VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA PODNIKATELSKÁ
FACULTY OF BUSINESS AND MANAGEMENT

ÚSTAV INFORMATIKY
INSTITUTE OF INFORMATICS

ZÁLOHOVÁNÍ DAT A DATOVÁ ÚLOŽIŠTĚ

DATA BACKUP AND DATA STORAGES

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE — Marian Paytok
AUTHOR

VEDOUCÍ PRÁCE — Ing. Jiří Kříž, Ph.D.
SUPERVISOR

BRNO 2016

ZADÁNÍ BAKALÁRSKÉ PRÁCE

Paytok Marian

Manažerská informatika (6209R021)

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách, Studijním a zkušebním řádem VUT v Brně a Směrnicí děkana pro realizaci bakalářských a magisterských studijních programu zadává bakalářskou práci s názvem:

Zálohování dat a datová úložiště

v anglickém jazyce:

Data backup and data storages

Pokyny pro vypracování:

Úvod
Cíle práce, metody a postupy zpracování
Teoretická východiska práce
Analýza současného stavu
Vlastní návrhy řešení
Závěr
Seznam použité literatury
Přílohy

Podle § 60 zákona c. 121/2000 Sb. (autorský zákon) v platném znění, je tato práce "Školním dílem". Využití této práce se řídí právním režimem autorského zákona. Citace povoluje Fakulta podnikatelská Vysokého učení technického v Brně.

Seznam odborné literatury:

GÁLA, L.; POUR, J.; TOMAN, P. Podniková informatika. Praha : Grada Publishing, a.s., 2006. 484 s. ISBN 80-247-1278-4.

POŽÁR, Josef. Manažerská informatika. Plzeň: Aleš Čeněk, 2010, 357 s. ISBN 978-80-7380-276-9.

SOSINSKY, Barrie. Mistrovství – počítačové sítě. 1. vyd. Brno: Computer Press, 2010, 840 s. ISBN 978-80-251-3363-7.

PECINOVSKÝ, P. Archivace a komprimace dat. Praha: Grada Publishing, 2003. 116 s. ISBN 80-247-0659-8.

Vedoucí bakalářské práce: Ing. Jiří Kříž, Ph.D.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2015/2016.

L.S.

doc. RNDr. Bedřich Půža, CSc.	doc. Ing. et Ing. Stanislav Škapa, Ph.D.
Ředitel ústavu	Děkan fakulty

V Brně, dne 29.2.2016

Abstrakt

Tato bakalářská práce se zaměřuje na zálohování, obnovu a ukládání dat a datová úložiště. Dále popisuje problematiku a způsob zálohování a obnovy dat v konkrétní firmě. Výsledkem by měl být návrh optimálního řešení problému zálohování a ukládání dat v této společnosti.

Abstract

This thesis focuses on backup, restore and data storage. It also describes the issues and method of data backup in a particular company. The result should be an optimal solution to the problem of data backup and data storage in this company.

Klíčová slova

zálohování, obnova, datová úložiště, data, deduplikace

Key words

backup, restore, data storage, data, deduplication

Bibliografická citace

PAYTOK, M. *Zálohování dat a datová úložiště.* Brno: Vysoké učení technické v Brně, Fakulta podnikatelská, 2016. 45s. Vedoucí bakalářské práce Ing. Jiří Kříž, Ph.D..

Čestné prohlášení

Prohlašuji, že předložená bakalářská práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem ve své práci neporušil autorská práva (ve smyslu zákona č. 121/2000Sb., o právu autorském a o právech souvisejících s právem autorským).

V Brně …….. podpis ……..

Poděkování

Rád bych poděkoval svému vedoucímu a všem, kteří mně v mé práci byli nápomocni.

OBSAH

ÚVOD 10

CÍL PRÁCE 11

1. Zálohování a obnova dat 12

1.1 Typy záloh 12

1.1.1 Plná záloha (full backup) 12

1.1.2 Inkrementální záloha (přírůstková) 12

1.1.3 Diferenciální záloha 12

1.2 Příčiny ztráty dat 13

1.3 Disková pole - RAID 14

1.3.1 RAID 0 – Striping 14

1.3.2 RAID 1 - Mirroring 15

1.3.3 RAID 5 – Striping with parity across drives 15

1.3.4 RAID 6 – Striping with dual parity across drives 16

1.3.5 RAID 10 17

1.4 Manipulace s daty 18

1.4.1 Komprese dat 18

1.4.2 Deduplikace dat 18

1.4.3 Replikace dat 20

1.5 Topologie připojení datového úložiště 20

1.5.1 DAS (Direct Attached Storrage) 20

1.5.2 NAS (Network Attached Storage) 21

1.5.3 SAN 22

2. ANALÝZA SOUČASNÉHO STAVU 23

2.1 Datacentrum 23

2.2 Rozdělení disků ... 24
2.3 Zálohovací systém ... 25
2.3.1 Zálohování souborového serveru ... 25
2.3.2 Zálohování Exchange serveru ... 25
2.4 EMC Avamar ... 27
2.5 Archivace ... 28
2.6 Zhodnocení současného stavu ... 29
3 VLASTNÍ NÁVRH ŘEŠENÍ ... 30
3.1 Hybridní diskové pole ... 30
3.2 EMC VNX 5200 ... 30
3.2.1 Zvolená konfigurace disků v datovém úložišti ... 32
3.3 Software ... 34
3.3.1 Base Software ... 34
3.3.2 Unisphere Management ... 34
3.3.3 Rozšiřující balíky ... 35
3.4 Hot and Cold data ... 36
3.5 Deduplikace diskového pole ... 37
3.6 Cenové shrnutí ... 38
3.6.1 Cenové shrnutí hardwaru ... 38
3.6.2 Cenové shrnutí softwaru ... 38
3.7 Zhodnocení vlastního návrhu ... 39
ZÁVĚR ... 40
SEZNAM POUŽITÉ LITERATURY ... 41
SEZNAM OBRÁZKŮ ... 44
SEZNAM TABULEK ... 45
SEZNAM GRAFŮ ... 45

ÚVOD

Objem ukládaných dat v domácnostech i firmách se neustále zvyšuje. Riziko ztráty dat si však připouští jen málokdo. Pokud však ztráta nebo poškození uložených dat potká společnost, ve které jsou veškeré dokumenty vedeny pouze v digitalizované podobě, mohou při nedůkladném zálohování, pro takovouto společnost nastat i existenční problémy.

Každá firma by měla mít svého správce firemní sítě, případně člověka, jehož ukládáním, zálohováním a ochranou dat pověří. Firmám se ale může zdát zbytečné zaměstnávat a platit za práci takovéhoto člověka a budou přehlížet rizika spjatá se ztrátou dat nebo ani s případnými riziky nemusí být obeznámeni.

CÍL PRÁCE

Bakalářská práce představí základní teoretická východiska používané pro bezpečné zálohování a skladování dat s ohledem na rozdílné potřeby firem, dále popisuje postupy při zálohování, technologie pro zálohování, typy záloh, používaný software a také rizika a problémy spjaté se zálohováním.

Cílem první praktické části je analýza současného stavu zálohovacího procesu v konkrétní firmě a pro druhou část je to vytvoření vlastního návrhu pro nové datové úložiště, které zajistí lepší provázanost se zálohovacím systémem, navýšení dostupné úložné kapacity a také větší výkon při provádění operací s daty v úložišti.

1. Zálohování a obnova dat

Zálohování dat se provádí proto, aby byla data ochráněna a mohla proběhnout jejich obnova v případě, že v jejich primárním umístění budou smazána, porušena nebo zničena. Zálohování by mělo probíhat pravidelně dle předem stanoveného plánu. Ve firemním prostředí například po konci každého pracovního dne (1,2).

1.1 Typy záloh

Zálohování dat může pobíhat v několika různých metodách, dle konkrétních potřeb osoby či subjektu.

1.1.1 Plná záloha (full backup)

Plná, tedy celková záloha obsahuje všechna data, která jsme určili pro zálohu. Je ideální pro obnovu dat, protože obsahuje všechny soubory a složky, které byly ztraceny v původním umístění (2). Její rozsáhlost je však i její nevýhodou, protože vytvoření tak velké zálohy zabere spoustu času. Proto se plné zálohy využívá spíše jen v týdenních či měsíčních cyklech a je pouštěna zejména přes noc (3).

1.1.2 Inkrementální záloha (přírůstková)

Inkrementální záloha začíná jako záloha plná, tedy vytvoří se kopie všech dat. Změna následuje v dalších zálohovacích cyklech, kdy dochází k záloze pouze těch dat, která se změnila v porovnání s předchozí zálohou. Výhodou je rychlejší záloha (ponechává nezměněná data) a totéž platí pro obnovu takto zálohovaných dat. Pokud by však došlo ke ztrátě dat v jakémkoliv přírůstku nelze již data obnovit z přírůstků následujících (2).

1.1.3 Diferenciální záloha

Je záloha, která obsahuje všechny soubory, které prošli jakoukoliv změnou od poslední úplné zálohy. Rozdíl mezi inkrementální zálohou je, že přírůstková ukládá nejdříve změnu oproti plné záloze a následně jen oproti přírůstkům a diferenciální ukládá všechny změny od poslední plné zálohy. Proto je její velikost násobně větší (2,3).

1.2 Příčiny ztráty dat

Ztráta dat dokáže negativně ovlivnit fungování celé firmy a podnikatelských aktivit, na které se soustřeďuje.

Nejhorší variantou je ztráta dat z primárního úložiště a neexistence žádné zálohy, případně existence pouze zálohy, která neobsahuje aktuální potřebná data z důvodu špatného nebo neexistujícího plánu zálohování.

Samotné příčiny, které vedou k selhání nosiče dat, zejména pevných disků, lze rozdělit na 2 skupiny (26).

a) Softwarové

Do této skupiny spadají závady způsobené neočekávaným chováním aplikace. Může se jednat o:

- Chybu či pád operačního systému
- Chybu či pád spuštěné aplikace
- Infikování dat virem, jehož cílem je znehodnocení dat (např.: zašifrováním) či destrukce disku
- Selhání lidského faktoru (úmyslné či neúmyslné)

b) Fyzikální

V této skupině se vyskytují typy závad, které mají přímý vliv na hardware datového úložiště.

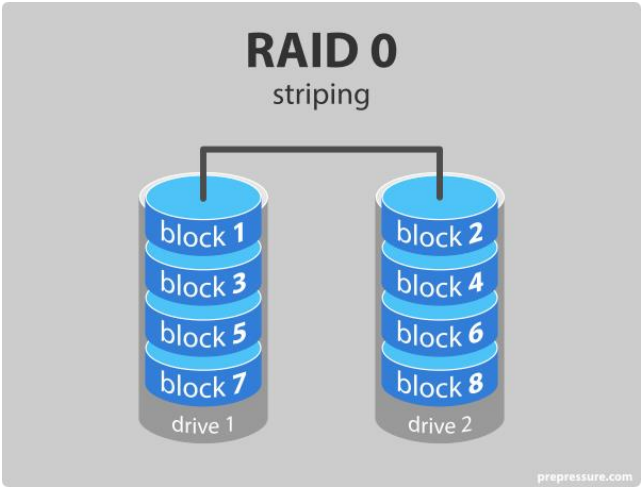
- Výrobní závada
- Překročení omezené životnosti
- Výpadek či výkyvy elektrického proudu
- Vniknutí tekutiny
- Přehřátí nebo dlouhodobé vystavení vysoké teplotě
- Poškození při nárazu způsobeného pádem či otřesy
- Poškození vlivem přírodní katastrofy

1.3 Disková pole - RAID

Jedná se o metodu zvýšení bezpečnosti za pomocí skupiny pevných disků, které jsou spojeny pomocí RAID řadiče a logicky vystupují jako jeden disk. Jedná se spíše o datové úložiště než o způsob zálohy dat, i když to je částečně také pravda. Disková pole ukládají stejná data nebo paritní informace na více disků, aby mohlo dojít k jejich obnově, pokud by ale došlo k poškození dat vlivem zásahu blesku nebo požáru, s velkou pravděpodobností budou zničena i data v diskovém poli. Jednou z možností ochrany proti ztrátě dat při mechanickém poškození úložiště je replikace dat mezi totožnými úložišti, umístěnými na dvou či více odlišných lokacích (4,6).

1.3.1 RAID 0 – Striping

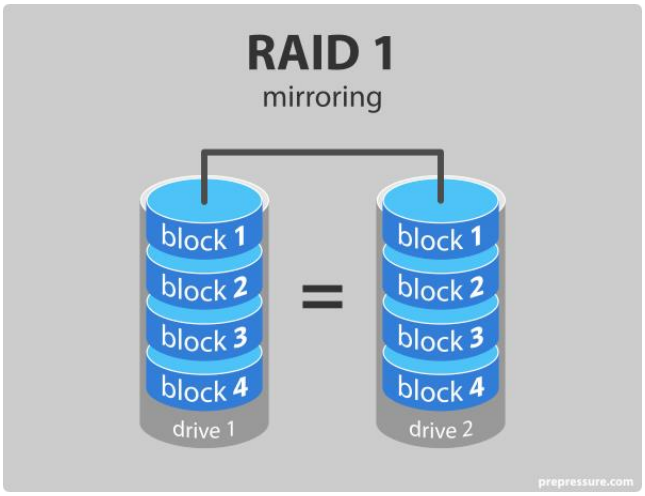
Metoda stripování zapisovaných dat mezi několik (2 a více) pevných disků nezajišťuje ochranu před ztrátou dat v případě selhání některého z disků v poli, podstatně však zvyšuje rychlost zápisu i čtení dat v takovémto úložišti. Data jsou při zápisu rozdělena na bloky, které jsou současně zapisovány na disky v úložišti, tím je dosaženo využití maximální rychlosti všech disků a výsledná rychlost je součtem těchto rychlostí. Obvyklá velikost bloků, na kterou jsou data dělena, jsou 4 – 128 KB. Hlavní nevýhodou, opomenu-li redundanci dat, kterou tato metoda nezajišťuje, je plýtvání volným místem. Pokud je zvoleno dělení na bloky o velikosti 128 KB, které bude zajišťovat dosahování nejvyšších rychlostí, a zapisován bude soubor o velikosti 100 KB, bude na obou discích alokováno 256 KB, z čehož 56 KB bude nevyužitý volný prostor (16).



Obrázek č. 1: RAID 0 (Zdroj: 17)

1.3.2 RAID 1 - Mirroring

Při této metodě je použit stejný počet primárních disků, jako disků, na které jsou data zrcadlena. Při zápisu dat na primární disk, jsou identická data zapisována i na druhý disk v poli. Výsledkem je jeden či více párů pevných disků se shodným obsahem, kdy při porušení dat v jednom z disků mohou být data přečtena z jeho zrcadlové kopie (8).

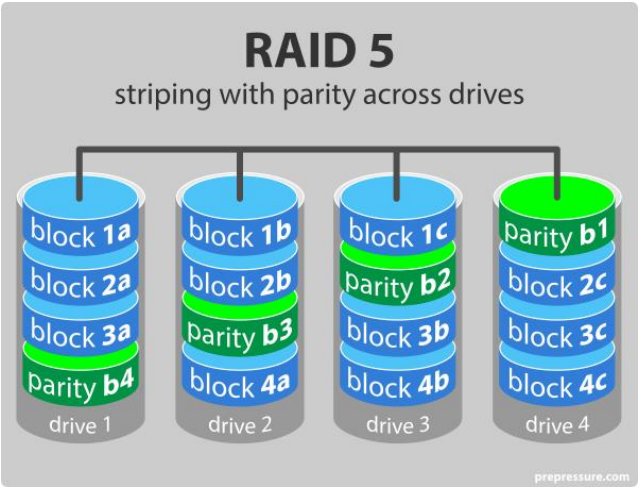


Obrázek č. 2: RAID 1 (Zdroj: 17)

1.3.3 RAID 5 – Striping with parity across drives

Při této metodě redundance, jsou data rozdělena do bloků, které jsou střídavě zapisovány mezi disky v poli. Na závěr se ze zapsaných dat vypočítá paritní kontrolní součet a ten je poté uložen na jeden z disků v poli. Při zápisu dalšího souboru je pro uložení parity zvolen jiný disk. Paritní kontrolní součty celkově odpovídají velikosti kapacity jednoho disku. Minimální počet disků v poli RAID 5 jsou tři. Efektivní využitelná kapacita bude vždy odpovídat n-1 diskům. Při selhání jednoho libovolného disku v tomto poli, proběhne po jeho výměně dopočítání dat, která byla uložena na disku původním.

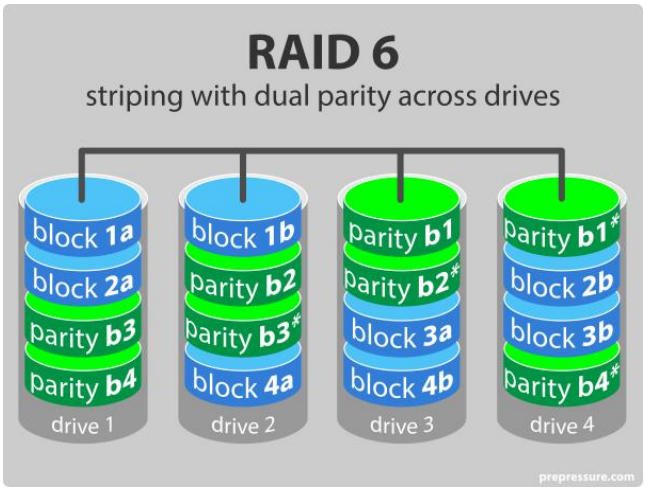
Nevýhodou této metody je pomalejší zápis, kvůli vytváření kontrolní parity zapisovaných dat a také dlouhá doba obnovy dat na nový disk v poli. Při obnovování dat také dochází ke snížení výkonnosti celého úložiště, z důvodu využití všech disků k dopočítání ztracených dat (17).



Obrázek č. 3: RAID 5 (Zdroj: 17)

1.3.4 RAID 6 – Striping with dual parity across drives

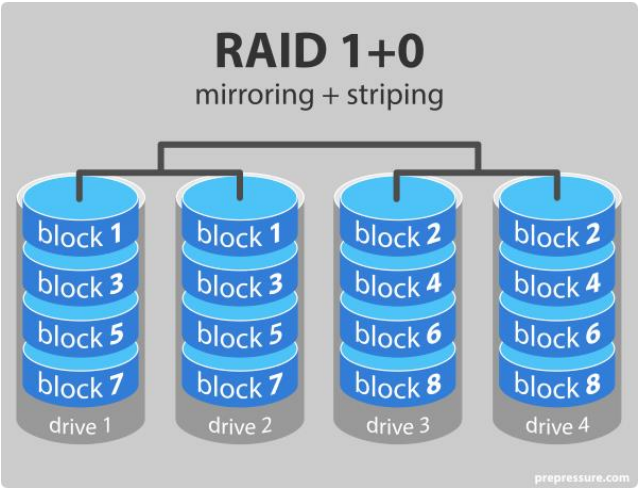
Princip této metody je podobný metodě RAID 5, zde je však paritní součet ze zapsaných dat uložen na dva různé disky v poli. Hlavní výhodou je zvýšená ochrana proti selhání až dvou disků v poli najednou. Tímto dochází ke zvýšení minimálního počtu disků na 4, s tím že kapacitu dvou disků bude zabírat právě parita. Efektivní využitelná kapacita tedy odpovídá n-2 diskům (17).



Obrázek č. 4: RAID 6 (Zdroj: 17)

1.3.5 RAID 10

Tato metoda redundance dat je kombinací zrcadlení u RAID 1 a stripování u RAID 0. Výhodou je vysoký výkon takového úložiště především při zápisu dat a obnově dat při selhání disku na nový. Nevýhodou jsou náklady na pořízení úložiště, protože kapacitu pro zápis poskytuje jen polovina disků v poli (17).



Obrázek č. 5: RAID 10 (Zdroj: 17)

1.4 Manipulace s daty

Následující kapitola popisuje nejčastější operace s daty ve formě digitální informace při jejich zálohování a archivaci.

1.4.1 Komprese dat

Jedná se o proces, jehož hlavním cílem, je snížit potřebu zdrojů pro uložení informace. Obecně se jedná o zmenšení velikosti datových souborů, což je vhodné při jejich archivaci, přenosu po sítí s omezenou datovou propustností nebo datovými limity.

Výhodou komprimace při archivaci, je možnost uložení celé datové struktury (souborů, složek a podsložek), která bude logicky vystupovat jako jeden archiv (30).

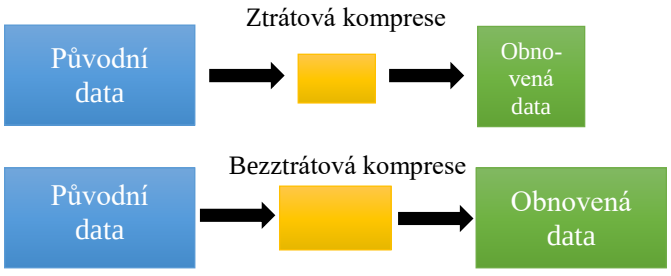
Dva základní typy komprese jsou:

a) Komprese ztrátová

Tento způsob komprese dat je použit v případě, kdy je možné tolerovat nenávratnou ztrátu části archivované informace. Například při kompresi fotografií dochází k úspoře úložné kapacity (dle nastavení kompresního poměru) vynecháním informací, které jsou mimo rozlišovací schopnosti lidského oka.

b) Komprese bezztrátová

Při této metodě nedochází při komprimaci a dekomprimaci ke ztrátě informací a zkreslení původních dat. Po dekomprimaci jsou data identická s původními.



Obrázek č. 6: Komprese dat (Zdroj: Vlastní zpracování)

1.4.2 Deduplikace dat

Deduplikace je nástroj, který pomocí algoritmů vyhledává ve zdrojovém nebo cílovém úložišti duplicitní soubory. Deduplikací dochází k uchování pouze jedné kopie unikátního zálohovaného souboru napříč celým úložištěm.

Pokud má více uživatelů na svém úložišti stejný soubor, je zálohován jen od prvního uživatele a při dalším jeho výskytu je uložen pouze odkaz na soubor v zálohovacím úložišti.

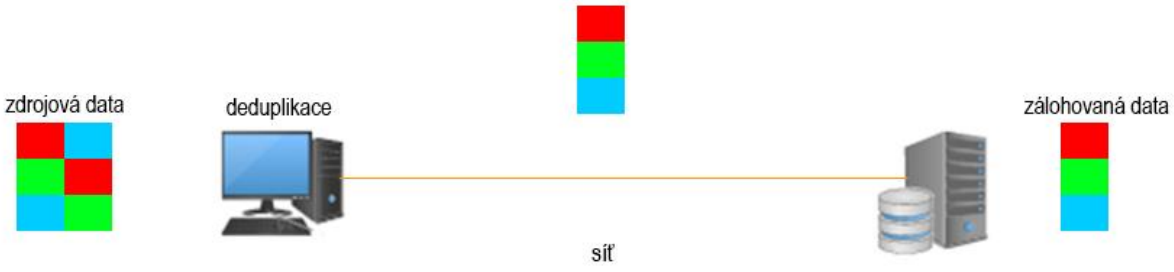
Při úpravě zdrojového souboru, se jeho změna uloží jako přírůstek k původnímu souboru (deduplikace na úrovni částí souboru). Tímto je možné obnovit původní soubor do podoby z libovolné zálohy ze zálohovacího rámce.

Deduplikaci je možno rozdělit na dva elementární typy:

a) Deduplikace na zdroji

   Při deduplikaci na zdrojovém úložišti je po inicializaci zálohování ze strany serveru provedena deduplikace pomocí klientské aplikace na koncovém zařízení. Do zálohovacího úložiště jsou přenášeny pouze unikátní bloky dat.

   Hlavní výhodou je významná redukce dat přenášených po LAN/WAN od klienta do zálohovacího úložiště. Zálohování tímto typem deduplikace může probíhat i několikrát během dne, protože přenos dat nebude vytěžovat síťovou infrastrukturu a tím zpomalovat ostatní komunikaci.

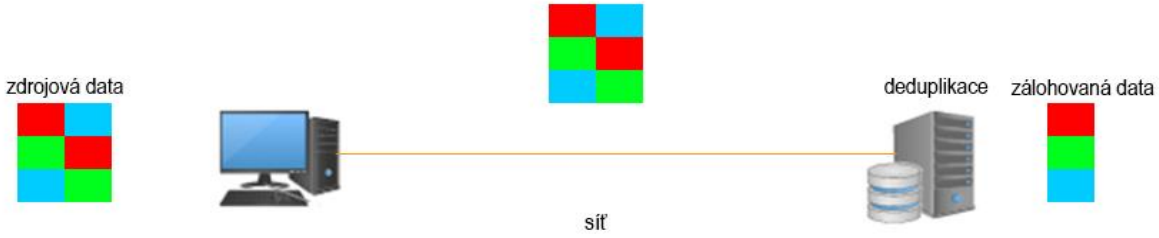


Obrázek č. 7: Deduplikace na zdroji (Zdroj: Vlastní zpracování)

b) Deduplikace na cíli

   Tato metoda byla prvním typem deduplikace, která se dostala na trh zálohování. Při inicializaci zálohování jsou všechna data ze zdrojového úložiště nejprve přenesena na zálohovací server, kde až následně probíhá kontrola dat na duplicity (19).

Nevýhodou je zatěžování sítě datovým přenosem z klienta na server a také mnohem vyšší nároky na kapacitu úložiště zálohovacího serveru.



Obrázek č. 8: Deduplikace na cíli (Zdroj: Vlastní zpracování)

1.4.3 Replikace dat

Replikace slouží k přenášení a ukládání identických dat ze zdrojového úložiště na jedno či více uložišť cílových. Cílová úložiště se obvykle nachází na geograficky odlišné lokaci. Tím jsou data chráněna v případě zničení primárního úložiště požárem či zemětřesením. Slouží však také jako záloha dat při selhání více disků, než kolik jsme schopni obnovit z redundantního pole.

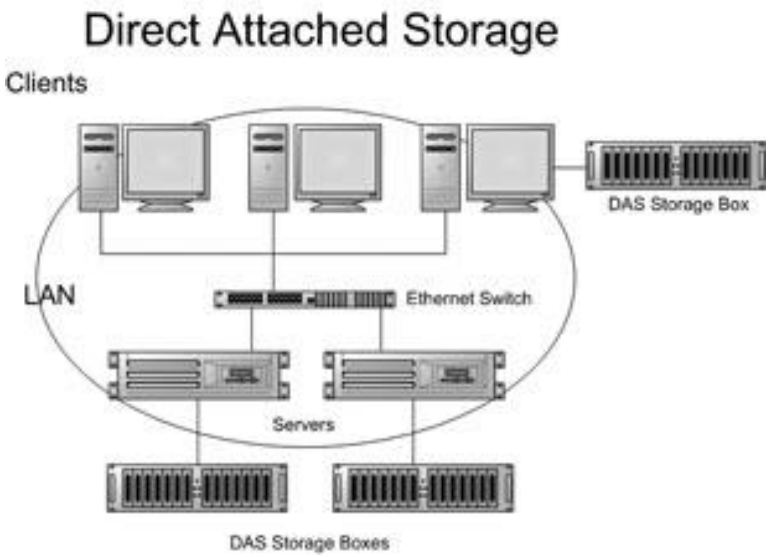
1.5 Topologie připojení datového úložiště

V této kapitole jsou teoreticky popsány jednotlivé architektury pro připojení datových úložišť. Obecně je lze členit na úložiště pro mále (DAS), střední (NAS) a velké (SAN) firmy.

1.5.1 DAS (Direct Attached Storrage)

V této architektuře je datové úložiště přímo připojeno ke koncovému zařízení (PC nebo server) pomocí host bus adapteru. Výhodami přímého připojení jsou dosažení vysoké rychlosti přenosu dat, jednoduchá implementace bez podrobného plánování a nízké požadavky na výkon síťové infrastruktury. Nevýhodou je dostupnost pouze ze zařízení, ke kterému je úložiště připojeno.

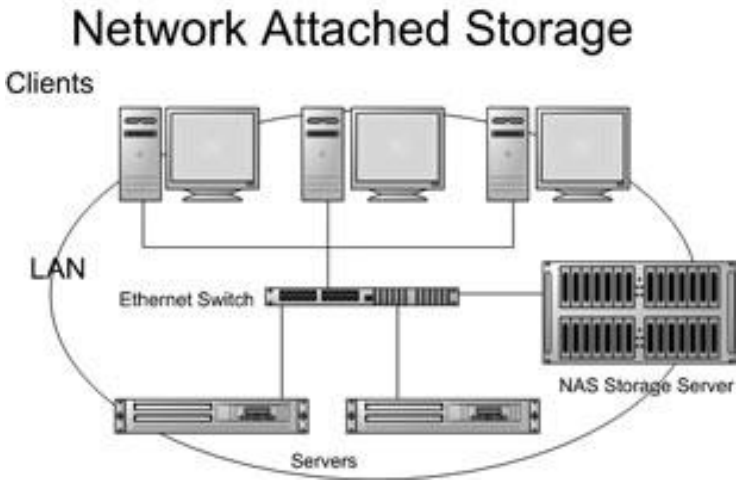
DAS nemá žádný síťový interface ani prostředky ke komunikaci v síti, může však být sdíleno zařízením, ke kterému je toto úložiště připojeno. DAS je vhodné pouze pro využití v malých firmách, kde není požadavek na neustálý přístup uloženým informacím.



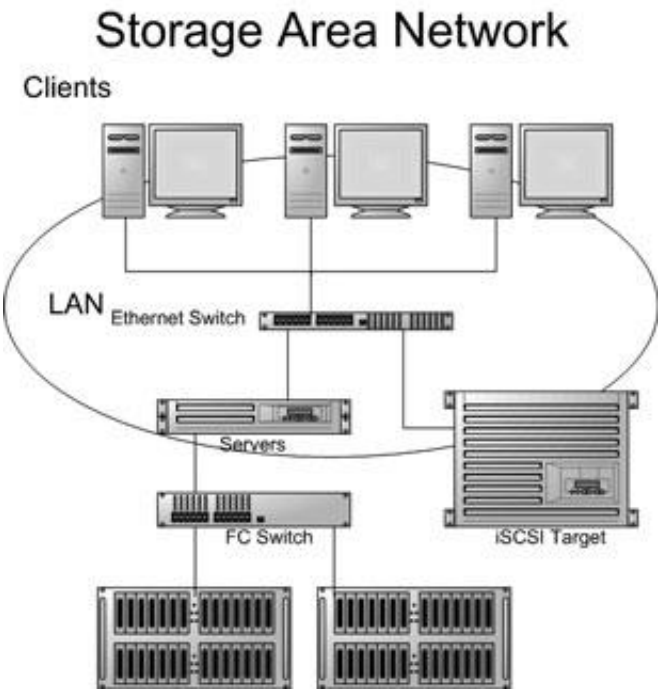
Obrázek č. 9: DAS (Zdroj: 21)

1.5.2 NAS (Network Attached Storage)

NAS je architektura, která využívá serveru jako sdíleného síťového úložiště, který je pomocí switche připojen k existující síti, a ke kterému mohou přistupovat všichni klienti a servery z vnitřní LAN případně z vnější WAN sítě bez ohledu na jejich architekturu nebo operační systém. Základem je jeden či více disků, většinou spojených do RAID pole. K datovému přenosu jsou využívány především protokoly NFS, FTP a AFP.



Obrázek č. 10: NAS (Zdroj: 21)



Obrázek č. 11: SAN (Zdroj: 21)

1.5.3 SAN

Architektura SAN je síť sloužící výhradně k datovým přenosům. Vzniká spojením datových úložišť se servery pomocí Fiber Channel switchů, připojených pomocí optických kabelů, čímž vytváří sdílené diskové úložiště (storage pool). SAN umožňuje přesouvání dat mezi jednotlivými úložišti v síti, sdílení dat mezi servery a slouží jako přenosové médium pro přístup k datům, případně pro zálohování a obnovu dat. Nejpoužívanějším komunikačním protokolem v síti SAN je iSCSI, který je zapouzdřen do IP paketů a také protokol FCoE (Fiber Channel over Ethernet), který zapouzdřuje FC rámce do rámců Ethernet. Rychlost přenosu dat dosahuje v síti SAN až 10 Gb/s. Sítě SAN jsou vhodné zejména pro velké firmy, využívající desítky serverů s velkými nároky na datové přenosy. Nevýhodou jsou vysoké náklady na vybudování infrastruktury této sítě (20).

2. ANALÝZA SOUČASNÉHO STAVU

V této kapitole je analyzován současný stav zálohování a ukládání záloh v prostředí nejmenované bankovní instituce.

Tato bankovní instituce zaměstnává přibližně 900 lidí v pobočkách po celé České republice. Každý zaměstnanec má k dispozici pracovní stanici, ať už desktop nebo laptop, emailový a doménový účet, které jsou spravovány systémem Microsoft Windows Server. Podle konkrétní náplně práce na dané pozici používají uživatelé specifické aplikace, které mohou běžet vzdáleně na serveru nebo lokálně na klientské stanici. Většina lokálně běžících aplikací ale také komunikuje se svou serverovou částí. Její hlavní datacentrum pro Českou republiku a převážná část IT oddělení je situováno v brněnské kancelářské budově.

2.1 Datacentrum

V datacentru je umístěno celkem 100 fyzických serverů, nad nimiž je vytvořeno dalších 150 virtuálních serverů pomocí virtualizačního nástroje VMware. V závislosti na využití serveru je volen operační systém Windows Server, případně některá z Linux distribucí. Tyto servery zajišťují chod firemních zařízení a aplikací. Windows Server je použit pro řadič domény, adresářový server, souborový server, DNS a DHCP servery, Exchange poštovní sever a Microsoft SQL Server. Na Linuxové platformě jsou založeny převážně databázové servery. Na těchto serverech jsou provozovány, případně testovány databáze typu IBM DB2 a Oracle. 8 serverů je vyhrazeno pro portál internetového bankovnictví.

Veškerá dat jsou ukládána do centrálního diskového pole od společnosti HP. Tvoří jej celkem 40 TB úložného prostoru. Diskové pole je osazeno SAS disky o velikosti 800 GB a 600 GB a rychlostech 10 či 15 tisíc RPM. Konfigurace diskového pole automaticky určuje, zda budou data uložena na pomalejších či rychlejších discích v závislosti na tom, jak často jsou tato data čtena či editována. Data jsou chráněna před případným selháním až dvou disků metodou redundance RAID 6.

2.2 Rozdělení disků

Celková disková kapacita je logicky rozdělena na několik úložišť, ze kterých jsou 3 dostupné všem uživatelům. Prvním z nich je disk, na kterém je každému uživateli vytvořena jeho soukromá složka a následně je namapována jako disk k jeho účtu. Uživatel ji má k dispozici na každém zařízení, na kterém je přihlášen svým doménovým účtem. Velikost této složky je vždy 20 GB.

Druhým logickým úložištěm je disk, který je dále dělen na složky pro všechny jednotlivá oddělení. Kapacita pro každé oddělení je přidělována individuálně dle potřeby, tak aby místo na disku nebylo zbytečně nevyužito. Velikost a pohybuje se v rozmezí 1-50 GB a v případě potřeby je možné ji dočasně či permanentně zvýšit. Uživatelé jsou rozděleni do doménových skupin, dle jejich pracovní pozice, což následně určuje jejich oprávnění ke čtení, změně či zápisu do složek jednotlivých oddělení. Individuálně je však možno udělit uživateli práva například pro zápis i do složky jiného oddělení.

Třetím diskem je společné úložiště. Tento disk má kapacitu 200 GB a je určen pro výměnu dat, především větších souborů, mezi všemi uživateli. Specifikem tohoto úložiště je, že pokud soubor není po delší dobu změněn, ať už editací či přejmenováním nebo přesunutím, je následně z disku smazán. Tímto způsobem je stále udržován dostatek prostoru pro všechny uživatele. Pro uživatele důležitá data by měla mít pro případ smazání svou kopii na soukromém disku.

2.3 Zálohovací systém

Banka využívá komplexního řešení zálohování i obnovy dat AVAMAR od společnosti EMC. Systém je tvořen jedním serverem, pracujícím na Linuxové platformě a diskovým polem. V diskovém poli se nacházejí disky s SAS řadičem a velikostí 600 GB a rychlostí 10 tisíc RPM. Banka disponuje v Brně dvěma těmito systémy, mezi kterými probíhá replikace dat. Jeden systém se nachází přímo v serverovně budovy a druhý je externě umístěn v datacentru firmy Faster, který zajišťuje offsite zálohu pro případ fyzického poškození zařízení v budově sídla firmy.

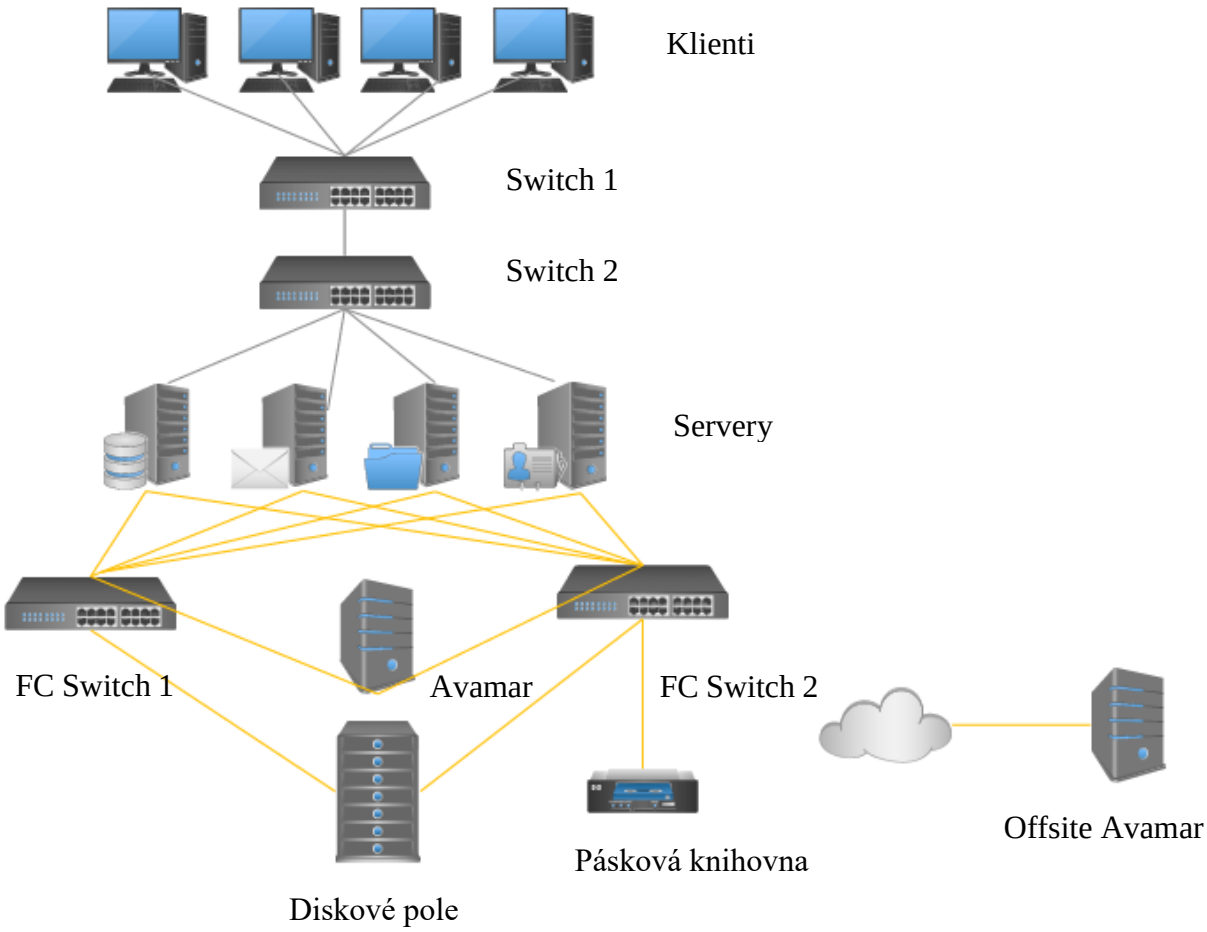
2.3.1 Zálohování souborového serveru

Zálohování souborového serveru probíhá 1x denně od 20ti hodin, kdy zálohovací server pošle požadavek klientu, který provede deduplikaci. Data, která jsou systémem vybrána k zálohování, jsou přenášena po vnitřní optické síti ze serveru na diskové úložiště zálohovacího systému. Zálohovací server následně inkrementálně připisuje nová data k předchozím zálohám. Každá sebemenší záloha tvoří v úložišti logicky plnou zálohu, ze které je možné provést obnovu do původního stavu v libovolný požadovaný den, přičemž je velmi efektivně využita disková kapacita.

2.3.2 Zálohování Exchange serveru

Záloha emailového serveru je rozdělena zálohování aktuální schránky a archivů. Schránky uživatelských účtů jsou zálohovány 5x denně, přičemž poslední záloha dne vzniká ve 20 hodin. Archivované emaily z těchto schránek se zálohují jednou denně, rovněž ve 20 hodin.

Po dokončení přenosu všech inkrementů záloh ze všech serverů na lokální Avamar zálohovací server, jsou následně nově zapsaná data replikována na druhý identický zálohovací server, umístěný v pronajatém prostoru datacentra firmy Faster. Při replikaci probíhá šifrování přenášených dat. Tuto funkci umožňuje software pro správu zálohovacího serveru Avamar. Přenos dat je skrze vyhrazenou optickou trasu mezi oběma body s rychlostí linky 1 Gb/s.



Obrázek č. 12: Schématické znázornění vnitřní sítě (Zdroj: Vlastní zpracování)

2.4 EMC Avamar

Toto řešení se skládá ze dvou částí. První částí je Avamar server s datovým úložištěm a druhou částí je aplikace klienta, která musí být instalována na každý server, který je zařazen do správy zálohování tímto řešením.

Aplikace přímo na zálohovaném serveru také zajišťuje hlavní funkci tohoto systému, a tou je deduplikace dat. Ta je provedena ještě před přenosem dat skrze síť do datového úložiště. Klient instalovaný na serveru po inicializaci zálohování pomocí deduplikačních algoritmů dělí data v úložišti na bloky a porovnává jejich duplicitní existenci v rámci celého systému. Unikátní soubor se tedy v záloze vyskytuje jen jednou napříč všemi jeho lokacemi. Pokud systém najde duplicitu tohoto souboru, tak se do zálohy uloží pouze odkaz na jeho umístění.

Avamar vyváří inkrementální deduplikované zálohy, které logicky vytváří plnou zálohu. Systém nejprve zjistí, která data byla změněna od poslední zálohy a poté jsou tyto nové změny rozděleny do bloků, které jsou porovnány se zálohami. Pokud není zjištěna shoda, pouze tyto nové bloky dat jsou určeny k přenosu do zálohovacího diskového pole, kde jsou spolu s metadaty, které časově označují novou zálohu, připsány k předchozí záloze. Tímto dochází k výraznému snížení jak objemu přenášených dat, tak i využití přenosové kapacity linky a také velmi výraznému zvýšení rychlosti provedení zálohy oproti plné záloze přenášené přes síť.

Obrázek č. 13: EMC Avamar (Zdroj: 15)

2.5 Archivace

Archivace – dlouhodobé uchování zálohovaných dat, je prováděna ukládáním dat na zálohovací pásky typu LTO-6. Všechny používané pásky mají shodnou kapacitu 3 TB. Zápis na tyto pásky zajišťuje pásková mechanika Hewllet-Packard Ultrium 6250. Archivace dat probíhá ve čtvrtletních intervalech. Každý 3. měsíc v roce jsou tedy zálohy z předchozího období zkopírovány ze systému Avamar na magnetické pásky. Tyto pásky jsou následně uskladněny a uchovávány v bezpečnostním trezoru. Pro případ potřeby obnovit data jsou tyto pásky skladovány po dobu deseti let. Po uplynutí této doby jsou pásky předány k bezpečnému znehodnocení. Z technologických důvodů zde jejich životnost končí, neboť může být omezena jejich kompatibilita s novou páskovou mechanikou, podporující zápis pouze na novější typ pásek standardu LTO. Dále také by vzrostla míra rizika ztráty nebo poškození dat na ní zapsaných. Zápis dat ze zálohovacího serveru tedy vždy probíhá na novou, čistou pásku. Na pásky neprobíhá dodatečný zápis pro využití celkové kapacity, pokud jsou na pásku již jednou zapsána data, je následně určena už pouze ke čtení.

Archivace probíhá způsobem Disk-to-Disk-to-Tape, kdy je nejprve zdrojový server zálohován na zálohovací server a z něj následně probíhá archivace na magnetickou pásku ve výše zmíněném intervalu. Pokud je nutné načíst data z důvodu jejich obnovy, je páska vložena do mechaniky a pomocí webového konfiguračního prostředí zálohovacího systému je záloha načtena a uložena na disky Avamaru a následně obnovena do hlavního diskového pole.

Obrázek č. 14: HPE StoreEver LTO-6 Ultrium 6250 TapeDrive (Zdroj: 16)

2.6 Zhodnocení současného stavu

Proces zálohování a obnovy dat je v této společnosti v současné době na velmi vysoké úrovni. V bankovní instituci ani nelze čekat opak. Použití zálohovacího řešení od EMC poskytuje velmi efektivní využití úložného prostoru pro ukládání záloh, stejně jako jejich přenos po vnitřní síti.

Proces archivace na pásková média osobně považuji v dnešní době za poměrně zastaralý. Zálohovací server by měl být schopen dlouhodobě uchovat data. Tato doba by mohla být místo deseti let například jen 8, ale byla by zajištěna maximální kompatibilita úložných médií, kterou standard LTO nezaručuje. Avšak vzhledem k dnešním cenám pásek a důležitosti dat uvnitř banky je to logická volba, která jen minimálně zatěžuje rozpočet.

Naopak v oblasti datového úložiště firmy je prostor ke zlepšení a modernizaci. Současné úložiště je zhruba na 88 % své kapacity. V současné době neustále zvětšujících se objemů zpracovávaných dat by tato kapacita zanedlouho nemusela dostačovat. Diskové pole by také mohlo být vylepšeno pomocí SSD disků, které by urychlily zpracování velkého množství transakčních dat, a tím zvýšily výkon celého úložiště. Společnost disponuje dostatečnými finančními prostředky pro investice do nových technologií, proto je možné navrhnout použití jedno z nejlepších řešení na trhu.

3 VLASTNÍ NÁVRH ŘEŠENÍ

3.1 Hybridní diskové pole

V první části se zaměřím na možnost pořízení nového diskového pole, ve kterém je možné používat současně rotační a SSD disky. Ty poskytují větší rychlost čtení i zápisu, v dnešní době srovnatelnou spolehlivost a kapacitu a nižší energetickou spotřebu.

Vyšší rychlost čtení se projeví zejména při náhodném přístupu k datům, kdy je přímý přístup k paměťové buňce a v ní uložené informaci. U klasického plotnového disku je omezujícím faktorem rotace ploten a pohyb čtecích hlav k nalezení a přečtení zapsaných dat, zejména pokud je čtený datový soubor rozdělen na několik částí napříč celým diskem, tzv. fragmentace.

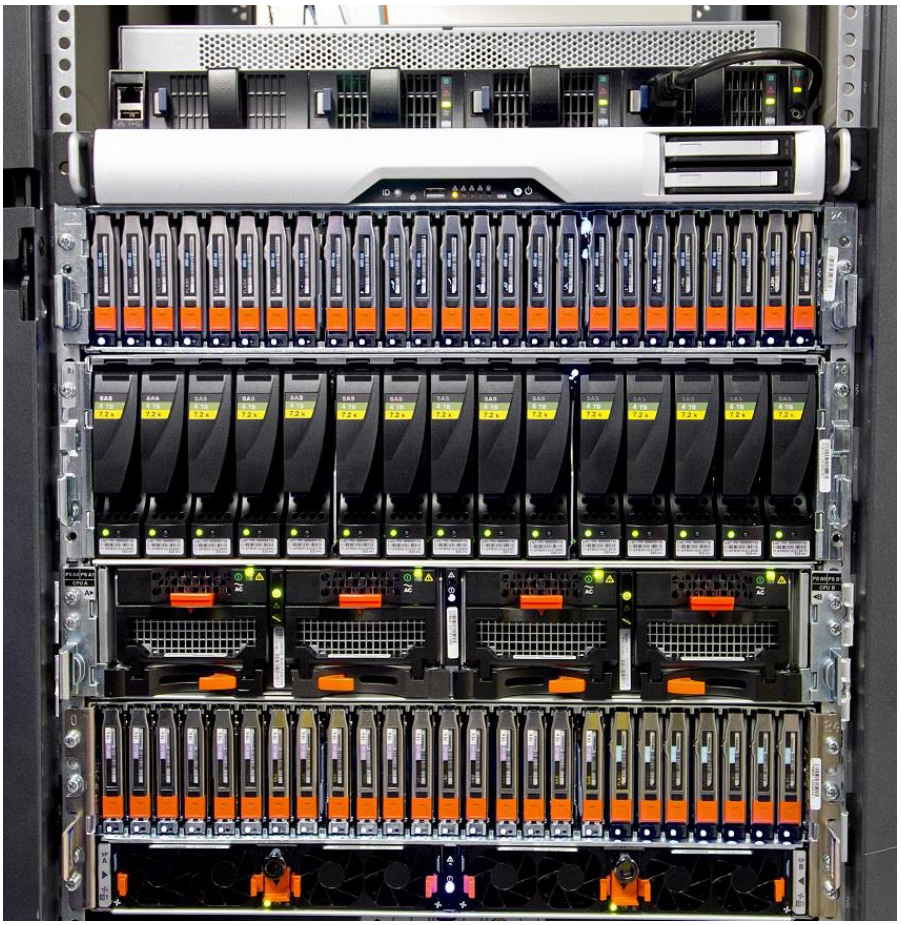
3.2 EMC VNX 5200

V návrhu nového řešení využiji kombinované – hybridní úložiště dat. To umožňuje použití 2,5“ SSD disků pro rychlou práci s daty, která jsou často využívána, dále pak 3,5 palcových plotnových disků s rychlostí 15k RPM a největší diskovou kapacitu poskytnou 2,5 palcové plotnové disky o rychlosti 10k RPM, na kterých budou uchována data, ke kterým není často přistupováno, ale je nutné jejich uchování v dlouhodobém časovém intervalu.

Pro toto kombinované diskové pole bude použito řešení od společnosti EMC, jenž zajistí nejvyšší možnou kompatibilitu se zálohovacím systémem Avamar od téže firmy, kterým již banka v současné době disponuje. Konkrétně se bude jednat o diskové pole řady VNX, konkrétně model 5200, které svými parametry a možností případného budoucího jednoduchého rozšíření plně vyhovuje požadavkům. Diskové pole v základní konfiguraci poskytuje 2 řadiče disků. Jeden z nich pro 25 disků velikosti 2,5“ ať už HDD nebo SSD s SAS rozhraním a druhý pro 15, 3,5“ SAS disků. K základní konfiguraci bude nutné, k dosažení požadované a dostatečné diskové kapacity, dokoupit rozšiřující řadič pro dalších 25 disků velikosti 2,5“.

Diskové pole VNX 5200 má následující parametry (12):

- **CPU:** 2x Intel Xeon E5-2600 2,4 GHz
- **RAM:** 32 GB DDR3
- **Maximální kapacita:** 500 TB
- **Podporované disky:** Flash SSD, SAS, NL-SAS
- **Podpora RAID:** 0, 1, 10, 3, 5, 6
- **Maximální počet disků:** 125
- **Konektivita:**
    - FC 8Gb/s
    - 1GBASE-T iSCSI
    - 10GbE iSCSI



Obrázek č. 15: Datové úložiště EMC VNX 5200 (zdroj: 10, upraveno)

3.2.1 Zvolená konfigurace disků v datovém úložišti

Diskové pole bude tvořit celkem 65 disků rozdělených na 3 skupiny. Jejich celková teoretická kapacita je 61 TB, ta však bude snížena tím, že disky budou nakonfigurovány v redundantním poli RAID. Další snížení využitelné kapacity zapříčiní použití některých disků v režimu hot spare. V tomto režimu není disk využíván k zápisu ani čtení, pokud by však systém detekoval selhání některého z disků v poli, okamžitě se inicializuje obnova dat z parity uložené napříč polem na „čekající“ disk. Tento disk však musí mít stejnou kapacitu i rychlost otáček jako disk, který selhal.

Specifickou vlastností datového úložiště VNX je Permanent Sparing. Standardně při dopočítaní dat na hot spare disk a následné výměně poškozeného disku, jsou data na nový disk překopírována a disk se vrací do režimu hot spare. U VNX je nový vložený disk automaticky přepnut do režimu hot spare.

Tabulka 1: Zvolená konfigurace diskového úložiště

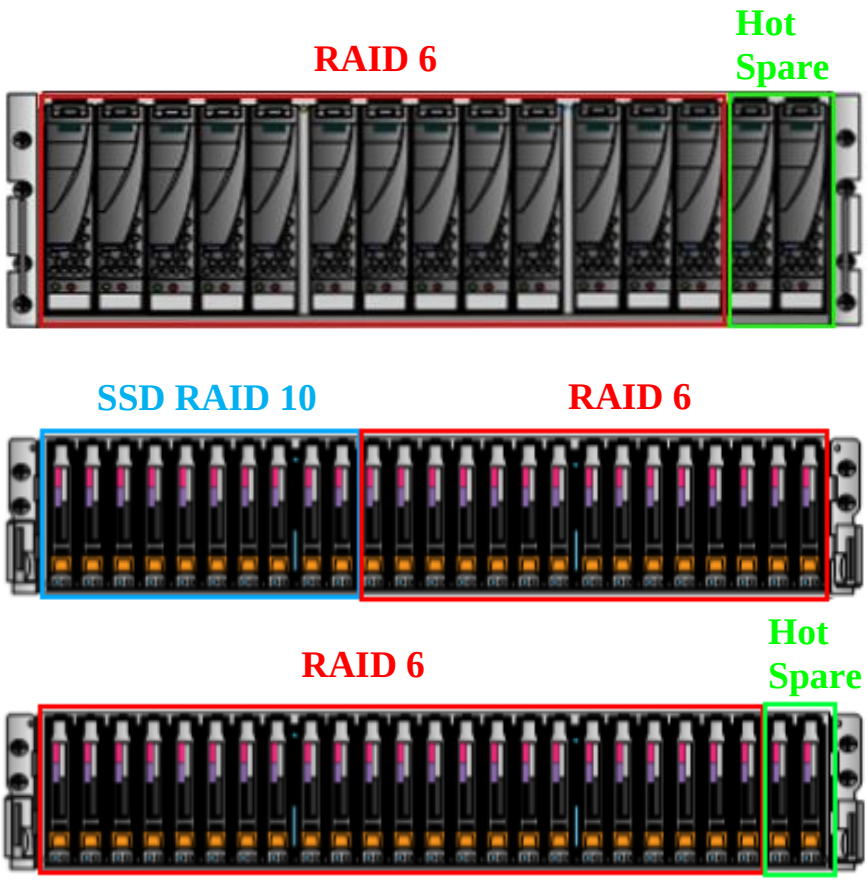
Velikost	Typ	Počet	Kapacita	Otáčky	Rozhraní
2,5"	SSD	10	400 GB	-	FC
3,5"	HDD	15	600 GB	15k	SAS 6Gb
2,5"	HDD	40	1200 GB	10k	SAS 6Gb
Celkem		65	61000 GB		

(Zdroj: Vlastní zpracování)

První skupinou bude 10 SSD disků o kapacitě 400 GB. Ty budou fungovat v režimu RAID 10. Tento režim umožní dosažení vysokých rychlostí čtení a zápisu a zároveň poskytne ochranu dat pomocí redundance, pro případ selhání disku. Celková užitná kapacita bude 2 TB. Komunikaci bude zajišťovat rozhraní FC.

Druhá skupina bude tvořena 15ti, 3,5“ pevnými disky. Každý z nich má kapacitu 600 GB. Při konfiguraci do RAID 6 bude využita kapacita jedenácti disků. Parita ukládaných dat bude zabírat kapacitu dvou disků. Další 2 disky budou nastaveny do módu hot spare.

Třetí a největší skupinu bude tvořit 40 pevných disků o velikosti 2,5“, s rychlostí otáček 10k RPM. Kapacita každého disku je 1,2 TB. Tato část úložiště bude pracovat rovněž v konfiguraci RAID 6, kdy použitelná kapacita bude odpovídat 36ti diskům, kapacitu dvou disků zabere parita a 2 disky budou opět v režimu hot spare.



Obrázek č. 16: Grafické znázornění konfigurace RAID (zdroj: 14, upraveno)

Tabulka 2: Teoretické maximální rychlosti v konfiguraci RAID

Disk	Otáčky	Raid	Konektivita	Čtení	Zápis
SSD	-	10	FC	900 MB/s	850 MB/s
HDD	15k	6	SAS	700 MB/s	600 MB/s
HDD	10k	6	SAS	600 MB/s	500 MB/s

(Zdroj: 11, Vlastní zpracování dat)

3.3 Software

Softwarové nástroje k tomuto diskovému poli lze rozdělit na 3 skupiny.

3.3.1 Base Software

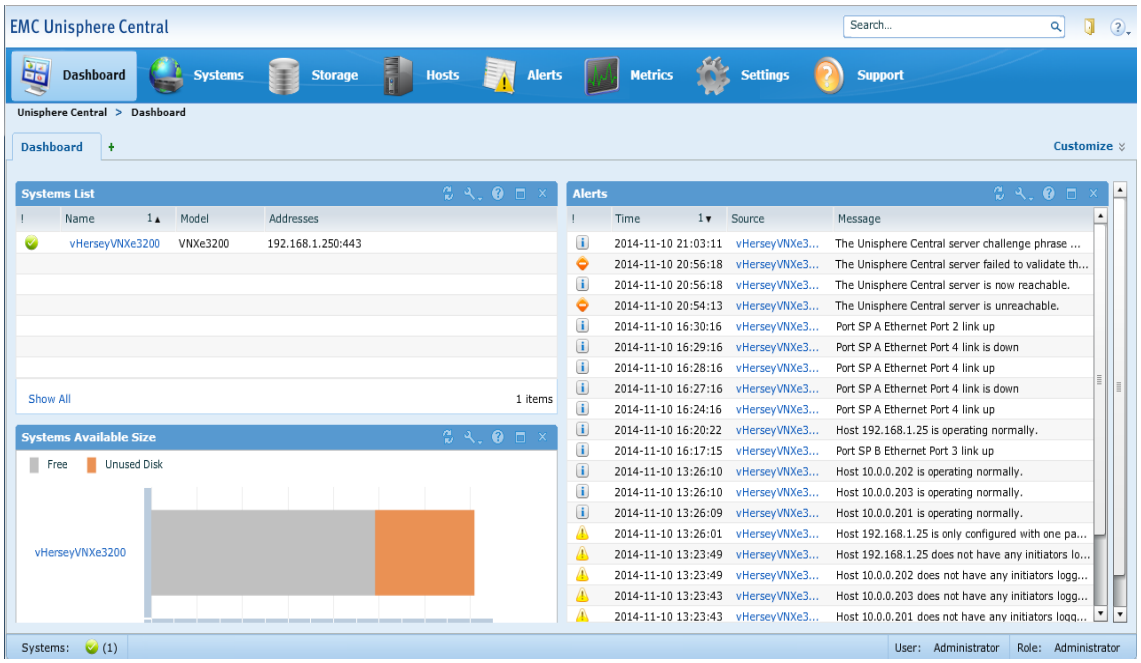
Do skupiny základního softwaru patří nástroje, zajišťující samotné fungování diskového pole. Tento software je již na zařízení instalován při jeho zakoupení.

Konkrétně se jedná o:

- VNX Operating Environment
    - Software pro řízení přístupu a správu úložiště na úrovni souborů
- Protokoly přístupu pro iSCSI, FC a FCoE
- Nástroj pro deduplikaci souborů a bloků dat
- Nástroj pro kompresi souborů a bloků dat

3.3.2 Unisphere Management

Jedná se o rozšiřitelnou platformu pro úložiště VNX, která poskytuje uživatelské rozhraní pro jeho kompletní správu. Software umožňuje správu jednoho i více VNX úložišť v síti, nastavení logického uspořádání disků a provádí jejich monitoring. K aplikaci lze přistupovat lokálně nebo skrze webové prostředí.



Obrázek č. 17: Prostředí softwaru Unisphere (zdroj: 22)

3.3.3 Rozšiřující balíky

EMC nabízí k úložištím VNX několik balíků doplňkového softwaru. Jedná se o rozšiřující softwarové balíky, které obsahují více specifické nástroje pro manipulaci s daty. Pro návrh vlastního řešení využiji základní balík VNX Software Essentials Pack. Součástí balíku jsou následující nástroje:

- FAST Suite
    - Kombinace softwaru FAST VP (Virtual Pool) pro tiering logických skupin disků a FAST Cache pro rychlejší přístup k informacím, pomocí cashování dat na rychlejší SSD disky. FAST Suite umožňuje současného využívání SSD a HDD disků v jednom úložišti. Podrobněji FAST rozeberu v následující podkapitole.

- VNX Events and Retention Suite
    - Tento software rozšiřuje Unisphere Management aplikaci o nástroje pro ochranu dat před nechtěnými změnami nebo smazáním.

- Local Protection Suite
    - Software přidává uživateli možnost vrátit se kdykoliv k předchozí verzi uložených dat.

- Remote Protection Suite
    - Nástroje pro ochranu dat v případě selhání disků, výpadku napájení nebo poruchám v datovém centru.

3.4 Hot and Cold data

Diskové pole VNX využívá strategii „flash first“ pomocí technologie FAST (Fully automated Storage Tiering) neboli plně automatizovaný tiering úložiště. Tato technologie konstantně provádí monitoring všech uložených dat a vyhodnocuje, která data jsou aktivně využívána, a přesouvá je na rychlé SSD disky, tak aby byla zkrácena přístupová doba a urychleno jejich načítání. Naopak dlouhodobě nepoužívaná data jsou přesunuta na HDD, jejichž kapacita je levnější.

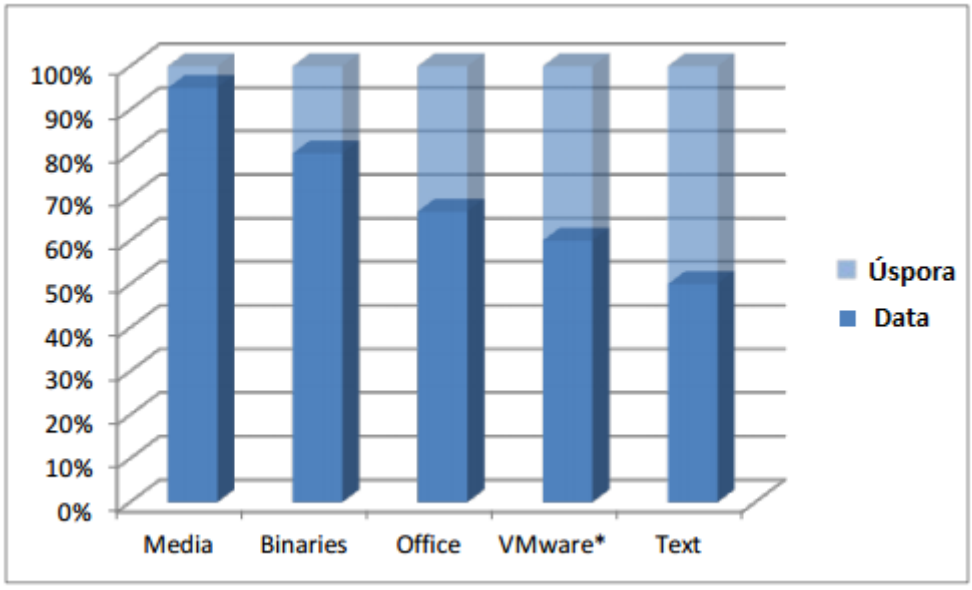
3.5 Deduplikace diskového pole

Diskové pole řady VNX je schopno provádět deduplikaci datových souborů přímo v úložišti. To jednoduše znamená, že žádná data nebudou uložena dvakrát.

Ve stávajícím řešení jsou data deduplikována až při inicializaci zálohování ze strany Avamar serveru a v této podobě jsou přenesena do zálohovacího úložiště. V primárním datovém poli však zůstávají ve své původní podobě.

V navrhovaném řešení bude nově provedena konfigurace datového úložiště tak, že systém bude vyhodnocovat a následně provádět deduplikaci datových souborů odpovídajících zadaným kritériím. Bude se jednat o soubory, které již byly systémem tieringu vyhodnoceny jako dlouhodobě nepoužívané a byly přesunuty na pomalejší diskové úložiště. Jedná se především o dokumenty, které má ve shodné nebo podobné podobě na svém úložišti několik uživatelů současně. Z deduplikace budou automaticky vyloučena veškerá operační a mission critical data, nacházející se na rychlejších SSD discích. Tato konfigurace je možná pomocí softwaru, se kterým bude toto řešení zakoupeno.

Graf 1: Potencionální možná úspora úložiště při použití deduplikace



(Zdroj: 13)

3.6 Cenové shrnutí

V následující kapitole provedu cenovou kalkulaci veškerého hardwarového a softwarového vybavení potřebného k použití navrhnutého řešení.

3.6.1 Cenové shrnutí hardwaru

Následující tabulka zobrazuje cenu jednotlivých hardwarových komponent tvořících úložiště dle zvolené konfigurace.

Tabulka 3: Cenová kalkulace hardwarových komponent

Zařízení	Množství (ks)	Cena za jednotku	Celková cena
VNX 5200 25x2,5" DISKLESS	1	234 562 Kč	234 562 Kč
Řadič disků 25x2,5"	1	120 458 Kč	120 458 Kč
Řadič disků 15x3,5"	1	80 452 Kč	80 452 Kč
EMC 400GB 4 Gb/s FC SSD	10	84 522 Kč	845 220 Kč
EMC 600GB SAS 6 Gb/s 15K 3,5" HDD	15	28 133 Kč	421 995 Kč
EMC 1,2TB SAS 6 Gb/s 10K 2,5" HDD	40	38 155 Kč	1 526 200 Kč
Celkem			3 228 887 Kč

(Zdroj: 23, 24, 25, Vlastní zpracování dat)

Z tabulky lze vyčíst, že nejdražším typem disku je SSD o kapacitě 400 GB, jehož cena převyšuje 80 tisíc za kus, což je poměrně dost, nicméně v oblasti enterprise takové ceny nikoho nepřekvapí. Nejdražší souhrnnou položkou jsou 1,2 TB disky, které budou tvořit základ úložiště.

3.6.2 Cenové shrnutí softwaru

V tabulce níže je soupis použitého softwaru a jeho cena. Nejnákladnější položkou je licence softwaru Unisphere pro správu úložiště.

Tabulka 4: Cenová kalkulace softwaru

Položka	Celková cena
VNX5200 Software Essentials Pack	145 258 Kč
Premium podpora 1 rok	28 272 Kč
VNX5200 Unisphere Management	170 349 Kč
Celkem	343 879 Kč

(Zdroj: 23, 24, 25, Vlastní zpracování dat)

3.7 Zhodnocení vlastního návrhu

Návrh na využití nového úložiště přináší dle mého názoru mnohá zlepšení, jak zálohovacího procesu, tak i ukládání, uchování a práci s daty.

V zálohovacím procesu je to vyšší integrita mezi úložištěm a zálohovacím systémem použitím hardwaru i softwaru od stejné firmy na obou zařízeních.

Nové úložiště řeší problém se současnou diskovou kapacitou, která se pomalu, ale jistě stává nedostatečnou. Přináší také zvýšení výkonu pro manipulaci se zapsanými daty, zejména při zpracovávání transakcí na databázových serverech.

Z informací získaných při zpracovávání této práce a osobními zkušenostmi považuji společnost EMC za lídra na trhu zálohovacích systémů a úložišť. Kladně je hodnocen zejména jejich dodávaný software, funkcionalita i technická podpora.

Jediným negativem mohou být celkové náklady na toto řešení, jak je zřejmé z následující tabulky.

Tabulka 5: Cenová kalkulace - souhrn

Položka	Cena
Hardware	3 228 887 Kč
Software	343 879 Kč
Celkem	3 572 766 Kč

(Zdroj: Vlastní zpracování dat)

ZÁVĚR

Cílem první části této práce bylo zpracování teoretických informací, zabývajících se procesem ukládání, uchování a zálohování dat. Tato teoretická část byla zpracovávána za pomocí literatury a informací dostupných na internetu.

Ve druhé části byla za pomoci konzultanta, působícího ve firmě, na kterou je tato práce zaměřena, provedena analýza infrastruktury a metod zálohovacího procesu. Výsledkem bylo zjištění velmi dobře navrženého procesu zálohování i použitého hardwaru a softwaru, což je v bankovní instituci, zpracovávající informace o tisících klientech předvídatelný fakt. Současným zjištěním bylo využití datového úložiště, které začínalo dosahovat limitu své úložné kapacity. Vlastní návrh byl proto zaměřen na výběr nového úložiště, které přinese kromě navýšení kapacity i mnohem vyšší výkon při zpracovávání informací.

SEZNAM POUŽITÉ LITERATURY

[1] Zálohujte svá data. *hurican.cz* [online]. 2009 [cit. 2015-04-24]. Dostupné z: http://www.hurican.cz/tipy-zakaznikom/zalohujte-sva-data

[2] JUNEK, Pavel. Zálohování a archivace dat v podnikovém prostředí – 5. díl, Typy záloh a jejich rotační schémata. Zalohovani.net [online].2013a [cit. 2015-04-26]. Dostupné z: http://www.zalohovani.net/zalohovani-a-archivace-dat-v-podnikovemprostredi-5-dil-typy-zaloh-a-jejich-rotacni-schemata/

[3] ŠULC, O. Zálohování dat a datová úložiště. Brno: Vysoké učení technické v Brně, Fakulta podnikatelská, 2014. 58 s. Vedoucí bakalářské práce Ing. Jiří Kříž, Ph.D.

[4] Disková pole – RAID. *pit.wz.cz* [online]. 2005 [cit. 2015-04-25]. Dostupné z: http://pit.wz.cz/Konstrukce/raid.php

[5] Co je to RAID?. *netcorp.cz/web/* [online]. [cit. 2015-04-22]. Dostupné z: http://www.netcorp.cz/storage_systems/raid_description.htm

[6] ŠULC, Tomáš. VelociRaptory a RAID - když jeden nestačí. *PCtuning.cz* [online]. 2008 [cit. 2015-04-26]. Dostupné z: http://pctuning.tyden.cz/component/content/11987?task=view&limit=1&start=1

[7] DOČEKAL, Michal. Správa linuxového serveru: RAID teoreticky. *Linuxexpres.cz* [online].2009 [cit. 2015-04-26]. Dostupné z: http://www.linuxexpres.cz/praxe/sprava-linuxoveho-serveru-raid-teoreticky

[8] RAID. *Prepressure.com* [online].©1997-2013 [cit. 2015-04-27]. Dostupné z: http://www.prepressure.com/library/technology/raid

[9] RAID 5 or RAID 6. VANOVER, Rick. *Techrepublic.com/* [online]. 2010 [cit. 2015-04-21]. Dostupné z: http://www.techrepublic.com/blog/the-enterprise-cloud/raid-5-or-raid-6-which-should-you-select/

[10] EMC VNX5200 Review. StorageReview Enterprise Lab. storagereview.com[online]2015[cit. 2016-03-02]. Dostupné z: http://www.storagereview.com/emc_vnx5200_review

[11] EMC VNX2 Unified Best Practices for Performance. EMC.com [online]2015[cit.2016-03-01]. Dostpné z: https://www.emc.com/collateral/software/white-papers/h10938-vnx-best-practices-wp.pdf

[12] EMC. VNX SERIES UNIFIED STORAGE SYSTEMS. [online]2015[cit. 2016-02-10]. Dostupné z: https://www.emc.com/collateral/software/specification-sheet/h8514-vnx-series-ss.pdf
[13] RADA, Petr. EMC deduplikace Data Domain a Avamar.[online]2010[cit 2016-03-01]. Dostupné z: http://www.slideserve.com/dixie/emc-deduplikace-data-domain-a-avamar
[14] JUSTIN, Paul. Anatomy of an VNX array. jpaul.me[online]2012[cit. 2016-03-02]. Dostupné z: http://www.jpaul.me/2012/11/anatomy-of-an-emc-vnx-array
[15] Data Archive Corporation.[online]2015[cit. 2016-03-05]. Dostupné z: http://dataarchivecorp.com/emc-avamar-backup-solutions/
[16] TORRES, Gabriel. Does RAID 0 really increase disk performance?[online]2006[cit. 2016-05-15]. Dostupné z: http://www.hardwaresecrets.com/does-raid0-really-increase-disk-performance/6/
[17] LEURS, Laurence. RAID.[online]2015[cit. 2016-05-15]. Dostupné z: http://www.prepressure.com/library/technology/raid#raid-5
[18] KREJČÍ, M. Komprese dat. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 89 s.
[19] KEEGAN, Colm. Source vs Target Based Data Deduplication .[online]2013[cit. 2016-05-15]. Dostupné z: http://www.storage-switzerland.com/Articles/Entries/2013/1/2_Source_vs_Target_Based_Data_Deduplication.html
[20] STOPKA, Marek. Storage area network. [online]2010[cit. 2016-05-15]. Dostupné z: http://www.abclinuxu.cz/clanky/storage-area-network-1-uvod
[21] http://www.itzone.co.ke/products/enterprise/storage/
[22] CARTWRIGHT, Hersey. VNXe3200 and Unisphere Central. [online]2014[cit. 2016-05-20]. Dostupné z: http://www.vhersey.com/2014/11/vnxe3200-and-unisphere-central
[23] EMC. VNX Boost Program. [online][cit. 2016-05-20]. Dostupné z: https://www.arrowthehub.co.uk/media/27168/vnx-boost-program.pdf
[24] Kernel Software. EMC price list. [online]2016[cit. 2016-05-20]. Dostupné z: http://www.kernelsoftware.com/products/catalog/emc.html ceník

[25] EMC. Price List. [online]2015[cit. 2016-05-20]. Dostupné z: http://www.emc.com/microsites/stateofmsepl3760/pdf/MS%20EPL%203760%20Price%20List%20-%202015-11.pdf
[26] RECOVERYCENTRUM.CZ. Příčiny ztráty dat. [online][cit. 2016-05-22]. Dostupné z: http://www.recoverycentrum.cz/priciny-ztraty-dat/
[27] GÁLA, L.; POUR, J.; TOMAN, P. Podniková informatika. Praha : Grada Publishing, a.s., 2006. 484 s. ISBN 80-247-1278-4.
[28] POŽÁR, Josef. Manažerská informatika. Plzeň: Aleš Čeněk, 2010, 357 s. ISBN 978-80-7380-276-9.
[29] SOSINSKY, Barrie. Mistrovství – počítačové sítě. 1. vyd. Brno: Computer Press, 2010, 840 s. ISBN 978-80-251-3363-7.
[30] PECINOVSKÝ, P. Archivace a komprimace dat. Praha: Grada Publishing, 2003. 116 s. ISBN 80-247-0659-8.

SEZNAM OBRÁZKŮ

Obrázek č. 1: RAID 0 (Zdroj: 17) 14
Obrázek č. 2: RAID 1 (Zdroj: 17) 15
Obrázek č. 3: RAID 5 (Zdroj: 17) 16
Obrázek č. 4: RAID 6 (Zdroj: 17) 16
Obrázek č. 5: RAID 10 (Zdroj: 17) 17
Obrázek č. 6: Komprese dat (Zdroj: Vlastní zpracování) 18
Obrázek č. 7: Deduplikace na zdroji (Zdroj: Vlastní zpracování) 19
Obrázek č. 8: Deduplikace na cíli (Zdroj: Vlastní zpracování) 20
Obrázek č. 9: DAS (Zdroj: 21) 21
Obrázek č. 10: NAS (Zdroj: 21) 22
Obrázek č. 11: SAN (Zdroj: 21) 22
Obrázek č. 12: Schématické znázornění vnitřní sítě (Zdroj: Vlastní zpracování) 26
Obrázek č. 13: EMC Avamar (Zdroj: 15) 27
Obrázek č. 14: HPE StoreEver LTO-6 Ultrium 6250 TapeDrive (Zdroj: 16) 28
Obrázek č. 15: Datové úložiště EMC VNX 5200 (zdroj: 10, upraveno) 31
Obrázek č. 16: Grafické znázornění konfigurace RAID (zdroj: 14, upraveno) 33
Obrázek č. 17: Prostředí softwaru Unisphere (zdroj: 22) 34

SEZNAM TABULEK

Tabulka 1: Zvolená konfigurace diskového úložiště 32

Tabulka 2: Teoretické maximální rychlosti v konfiguraci RAID 33

Tabulka 3: Cenová kalkulace hardwarových komponent 38

Tabulka 4: Cenová kalkulace softwaru 38

Tabulka 5: Cenová kalkulace - souhrn 39

SEZNAM GRAFŮ

Graf 1: Potencionální možná úspora úložiště při použití deduplikace 37