb/s} = 375\text{ MBps}$. [5]

1.2.3 SCSI - Small Computer Systems Interface

Jedná se o skupinu standardů pro fyzické připojení a přenos dat mezi počítačem a periferních zařízení. Používá paralelní sběrníkový port. Ultra640 SCSI teoreticky 640 MBps[5]

1.2.4 SAS - Serial Attached SCSI

SAS vzniklo, stejně jako SATA, jako moderní rozšíření sběrníkové technologie pro připojení pevných disků a optických mechanik. Jedná se o sériovou sběrnici vycházející z SCSI, která je zpětně kompatibilní se SATA disky. Teoreticky 3Gbps. Interně disky buď SATA 2 nebo SCSI. [5]

1.2.5 Fiber Channel - FC

Jedná se o síťovou technologii primárně určenou pro datové sítě (storage networking), stalo se standardem pro SAN úložiště. Tradičně se používá optických kabelů, které mohou dosahovat rychlosti až 20 Gbps. Používá se pro vedení mezi servery a SAN kontrolery, ale také jako rozhraní na discích uvnitř SAN diskových polic. [5]

1.3 Protokoly

1.3.1 CIFS

CIFS neboli Server Message Block (SMB) je síťový komunikační protokol aplikační vrstvy, který slouží ke sdílenému přístupu k souborům, tiskárnám, sériovým portům a další komunikaci mezi uzly na síti. Poskytuje také autentizovaný mechanismus pro meziprocessovou komunikaci. Je využíván hlavně na počítačích s operačními systémy rodiny Windows.

Protokol díky firmě Microsoft je rozšířený v oblasti komunikace souborových serverů a klientů sítí LAN. Protokol pracuje na principu klient-server. Server umožňuje klientům síť přistupovat k tzv. sdíleným prostředkům, např. sdíleným diskům, adresářům, tiskovým frontám nebo pojmenovaným kanálům. Sdílené prostředky jsou rozpoznávány pomocí síťové adresy UNC (\\jméno_serveru\jméno_zdroje). Klientská část má za úkol definovat požadavky na sdílené prostředky serveru. Server vykonává rozbor požadavků odeslané klientem ve tvaru bloku (paketu) SMB, porovnává přístupová práva a na podkladě přístupových práv zahajuje požadovanou operaci (vytvoření adresáře, spuštění souboru atd.). Výsledek je s odpovědí poslán klientu identickým blokem SMB.

Z pohledu serveru SMB k přístupu ke sdíleným prostředkům máme dva pohledy: řízení přístupu z pohledu na úroveň sdíleného prostředku (share level), řízení přístupu z pohledu na uživatelskou úroveň (user level).

V prvním případě server povoluje přístup ke sdíleným prostředkům na základě správného hesla, které je přiřazeno k jednotlivému sdílenému prostředku. Po zadání správného hesla klientem je klientu přidělen identifikátor prostředku NID (Network ID), pomocí něhož přistupuje k prostředku.

V druhém případě se klient přihlašuje na server hned pomocí jména a hesla a v případě správnosti server přiřadí klientu uživatelský identifikátor UID (User ID), pomocí něhož server odvozuje přístupová práva při manipulaci se sdíleným prostředkem.

Princip komunikace je následující: Klient zasílá serveru požadavek — dochází k specifikaci parametru spojení a verzi protokolu mezi serverem a klientem.

Klient odešle požadavek, ve kterém je uvedeno jméno a heslo uživatele. Pokud je server v režimu user-level, přidělí uživateli UID.[4]

1.3.2 NFS

Network File System (NFS) je internetový protokol pro vzdálený přístup k souborům přes počítačovou síť. Protokol byl původně vyvinut společností Sun Microsystems v roce 1984, v současné době má jeho další vývoj na starosti organizace Internet Engineering Task Force (IETF). Funguje především nad transportním protokolem UDP, avšak od verze 3 je možné ho provozovat také nad protokolem TCP.

V praxi si můžete prostřednictvím NFS klienta připojit disk ze vzdáleného serveru a pracovat s ním jako s lokálním. V prostředí Linuxu se jedná asi o nejpoužívanější protokol pro tyto účely.[4]

1.3.3 FC

Fibre Channel využívá se optických vedení Fibre Optics (v současnosti o rychlosti 4Gbps) a protokolu FCP. Jsou potřeba speciální FC switche (pokud nemáme pouze spojení Point-to-Point) a servery musí mít Fibre Channel Host Bus Adapter (HBA) .[4]

1.3.4 FcoE

Fibre Channel over Ethernet (FCoE) je zapouzdření FC přes Ethernetové síť. To umožňuje FC používat 10Gbps Ethernetové síť (nebo vyšší rychlost), při zachování Fibre Channel protokol.[4]

1.3.5 iSCSI

FC je značně drahá záležitost, takže se v době, kdy se začal používat gigabitový ethernet (v té době se používalo FC s rychlostí 2Gbps) objevil standardizovaný protokol Internet Small Computer System Protocol (iSCSI). Ten zapouzdřuje SCSI do TCP/IP protokolu, takže může fungovat na běžné ethernetové síti (podmínkou je minimální rychlost 1 Gbps). Teoreticky může fungovat na již existující síti (switchích a síťových kartách), ale v praxi se rozhodně doporučuje vybudovat oddělenou síť pro SAN (pro datový provoz není problém vytížit 1Gbps síť).

iSCSI protokol sice není pro datový přenos, tak dobrý jako FCP protokol, který byl speciálně pro tento účel vyvinut, ale je dostatečný. V dnešní době jsou SAN sítě založené na iSCSI značně rozšířené. Je tu ovšem otázka síťové karty. iSCSI se může provozovat přes klasickou síťovou kartu, ale pak veškeré zpracování (zapouzdřování SCSI do iSCSI) provádí procesor počítače. Druhá možnost je zakoupit speciální iSCSI HBA, který provádí tyto úkony a počítači již předává normální SCSI. [4]

1.3.6 InfiniBand

Infiniband je vysokorychlostní nízkolatenční síťové propojení. Poskytuje obvykle výrazně lepší propustnost než standardní gigabitový Ethernet. [6]

1.3.7 NDMP

NDMP (Network Data Management Protocol) protokol je navržen k zálohování systémů NAS bez potřeby transportu dat přes záložní server, čímž dochází ke zvýšení rychlosti zálohy a snížení zatížení záložního serveru. Tento protokol byl vyvinut v kooperaci mezi společnostmi Legato a NetApp. NDMP je v současné době široce podporován napříč zálohovacím softwarem i NAS produkty.[7]

1.4 Ochrana dat

1.4.1 RAID

Vzhledem k tomu, že pevný disk je složité zařízení kombinující elektroniku a jemnou mechaniku, je již ze svého principu náchylný k poruše. Toto je nepříjemné zejména u serverů, kde jednak cena uložených dat může představovat mnohamilionové částky, jednak - i při pravidelném zálohování - jen pouhá odstávka serveru spojená s opravou a obnovou dat představuje značnou ztrátu na prostojích mnoha uživatelů. Proto byla zkonstruována disková pole, kde se pomocí speciálního řadiče více disků fyzických navenek jeví jako jeden disk logický. Dalším důvodem použití pole je vytvoření větší diskové kapacity, než se v dané době vyrábí v podobě samostatného disku. Pole se ve zkratce nazývají RAID. Nejprve to znamenalo Redundant Array of Inexpensive Disks,

dnes se zkratka překládá spíše jako Redundant Array of Independent Disks. Pro vyjasnění terminologie byly definovány tzv. úrovně RAID. Některé se s vývojem techniky již přestaly používat. V současné době jsou nejpoužívanější úrovně RAID 0,1 a 5. Kombinací typu 0 a 1 vznikne pole typu RAID 10. V poslední době se začínají používat také pole typu RAID 6 (zdvojená parita).[8]

1.4.1.1 Raid 0

Především je třeba říci, že toto pole není redundantní. Ztráta jednoho disku znamená ztrátu celého pole. Důvodem použití je výkon, tedy zvýšení přenosové rychlosti nebo propustnosti dat tam, kde na uchování dat nezáleží tak, jako na rychlosti, například při střihání videa. [8]

1.4.1.2 Raid 1

Raid 1 je známé “zrcadlení”, kdy se na dva disky stejných kapacit ukládají totožné informace a při výpadku jednoho disku se bez přerušení pokračuje v činnosti. Jednoduchá implementace, často čistě softwarová, zato je potřeba 100% diskové kapacity navíc. Z hlediska výkonu pomalejší zápis (zapisuje se 2x), rychlejší čtení (řadič může střídat požadavky mezi disky, “rozdávat práci”).[8]

1.4.1.3 Raid 3

Je použito N+1 stejných disků. Na N disků se zapisují data, na poslední disk paritní informace. V případě výpadku disku jsou chybějící informace dopočítány z ostatních datových disků a disku paritního. Každý zápis na pole znamená práci i pro paritní disk, proto je více namáhán a lze u něj očekávat vyšší poruchovost. Proto byla tato úroveň již téměř opuštěna ve prospěch Raid 5 a vyšších, na druhé straně se objevila hardwarová implementace “přímo v křemíku”, kdy hardwarová implementace je snadnější než u jiných úrovních. [8]

1.4.1.4 Raid 5

Redundantní pole s distribuovanou paritou. Minimální počet disků jsou 3. Režie je 1 disk z n-diskového pole. Máme-li například pole RAID 5 z 5 disků o kapacitě 36 GB, užitečná kapacita je $4 \times 36 = 144$ GB a 1 disk je režie. Data se zapisují postupně na disky 0,1... až na poslední disk se zapíše parita. Při výpadku některého disku pak máme buď všechna data (a nepotřebujeme paritu), nebo máme část dat a paritu a chybějící data ze ztraceného disku umíme dopočítat z dat, která máme a parity. Výkon při čtení dobrý, zápis je pomalejší (jeden zápis je čtení starých dat, čtení staré parity, výpočet nové parity a zápis nových dat a nové parity). [8]

1.4.1.5 Raid 6

Pole typu RAID 5 ještě s jedním paritním diskem navíc. Je odolné proti výpadku dvou disků. Důvodem použití je ta skutečnost, že při obrovských kapacitách dnešních disků trvá rekonstrukce pole při výpadku disku dosti dlouho, a po dobu rekonstrukce již pole není chráněno proti výpadku dalšího disku. Navíc se u RAID 5 může stát, že právě při rekonstrukci, kdy se kvůli rekonstrukci chybějících dat čtou kompletní povrchy všech zbývajících disků pole, se na některém z těchto disků narazí na chybu čtení, která se dosud v provozu nemusela projevit, řadič takový disk taktéž odpojí a neštěstí je hotovo - úplná ztráta dat celého pole. [8]

1.4.1.6 Raid 10

Kombinace RAID 0 (stripe) a RAID 1 (zrcadlo). Jedná se vlastně o zrcadlený stripe. Minimální počet disků 4, režie 100% diskové kapacity navíc. Poskytuje nejvyšší výkon v bezpečných typech polí, podstatně rychlejší než RAID 5 zejména při zápisu. Další výhodou je odolnost proti ztrátě až 50% disků (naproti tomu RAID 5 odolává ztrátě pouze jednoho disku). [8]

1.4.1.7 Hot-spare

Pro aplikace, kde je velmi důležitá nepřetržitost provozu, se používá technologie hot-spare. Ze čtyř disků například vytvoříme pole RAID 5 (jeho užitečná kapacita je tedy

3xkapacita disku). Pátý disk je zasunut a zkonfigurován jako hot-spare, což znamená, že v případě výpadku jednoho z disků pole je okamžitě automaticky aktivován a jsou na něj dopočítána chybějící data za vypadlý disk. Tak je minimalizováno rizikové časové okno, kdy pole sice funguje, ale výpadek dalšího disku již znamená ztrátu všech dat pole, protože není třeba čekat na příjezd technika a výměnu disku. Disk Hot-spare je možno sdílet pro více polí - jsou-li v serveru například 2 pole, jedno ze tří disků a jedno z pěti, je disk hot-spare aktivován při výpadku jak disku z prvního tak druhého pole. [8]

1.4.2 SnapShot

Snapshot je standardní termín pro schopnost zaznamenat stav paměťového zařízení k danému okamžiku, určená pro rychlou obnovu dat. Snapshot je kopií dat k určitému času (Point-in-time-copy). Většinou je Snapshot po vytvoření okamžitě k dispozici pro použití v jiných aplikacích, jako je zálohování, testování či analýza dat. Originální data jsou i nadále k dispozici aplikaci bez přerušení provozu, pouze s minimálním pozastavením datového toku. Díky snapshotům je možné rychle obnovit například omylem smazaná data bez nutnosti použití obnovy pomocí zálohovacího software. To je obzvláště přínosné, pokud úložiště umožňuje vytvořit velké množství snapshotů v krátkých časových intervalech. Vytváření snapshotů s sebou nese nároky jak na výkon, tak na kapacitu datového úložiště v závislosti na použité metodě tvorby snapshotů.[9]

1.4.3 SnapVault

NetApp Snap Vault je novou generací zálohování dat. Je založen na blokových změnách dat a minimalizaci přenosu dat po LAN/WAN linkách. Umožňuje rychlou obnovu dat, je vhodný jako levné řešení typu Disaster Recovery. Pomocí NetApp SnapVault urychlíte a zjednodušíte vaše řešení zálohování a archivace dat.

Na rozdíl od snapshotů technologie SnapVault, poskytuje ochranu, která v případě selhání hardwaru určuje tzv. primární a sekundární úložiště (primary a secondary storage). Snapshoty jsou vytvářeny jak na primárním, tak na sekundárním úložišti. Sekundární úložiště je vzdálené úložné zařízení, které poskytuje ochranu pro případ selhání nebo havárie hardwaru. Počty primárních a sekundárních snapshotů se mohou lišit. Na primárním úložišti můžete mít zálohy za první týden (pro případ rychlé obnovy; uživatelé

mohou velice jednoduše obnovovat soubory z primárního úložiště sami pomocí Windows funkce „Previous versions“, případně otevřít skrytý adresář .snapshot přes NFS nebo ~snapshot přes CIFS).[10]

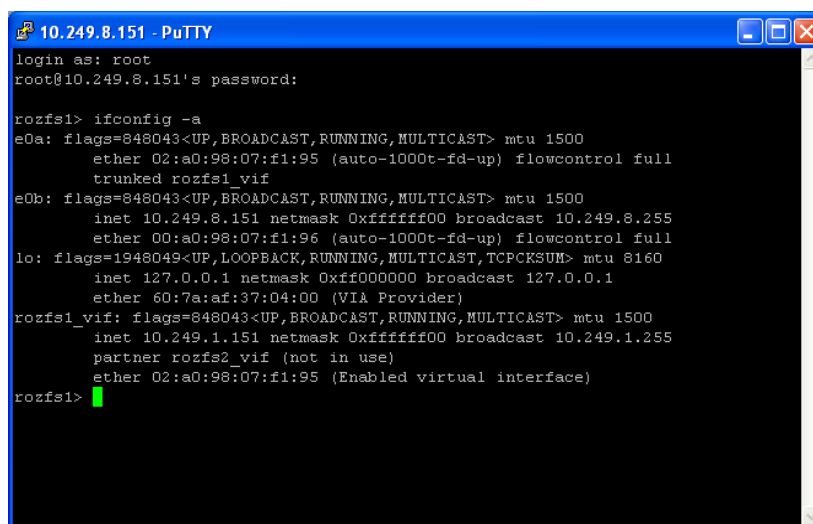
1.4.4 SnapMirror (Replikace)

SnapMirror je software určený k replikaci oddílů nebo QTree napříč sítí mezi jednotlivými filery. SnapMirror je, jak již název napovídá, založen na vytváření snapshotů. V případě, že inicializujete relaci SnapMirror, tak se vytvoří pracovní snapshot a všechno starší, než je tento snapshot, se zreplikuje na cílový filer (včetně všech starších snapshotů). Při příštím spuštění SnapMirror se vytvoří nový snapshot a přenesou se všechny změny oproti předchozímu snapshotu.[10]

1.5 Pojmy datových úložišť

1.5.1 Filer

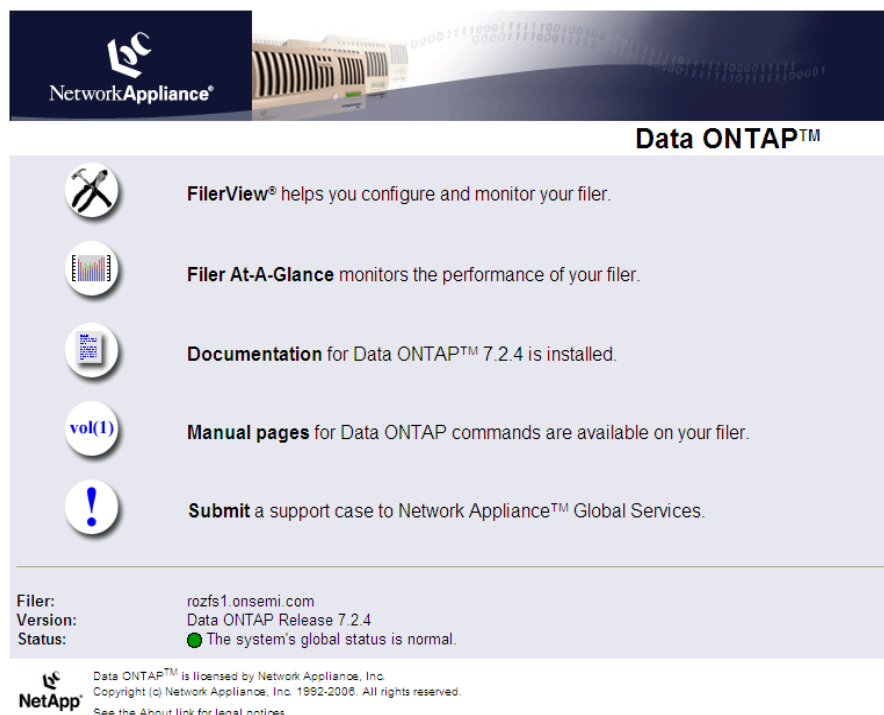
Většinou se tímto slovem označují zařízení FAS (NetApp Fabric-Attached Storage). Jedná se o Storage box, který kombinuje vlastnosti zařízení SAN a NAS spolu s filesystémem, který se nazývá WAFL. Filery mají baterií zálohovaná NVRAM a operační systém, kterému se říká Data ONTAP. Správa Fileru se může provádět pomocí webového rozhraní nebo SSH klienta.[10]



```
10.249.8.151 - PuTTY
login as: root
root@10.249.8.151's password:

rozfs1> ifconfig -a
e0a: flags=848043<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    ether 02:a0:98:07:f1:95 (auto-1000t-fd-up) flowcontrol full
    trunked rozfs1_vif
e0b: flags=848043<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 10.249.8.151 netmask 0xfffff00 broadcast 10.249.8.255
    ether 00:a0:98:07:f1:96 (auto-1000t-fd-up) flowcontrol full
lo: flags=1948049<UP,LOOPBACK,RUNNING,MULTICAST,TCPCKSUM> mtu 8160
    inet 127.0.0.1 netmask 0xff000000 broadcast 127.0.0.1
    ether 60:7a:af:37:04:00 (VIA Provider)
rozfs1_vif: flags=848043<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 10.249.1.151 netmask 0xfffff00 broadcast 10.249.1.255
    partner rozfs2_vif (not in use)
    ether 02:a0:98:07:f1:95 (Enabled virtual interface)
rozfs1>
```

Obrázek 4 Správa fileru pomocí SSH



Obrázek 5 Správa fileru pomocí webového rozhraní

1.5.2 Qtree

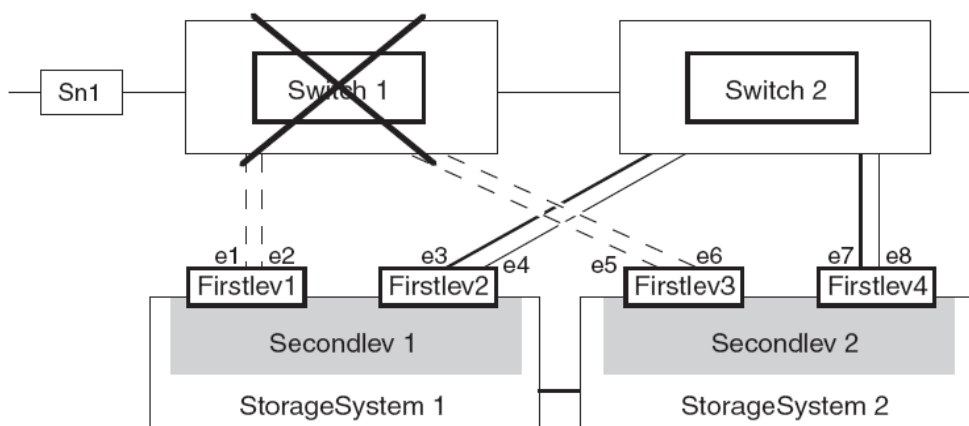
QTree je logický prvek jenž stojí nad každým oddílem. Ten z hlediska klientů vypadá jako běžný adresář, z hlediska filesystému můžete pro každý QTree definovat například různé limity (velikost, masku pro jména souborů – pokusit se tak uživatelům zabránit v tom, aby si na firemní storage servery ukládali multimediální obsah nebo potenciálně nebezpečné spustitelné soubory), případně můžete pro každý takový QTree vytvořit Snapshot nebo relaci SnapMirror.[10]

1.5.3 HBA - Hot Bus Adapter

HBA je o adaptér (karta), která je komunikačním rozhraním mezi sítí a například počítačem. Fibre channel HBA jsou dostupné pro většinu předních open systémů, počítačových architektur a sběrnic včetně PCI a Sbus. Každý HBA má unikátní WWN (World Wide Name), které je podobné Ethernetovým MAC adresám. Používají se dva druhy WWN - pro celý adaptér a pro každý port.[10]

1.5.4 FileOver (TakeOver)

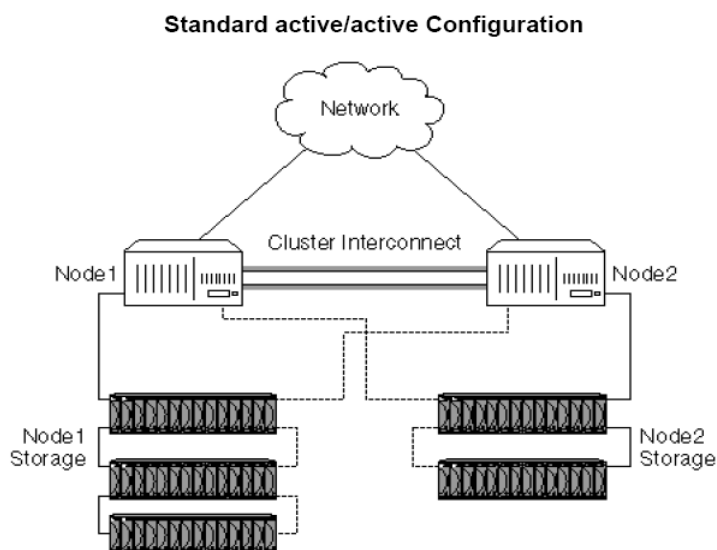
Failover je automatické přepnutí na redundantní nebo pohotovostní server, systém nebo síť při poruše nebo výpadku aktivního serveru, systému nebo sítě. Technologie failover se obvykle využívá na serverech, systémech nebo sítích, které vyžadují nepřetržitou dostupnost a vysoký stupeň spolehlivosti.[4]



Obrázek 6 FileOver [4]

1.5.5 Cluster active/active

Cluster A/A znamená, že služby běží na obou nodech, při takeoveru jsou všechny provozovány na jednom nodu. Po obnovení funkčnosti lze služby přesunout na původní node.[4]



Obrázek 7 Cluster Active/Active [4]

1.6 Situace na trhu datových úložišť

Na trhu datových úložišť se pohybuje mnoho firem Fujitsu, Hewlett Packard, IBM, SUN, Dell. Firma OnSemiconductor využívá zejména v USA outsourcingu společnosti Hewlett Packard. V EMEA regionu ale dává přednost osvědčenému výrobcí datových úložišť společnosti NetApp.

1.6.1 NetApp

Společnost NetApp, Inc. (dříve Network Appliance, Inc.) byla založena v roce 1992 v USA. Od svého vzniku se soustředí na vývoj a výrobu datových úložišť a v této oblasti patří k průkopníkům a jednoznačně ke světové špičce. Díky svému zaměření na jednoduchost ovládání, spolehlivost a výkon datových úložišť uspokojí produkty společnosti NetApp potřeby každého zákazníka. Z produktů lze vybrat od malých datových úložišť s kapacitou do 4 TB, přes enterprise úložiště s kapacitou kolem 100 TB, po zařízení s kapacitou až 14 PB.

Kde ale NetApp výrazně vybočuje z řady producentů diskových polí, jsou specifické funkcionality a softwarové vybavení diskových polí, které přinášejí velkou přidanou hodnotu do všech myslitelných řešení. Zde produkty NetApp nedávají konkurenci šanci. Mezi základní funkcionality všech diskových polí NetApp patří například:

- integrovaný automatický RAID manager
- Snapshot (až 250 snapshotů na LUN)
- Fast Boot
- telnet
- e-mailová varování
- FilerView
- NDMP
- FlexVol
- FlexCache
- Thin Provisioning
- deduplikace

Samozřejmostí je také podpora téměř všech komunikačních a přístupových protokolů, které se pro práci s datovými úložišti používají:

Síťové protokoly:

- NFS V2/V3/V4 přes UDP nebo TCP
- PCNFSD V1/V2 pro (PC) NFS autentizaci klientů
- Microsoft CIFS
- VLD
- HTTP 1.0
- HTTP 1.1 Virtual Hosts

SAN protokoly:

- Fibre Channel Protocol (FCP) pro SCSI; "fabric-attached" a "direct-attached"
- iSCSI

Kromě jiného nabízí řešení diskových polí v konfiguracích Active/Active Cluster a MetroCluster pro maximální dostupnost dat za všech okolností.

Datové úložiště od společnosti NetApp se pyšní oproti konkurenci různými vylepšeními jako je deduplikace a Thin Provisioning. Soubory uložené na WAFL filesystému jdou rozdělené do bloků po 4 kB a WAFL v sobě obsahuje deduplikaci právě na úrovni těchto bloků. Deduplikace v podstatě znamená, že pokud máte na disku dva stejné soubory (respektive zmiňované bloky), tak je ve skutečnosti na disk uložen blok jen jeden a pro druhý existují jen metadata. Z toho plyne nějaká úspora nákladů na ukládání dat. Úspory budou větší ve virtualizovaném prostředí, kde je více VMWare image disků na NetAppu se stejným operačním systémem a velice podobným aplikačním setem, mnohem menší budou v nevirtualizovaném prostředí, kde je většina vašich dat unikátních.

V systémech, které podporují Thin Provisioning, je možné oddíl kdykoli zvětšovat bez nutnosti odstávky služby, a to během několika sekund. Případně se dá zapnout tzv. auto grow, což znamená, že ve chvíli, kdy začne na nějakém oddílu docházet místo, systém se sám pokusí jeho velikost zvětšit – a pokud je na agregátu volné místo, tak se to povede.[11]

2 ANALÝZA DATOVÉHO ÚLOŽIŠTĚ ON SEMICONDUCTOR

2.1 On Semiconductor

Společnost ON Semiconductor patří mezi přední světově výrobce polovodičů. Společnost sídlí ve Phoenixu v americkém státě Arizona. ON Semiconductor má dlouholetou tradici a téměř 1 500 zaměstnanců v České republice. Produkty této společnosti mají široké využití v různých oblastech, jako jsou spotřební a průmyslová elektronika, telekomunikační technologie, automobilový průmysl a vypočtení technika.

Moderní integrované obvody jsou dnes navrhovány a vyráběny v rožnovských společnostech SCG Czech Design Center a ON Semiconductor Czech Republic v Rožnově pod Radhoštěm a dále v brněnské společnosti ON Design Czech. Mezi hlavní aktivity společnosti ON Semiconductor v České republice patří návrh integrovaných obvodů (Rožnov p. R. a Brno), výzkum a vývoj, výroba křemíku, výroba čipů, sdílené IT služby (Rožnov p. R.).

IT oddělení společnosti ON Semiconductor používala diskové pole NetApp řady 2000 pro svůj produkční provoz již delší dobu, takže s technologiemi NetApp již má zkušenosti a je s nimi spokojeno. A protože tento systém byl ale na hranici své životnosti, začalo se uvažovat o upgradu toho systému za nový systém NetApp s větším výkonem i možností rozšíření. Další úvahy směřovaly ke zvýšení zálohovací kapacity a implementaci zálohování pomocí software Symantec Netbackup.

2.2 Typy datových úložišť v ON Semiconductor

2.2.1 Win cluster + IBM FastT200

Storage IBM FastT200 o velikosti 0,9 TB bylo využíváno od roku 2000 do roku 2008 servery Mirka (Oracle databáze), Lucka (Aplikační server), Lenka a Hanka (Windows 2000 server – sdílené složky pomocí CIFS protokolu)



Obrázek 8 IBM FastT200

2.2.2 EMC CX500

Úložiště EMC CX500 o velikosti 0,5TB je využíváno od roku 2006 dodnes. Je využíváno SQL servery S-Dana a S-Denisa pro SQL databáze pro lokální aplikace a dále Windows 2003 servery S-Olga a S-Svatava na kterých běží lokální virtuální aplikace.



Obrázek 9 EMC CX500

2.2.3 EMC CX3-20

Další úložiště EMC CX3-20 má velikost 0,9TB a využívaly ho v letech 2007 až 2011 servery S-Jitka a S-Blanka (Oracle databáze).



Obrázek 10 EMCCX3-20

2.2.4 NetApp FAS270

První úložiště od firmy NetApp je využíváno od roku 2008 a má velikost 0,9TB. Jsou na něm uložena sdílená data (U-drive a G-drive). Původně byl v Rožnově, nyní po implementaci nového datového úložiště byl přesunut do sesterské společnosti do Piešťan.



Obrázek 11 NetApp FAS270

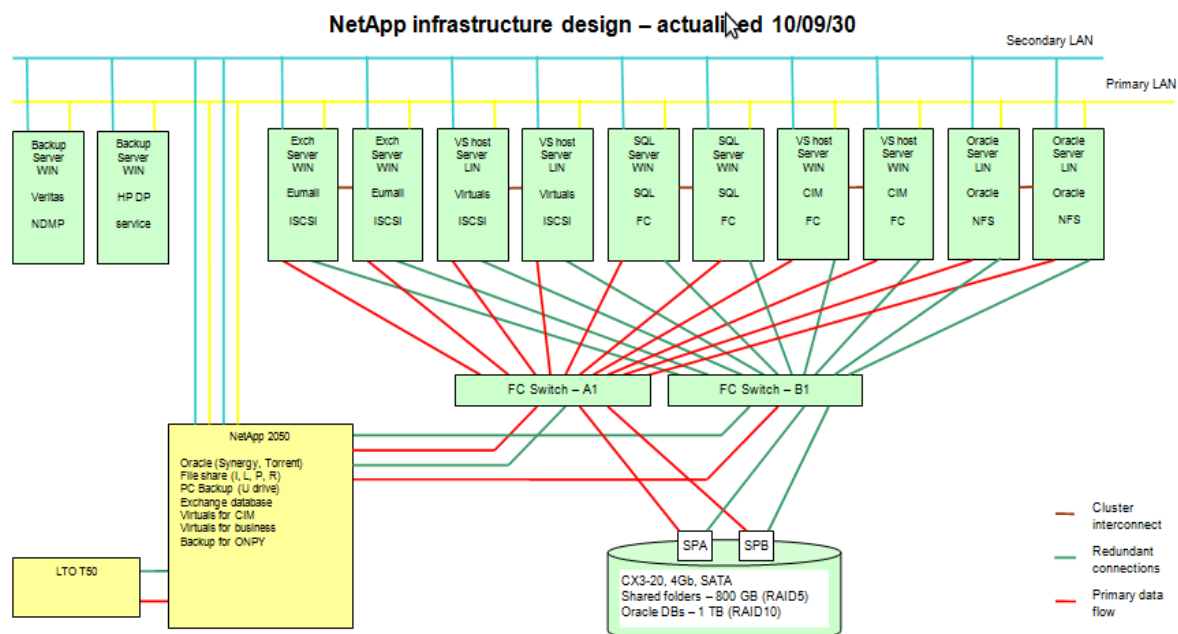
2.2.5 NetApp FAS2050

Druhé úložiště od firmy NetApp je o velikosti 15TB a bylo zakoupeno do společnosti v roce 2008. A slouží jako NAS úložiště pro produkční sdílené data (G-drive), a pro část Exchange databáze.



Obrázek 12 NetApp FAS2050

2.3 Zapojení v síťové infrastruktuře společnosti



Obrázek 13 Zapojení NetApp FAS2050 a EMC CX3-20

2.4 Hodnocení stavu

Stávající stav již nesplňoval nároky na výkon a s vidinou rozšíření kapacitních možností byl proveden upgrade (vylepšení) na novější verzi NetAppu.

3 VÝBĚR NOVÉHO DATOVÉHO ÚLOŽIŠTĚ

Je vyžadována změna v oblasti systému ukládání dat a infrastruktury serverů. Předmětem projektu je implementace storage systému NetApp včetně potřebné infrastruktury a systému zálohování pomocí zálohovacího SW Symantec NetBackup.

3.1 Požadavky na nové datové úložiště

3.1.1 Kapacitní nároky

Kapacitní nároky byly vypočítány z dosavadních úložišť a z předpovědí nárůstu objemu dat na dalších 5 let.

Tabulka 2 Kapacitní nároky na nové datové úložiště

Aplikace	Typ Dat	Server	Komunikační protokol	Čistá kapacita [GB]	Roční nárůst [%]	Plán na období [roky]	Celková plánovaná čistá kapacita [GB]
Oracle 1	DB	ORA1	NFS	1750	10%	5	2 818,39
Oracle 2	DB	ORA2	NFS	750	10%	5	1 207,88
Linux virtuals	File	RHEL cluster	FC	2000	0%	5	2 000,00
G-drive	File	NAS	CIFS	2000	10%	5	3 221,02
U-drive	File	NAS	CIFS	5000	10%	5	8 052,55
MS SQL	DB	Win cluster	FC	1500	10%	5	2 415,77
LB aplikace	DB	Win cluster	FC	1000	0%	5	1 000,00
CIM aplikace	File	Lin cluster	NFS	300	0%	5	300,00
CIM data	File	Lin cluster	NFS	700	10%	5	1 127,36
Restore pool	File	NAS	CIFS/NFS	800	0%	5	800,00
Celkem				15 800,00			22 942,97

3.1.2 Nároky na aplikační prostředí

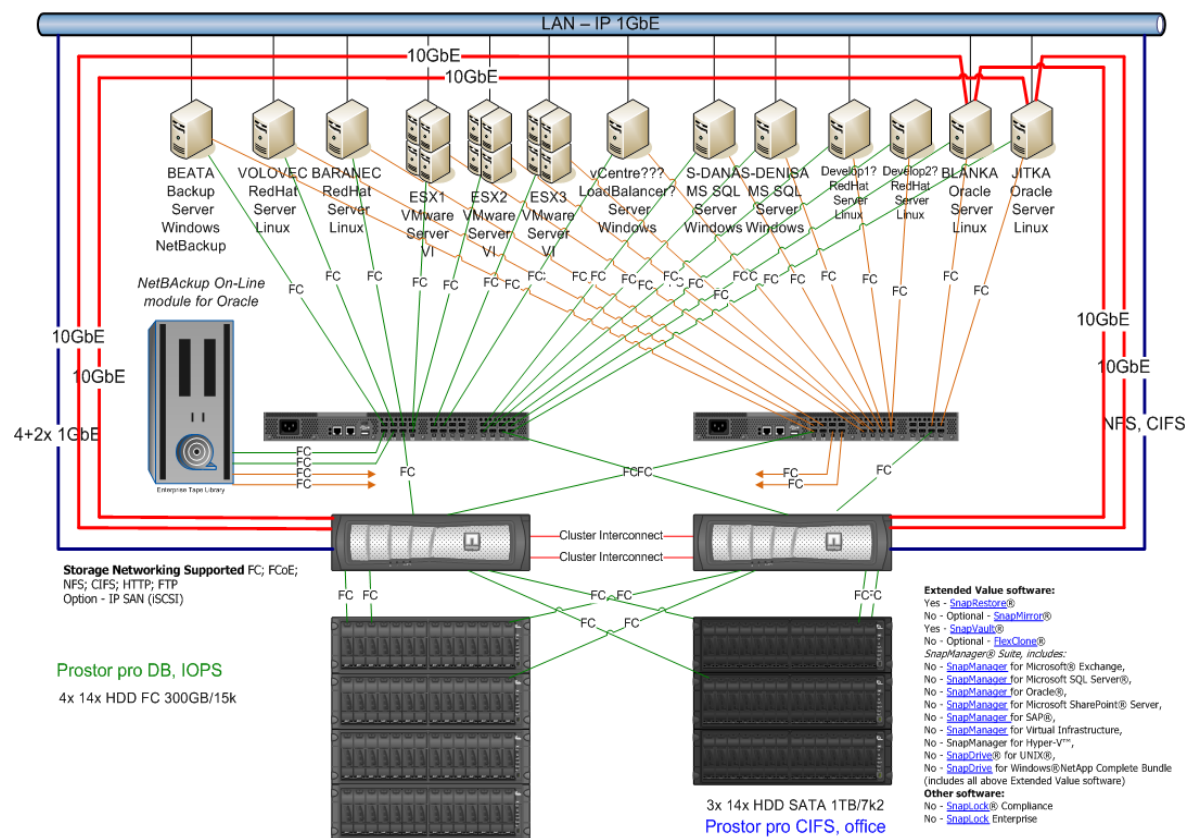
Nároky na aplikační prostředí se týkaly virtualizace, kdy pro virtualizaci bude použit VMWare. Systémy VMware budou k diskovému poli připojeny prostřednictvím FC konektivity. Ve virtuálním prostředí poběží celkem 4x virtuální servery 2x VMware s operačním systémem Windows 2008 Enterprise a 2x VMware s operačním systémem Windows 2008 Standard.

Dále oracle databáze počítá s primárním připojením prostřednictvím NFS s možností přechodu k sekundární variantě v případě potřeby, to znamená připojení prostřednictvím FC.

SQL databáze – servery jsou připojeny prostřednictvím FC rozhraní a jedná se o samostatné servery běžící v MS active/active Clusteru.

3.2 Nové datové úložiště

V rámci realizačního projektu společnost Proact nabídla řešení diskového pole NetApp FAS3240 s diskovými policemi s rychlými SAS i kapacitními SATA disky o velikosti 35 TB. Součástí řešení bylo i vybudování nové SAN infrastruktury.



Obrázek 14 Schéma zapojené nového datového úložiště

Firmou Proact byly dodány následující komponenty:



Obrázek 15 NetApp FAS 3240EAX

Tabulka 3 Seznam komponentů dodaných firmou Proact

Disk storage NetApp FAS 3240EAX
2x box DS4243 24x 300GB/15k SAS
2x box DS4243 24x 1TB/7k2 SATA
Warranty and SW subscription 5 year
8Gb Dual Port FC HBA, PCIe, LC multi-mode optic for Backup server
Backup SW
6x Symantec NetBackup Standard Client 7.1 WIN/LIN
5x Symantec NetBackup Enterprise Client 7.1 WIN/LIN
2x On-line module for Oracle
1x Symantec NetBackup NDMP licence
1x Symantec NetBackup DRIVE licence
1x Symantec NetBackup Upgrade to Enterprise edition
Backup Library
NEO8000
4x LTO5 FC Drive, 2/4/8Gb, 1,5/3 TB
100 Slot for LTO media
Bar code scanner
Warranty 3y
100x Media LTO5
1x cleaning cartridge
Services
Installation, integration, implementation
training for administrators and operators
14 days support after finish of implementation

A standardní softwarový balíček:

SAN nebo NAS protocol (iSCSI, FC, NFS, nebo CIFS), RAID manager, obsahující RAID-DP®, Snapshot™, FlexVol®, FlexShare®, Deduplication, Thin provisioning, Operations Manager, Provisioning Manager, Protection Manager, Protection Manager, Management Server (optional), System Manager, FilerView®, MultiStore®, FlexCache™, NearStore®, Cluster failover, MetroCluster, SyncMirror®, Open Systems SnapVault.

V rámci tohoto projektu byla dodána i nová pásková knihovna Overland NEO8000 s LTO-5 mechanikami. Na tyto mechaniky jsou pomocí implementace Symantec Netbackup prováděny zálohy produkčních dat.

Plně se využívá jak bloková záloha snapshotů pomocí protokolu NDMP a to jak File systémů CIFS, NFS, tak i prostředí Oracle DB. Došlo k desetinásobnému urychlení procesů zálohování a obnov, což potvrdil nejen zkušební provoz zálohovacího prostředí založeného na inteligentním řešení Symantec Netbackup Enterprise 7.x s Overland Enterprise library Neo8000e.

II. PRAKTICKÁ ČÁST

4 MIGRACE DAT

Migrate dat byla provedena z datového úložiště NetApp FAS2050 na nové NetApp FAS3420. Jednalo se o data se sdílených disku G-drive a U-drive pomocí protokolu CIFS, který je standardem pro sdílení souborů ve Windows platformách.

Dále byla provedena migrace pomocí technologie SnapVault, která se stará o zálohování dat pomocí snapshotů z datového úložiště NetApp FAS270, které se nachází v dceřiné společnosti v Piešťanech na Slovensku a je přenášena pomocí LAN

4.1 Konfigurace hardwaru

4.1.1 Konfigurace oddílů SATA disků

Disky jsou spojovány do RAID skupin, kdy každá skupina obsahuje 20 fyzických disků. Toto rozmístění je doporučeno pro dosažení optimálního volného prostoru. Osm dalších disků je nepřirazeno a slouží jako náhradní disky, které mohou být připojeny, pokud je to nutné.

Konfigurace SATA disků:

```
Aggregate aggr0 (online, raid_dp) (block checksums)
```

```
  Plex /aggr0/plex0 (online, normal, active)
```

```
    RAID group /aggr0/plex0/rg0 (normal)
```

RAID	Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
dparity	4b.01.4		4b	1	4	SA:A	-	BSAS	7200
parity	4b.01.5		4b	1	5	SA:A	-	BSAS	7200
data	4b.01.6		4b	1	6	SA:A	-	BSAS	7200
data	4b.01.7		4b	1	7	SA:A	-	BSAS	7200
data	4b.01.8		4b	1	8	SA:A	-	BSAS	7200
data	4b.01.9		4b	1	9	SA:A	-	BSAS	7200
data	4b.01.10		4b	1	10	SA:A	-	BSAS	7200

data	4b.01.11	4b	1	11	SA:A	-	BSAS	7200
data	4b.01.12	4b	1	12	SA:A	-	BSAS	7200
data	4b.01.13	4b	1	13	SA:A	-	BSAS	7200
data	4b.01.14	4b	1	14	SA:A	-	BSAS	7200
data	4b.01.15	4b	1	15	SA:A	-	BSAS	7200
data	4b.01.16	4b	1	16	SA:A	-	BSAS	7200
data	4b.01.17	4b	1	17	SA:A	-	BSAS	7200
data	4b.01.18	4b	1	18	SA:A	-	BSAS	7200
data	4b.01.19	4b	1	19	SA:A	-	BSAS	7200
data	4b.01.20	4b	1	20	SA:A	-	BSAS	7200
data	4b.01.21	4b	1	21	SA:A	-	BSAS	7200
data	4b.01.22	4b	1	22	SA:A	-	BSAS	7200
data	4b.01.23	4b	1	23	SA:A	-	BSAS	7200

RAID group /aggr0/plex0/rg1 (normal)

RAID	Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----		-----			-----	-----	-----	-----
dparity	4a.02.4		4a	2	4	SA:B	-	BSAS	7200
parity	4a.02.5		4a	2	5	SA:B	-	BSAS	7200
data	4a.02.6		4a	2	6	SA:B	-	BSAS	7200
data	4a.02.7		4a	2	7	SA:B	-	BSAS	7200
data	4a.02.8		4a	2	8	SA:B	-	BSAS	7200
data	4a.02.9		4a	2	9	SA:B	-	BSAS	7200
data	4a.02.10		4a	2	10	SA:B	-	BSAS	7200
data	4a.02.11		4a	2	11	SA:B	-	BSAS	7200
data	4a.02.12		4a	2	12	SA:B	-	BSAS	7200

data	4a.02.13	4a	2	13	SA:B	-	BSAS	7200
data	4a.02.14	4a	2	14	SA:B	-	BSAS	7200
data	4a.02.15	4a	2	15	SA:B	-	BSAS	7200
data	4a.02.16	4a	2	16	SA:B	-	BSAS	7200
data	4a.02.17	4a	2	17	SA:B	-	BSAS	7200
data	4a.02.18	4a	2	18	SA:B	-	BSAS	7200
data	4a.02.19	4a	2	19	SA:B	-	BSAS	7200
data	4a.02.20	4a	2	20	SA:B	-	BSAS	7200
data	4a.02.21	4a	2	21	SA:B	-	BSAS	7200
data	4a.02.22	4a	2	22	SA:B	-	BSAS	7200
data	4a.02.23	4a	2	23	SA:B	-	BSAS	7200
Spare disks								
RAID Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----	-----	-----	-----	-----	-----	-----	-----
Spare disks for block or zoned checksum traditional volumes or aggregates								
spare	4a.02.0	4a	2	0	SA:B	-	BSAS	7200
spare	4a.02.1	4a	2	1	SA:B	-	BSAS	7200
spare	4a.02.2	4a	2	2	SA:B	-	BSAS	7200
spare	4a.02.3	4a	2	3	SA:B	-	BSAS	7200
spare	4b.01.0	4b	1	0	SA:A	-	BSAS	7200
spare	4b.01.1	4b	1	1	SA:A	-	BSAS	7200
spare	4b.01.2	4b	1	2	SA:A	-	BSAS	7200
spare	4b.01.3	4b	1	3	SA:A	-	BSAS	7200

4.1.2 Konfigurace oddílů SAS disků

SAS disky jsou spojovány do tří RAID skupin po 16 fyzických discích. Toto rozmístění je doporučeno pro dosažení optimálního volného prostoru. Pouze dva disky zůstávají jako spare (náhradní) disky.

Konfigurace SAS disků:

```
Aggregate aggr0 (online, raid_dp) (block checksums)
```

```
Plex /aggr0/plex0 (online, normal, active)
```

```
RAID group /aggr0/plex0/rg0 (normal)
```

RAID	Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----		-----			-----	-----	-----	-----
dparity	0b.11.0		0b	11	0	SA:B	-	SAS	15000
parity	0b.11.1		0b	11	1	SA:B	-	SAS	15000
data	0b.11.2		0b	11	2	SA:B	-	SAS	15000
data	0b.11.3		0b	11	3	SA:B	-	SAS	15000
data	0b.11.4		0b	11	4	SA:B	-	SAS	15000
data	0b.11.5		0b	11	5	SA:B	-	SAS	15000
data	0b.11.6		0b	11	6	SA:B	-	SAS	15000
data	0b.11.7		0b	11	7	SA:B	-	SAS	15000
data	0b.11.8		0b	11	8	SA:B	-	SAS	15000
data	0b.11.9		0b	11	9	SA:B	-	SAS	15000
data	0b.11.10		0b	11	10	SA:B	-	SAS	15000
data	0b.11.11		0b	11	11	SA:B	-	SAS	15000
data	0b.11.12		0b	11	12	SA:B	-	SAS	15000
data	0b.11.13		0b	11	13	SA:B	-	SAS	15000
data	0b.11.14		0b	11	14	SA:B	-	SAS	15000
data	0b.11.15		0b	11	15	SA:B	-	SAS	15000

RAID group /aggr0/plex0/rg1 (normal)

RAID	Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----		-----			-----	-----	-----	-----
dparity	0b.11.16		0b	11	16	SA:B	-	SAS	15000
parity	0b.11.17		0b	11	17	SA:B	-	SAS	15000
data	0b.11.18		0b	11	18	SA:B	-	SAS	15000
data	0b.11.19		0b	11	19	SA:B	-	SAS	15000
data	0b.11.20		0b	11	20	SA:B	-	SAS	15000
data	0b.11.21		0b	11	21	SA:B	-	SAS	15000
data	0b.11.22		0b	11	22	SA:B	-	SAS	15000
data	0b.11.23		0b	11	23	SA:B	-	SAS	15000
data	4c.12.0		4c	12	0	SA:A	-	SAS	15000
data	4c.12.1		4c	12	1	SA:A	-	SAS	15000
data	4c.12.2		4c	12	2	SA:A	-	SAS	15000
data	4c.12.3		4c	12	3	SA:A	-	SAS	15000
data	4c.12.4		4c	12	4	SA:A	-	SAS	15000
data	4c.12.5		4c	12	5	SA:A	-	SAS	15000
data	4c.12.6		4c	12	6	SA:A	-	SAS	15000
data	4c.12.7		4c	12	7	SA:A	-	SAS	15000

RAID group /aggr0/plex0/rg2 (normal)

RAID	Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----		-----			-----	-----	-----	-----
dparity	4c.12.8		4c	12	8	SA:A	-	SAS	15000
parity	4c.12.9		4c	12	9	SA:A	-	SAS	15000

data	4c.12.10	4c	12	10	SA:A	-	SAS	15000
data	4c.12.11	4c	12	11	SA:A	-	SAS	15000
data	4c.12.12	4c	12	12	SA:A	-	SAS	15000
data	4c.12.13	4c	12	13	SA:A	-	SAS	15000
data	4c.12.14	4c	12	14	SA:A	-	SAS	15000
data	4c.12.15	4c	12	15	SA:A	-	SAS	15000
data	4c.12.16	4c	12	16	SA:A	-	SAS	15000
data	4c.12.17	4c	12	17	SA:A	-	SAS	15000
data	4c.12.18	4c	12	18	SA:A	-	SAS	15000
data	4c.12.19	4c	12	19	SA:A	-	SAS	15000
data	4c.12.20	4c	12	20	SA:A	-	SAS	15000
data	4c.12.21	4c	12	21	SA:A	-	SAS	15000
Spare disks								
RAID Disk	Device	HA	SHELF	BAY	CHAN	Pool	Type	RPM
-----	-----	-----	-----	-----	-----	-----	-----	-----
Spare disks for block or zoned checksum traditional volumes or aggregates								
spare	4c.12.22	4c	12	22	SA:A	-	SAS	15000
spare	4c.12.23	4c	12	23	SA:A	-	SAS	15000

4.2 Migrace pomocí CIFS

Migrace pomocí CIFS se týká U-drive a G-drive, tedy síťovým diskům na kterých jsou uložena sdílená data. Nejprve se vytvořily oddíly na datovém úložišti, poté se vytvořily sdílené disky, kterým se určily kvóty a překopírovaly se data pomocí softwaru FastCopy ze starých sdílených disků. Následně se musel změnit DNS záznam ukazující na sdílené disky.

Vytvoření oddílu pro sdílení pomocí CIFS:

```
# creation of volume

vol create "volume" -s volume aggr0 "size"

mtree security /vol/"volume" ntfs

cifs shares -add "share-name"

vol options "volume" nosnap on

vol options "volume" fractional_reserve 0

vol options "volume" no_atime_update on

vol options "volume" try_first volume_grow

# volume increasing method

vol autosize "volume" on

vol autosize "volume" -m "size"

# autodelete policy

snap reserve "volume" 0

snap autodelete "volume" on

snap autodelete "volume" trigger volume

snap autodelete "volume" target_free_space 5

snap autodelete "volume" commitment try

snap autodelete "volume" defer_delete none

snap autodelete "volume" delete_order oldest_first

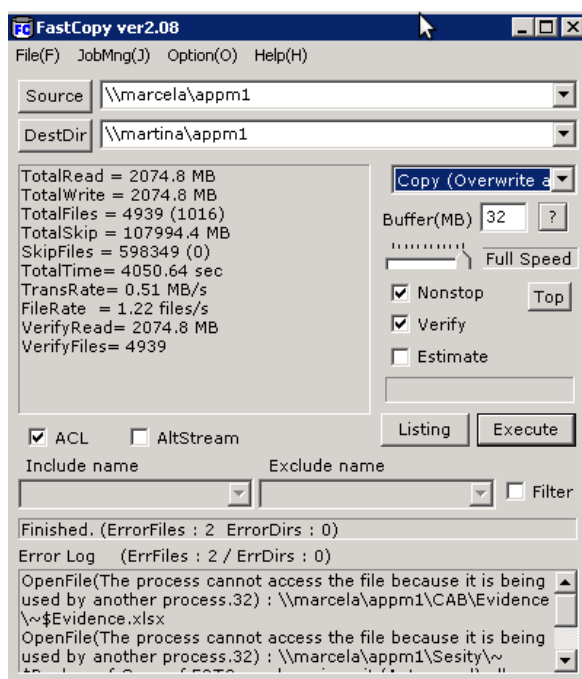
# creation of LUN
```

```
lun create -s "size" -t "type" /vol/"volume"/"lun".lun

# mapping of LUN

Lun map /vol/"volume"/"lun".lun "server" 1
```

Pomocí softwaru FastCopy se podařilo zkopírovat sdílená data společně i s jejími přístupovými právy.



Obrázek 16 CIFS - FastCopy

4.3 Migrace SnapVault

U migrace dat pomocí technologie SnapVault bylo nejprve nutné nakonfigurovat primární storage (ten na který se budou dělat zálohy) poté se nastavil sekundární storage a vzájemný vztah mezi nimi. A nastavení pravidelného zálohování přesunu snapshotu z primárního na sekundární storage.

Zastavení zrcadlení:

```
rzfs201> license delete cf_remote

unlicensed cf_remote.
```

```
Fri Sep 9 10:41:08 CEST [rzfs201: telnet_0:notice]: cf_remote
unlicensed

Fri Sep 9 10:41:08 CEST [rzfs201: wafl.groupCP.disable:info]:
WAFL Group-CP has been disabled.

cf_remote disabled.

rzfs201> license delete syncmirror_local

unlicensed syncmirror_local.

Local synchronous mirror support is no longer licensed.
You must reboot your system to disable this feature.
Any existing mirror volumes will not be usable.
Local synchronous mirror will be disabled after reboot.
```

Ověření SnapVault:

```
rzfs201> options snapvault

snapvault.access          none
snapvault.enable          off
snapvault.lockvault_log_volume
snapvault.nbu.archival_snap_default on
snapvault.ossv.compression off
snapvault.preservesnap     off
snapvault.snapshot_for_dr_backup vsm_base_only
```

Nastavení primárního úložiště a spuštění funkce SnapVault: viz příloha P1

5 HODNOCENÍ PRŮBĚHŮ MIGRACÍ

Provedené migrace proběhly bez sebemenšího problému, nejdůležitější bylo si připravit a nakonfigurovat datové úložiště a jeho oddíly. Nastavit jim kvóty a veškeré politiky jak byly na původních úložištích. Samotná migrace dat musela být situována do období výrobní odstávky, kdy bylo možné krátkým výpadkem přednastavit DNS záznamy.

Důležité také bylo pečlivě nastavit funkci SnapVault na NetAppu FAS270 v Piešťanech.

ZÁVĚR

Cílem této bakalářské práce bylo popsat situaci v oblasti datových úložišť a v technologiích využívaných v této oblasti. Dále v zanalyzování stávajícího datového úložiště firmy ON Semiconductor Czech Republic, s.r.o. a výběrem nového datového úložiště. Nové datové úložiště od společnosti NetApp, které je zapojeno do stávající infrastruktury, bylo úspěšně zprovozněno. Data byla zkonsolidována z ostatních úložišť právě na nový NetApp FAS 3420.

Tento projekt byl přínosem pro celou firmu, zejména zvětšení kapacity diskových úložišť především pro blokové prostory pro operační systémy Windows, Linux a VMware a souborové prostory pro databáze ORACLE propagované protokolem NFS.

Zároveň bylo toto datové úložiště využito pro vzdálené zálohování dat z diskového pole NetApp ze sesterské společnosti v Piešťanech na Slovensku pomocí technologie SnapVault.

Hlavní přínos je tato bakalářská práce samozřejmě pro mě samotnému. Seznámil jsem se s budováním SAN architektury datových úložišť a jejich fyzickém zapojení. Implementaci, ladění a provozování datových úložišť a získání praktických zkušeností v tomto odvětví.

ZÁVĚR V ANGLIČTINĚ

The goal of this thesis is to describe the current situation and the technologies used in the field of data storage as well as the analysis of the (now already) previous data storage solution of the ON Semiconductor Czech Republic, s. r. o. company followed by a selection of a new solution. The discussed new NetApp data storage solution has already been successfully deployed and integrated into the existing infrastructure of the company. All the data were consolidated from the previous data storage units into the new NetApp FAS 3420 system.

This project was beneficial for the whole company. Especially the block storage capacity increase (for the Windows and Linux operating systems and the VMWare virtualization images) and the capacity increase of the file systems used for Oracle databases exported through NFS (network file system) shares were perceived as particularly beneficial.

At the same time, the new data storage system is used as a remote data backup storage space for a NetApp system of a company subsidiary based in Piestany, Slovakia. This remote backup has been implemented using SnapVault technology.

The experience that preceded this thesis was very beneficial to me since I had the opportunity to be closely involved in the design

SEZNAM POUŽITÉ LITERATURY

- [1] Datová úložiště. *Datová úložiště* [online]. 2012 [cit. 2012-09-03]. Dostupné z: <http://www.storyflex.cz/128-datova-uloziste.html>
- [2] Ukládání dat - přímo připojená úložiště. *Ukládání dat DAS* [online]. 2012 [cit. 2012-09-03]. Dostupné z: <http://www.vahal.cz/cz/podpora/technicke-okenko/ukladani-dat-das.html>
- [3] STANĚK, Pavel. *Datová úložiště a nabídka jejich outsourcingových služeb*. Vysoká škola ekonomická v Praze, 2006. Bakalářská práce. Vysoká škola ekonomická v Praze.
- [4] SOMASUNDARAM, G a Alok SHRIVASTAVA. *Information storage and management: storing, managing, and protecting digital information*. Indianapolis, Ind.: Wiley Publishing, 2009, xx, 455 p. ISBN 978-0-470-29421-5.
- [5] Computer Storage - architektury, protokoly, rozhraní. *Disková rozhraní* [online]. 2007 [cit. 2012-09-03]. Dostupné z: <http://www.samuraj-cz.com/clanek/computer-storage-architektury-protokoly-rozhranni/>
- [6] Paralelní aplikace a InfiniBand. *InfiniBand* [online]. 2011 [cit. 2012-09-03]. Dostupné z: http://meta.cesnet.cz/wiki/Paraleln%C3%AD_aplikace_a_InfiniBand
- [7] NDMP. *NDMP* [online]. 2010 [cit. 2012-09-03]. Dostupné z: <http://searchstorage.techtarget.com/definition/NDMP>
- [8] Technologie diskových polí. *Technologie diskových polí* [online]. 2010 [cit. 2012-09-03]. Dostupné z: <http://www.vahal.cz/cz/podpora/technicke-okenko/diskova-pole.html>
- [9] Snapshoty - způsob jak posunout zálohování a obnovu dále. *Snapshoty* [online]. 2010 [cit. 2012-09-03]. Dostupné z: <http://www.netguru.cz/odborne-clanky/snapshots-zpsob-jak-posunout-zalohovani-a-obnovu-dale.html>
- [10] NetApp – úložná řešení pro podniky. *NetApp* [online]. 2010 [cit. 2012-09-03]. Dostupné z: <http://www.abclinuxu.cz/clanky/hardware/netapp-ulozna-reseni-pro-podniky>
- [11] Datová úložiště. *Datová úložiště* [online]. 2012 [cit. 2012-09-03]. Dostupné z: <http://www.xanadu.cz/cs/it-produkty/serverova-reseni/datova-uloziste/>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

Agregate	Kombinace RAID skupin.
Volume	Logický oddíl
Node	Server, který je částí clusteru - uzel.
WAFL	správce oddílů a souborového systému v zařízeních FAS (filerech)
FAS	Fabric-Attached Storage
SCSI	Small Computer System Interface
iSCSI	Internet Small Computer System Protocol
FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
SSH	Secure Shell
MB	Megabyte
GB	Gigabyte
TB	Terabyte
PB	Petabyte
SQL	Structured Query Language
NAS	Network-attached storage
DAS	Direct-attached storage
SAN	storage area network
NDMP	Network Data Management Protocol
HBA	host bus adapter
EMEA	Europe, the Middle East and Africa
RAID	redundant array of independent disks
NFS	Network File System
CIFS	Common Internet File System

SMB Server Message Block

FlexVol Flexible volume

DNS Domain Name System

SAS Serial Attached SCSI

SEZNAM OBRÁZKŮ

<i>Obrázek 1 Řešení typu DAS [4]</i>	<i>12</i>
<i>Obrázek 2 Řešení typu NAS [4]</i>	<i>13</i>
<i>Obrázek 3 Řešení typu SAN [4]</i>	<i>14</i>
<i>Obrázek 4 Správa fileru pomocí SSH</i>	<i>23</i>
<i>Obrázek 5 Správa fileru pomocí webového rozhraní</i>	<i>24</i>
<i>Obrázek 6 FileOver [4]</i>	<i>25</i>
<i>Obrázek 7 Cluster Active/Active [4]</i>	<i>25</i>
<i>Obrázek 8 IBM FastT200</i>	<i>29</i>
<i>Obrázek 9 EMC CX500</i>	<i>29</i>
<i>Obrázek 10 EMCCX3-20</i>	<i>30</i>
<i>Obrázek 11 NetApp FAS270</i>	<i>30</i>
<i>Obrázek 12 NetApp FAS2050</i>	<i>30</i>
<i>Obrázek 13 Zapojení NetApp FAS2050 a EMC CX3-20</i>	<i>31</i>
<i>Obrázek 14 Schéma zapojené nového datového úložiště</i>	<i>33</i>
<i>Obrázek 15 NetApp FAS 3240EAX</i>	<i>33</i>
<i>Obrázek 16 CIFS - FastCopy</i>	<i>44</i>

SEZNAM TABULEK

<i>Tabulka 1 Porovnání DAS, NAS, SAN</i>	<i>15</i>
<i>Tabulka 2 Kapacitní nároky na nové datové úložiště</i>	<i>32</i>
<i>Tabulka 3 Seznam komponentů dodaných firmou Proact</i>	<i>34</i>

SEZNAM PŘÍLOH

Příloha P I: CD Obsahující kód nastavení primárního úložiště a spuštění funkce SnapvVault