Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií

Bakalářská práce

Zálohování a archivace dat

Martin Volek

© 2015 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Katedra informačních technologií

Provozně ekonomická fakulta

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Martin Volek

Veřejná správa a regionální rozvoj

Název práce

Zálohování a archivace dat

Název anglicky

Backup and archiving as a tool of data protection

Cíle práce

Cílem této práce je porovnání možností způsobu zálohování a archivace dat na dostupná zálohovací média. Práce se soustředí na metody a strategie zálohování a také na jejich specifika, dále zahrnuje analýzu datových a online úložišť a jejich výhod a nevýhod. V praktické části je cílem analyzovat způsoby datové problematiky ve veřejné správě za účelem zefektivnění zálohovacích procesů.

Metodika

Metodika této práce je založena na výzkumu a analýze dané problematiky. Odborné poznatky a informace budou čerpány z literárních tištěných zdrojů a ověřených internetových stránek. Praktická část bude realizována ve spolupráci s IT technikem v sektoru veřejné správy. Na základě získaných informací budou formulovány závěry práce.

Doporučený rozsah práce

30 – 40 stran

Doporučené zdroje informací

DE GUISE, Preston. Enterprise Systems Backup and Recovery: A Corporate Insurance Policy. New York: CRC Press , 2008. ISBN 987-1-4200-7639-4.
Internetové zdroje.
KŘÍŽ, Libor. Komprimační a archivační programy. Praha: Computer Press, 2002. ISBN 80-7226-757-4
LEIXNER, Miroslav. PC – zálohování a archivace dat. Praha: Grada, 1993. ISBN 80-85424-73-8.

Předběžný termín obhajoby

2015/06 (červen)

Vedoucí práce

Ing. Martin Havránek, Ph.D.

Elektronicky schváleno dne 31. 10. 2014

Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 11. 11. 2014

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 07. 03. 2015

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci "Zálohování a archivace dat" jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 11. 3. 2015 _____

Poděkování

Rád bych touto cestou poděkoval panu Ing. Martinu Havránkovi za cenné rady a odborné vedení, které mi poskytl při zpracovávání bakalářské práce. Dále bych rád poděkoval panu Ing. Miloslavovi Klimešovi, který mi poskytl profesionální asistenci pro získání informací při zpracovávání analytické části.

Zálohování a archivace dat

Backup and archiving as a tool of data protection

Souhrn

Tato bakalářská práce pojednává o problematice zálohování a archivaci dat. Je rozdělena do dvou částí. První – teoretická část je zaměřena na definování pojmů a seznámení se s procesem zálohování, s procesem archivace, dále na kritéria a výběr úložného média, příčiny ztráty dat a také na strategii zálohování. Na závěr jsou zmíněny postupy, jak zpracovávat data před archivací. Druhá – praktická část je zaměřena na politiku zálohování a archivaci dat v sektoru veřejné správy s odbornou asistencí IT technika. Hlavní částí je zálohování, zabezpečení a skladování dat ve veřejné správě. Dále je pojednáno o nákladech zálohovacího procesu a na závěr jsou stanoveny návrhy řešení a vyhodnocení.

Klíčová slova: zálohování, archivace, software, backup, cloudová úložiště, datové média, RAID, pevný disk, komprese, ztráta dat.

Summary

This thesis deals with the topic of backup and data archiving. It is divided into two parts. First - the theoretical part focuses on defining the key words and concepts of the backup process, the process of archiving, as well as the criteria and selection of storage media, causes of data loss and also backup strategy. At the end of this part there are mentioned procedures how to process data before archiving. The second - practical part is focused on backup and archiving in the public administration sector with support of IT technician. The main part is the backup, security and storage of data in public administration. Also the cost of the backup process is discussed and at the end design solutions and evaluation are set.

Keywords: backup, archiving, software, cloud storages, RAID, hard disk, data lost.

Osnova

1 Úvod 5

2 Cíl práce a metodika 6

2.1 Cíl práce 6

2.2 Metodika 6

3 Teoretická část 7

3.1 Pojem zálohování 7

3.2 Pojem archivace 7

3.3 Proč zálohovat? 8

3.4 Co je předmětem zálohování? 9

3.5 Strategie zálohování 9

3.5.1 Plné zálohování (full backup) 10

3.5.2 Částečné zálohování 11

3.5.3 Přírůstkové zálohování (Inkrementální zálohování) 11

3.5.4 Rozdílové zálohování (Diferenciální zálohování) 12

3.6 Zálohovací rutiny 13

3.7 RAID 15

3.7.1 RAID 1 16

3.7.2 RAID 5 16

3.8 Jak správně zálohy skladovat? 17

3.9 Nejdůležitější rozhodovací kritéria 18

3.9.1 Kapacita 18

3.9.2 Rychlost zápisu a čtení 18

3.9.3 Životnost a odolnost 18

3.9.4 Cena 19

3.10 Rozdělení datových nosičů (médií) 19

3.10.1 Magnetická média 19
3.10.2 Optická média 24
3.10.3 Shrnutí optických médií 26
3.10.4 Flash paměti a SSD disk 27
3.10.5 Cloudová úložiště 28
3.10.6 Webová úložiště pro zálohování dat 29
3.11 Software pro zálohování digitálních dokumentů 30
3.11.1 Acronis True Image 30
3.11.2 Cobian Backup 30
3.12 Software pro šifrování médií 31
3.12.1 TrueCrypt 31
3.12.2 BitLocker Drive Encryption 32
3.13 Ochrana a uchování digitálních dokumentů 33
3.14 Archivace dat a digitální dědictví 34
3.15 Zpracování dat před archivací 35
4 Analytická část 37
4.1 Úvod 37
4.2 Legislativa 38
4.2.1 Právní předpisy 38
4.2.2 Ochrana archiválií 38
4.2.3 Elektronický dokument jako archiválie 39
4.2.4 Základní technické podmínky archivu 39
4.3 Zálohovací software 39
4.3.1 Veeam Backup 40
4.3.2 Prostředí WMware a virtualizace 41
4.4 Princip zálohování ve veřejné správě 42

4.5 Digitální spisová služba - eSpis 43
4.6 Deduplikace (komprese) dat 44
4.7 Bezpečnost a skladování dat 44
4.8 Vyhodnocení analytické části 45
5 Závěr 46
6 Seznam použitých zdrojů 47
6.1 Knižní publikace 47
6.2 Články v seriálové publikaci 47
6.3 Elektronické dokumenty 48
7 Seznam obrázků 51

1 Úvod

Píše se 21. století, v historii lidstva století všestranně technologicky nejvyspělejší. Počítačová technika se rozšířila do všech odvětví lidské činnosti a ve vyspělém světě je dnes alespoň jedním počítačem vybavena téměř každá domácnost. Počet uživatelů, kteří využívají jeho takřka nekonečné schopnosti a možnosti, neustále roste. S každým dalším okamžikem, kdy počítač lidstvo obohacuje, důvěra v něj stoupá. Avšak je nutné brát ohledy na problematiku zálohování a chránit svá data před kolizí počítačové stanice. Někteří uživatelé data nezálohují vůbec, neboť si myslí, že počítač je a bude bezchybný. Opak je pravdou, hlavní nosič uživatelských dat - pevný disk - je mechanická periferie stanice, která pod nepřetržitým napájením a rotováním hliníkových ploten skladuje naše data. Jedná se tedy o velmi zaneprázdněnou periferii, která má omezenou životnost. Zálohování je v dnešním IT světě nezbytnou součástí, pokud nechceme jednoho dne přijít o důležitá osobní data, např. o výsledky své dlouhodobé tvůrčí práce či o rodinné fotografie či o jiná citlivá data v souvislosti s tím, že byl počítač napaden viry, nebo vypověděl službu stářím. Počítač není bezchybný společník, ale pouze stroj, který má i své chyby. Zkrátka není dokonalý a je třeba být se zálohovanými daty o krok napřed.

Nebezpečí kolize počítačové stanice nehrozí jen běžným uživatelům, nýbrž i sestavě podnikových stanic, které uchovávají velmi citlivá a důležitá data, o které by za žádných okolností neměli přijít z důvodu potencionálních negativních následků, např. ztráta důvěry klientely v podnik či firmu. Z těchto důvodů bych v mé bakalářské práci chtěl blíže specifikovat a objasnit, proč je důležité zálohovat (archivovat) data a nepodceňovat bezchybnost počítačové stanice. Doufám, že tato práce pomůže přispět k tomu, aby se počet uživatelů, kteří o svá data přichází, snížil.

2 Cíl práce a metodika

2.1 Cíl práce

Cílem této práce je porovnání možností způsobu zálohování a archivace dat na dostupná zálohovací média. Práce se soustředí na metody a strategie zálohování a také na jejich specifika, dále zahrnuje analýzu datových a online úložišť a jejich výhod a nevýhod. V praktické části je cílem analyzovat způsoby datové problematiky ve veřejné správě za účelem zefektivnění zálohovacích procesů.

2.2 Metodika

Metodika této práce je založena na výzkumu a analýze dané problematiky. Odborné poznatky a informace budou čerpány z literárních tištěných zdrojů a ověřených internetových stránek. Praktická část bude realizována ve spolupráci s IT technikem v sektoru veřejné správy. Na základě získaných informací budou formulovány závěry práce.

3 Teoretická část

3.1 Pojem zálohování

Zálohovací proces je takový mechanismus, během kterého jsou selektována data či soubory, které jsou následně uloženy na jiném datovém médiu, případně na jiném místě z důvodu předcházení situací, kdy dojde nepředvídaně ke ztrátě dat. Jestliže dojde ke zničení prvotního média (například pevný disk), data jsou následně obnovena z předem vytvořené zálohy. Zálohovaná data můžou pomoci například při obnově databázového systému, pro obnovení původního nastavení operačního systému či pro obnovu nechtěně smazaných dokumentů uživatelem. Zálohovací proces může probíhat nepravidelně (např. v domácnostech) nebo pravidelně (např. v sektoru veřejné správy či v podnicích). Hlavní důvod zálohovacího procesu spočívá v ochraně dat před ztrátou v souvislosti s poškozením datového média.

3.2 Pojem archivace

„Archivace představuje trvalé odstranění dat ze systému, které již nemusí být dostupné on-line“ [1].

Základní myšlenkou archivačního procesu je uchování a zajištění digitálních dat pro dlouhodobou dostupnost tak, aby byla uchována bez jakéhokoliv poškození. Může to být v časovém horizontu řádově od několika měsíců až po několik let. Data jsou uložena na nosiče s dlouhodobou životností a bezztrátovým uchováním souborů, například na magnetické pásky, nebo na vybrané pevné disky. Archivace dat je levnější než proces zálohování, jelikož u médií, na které jsou zálohována data, se musí držet krok s technologickým vývojem v oblasti rychlosti přístupu k datům a velikosti úložného prostoru, zatímco pro účely archivace stačí mít starší a pomalejší úložná média, která svůj účel archivace splní podle předpokladů.

[1] BLŠTÁK, O., Archivace – jediný možný cíl vašich dat. 1.díl. [online]. 2012 [cit. 2014-11-06]. Dostupné z: http://bit.ly/1LYAALa

3.3 Proč zálohovat?

Podle M. Leixnera (2003, s. 9) se za posledních několik let stal personální počítač tzv. „lepší kancelářskou výbavou“ a téměř povinnou výbavou domácností, či nezbytně nutnou pracovní pomůckou. Leixner dále uvádí, že ne každý si však uvědomuje, že počítač sice opravdu velmi zjednodušuje a zrychluje nejrůznější činnosti, ale má také svá velice nepříjemná rizika. V okamžiku, kdy počítači svěříme své údaje, s nimiž denně pracujeme, přebíráme zároveň svůj díl zodpovědnosti za to, že tyto údaje jsou všestranně zabezpečeny a že se tudíž nestanou příčinou kolize celého soukolí.

„Pevný disk se může poškodit v kterémkoli okamžiku a potřebná data se mohou stát nedostupnými. Může jít o mechanické poškození, výpadek elektrického proudu v okamžiku zápisu do systémové oblasti, o prach, či dokonce o cigaretový kouř, jejichž vlivem dojde k poškození“ [2].

Nemusí se jednat pouze i mechanické poškození, které při běžném používání na základě statistik dochází po 4 letech, nýbrž roli může hrát i napadení systému virusem, který může udělat velké škody, např. vymazat či přeformátovat harddisk. Může jít také o uživatelsky nechtěnou manipulaci s daty. Mezi ty méně časté případy patří krádež, požár či zásah bleskem. Cílem zálohování je rychle obnovit plně funkční stav informačního systému takový, jaký byl těsně před poruchou. Zálohování je možné popsat jako vytváření bezpečnostní kopie dat nebo celého operačního systému tak, abychom mohli v případě havárie obnovit stav, který existoval těsně před vznikem poruchy. Důvod, proč zálohovat svá data je prostý, nepřijít o citlivá data. Nikdo nechce přijít o své jedinečné soubory, které už by poté velmi těžce získal zpátky, může jít například o důležité projekty, fotografie, či jakékoliv výsledky vlastní práce.

„Problémem zálohování i archivace je nedostatek vhodných časových oken, kdy by tyto operace měly být provedeny. Při obou činnostech může současně dojít jak ke snížení výkonu sítě či počítače, tak i k nedostupnosti některých dat. Vzhledem k tomu, že stále více aplikací je provozováno v režimu 24x7, zůstává jen minimum příležitostí realizovat zálohování či archivaci bez vlivu na ostatní prvky podnikové architektury.“ [3]

[2] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 11
[3] KYSELA, F., časopis Technik 12/2013, vydavatelství Economia, str. 40

3.4 Co je předmětem zálohování?

U otázky předmětu zálohování činí rozhodující faktor objem dat. Pokud objem dat určený k záloze nepůsobí problémy vzhledem ke kapacitě záložního média, lze v zásadě provádět pokaždé zálohu kompletně všech dat. Obnova pak bude daleko jednodušší a intuitivní. Musí se ale počítat s vyšší časovou náročností na celý proces zálohování i obnovy. V případě větších datových nároků je nutné data selektovat a vytvořit priority souborů. Hlavním důvodem pro tento krok je úspora úložného prostoru i času. Ve výsledku se sníží i celkové náklady na zálohování.

Data lze z hlediska důležitosti rozdělit do tří skupin. Nekritická data, jejichž ztráta nepřinese výraznější problémy a je možné je obnovit i v případě neexistence zálohy (např. instalace operačního systému, nebo jiného softwaru), stačí zálohovat jednou měsíčně. Do skupiny méně kritických dat patří data, *„která se mění jen velmi málo a změny se dají s určitou námahou dohledat a obnovit“* [4]. Do této skupiny patří například fotografie a videosoubory, se kterými je pracováno (střih videa apod.). Ty pak zálohujeme alespoň jednou týdně. Poslední skupinu tvoří data kritická, která se neustále vytváří a jejich ztráta může způsobit nestabilitu až pád systému. Kritická data se zálohují i několikrát denně (databázové soubory).

3.5 Strategie zálohování

Zálohovací strategie se obecně rozlišují podle toho, kam jsou zálohované soubory či data ukládána, jak je záloha velikostně náročná a jak často probíhá proces zálohování. Typicky se rozdělují do tří základních skupin. První je plné zálohování, což znamená, že jsou vytvářeny identické kopie například celého oddílu pevného disku. Druhý typ se nazývá přírůstkové zálohování, kdy se zálohují jen přírůstky nově vytvořených dat a poslední typ – rozdílové zálohování, které zálohuje všechna data, která byla aktualizována v době od posledního plného zálohování.

[4] ČANDÍK, Marek. Zálohování dat v informačních systémech. *Security magazín*. 2004, č. 2.

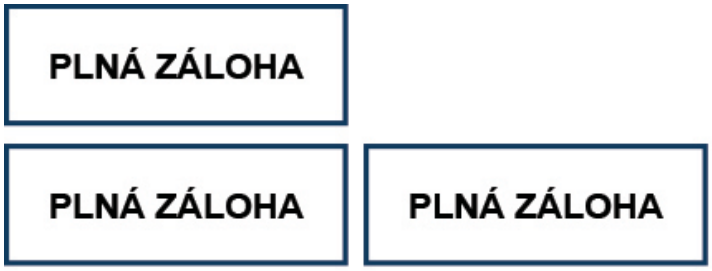
3.5.1 Plné zálohování (full backup)

„Plné zálohování je nejzákladnějším typem zálohování a hraje roli vlastně ve všech zálohovacích rutinách. V průběhu plného zálohování jsou uloženy do zálohového setu všechny zadané soubory bez ohledu na to, zda byly či nikoli v době od posledního zálohování aktualizovány“ [5].

Plné zálohování je například záloha všech důležitých souborů (například softwarové vybavení počítače), která po neočekávané havárii počítačových stanic umožní provést kompletní obnovení souborů a dat do původního stavu před provedením zálohy. Frekvence plného zálohování záleží na tom, jak jsou data důležitá a jak moc se nesmí dovolit o ně přijít. Doporučení je zálohu provést jednou týdně (u důležitých dat), nebo i v delším časovém horizontu, např. měsíčně (pro méně důležitá data, která v případě ztráty nahradíme jinými).

Nevýhoda spočívá ve velkém objemu zálohovaných dat, nicméně pokud se pracuje jen s omezenými datovými objemy, řádově několik desítek megabajtů, může se zálohovat každý den. Obnova je v takovém případě rychlá a bezproblémová.

Komentář k obrázku: každá plná záloha je samostatná a nezávislá.



Obrázek 1 - Schéma plného zálohování, vlastní tvorba

[5] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 17

3.5.2 Částečné zálohování

Tato strategie bere v potaz všechny soubory, které byly přidány nebo změněny od provedení poslední zálohy. Mezi částečnými zálohami lze rozlišovat dva základní druhy, přírůstkové neboli inkrementální a rozdílové neboli diferenciální. První z nich se zaměřuje na soubory přidané či změněné od poslední částečné zálohy, v druhém případě pak jde o všechny soubory změněné od poslední kompletní zálohy.

3.5.3 Přírůstkové zálohování (Inkrementální zálohování)

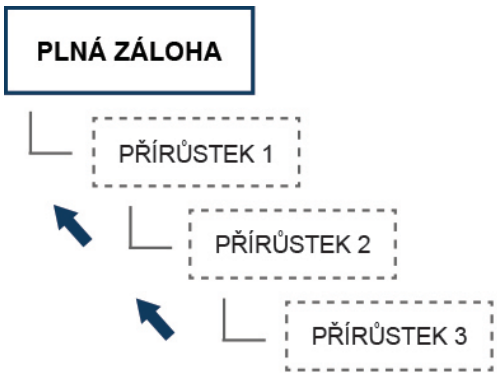
„Přírůstkové zálohování umísťuje aktualizovaná data (soubory) jako přírůstky zálohových setů“[6], která jsou ideální součástí plánu zahrnujícího kombinaci kompletních záloh a částečných denních záloh. V rámci této metody se kopíruje jen malé množství souborů. Velké množství denních inkrementálních záloh může znamenat pomalou a komplikovanou obnovu systému.

Podle M. Leixnera (2003, s. 18) přírůstkové zálohování zaznamenává jen poslední stavy aktualizovaných souborů, proto je důležité tuto metodu kombinovat s plným zálohováním, aby byla zajištěna skutečně stoprocentní ochrana. Mezi dvěma plnými zálohami je nezbytně nutné uchovávat stále všechny přírůstkové sety a to z důvodu, že pokud by došlo k poruše pevného disku, je možné obnovit jeho stav z doby, kdy byla provedena poslední plná záloha, poslední aktuální však jen tehdy, pokud jsou všechny navazující přírůstky k dispozici.

Přírůstkový mechanismus zálohování je prakticky nejrychlejší zálohovací metoda, jelikož se při tomto procesu vždy zaznamenávají jen poslední aktualizace známých vytvořených souborů.

[6] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 18

Komentář k obrázku: Pokud je poškozen přírůstek č. 2, přírůstek č. 3 není následně možné obnovit.



Obrázek 2 - Schéma přírůstkového zálohování, vlastní tvorba

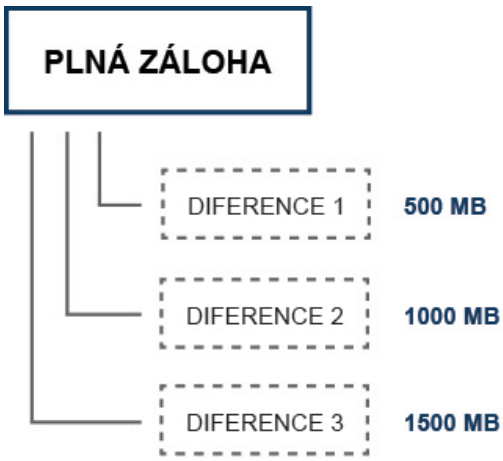
3.5.4 Rozdílové zálohování (Diferenciální zálohování)

„Rozdílové zálohování je metoda částečného zálohování.“[7] Diferenciální záloha zaznamenává změny a zálohuje všechny ty soubory, které byly aktualizovány v době od posledního plného zálohování. Na rozdíl od přírůstkové (inkrementální) zálohy, kde každý vytvořený záložní soubor je součástí tzv. "řetězce“, rozdílová záloha vytvoří zcela nezávislý soubor, který obsahuje všechny aktualizované soubory od vytvoření původní plné zálohy. V případě používání této metody není nutné udržovat v období mezi dvěma plnými zálohami všechny rozdílové sety, ale stačí vždy jen ten poslední. Diferenciální zálohy na sobě proto nejsou závislé a ani nezabírají příliš místa. Pokud se jedna z diferenciálních záloh poškodí, vliv na žádnou jinou diferenciální zálohu to nemá. *„K úplné obnově poškozeného disku totiž v tomto případě stačí poslední plná záloha a poslední rozdílový set.“*

Rozdílové zálohy jsou vhodné v případě, kdy jde o co nejkratší čas obnovy (protože se zde nemusí zpracovávat dlouhý řetězec předchozích přírůstkových záloh).

[7] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 18

Komentář k obrázku: Pokud je poškozena rozdílová záloha (diference) č. 2., nemá to žádný vliv na ostatní diference.



Obrázek 3 - Schéma rozdílového zálohování, vlastní tvorba

3.6 Zálohovací rutiny

Vzhledem k tomu, že existují 3 hlavní zálohovací metody – plná, přírůstková a rozdílová, je na každém uživateli aby si vybral, kterou bude volit nejčastěji či které bude kombinovat. Protože uživatel během času, který tráví u počítačové stanice, nepracuje se všemi daty, není potřeba zálohovat kompletně všechna data každý den či týden. *„Nejlepším řešením bude provádění denních částečných záloh, zaznamenávající pouze aktualizační změny či souborové přírůstky. Takové částečné zálohování však musí být součástí zálohovacího cyklu, v němž jsou plné zálohy prováděny podle naplánovaného schématu.“*[8]

Podle M. Leixnera (2003, s. 21) rozhodnutí o strategii zálohování představuje více než stanovení cyklu záznamů, musí se rozhodnout o nejvhodnější zálohovací metodě pro danou aplikaci a jaké je například optimum periody pro plné zálohování, či jak budou skladovány zálohovací média. Dále uvádí, pokud uživatel denně pracuje převážně se **stejnými soubory** (například se stejnými databázemi nebo s textovým editorem, kde není třeba uchovávat starší verze souborů), bude nejlepší volbou rozdílové zálohování. Následující obrázkové schéma je příkladem aplikace této strategie.

[8] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 21

Obrázek 4 - Schéma zálohovací rutiny 1, vlastní tvorba

Výhodou je udržení evidence jen dvou zálohových setů, jedním z nich je poslední plná záloha a druhý set obsahuje poslední rozdílovou zálohu - tím se zároveň šetří výrazně místo na zálohovacím médiu. Mezi další výhodu patří i zbavení nutnosti udržovat na disku více kopií téhož souboru, protože každá nová verze stejnojmenného souboru přepisuje verzi předešlou.

Leixner dále pojednává o přírůstkovém zálohování, jež spojuje s denní prací, která zahrnuje **různé soubory**. Tento způsob zálohování spojuje s prostředím, kde pravidelně vznikají nové soubory, například tam, kde se zpracovává kancelářská administrativa. Při této strategii je denní zálohování rychlejší a efektivní, jelikož se zaznamenávají jen naposled provedené změny. Soubory, které nebyly od posledního přírůstkového zálohování aktualizovány, nejsou do přírůstkových setů zahrnovány. Příkladem je následující obrázkové schéma.

Obrázek 5 - Schéma zálohovací rutiny 2, vlastní tvorba

Při této strategii je nutnost udržovat poslední plnou zálohu a všechny přírůstkové sety z důvodu záznamů nových aktualizací. *„Teoreticky platí, že nepracujete-li se stejnými soubory, je přírůstkové zálohování efektivnější, protože při něm neopakujete zaznamenávání ekvivalentních souborů.“*[9]

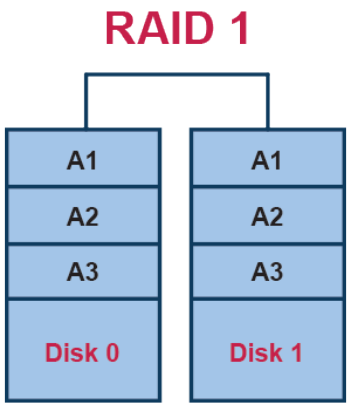
3.7 RAID

RAID je zkratka pro Redundant Array of Independent Disks. Jedná se o metodu propojení jednotlivých stejných disků v jeden celek pomocí diskového řadiče tak, že je počítač chápe jako jeden pevný disk. Spojení disků je označováno obecně jako dynamický disk. Účelem takového propojení ve většině případů není zvýšení kapacity, ale zajištění rychlejšího přístupu k datům a vyšší bezpečnosti dat. Té je dosaženo díky tzv. redundanci (nadbytečnosti) dat. *„Při havárii se z nadbytečných dat doplní chybějící údaje na vadném disku.“*[10] RAID se vyskytuje v mnoha typech, které se označují různými čísly (nejčastější 0, 1, 5 ale i méně časté 2, 3, 10, 53 apod.). Ne všechny tyto metody se hodí pro zálohování a ne všechny lze zavést s nízkými náklady. Ze všech typů RAID na základě kritéria kompletnosti zálohy (lze obnovit všechna data) bude v následujících odstavcích pojednáno o dvou typech polí a sice RAID 1 a RAID 5.

[9] LEIXNER, Miroslav. PC - zálohování a archivace dat. V Praze: Grada, 1993. (Nestůjte za dveřmi) ISBN 80-854-2473-8, s. 25.
[10] HORÁK, Jaroslav. Hardware: učebnice pro pokročilé. 3. aktualiz. vyd. Brno: CP Books, 2005, 344 s. ISBN 80-251-0647-0, s. 193.

3.7.1 RAID 1

Takovéto pole musí být tvořeno nejméně dvěma stejnými disky. Na oba tyto disky se pak ukládají stejná data, tj. jeden disk je naprostou kopií (zálohou) druhého. Tím je dosaženo 100% redundance, což můžeme označit za největší výhodu tohoto typu RAID. Dochází k vysoké ochraně dat, která jsou v případě poruchy jednoho z disků velmi snadno obnovitelná. Zároveň se zvyšuje rychlost čtení. Mezi hlavní nevýhody bezesporu patří hardwarová náročnost, neboť potřebujeme dvojnásobnou kapacitu disků oproti objemu dat.

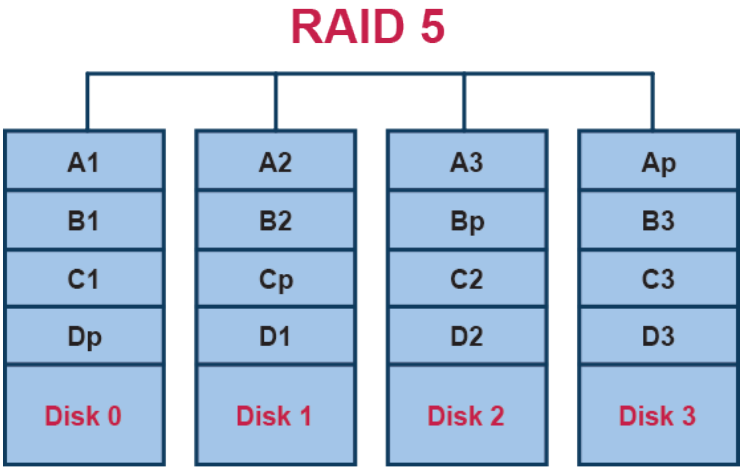


Obrázek 6 - Schéma pole RAID 1, vlastní tvorba

3.7.2 RAID 5

Na vytvoření tohoto typu RAID je zapotřebí nejméně tří disků, v praxi se však využívá alespoň čtveřice. *„První část dat se rozdělí na tři části a ty se zapíší na tři disky. Na čtvrtý disk se pak zapíše tzv. paritní informace. Je to jakási záloha všech tří datových sektorů. Stejně tak se zapíše další sektor dat s tím, že paritní informace se zapíše na jiný disk a takto se to střídá.“* [11]Tím se dosáhne vysoké bezpečnosti dat v případě výpadku jednoho z disků. Nepočítá se však s výpadkem více z nich. Znovu se disponuje vysokou rychlostí čtení, ale kvůli vytváření paritních informací nižší rychlostí zápisu.

[11] ŠULC, Tomáš. VelociRaptory a RAID - když jeden nestačí. PCTuning [online]. 2008 [cit. 2014-11-04]. Dostupné z: http://bit.ly/1EO0vDe



Obrázek 7 - Schéma pole RAID 5, vlastní tvorba

Obecně lze konstatovat, že využití RAIDU jako metody zálohování může být dobrá alternativa, neboť zálohována budou neustále všechna data. Tato metoda jednoznačně neodstraní všechny faktory, které mohou způsobit ztrátu dat, protože disky spojené v RAID se nachází na stejném místě. Důležité je i zdůraznit, že zprovoznění samotného RAIDového diskového pole není jednoduchý proces, proto vyžaduje jisté zkušenosti a schopnosti.

3.8 Jak správně zálohy skladovat?

Zálohovaná data by neměla být uložena na stejném místě (například na pracovišti, v budově), kde se nachází zálohované stanice. Jsou pro to jisté důvody, jedním z nich je hrozba požáru. Pokud nejsou datová média zabezpečena v nehořlavých skříních, jednoduše dojde ke zničení zálohovaných setů. Jedním z dalších případů může být např. krádež médií. U společností by měl být postup takový, že přístup k zalohovacím setům by měl být omezen a záložních kopií by se mělo tvořit více kusů, z nichž každá by měla být uložena na jiném místě, protože z principu obsahují zálohy citlivější data a proto je při jejich úschově nutno dbát přísnějších pravidel.

3.9 Nejdůležitější rozhodovací kritéria

3.9.1 Kapacita

Mezi nejhlavnější kritérium datového nosiče patří kapacita. Jedná se o virtuální a zároveň fyzické místo na nosiči či médiu, které slouží k zápisu dat. Při výběrů nosiče se musí brát ohled nejen na aktuální nároky, ale také zdali budou nároky na kapacitu do budoucna narůstat či nikoliv.

3.9.2 Rychlost zápisu a čtení

Pro účel zálohování dat je zcela jasně rozhodující faktor rychlost zápisu, který určuje, jaký objem dat je možné zapsat za jednu časovou jednotku na dané médium (MB/s apod.). Vysoká rychlost čtení bude oceněna v případě u datového nosiče, z něhož je záloha prováděna a především při obnově zálohy. Určuje, jaký objem dat je možné přečíst z daného média za časovou jednotku, tedy, přesný opak zápisu. V obou případech se nároky odvíjí od velikosti zálohy. V případě zálohy pouze velmi malého objemu dat se mohou tyto faktory prakticky ignorovat. Dalším důležitým faktorem ovlivňující rychlost média je přístupová doba. Nejčastěji se udává v milisekundách a stanovuje dobu, kterou je nutné čekat od zadání požadavku až ke zpřístupnění požadované informace.

3.9.3 Životnost a odolnost

Životnost a odolnost média je jeden z důležitějších faktorů při rozhodování výběru média. Princip zálohy spočívá v tom, že předchozí zálohu přemažeme zálohou novou. Zpravidla používáme stejný datový nosič a dochází k opotřebovávání. Z hlediska zálohování se doporučuje datové nosiče ve vhodných časových intervalech (kratších než je udávaná životnost) obměňovat, aby nedošlo ke zpomalení či znemožnění procesu zálohování nebo obnovy dat.

Otázka odolnosti pak velmi blízko souvisí s životností nosiče. V podstatě se jedná o schopnost nosiče chránit zapsaná dat vůči vnějším vlivům. Nejčastěji hovoříme o pádech nebo jiných mechanických poškozeních, či dokonce o zmagnetizování. Odolnost i životnost média lze prodloužit vhodným skladováním.

3.9.4 Cena

Přes všechna výše uvedená kritéria bude nakonec nejdůležitějším rozhodovacím faktorem cena, která se odvíjí od ceny za jednotku úložného prostoru (nejčastěji se udává za 1MB). Prakticky platí lineární závislost, čím více prostoru bude potřeba, tím větší bude pořizovací cena zálohovacího média.

3.10 Rozdělení datových nosičů (médií)

Obecně lze rozdělit datová média pro zálohování do více skupin. První je skupina, která zahrnuje magnetická média, jako jsou například magnetické pásky nebo pevné disky, další skupinou jsou optická média, například DVD nebo Blu-ray. Třetí skupinu zaujímají flash paměti, do které patří například SSD disky a poslední skupinu zaujímají Cloudová úložiště (Dropbox či Google Drive) a webové servery pro úschovu dat (Rapidshare či Ulož.to).

3.10.1 Magnetická média

3.10.1.1 Magnetická páska

Základním zástupcem magnetického nosiče je magnetická páska. Počátky vývoje magnetických pásek směřují do třicátých let 20. století. Běžná situace, kdy se lze setkat s magnetickou páskou, je videokazeta. Lze ji však použít i pro uchování enormního množství dat. Samotná páska se většinou vyrábí z polyesterové fólie nebo materiálu s podobnými vlastnostmi, na kterou je z jedné či z obou stran nanesena magnetická vrstva. Může dosahovat délky až několik stovek metrů, tím je dosaženo kapacity pro ukládání velkého množství dat. Páska je navinuta na cívce a s výjimkou pásek do starých sálových počítačů je obvykle schována v kazetě. Pod čtecí a záznamovou hlavou se pohybuje stálou rychlostí.

„Zápis probíhá tak, že se podélně či příčně zmagnetizuje malá část pásky pomocí zápisové hlavy, která obsahuje cívku s jádrem přerušeným úzkou mezerou. Mezera se nachází přesně nad páskou, tj. magnetické pole se v tomto místě z jádra rozšiřuje i přes pásku, jejíž magnetické dipóly jsou tímto polem natáčeny žádoucím směrem a díky tomu, že oxid železitý je feromagnetická látka, je orientace dipólů zachována i po posunu dotčeného

místa pásky mimo dosah zápisové hlavy“.[12] Čtení pak probíhá analogicky. Páska vybudí ve čtecí hlavě impuls napětí, z jehož polarity je zjištěno, jaká bitová informace byla zapsána (0,1). Nevýhodou pásek je, že jsou na ně data ukládána sekvenčně. To znamená, že pokud je zapotřebí z pásky přečíst konkrétní data, musí se počkat, než se páska natočí do požadované pozice.

3.10.1.2 Pásková knihovna

Magnetické pásky se obvykle sdružují do tzv. páskových knihoven. Ty jsou hojně využívány středními a většími společnostmi. Jedná se o automatizovaná zařízení, ve kterých je uloženo velké množství pásek, z nichž většina je v nečinnosti. V případě potřeby vyvolání dat z jakékoliv pásky systém danou pásku najde, zavede do čtecího slotu a data zpřístupní. Formátů magnetických pásek je celá řada, proto je stručně pojednáno jen o nejpoužívanější pásce druhu LTO.

Obrázek 8 - Pásková knihovna[13]

Páska Linear Tape Open (LTO) vznikla ve spolupráci firem IBM, Hewlett-Packard a Seagate. *„LTO je otevřený formát pro páskové produkty určené pro použití ve středně náročných systémech.“* [14]

[12] TIŠNOVSKÝ, Pavel. Magnetické paměti pro trvalý záznam dat. ROOT.CZ [online]. 2008 [cit. 2014-11-08]. Dostupné z: http://bit.ly/1DpJeyF
[13] FUJITSU. [online]. 2014 [cit. 2014-11-08]. Dostupné z: http://bit.ly/1DMATmb
[14] Linear Tape Open. Wikipedia. [online]. 2014 [cit. 2014-11-08]. Dostupné z: http://cs.wikipedia.org/wiki/Linear_Tape_Open

Přístupová doba se pohybuje mezi 50 - 70 sekundami. Důležitý z hlediska ochrany dat na magnetických páskách je systém WORM, který umožňuje nastavit ochranu proti přepisu či smazání jednotlivé pásky. Díky poměrně dlouhé životnosti (cca 15 - 30 let) se magnetické pásky využívají i pro archivaci. Nebezpečí této technologie spočívá (ostatně jako u všech magnetických medií) v působení silného magnetického pole, které by mohlo médium poškodit.

3.10.1.3 Pevný disk

Jednoznačně nejznámější zařízení pro běžné zálohování a uchování dat je pevný disk, zkratkou HDD (z anglického jazyka Hard Disk Drive). Jako první ho vyvinula společnost IBM již v roce 1956.

Zjednodušený princip fungování pevného disku lze popsat jako systém plotny, neboli kotouče, na kterém jsou zapisována data. Magnetická hlava má pak za úkol zapsání a čtení dat. Ploten je v pevném disku několik (jejich počet je jedním z faktorů kapacity). Hlavy jsou umístěny z obou stran, protože zápis je prováděn z obou stran plotny. Kotouče se otáčí, nejčastěji maximální rychlostí 7200 otáček za minutu. Jsou ale i serverové disky s rychlostí až 10000 otáček za minutu, či naopak disky do notebooků, kde plotny dosahují rychlosti maximálně 5400 otáček za minutu z důvodu snížení spotřeby baterie notebooku. Standardní provedení disku má velikost 3,5 palce, v případě disků v noteboocích nebo v případě SSD disků mohou mít rozměr 2,5 palce. Hlava není v kontaktu s plotnou, nicméně se nad ní pohybuje ve vzdálenosti pár mikrometrů. *„I drobné zrnko prachu mezi hlavou a plotnou by pak mohlo způsobit rýhu v disku a znehodnocení dat.“*[15] Proto je disk uložen ve vzduchotěsném pouzdře.

[15] HORÁK, Jaroslav. Hardware: učebnice pro pokročilé. 3. aktualiz. vyd. Brno: CP Books, 2005, 344 s. ISBN 80-251-0647-0, s. 132.

Obrázek 9 - Pevný disk[16]

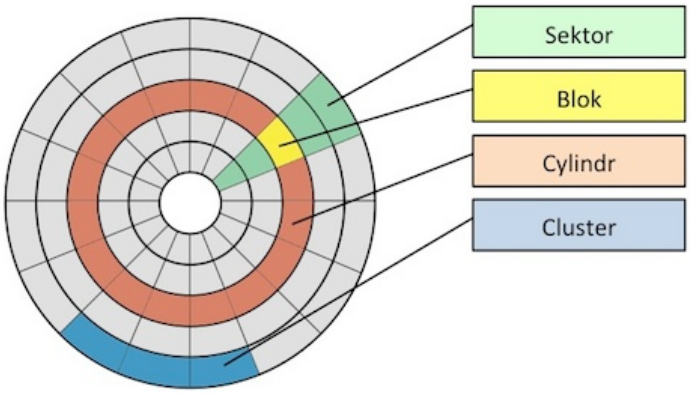
Pevný disk je vzhledem ke své konstrukci citlivé zařízení, o které se musí uživatel dobře starat, aby dlouho sloužilo. *„Obecně se doporučuje nevystavovat pevný disk prašnému prostředí, kouři, otřesům a prudkým výkyvům teplot a napětí.“*[17]

Kapacita pevných disků je různá, běžný počítačový standard dnešní doby je 500 GB až 1 TB, cenově je zařízení velmi dostupné. Existují samozřejmě i větší disky, například firma Western Digital má aktuálně k měsíci listopad (2014) největší disk o velikost 6 TB, avšak cena se pochopitelně pohybuje velmi vysoko. Oproti magnetickým páskám vyniká rychlostí čtení dat. Hlava nemusí postupně projíždět povrch celé plotny, aby našla, kde se daná informace nachází. Plotna je rozdělena do stop a sektorů. Řadič pak po požadavku operačního systému posune hlavu do přesné polohy (tzv. doba vystavení). Dále už řadič jen vyčká na vhodné natočení plotny (tzv. doba čekání). V tu chvíli jsou data nalezena a přečtena. To vše se odráží v nízké přístupové době pevného disku, která je rovna součtu doby vystavení a doby čekání a u disků se 7200 otáčkami za minutu se pohybuje do 9 milisekund. Mezi důležité údaje patří také teoretická přenosová rychlost, která se liší podle použitého rozhraní. Dnešní nejmodernější rozhraní SATA 3.0 nabízí teoretickou přenosovou rychlost až 6 Gbit/s čili 600 MB/s. Vyrovnávací paměť (cache) pevného disku bývá většinou 32 nebo 64 MB a slouží k vyrovnání rychlosti mezi pomalejším a rychlejším zařízením. Rychlejší zařízení je zpravidla procesor a pomalejší médium je pevný disk.

[16] Pevné disky. Alza.cz. [online]. 2014 [cit. 2014-11-08]. Dostupné z: http://bit.ly/1A14r0N
[17] MINASI, Mark. Pevné disky od A do Z. Praha: Grada, 1992, 471 s. ISBN 80-856-2335-8, s. 197.

„Cache umožňuje uchovat data určená k zápisu do doby, než je mechanická část pevného disku zpracuje a naopak uchovává přečtená data do doby, než je obsluha přerušení přesune do operační paměti.“ [18]

Z hlediska zálohování na pevný disk je považováno za standard provádět zálohy na disky interní a externí. Provádět zálohu na interní disk, ze kterého data pochází, se nemusí vyplatit. A to ani v případě, kdy je disk rozdělen do více logických jednotek neboli oddílů. Tento způsob neeliminuje téměř žádné faktory, které ohrožují data. Jen o něco málo lepší způsob zálohování je, pokud je zálohováno na další interní pevný disk stejného počítače. Ještě lepší způsob je zálohovat data na pevný disk v jiném počítači, například na síťový disk anebo na disk externí. Provádět zálohy na externí pevný disk je jedna z nejjednodušších, poměrně levných a spolehlivých metod zálohování. Data se nacházejí v jiném prostředí, a pokud je zároveň zamezen přístup nežádoucích osob k datům, je eliminována většina nebezpečí ztráty dat.



Obrázek 10 - Schéma plotny pevného disku[19]

[18] Cache. Wikipedia. [online]. 2014 [cit. 2014-11-08]. Dostupné z: http://cs.wikipedia.org/wiki/Cache
[19] Pevný disk – schéma plotny pevného disku. GJSZlin. [online]. 2014 [cit. 2014-11-08]. Dostupné z: http://bit.ly/1LYrPkm

3.10.2 Optická média

Zálohování na optická média v dnešní době je zcela již proces ne úplně efektivní. Optická média jsou postupně nahrazována flash (USB) pamětmi, externími HDD nebo způsobem zálohování dat na cloud úložiště. Následující kapitola vymezuje jednotlivé nejpoužívanější typy optických médií.

3.10.2.1 CD (Kompaktní disk)

Kompaktní disk byl vyvinut společnostmi Sony a Philips roku 1979 za účelem věrného uchování a reprodukce hudby. Jeho průměr byl 120 mm s 15 mm otvorem uprostřed a své rozměry si zanechal do dnes. *„Na rozdíl od pevných disků není povrch kompaktního disku rozdělen do sektorů, ale tvoří ho pouze jedna spirálovitá stopa (podobně jako gramofonová deska), která se odvíjí od středu disku směrem ven.“*[20]

Spirála je umístěna pouze na spodní straně disku (na disk jsou data zapisována z jedné strany). Její délka dosahuje 5-6 km a je rozdělena do bloků, které obsahují buď pole (landy), nebo tzv. pity (prohlubně nestejné délky). Při zápisu dat laserový paprsek o vlnové délce 780 nm projde z čtecí hlavy přes polopropustné zrcadlo a soustavu čoček až na disk. V případě, že narazí na pole, odrazí se zpět. Pokud se potká s pitem, je paprsek rozptýlen. Přes čočky a zrcátko se paprsek vrací zpět až na fotodiodu, která ho převede na elektrické impulsy, které jsou dále zpracovány v řadiči disku.

Kapacita kompaktních disků je různá, přičemž nejčastější standard je 700 MB. Přístupová doba se pohybuje mezi 100 až 300 ms. Přenosová rychlost se odvíjí od původního standardu Červená knihy (Digital Audio) ISO 908 – (tedy 150 kB/s) a je uváděna v násobcích rychlosti otáčení (1x je 150 kB/s, 4x je 600 kB/s atd.).

Od původního kompaktního disku byla rozvinuta celá řada obdobných formátů, které vylepšovaly původní disk. Příkladem je například formát CD-R, který umožňuje i zápis na disk nebo CD-RW, který umožňuje i přepis již zapsaných dat (tedy i jejich smazání).

[20] HORÁK, Jaroslav. Hardware: učebnice pro pokročilé. 3. aktualiz. vyd. Brno: CP Books, 2005, 344 s. ISBN 80-251-0647-0, s. 206.

3.10.2.2 DVD

Princip DVD je v základu věci stejný jako u CD a vychází z požadavku na zpětnou kompatibilitu s CD. Jeho rozměry se neliší a jsou stejné jako u CD. Základní a podstatný rozdíl lze však spatřit v kapacitě disku. Ta byla zvětšena až na maximálních 17,1 GB u dvouvrstvého oboustranného média (z důvodu zvýšení hustoty zápisu, čímž došlo zmenšením rozestupu landů a zkrácením délky pitu). U standardu DVD-5 je velikost udávána 4,7 GB (použitelné místo je však vždy menší, v tomto případě je 4,38 GB). DVD-9 nabízí jednu stranu zápisu a dvě vrstvy – udávaná velikost je 8,5 GB, použitelná pak 7,92 GB. DVD-10 nabídne dvě strany a z každé strany jednu vrstvu, použitelné místo k zápisu je 8,75 GB, oficiálně udávaná kapacita je udána 9,4 GB. DVD-14 disponuje velikostí 13,2 GB (použitelná pak 12,3 GB) – zápis proveden z obou stran disku - jedna strana má dvě vrstvy, druhá strana má poté jednu. Poslední standard DVD-18 má dvě strany a z každé strany nabídne dvě vrstvy, kapacita je 17,1 GB (použitelná je 15,9 GB)

Přístupová doba u DVD je okolo 250 ms. Přenosová rychlost mechaniky typu DVD se udává jako násobek základní rychlosti 1350 kB/s. Mechanika s rychlostí 8× umožňuje teoretickou přenosovou rychlost až 8 × 1350 = 10800 kB/s (tj. 10,8 MB/s). DVD se postupně vyvinulo tak, že se na trhu objevily přepisovatelné disky (DVD-RW a DVD+RW) a disky dvouvrstvé (DL – Dual Layer) a také oboustranné.

„*Za nejvhodnější typ pro práci s daty a pro zálohování uvádějí autoři knihy HW pro pokročilé DVD+RW. Jako hlavní výhody zmiňují vysokou rychlost přepisu, dobrou dostupnost a možnost ho zhruba tisíckrát přepsat*.“ [21]

Konkurence DVD+RW je DVD-RAM a hlavní výhoda, kterou konkuruje je možnost přepisu až sto tisíc krát. Mezi další výhody patří, že sním lze při kompatibilním systému zacházet jako s dalším pevným diskem. Z tohoto pohledu se DVD-RAM jeví jako ideální volba pro zálohování.

[21] HORÁK, Jaroslav. Hardware: učebnice pro pokročilé. 3. aktualiz. vyd. Brno: CP Books, 2005, 344 s. ISBN 80-251-0647-0, s. 230.

3.10.2.3 Blu-ray

„V Blu-ray mechanikách se používá laserové světlo s vlnovou délkou 405 nm.“[22] Podle jeho modro fialové barvy je disk také pojmenován. Disky umožňují záznam dat s celkovou kapacitou až 25 GB u jednovrstvého disku, 50 GB u dvouvrstvého disku až po 80 GB u oboustranné dvouvrstvé varianty. Nyní jsou sice schváleny i nové standardy, které umožní další zvětšování kapacity, tyto disky již však nebudou zpětně kompatibilní a bude zapotřebí nákup nové mechaniky. Prosazení na trhu tedy bude složité.

3.10.3 Shrnutí optických médií

Jelikož počítačové stanice bývají nějakým typem CD/DVD/Blu-Ray mechaniky schopné zápisu vybaveny, dá se konstatovat, že zálohování na optická média bude pravděpodobně cenově nejdostupnější metoda zálohování pro běžného uživatele. Z hlediska podniků a veřejné správy rozhodně ne. Zároveň je ale třeba si uvědomit, že životnost těchto médií v podstatě není osvědčena. Údaje, které uvádí výrobci (až 100 let), mohou být zavádějící. Dle současných zkušeností lze říci, že životnost přepisovatelných médií je obecně mnohokrát nižší, než u médií pouze zapisovatelných a může se pohybovat řádově pouze v rocích. Na trhu je možné zakoupit média přímo určená pro archivaci, jejichž životnost by měla být vyšší. Hlavním nepřítelem optických médií je nesprávná a častá manipulace s nimi, kdy dochází k opotřebení a zejména poškrábání povrchu disku. To přímo souvisí i se špatným skladováním (nepoužitím obalu apod.).

[22] TIŠNOVSKÝ, Pavel. Vývoj optických pamětí: od DVD k Blu-ray. ROOT.CZ [online]. 2008 [cit. 2014-12-02]. Dostupné z: http://bit.ly/1Iorkk2

3.10.4 Flash paměti a SSD disk

Zvláštním druhem média, jehož spektrum použití je velmi široké a rozšířené, jsou flash paměti a SSD disky. S komerční výrobou flash disků přišly v roce 2000 firmy IBM a Trek Technology. Původně byla média určená pro ukládání dat například v digitálních fotoaparátech (ve formě paměťových karet). Ovšem díky jejich výhodám se rychle rozšířily a začaly se používat jako datová úložiště. Mezi hlavní výhody patří absence mechanických součástek, vysoká spolehlivost, skladnost, přenosnost, stále klesající cena, rostoucí kapacita a velmi rychlá přístupová doba (obvykle do 1 ms). *„Hlavní nevýhodu tvoří především omezený počet zápisů, který je způsoben degradací jednotlivých paměťových buněk.“*[23]. V porovnání s životností magnetických pevných disků se moderní flash paměti mohou směle rovnat. SSD disky v brzké době pravděpodobně převezmou dosavadní postavení magnetických pevných disků. Disponují oproti nim nižší spotřebou, větší rychlostí zápisu a čtení, nižší hmotností a větší odolností. Kapacitně se nejnovější modely rovnají magnetickým pevným diskům. Pro zvolení systémového disku je skvělá volba právě SSD disk, který zajistí uživateli opravdu velmi rychlý chod počítačové stanice. Za hlavní překážky bránící širšímu použití lze považovat především stále vysokou cenu, která u vysokých kapacit dosahuje i 10 násobků ceny magnetických disků a také výše zmíněný problém s omezeným počtem přepisů, který by však měl výraznější dopad pouze při využití disku jako síťového.

Obrázek 11 - Porovnání HDD a SDD [24]

[23] BRECHLEROVÁ, Dagmar. Paměťová média: Informace, aneb kam s nimi. Security magazín. 2012, č. 2, s. 14-18.

[24] HORT, Tomáš. Technologie a zajímavosti z oblasti SSD disků II. PCTuning. [online]. 2014 [cit. 2014-12-02]. Dostupné z: http://bit.ly/1IojB5A

3.10.5 Cloudová úložiště

Cloudové služby se považují za budoucnost zálohování dat v informačních technologiích a v posledních letech zaznamenaly obrovský nárůst. Princip je jednoduchý, uživatel pomocí svého internetového prohlížeče nebo speciálního softwaru nahraje svá data na cloud server, kde jsou následně uložena a mohou být zobrazena (využita) z jakéhokoliv zařízení po celém světě, kde je možné se připojit k internetu. Přístup je možný ze všech zařízení, pro které je daný software vyvinutý. Lze pohodlně spravovat své dokumenty ze smartphonů, tabletů a počítačů.

„Záloha na lokální disk nebo DVD je snadná a rychlá, internetová úložiště tedy musí nabídnout mimořádně kvalitní a bezpečné služby, pokud chtějí těmto médiím konkurovat.“[25]

Mezi nejznámější cloudové úložiště patří Dropbox, Google Drive, OneDrive, Box, Apple iCloud Drive, Amazon Cloud Drive, SugarSync a další. Tabulka níže přináší porovnání 5 nejvíce známých cloud úložišť.

Dropbox	Google Drive	OneDrive	Apple iCloud	Box
basic 2 gb / free	basic 15 gb / free	basic 15 gb / free	basic 5 gb / free	personal 10 gb / free
dropbox pro 1 tb / 9,99€	100 gb / 1,99$	100 gb / 1,99$	20 gb / 0,99$	starter 100 gb / 4€
for business no limit / 12€	1 tb / 9,99$	200 gb / 3,99$	200 gb / 3,99$	business no limit / 12€
	10 tb / 99,99$	1 tb / 6,99$	500 gb / 9,99$	enterprise customised
	20 tb / 199,99$ 30 tb / 299,99$		1 gb / 19,99$	

Obrázek 12 - Porovnání cloudových úložišť, vlastní tvorba

[25] SCHREIBER, Manuel. Data v cloudu. CHIP. 2011, č. 12, s. 58-61.

Faktory typu kapacita a cena samozřejmě netvoří jedinou část, která může při rozhodování ovlivnit výběr úložiště. Jeden z dalších důležitých faktorů je například podpora rozdílných platforem tak, aby bylo možné se ke svým zálohám dostat přímo ze smartphonu s nejrůznějším systémem (iOS, Android, Windows Phone apod.). Dalším faktorem může být dostupnost software, který správu záloh usnadní a dovolí tak spravovat účet bez využití klasického webového rozhraní. Obecně se zálohování pomocí cloudových služeb jeví jako velmi výhodné (základní tarify jsou bezplatné, také není třeba nákup nové technologie). Je třeba ale upozornit na potenciální rizika, která se při tomto způsobu zálohování dat mohou vyskytnout. Zaprvé se neví, kdo má k daným zálohám přístup. Dále není zcela určité, po jaký časový horizont bude společnost danou službu poskytovat. V neposlední řadě se jeví riziko v odposlechu dat během přenosu.

3.10.6 Webová úložiště pro zálohování dat

Webová úložiště jsou v dnešní době už poměrně zastaralý trend, jak zálohovat svá data. Jedná se o webové servery, která nabízí možnost zálohy dat (někdy až po zaplacení určitého tarifu). Tarif je cena, která je úměrná prostoru, který lze využít k zálohám. Tedy například kdysi velmi populární server Rapidshare.com nyní nabízí tarif o velikosti 300 GB za měsíční poplatek zhruba 1400 Kč. Větší tarif 700 GB je pak za měsíční cenu 2800 Kč. Dalším zajímavým webovým storagem je MEGA.co.nz, jedná se o následníka Megauploadu, který byl nejrozšířenějším a nejznámějším úložištěm pro nelegální obsah.

„V lednu 2012 byl zakladatel Megauploadu Kim Dotcom zatčen novozélandskou policií v souladu s americkým obviněním za porušení autorských práv na jeho stránkách Megaupload.“ [26] MEGA (nástupce Megauploadu) nabízí nyní základní úložiště o velikosti až 50 GB, které je zdarma, dále je možné si zaplatit daleko větší prostor, například až úctyhodných 4 TB. Webových úschoven existuje další celá řada, zde jsou zmíněny jen ty nejvíce známé. Protože novějším trendem jsou již Cloudová úložiště, je na zamyšlenou, jaká budoucnost čeká webové úschovny, jestli si udrží nějakou pozici, nebo ne. Vzhledem k tarifním cenám je zřejmé, že například u služby Rapidshare se absolutně nevyplatí využívat služby, výhodnější je již zakoupit samostatný pevný disk či SSD disk.

[26] Kim Dotcom. Wikipedia. [online]. 2014 [cit. 2015-03-07]. Dostupné z: http://cs.wikipedia.org/wiki/Kim_Dotcom

3.11 Software pro zálohování digitálních dokumentů

Data, která jsou potřeba zálohovat, mohou být zálohována buď na externí pevné disky, nebo do cloudu, nebo na další média. Tento proces lze automatizovat pomocí určitého softwaru, který může zajistit velmi efektivní zálohování při správném nastavení programu. Mezi nejvíce dostupně známý software se řadí Acronis True Image a Cobian Backup.

3.11.1 Acronis True Image

Produkty se jménem Acronis patří právem mezi špičku zálohovacích systémů a nástrojů pro správu diskových oddílů. Tento software umožňuje vytvářet a ukládat zálohy celého počítače, souborů a adresářů lokálně a do cloudu. Zálohovací proces spočívá v zálohování aplikací (včetně obsažených dat), nastavení a uživatelských účtů. Disponuje snadnou obnovou, systém lze vrátit do jakéhokoliv bodu, ve kterém byl stabilní, jednoduše a bez instalace operačního systému. Obnovení zálohy celého PC lze provést kdykoli na jakýkoli jiný, hardwarově odlišný PC. Obnovit zálohu, která je uložená na externím disku, síťovém úložišti či na FTP serveru lze i bez instalace softwaru. Zálohovat lze i při vypnutí počítače, objem dat o velikosti 120 GB lze zálohovat zhruba kolem 8 minut. Software je dostupný v české lokalizaci.

3.11.2 Cobian Backup

Na internetovém portálu živě.cz byly zmíněny programy, které nabízí nejlepší funkce pro zálohování. Jedním z nejčastěji doporučovaných programů je Cobian Backup. „S*louží k zálohování souborů a adresářů z jejich původních umístění na jiné. To zvládá jak na stejném, tak i na vzdáleném počítači v síti. Podporuje i zálohování na FTP, kompresi ZIP pro úsporu místa v úložišti, šifrování, zálohování na více míst zároveň a pochopitelně zautomatizování zálohování v nastaveném čase. Hlavní přednosti programu jsou česká lokalizace, jednoduchost a zároveň funkčnost a cena.“*[27]

[27] KRAUS, Josef. Sedm nejlepších programů pro zálohování dat. ZIVE. [online]. 2012 [cit. 2014-12-18]. Dostupné z: http://bit.ly/1unEBD4

3.12 Software pro šifrování médií

3.12.1 TrueCrypt

TrueCrypt je kryptografický software, jež umí vytvořit virtuální disky, které jsou chráněné šifrováním (například AES). Tyto virtuální disky lze tvořit v souboru, anebo lze zašifrovat celý (systémový) disk. Virtuální disky v souboru lze libovolně přenášet, případně i nahrát na cloudová úložiště. Primární oddíl je navržený takovým způsobem, že při zápisu souborů sdílí prostor se skrytým oddílem, takže může data ve skrytém oddílu zároveň přepsat.

V průběhu vytváření disku je celý prostor zaplněn náhodnými daty, což znemožňuje zjistit, kolik souborů disk obsahuje a zda například obsahuje anebo neobsahuje skrytý oddíl.

V současnosti jsou známé pouze čtyři efektivní útoky na TrueCrypt: Slabé slovníkové heslo, získání hesla od uživatele (výslechem, kamerovým záznamem jeho práce s počítačem či softwarem monitorujícím stisky kláves), získání klíče z hibernačního souboru (pokud byl systém hibernován ve chvíli, kdy měl TrueCrypt připojené chráněné disky) anebo použití nástroje, který dokáže správný klíč získat z paměti počítače. Takový útok nabízí například Passware, který se chlubí tím, že dokáže prolomit jak BitLocker, tak TrueCrypt „během pár minut".[28]

V rámci použití softwaru TrueCrypt je možné zajistit velmi efektivní šifrování citlivých dat na pevných discích, na flash pamětích či na SSD discích za účelem předejití neoprávněného zneužití dat při ztrátě či krádeži těchto paměťových médií. Výhodou je možnost využití širokého spektra datových médií. Mezi nevýhody lze zařadit konec podpory softwaru (aktualizace) ze strany vývojářů v červnu 2014.

[28] RYBKA, Michal. Kdo a proč zabil TrueCrypt? PCTuning. [online]. 2014 [cit. 2015-03-07]. Dostupné z: http://bit.ly/1GsTvMO

3.12.2 BitLocker Drive Encryption

Šifrovací sofware od firmy Microsoft jménem BitLocker Drive Encryption lze použít pro ochranu všech digitálních souborů uložených na jednotce, kde je nainstalován operační systém Windows (7 a novější) a na ostatních pevných datových jednotkách, jedná se tedy například o interní disky. Dále je možné zajistit i ochranu na USB flash pamětích pomocí nástroje BitLocker To Go.

BitLocker **šifruje celou jednotku**. To je velmi dobrá funkce, která může zabránit hackerům v přístupu k systémovým souborům, které používají k získání hesla.

„Po přidání nových souborů na jednotku šifrovanou pomocí nástroje BitLocker je tento nástroj automaticky zašifruje. Soubory zůstávají šifrovány, pouze pokud jsou uloženy na šifrované jednotce. Soubory zkopírované na jinou jednotku nebo počítač jsou dešifrovány. Pokud sdílíte soubory s jinými uživateli, například pomocí sítě, jsou soubory uložené na šifrované jednotce šifrovány, avšak pro oprávněné uživatele jsou normálně přístupné.“[29]

Pokud je zašifrován oddíl operačního systému, BitLocker během zavádění OS zkontroluje, jestli jsou procesy spouštěcích a zaváděcích souborů správně zabezpečeny. Pokud je narušena bezpečnost, BitLocker ihned zamkne systémový oddíl a následně k jejímu úspěšnému zavedení potřebuje speciální obnovovací klíč, který je vytvořen po prvním spuštění nástroje BitLocker v rámci konfigurace. Pokud klíč není vytvořen, dojde k trvalému ztracení přístupu k datům.

Nevýhodou BitLockeru je zcela jistě podpora OS. Možnost jeho využití se vztahuje jen na Windows 7 Ultimate a Enterprise a dále na novější Windows 8.

[29] Microsoft. Ochrana souborů nástrojem BitLocker Drive Encryption. [online]. 2013 [cit. 2015-03-07]. Dostupné z: http://bit.ly/1BdAniA

3.13 Ochrana a uchování digitálních dokumentů

Ochrana digitálních dokumentů je proces, při kterém je potřeba zajistit, aby datové nosiče byly správně uchovány a byly zajištěny nezbytné podmínky pro potenciální archivaci. *„Ochrana digitálních dokumentů mění tradiční paradigma na nosič zaměřené ochrany dokumentů, a proto vyžaduje nové zkušenosti, jinou konceptualizaci a odlišná ochranná opatření. V digitálním světě je totiž možné bez jakéhokoli narušení autenticity oddělit nosič od informace (protože nosič sám žádnou nenahraditelnou informaci nepředstavuje) a přenesením na jiný elektronický nosič uchovávat pouze samotné informace. Z tohoto hlediska přestává být hlavním cílem ochrany životnost konkrétního nosiče. Elektronický nosiš sám obvykle není kulturním artefaktem a náležité zkopírování informací na nový nosič není vytvářením padělků nebo druhořadých napodobenin, ale představuje plnohodnotné zajištění ochrany. V tomto smyslu lze také říci, že digitální informace jsou teoreticky vzato věcné, protože je lze stále znovu a znovu překopírovat jinam, a tak prodlužovat jejich životnost donekonečna.“*[30]

Jelikož se příslušné technologie pro uchování a ochranu dat mění vysokým tempem, je potřeba brát ohled na zastarávání hardwarových a softwarových systémů, například data, která jsou nyní na optických nosičích, nemusí být za 10-15 let přečtena z důvodu, že optické mechaniky již budou přežitkem a do počítačových stanic se nebudou hardwarově instalovat anebo nebudou moct být data přečtena a otevřena v žádném existujícím softwaru (například zastaralost formátu).

„Exponenciální nárůst objemu digitálních dat, zvýšení jejich významu ve společnosti a nárůst jejich komplexity a diverzity, jakož i čím dál rychlejší zastarávání datových nosičů, softwaru, hardwaru a dalších technologických komponentů značně komplikují původní ochranná opatření aplikovaná v počátcích digitálního světa. Pouhé zálohování přestává být dostatečným řešením. „[31]

Z tohoto důvodu musí instituce od 90. let vymýšlet nová praktická opatření, která zajistí dostatečnou ochranu a bezpečnost dat. Hlavní obor, který vznikl, se nazývá digitální archivace dat.

[30] CUBR, Ladislav. Dlouhodobá ochrana digitálních dokumentů, Praha, 2010. ISBN 978-80-7050-588-5, str. 10-11
[31] Tamtéž, str. 15-16

3.14 Archivace dat a digitální dědictví

Dlouhodobé uchování digitálních dokumentů dat neboli digitální archivace dat je hlavním principem, jak trvale uchovat data bez následného poškození a ztráty. Jedna z nejznámějších definic oboru digitální archivace uvádí, že jde o všechna opatření, která jsou nutná pro udržení přístupu k digitálnímu navzdory degradaci datových nosičů. Dle Henryho M. Gladney je obor definován jako systematická opatření zajišťující mnohaletou využitelnost digitálních objektů. Takováto opatření musejí podle Gladneyho zajistit, aby digitální objekty nebyly nikdy ztraceny nebo zničeny, aby byly důvěryhodné a trvale vyhledatelné, a to bez ohledu na negativní dopady technologické zastaralosti nebo lidského faktoru (jako je např. slábnoucí lidská paměť).

S digitální archivací se spojuje i oblast digitálního dědictví. *„Jednou z mnoha oblastí, kterou výrazně ovlivnil nárůst objemu a komplexity dat je komunita kulturního dědictví. Lze říci, že do 90.let obsahovaly sbírky paměťových institucí téměr výhradně klasické dokumenty. V průběhu 90.let se však tento stav výrazně změnil a jejich sbírky se začaly rychle rozšiřovat o digitální dokumenty.“*[32] Mezi významné a hlavní důvody patřily například masová digitalizace knihovních sbírek nebo produkce nových typů digitálních dokumentů jako je například multimediální encyklopedie na optických nosičích.

Australská národní knihovna (National Library of Australia, 2003, s. 157) definuje digitální dědictví jako takové digitální dokumenty, které jsou dostatečně cenné na to, aby byly uchovávány pro účely budoucího zpřístupňování a využívání. UNESCO uvádí, že digitální dědictví jsou v digitální podobě vytvořené nebo digitalizované zdroje lidského poznání nebo vyjádření, jež mají trvalou hodnotu a význam.

[32] CUBR, Ladislav. Dlouhodobá ochrana digitálních dokumentů, Praha, 2010, ISBN 978-80-7050-588-5, str. 24

Trvalé uchování a archivace dat *předpokládá procesy soustavného zálohování a údržby nejen digitálních dokumentů, ale také přidružených počítačových technologií a datových nosičů (Verheul, 2006, s. 20). Dlouhodobé uchovávání musí čelit zejména rychlým změnám hardwarových technologií a nosičů a zajistit, že budou stále k dispozici vhodná čtecí a další elektronická zařízení, která umožní digitální objekty z elektronických nosičů načítat a kopírovat. Tento úkol není jednoduchý. Stačí si uvědomit, jak je dnes již obtížné nalézt zařízení pro čtení disket a přitom ještě před několika lety byly disketové mechaniky nedílnou součástí většiny osobních počítačů.*[33]

Předpokladem pro zpřístupnění digitálního dokumentu za nějaký časový horizont do budoucnosti je zálohovat data adekvátním příslušným softwarem. Každý digitální dokument je prvně vytvořen a zpracován určitou softwarovou aplikací, poté může být zpřístupněn uživateli v té podobě, ve které byl vytvořen. Podle L. Cubra například článek vytvořený ve formátu PDF bude za sto let nutné zobrazit stejným způsobem (zachování původního typu písma, rozložení textu apod.), jakým jej dnes zobrazí aplikace Adobe Reader.

3.15 Zpracování dat před archivací

Data, která mají být archivována, musejí splňovat určité normy a standardy podle zákonů či směrnic daného podniku, nebo sektoru veřejné správy. Jeden z nejdůležitějších pohledů všech organizací pro zálohování a archivaci dat je finanční hledisko. Objem dat neustále roste. „*Část dat jsou data živá, se kterými se pracuje, která se mění – například vznikající a editované dokumenty, živá data z ERP a CRM aplikací. Většina dat je ale již ze své podstaty neměnného charakteru. Jsou to data jako například účetní dokumenty, fotodokumentace, protokoly, smlouvy. Těchto dat, která se již dále nemění, je v organizacích přibližně 75 %. Přístup k nim není tak častý, vzhledem k jejich charakteru není ani problémem pomalejší odezva. Obnova archivních dat není tak urgentní jako obnova dat produkčních.*“[34]

[33] CUBR, Ladislav. Dlouhodobá ochrana digitálních dokumentů, Praha, 2010, ISBN 978-80-7050-588-5, str. 29

[34] TEICHMAN, Miroslav. Proč vlastně archivovat data? Data v péči. [online]. 2009 [cit. 2014-12-19]. Dostupné z: http://bit.ly/1EO1G5y

Obnova produkčních dat je otázkou řádu několika hodin. Obnovit archivní data je pak otázkou několika dnů až týdnů. Požadavky na objem dat zavádějí myšlenku uložit archivní data jiným způsobem než ty produkční. Pro snížení nákladů je velmi vhodné a efektivní uložit je levněji než živá data. Podle M. Teichmana lze tím zároveň vyřešit obnovu archivních dat pomocí levnějších technologií, protože čas na případnou obnovu bude mnohonásobně delší. Teichman dále zmiňuje, že pokud je objem dat v organizaci větší, alespoň v řádu TB, může znamenat oddělení archivních dat snížení provozních nákladů.

4 Analytická část

4.1 Úvod

Cílem praktické části bylo zjistit a definovat, na jakém principu funguje proces zálohování a archivace dat ve veřejné správě. V rámci této bakalářské práce byl vybrán Úřad městské části v Praze, kde byla za odborné asistence IT technika přiblížena legislativa zálohování a archivace, dále celý proces zálohování a využití softwaru k zálohování a archivaci podnikových dat.

V první části je převzata legislativa a působnost archivace dokumentů z webové stránky ministerstva práce a sociálních věcí, která má za úkol přiblížit normy a zákony, které tuto problematiku regulují a definují. V druhé části je pojednáno o samotných principech zálohování a archivace, o užití zálohovacího softwaru, o virtualizaci prostředí, o digitální spisové službě (eSpis), a na závěr o deduplikaci a bezpečnosti dat v podnikovém prostředí veřejné správy.

4.2 Legislativa

4.2.1 Právní předpisy

Pojem archivace ve státní (veřejné) správě zaujímá problematiku, jak bude nakládáno se spisy na základě legislativy. Definuje jí zákon o archivnictví a spisové službě - č. 499/2004 Sb. *„Skartace písemností a vyřazování písemností se provádí průběžně na základě souhlasu Státního oblastního archivu Praha, resp. územně příslušných Státních oblastních archivů. Určení původci dle § 3 zákona jsou povinni zajistit řádný příjem, evidenci, třídění, vyřizování, archivaci i skartaci dokumentů. Za tímto účelem jsou povinni vydat vnitřní předpis „Spisový a skartační řád“, který obsahuje podrobná pravidla pro manipulaci s dokumenty, vycházející z obecných pravidel stanovených zákonem o archivnictví a spisové službě a jeho prováděcími vyhláškami.“* [35]

Povinnost uchovávat dokumenty a umožnit výběr archiválií mají dle § 3 zákona:

- organizační složky státu, státní příspěvkové organizace
- státní podniky, územní samosprávné celky
- školy a vysoké školy
- právnické osoby zřízené zákonem, zdravotnická zařízení

4.2.2 Ochrana archiválií

Archiválie jsou chráněny státem. Při nahlížení do archiválií se provádí o každém nahlédnutí řádná evidence. Vstup zaměstnanců do archivu je dovolen pouze v doprovodu archiváře.

[35] Křížová. Příručka pro archivaci dokumentů. ESF v ČR. [online]. 2008 [cit. 2015-01-05]. Dostupné z: http://bit.ly/1AFApRo

4.2.3 Elektronický dokument jako archiválie

Elektronický dokument je zvláštní druh archiválie. V rámci předarchivní péče je nutné stanovit analýzu elektronického dokumentu, posouzení provádí pracovník archivu s původcem elektronického dokumentu. Ve věci je následně rozhodnuto, zda archiválie bude uchovávána v elektronické podobě, nebo zda bude převedena do písemné formy.

4.2.4 Základní technické podmínky archivu

- *Vstupní dveře do archivu jsou oplechované, protipožární.*
- *Uličky mezi regály jsou široké 80 – 90 cm.*
- *Relativní vlhkost ovzduší 45 – 65 %. Archiv je vybaven vlhkoměrem. Do deníku se pravidelně zapisují vlhkosti ovzduší. Teplota se pohybuje v rozmezí 14 °C – 18 °C.*
- *V archivu je zakázáno kouřit a pracovat s otevřeným ohněm. Archiv je vybaven práškovými hasicími přístroji.*[36]

4.3 Zálohovací software

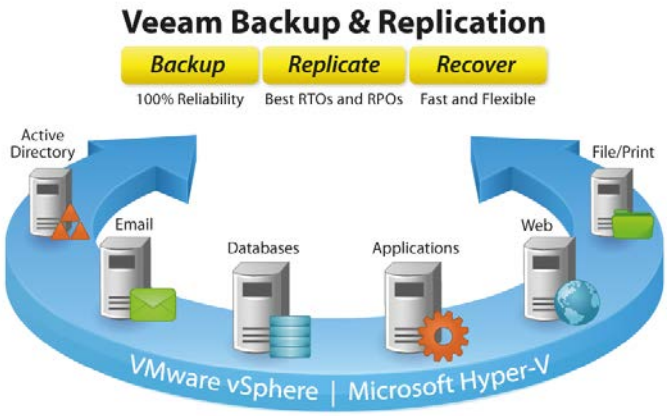
Software, který na Úřadě městské části v Praze zabezpečuje zálohování aplikací a dat se nazývá Veeam. Je to tzv. sofistikovaný systém, který má za úkol provádět zálohy v rámci požadovaného nastavení (záloha produkčního prostředí, nebo kompletní záloha diskového image serverů, dále například i přírůstkové zálohy) a v různých časových intervalech (každou hodinu – nejvíce důležitá data, nebo jednou týdně – podstatně méně důležitá data). Veeam Backup má spoustu zálohovacích funkcí, které jsou popsány níže. Celý tento systém běží v serverovém prostředí WMWare. Všechna data (aplikace, elektronické dokumenty, digitální spisová agenda) jsou zálohována na dynamické pevné disky, které jsou vzájemně propojeny a tvoří pole RAID. Magnetické pásky se nevyužívají.

[36] Křížová. Příručka pro archivaci dokumentů. ESF v ČR. [online]. 2008 [cit. 2015-01-05]. Dostupné z: http://bit.ly/1AFApRo

4.3.1 Veeam Backup

Software od firmy Veeam je velmi rychlé a spolehlivé řešení pro virtualizaci aplikací a dat v podniku. Produkty Veeam nabízí mnohem více než jen zálohování dat. Cíl je kladen na využití virtualizace, nebo úložiště, či cloudové technologie, *„jejich prostřednictvím zajistit funkce velmi rychlého obnovení, eliminace ztráty dat, ověřené ochrany, deduplikace dat a úplného přehledu a s jejich pomocí zajistit u VŠECH aplikací a dat cíle doby/bodu obnovení pod 15 minut.“* [37] Veeam se vyznačuje efektivní centralizovanou správou, kde lze pomocí webového klienta Veeam Enterprise Manager spojit až několik zálohovacích serverů z různých umístění a to bez nutnosti přihlašování se do podnikové sítě.

Na Úřadu MČ bylo zjištěno, že v rámci zálohovacího procesu systémem Veeam probíhají full backup zálohy 1x denně automatizovaně + přírůstkové zálohy průběžně, zároveň je občas provedeno nárazové obnovení dat v nepravidelných intervalech. Nepravidelné intervaly obnovení mají za úkol zjistit, jestli jsou data a aplikační/produkční prostředí spolehlivě zálohované a do budoucna půjdou pro případ nutnosti obnovit. Je kladen velký důraz na pravidelné kontroly záloh – buď automatizované, nebo ruční. V tomto případě administrátor datového centra provádí kontrolu a celistvost uložených dat.



Obrázek 13 - Schéma Veeam Backup [38]

[37] VEEAM. [online]. 2014 [cit. 2014-12-07] Ochrana před ztrátou dat. Dostupné z: http://www.veeam.com/cz/data-loss-avoidance.html
[38] VEEAM. [online]. 2014 [cit. 2015-02-07]. Veeam Backup & Replication . Dostupné z: http://bit.ly/16usyIX

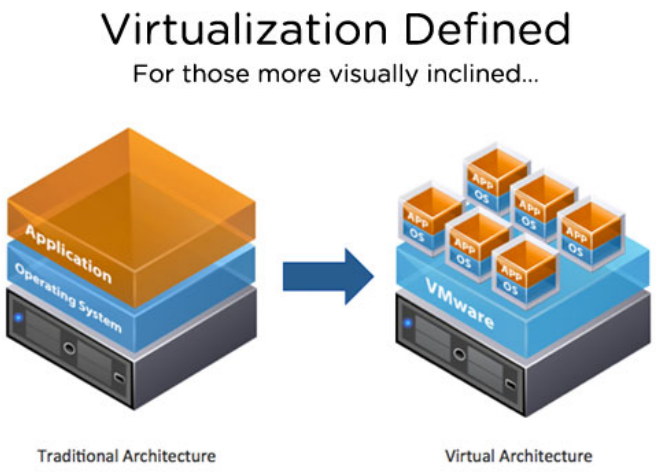
4.3.1.1 Funkce Veeam Backup

- Vysoká rychlost obnovení - obnovení virtuálního počítače po havárii za 2 minuty, okamžité obnovení souborů a složek operačního systému hosta.
- Ochrana před ztrátou dat – úplné šifrování - lze ochránit data během zálohování (při přenosu i v úložišti).
- Prověřená ochrana - Umožňuje automatické ověření možnosti obnovení každé zálohy při každém zálohování.
- Snížení rizik při nasazení – jde o tzv. virtuální laboratoř, která slouží při zavádění aplikací k otestování, aniž by to mělo pádný vliv na provozní infrastrukturu. K testování dochází v izolovaném prostředí, ve kterém se aplikace testuje, a zároveň se odstraňují nahodilé technické problémy.

4.3.2 Prostředí WMware a virtualizace

Na Úřadě MČ je virtualizace VMware využívána zejména z toho důvodu, neboť umožňuje snižovat kapitálové výdaje pomocí konsolidace serverů [39] a snižovat provozní náklady díky automatizaci. Mezi další hlavní důvody patří zefektivnění správy a řízení desktopů na všech odborech úřadu a zefektivnění práce koncových uživatelů. Lze virtualizovat nejdůležitější podnikové aplikace 1. vrstvy včetně databází (např. Oracle Databáze, SQL Servery) a některé další aplikace (například Microsoft Exchange). Srdcem virtualizace je virtuální stroj, samostatná výpočetní jednotka s operačním systémem a aplikacemi. Protože každý virtuální stroj je zcela samostatný a nezávislý, mnoho z nich může běžet současně na jednom počítači. „*Virtualizace desktopů šetří náklady na správu distribuovaného IT prostředí, ve kterém běží aplikace na koncových stanicích. Dalším přínosem je zvýšení bezpečnosti firemních dat. Všechny údaje mohou být centralizovány ve virtuálním privátním datovém centru a nemusejí být uloženy v koncových stanicích.*“ [40]

[39] Vysvětlivka - sjednocení serverů do jednoho, konsolidačního serveru (např. mail server + DHCP server)
[40] WMWare. Virtualizace. [online]. 2014 [cit. 2015-01-15]. Dostupné z: http://vmw.re/1xIfRRl



Obrázek 14 - Porovnání klasické a virtuální architektury [41]

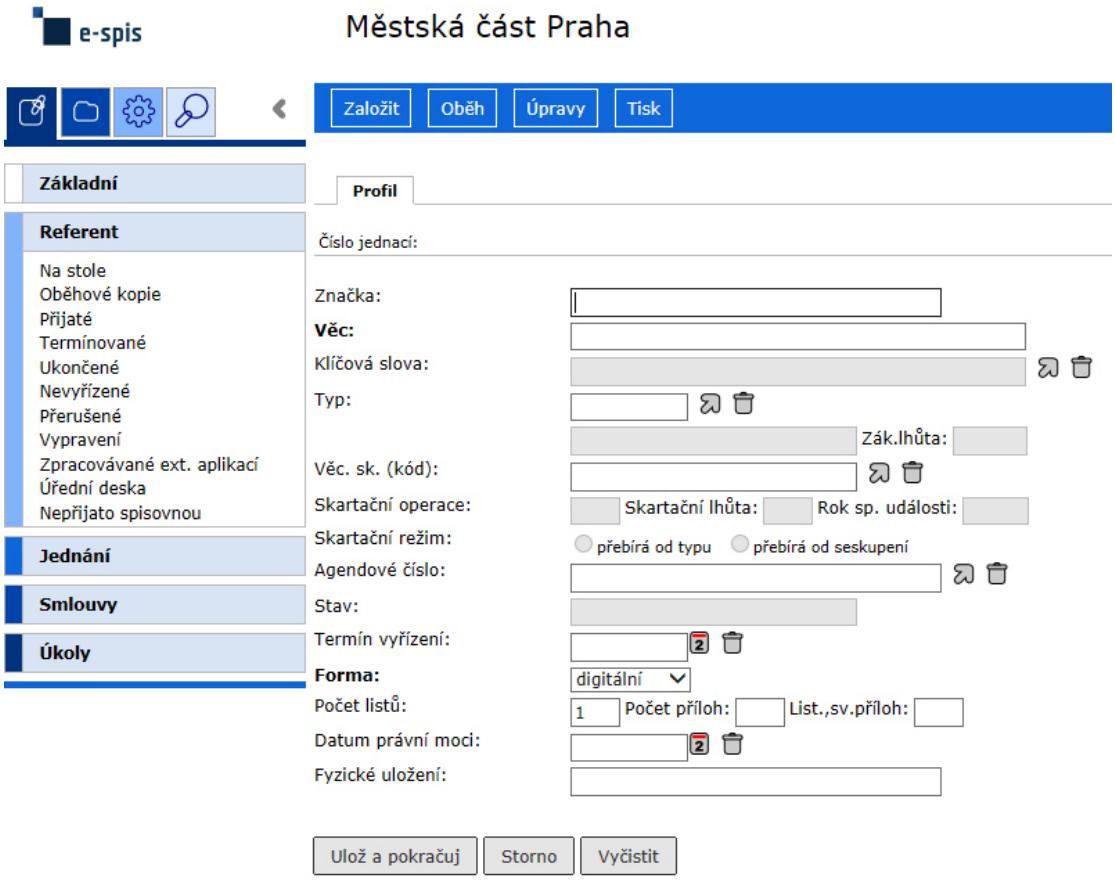
4.4 Princip zálohování ve veřejné správě

Samotný princip zálohování na Úřadu MČ probíhá ve virtualizovaném prostředí WMware, kde se zálohuje celé produkční prostředí, probíhají zálohy celých serverů, včetně databázových, aplikačních a dalších běžících serverů. Zálohování běží jednou denně plnou zálohou, dále probíhají přírůstkové zálohy (několikrát denně). Zálohování probíhá ve dvou rovinách, první je tzv. mirroring, neboli je vyžíváno pole RAID 1, které provádí zrcadlení disku, obsah se zaznamenává na dva disky zároveň, v případě výpadku jednoho disku je k dispozici ještě druhý, na kterém je kopie. Druhá rovina zálohování zaujímá princip přenosu dat po optickém vlákně mimo budovu úřadu, do jiného datového centra. Ve státní správě je velmi důležité zohlednit problematiku cloudového řešení, které je ze zákona zakázáno, jelikož zákon stanovuje a říká, že je nutné znát místo, kde se všechna data nachází.

[41] WMWARE. How virtualization works. [online]. 2014 [cit. 2015-01-15]. Dostupné z: http://vmw.re/1CEmoEd

4.5 Digitální spisová služba - eSpis

Digitální spisová služba (eSpis) představuje na Úřadu MČ sofware, pomocí kterého se spisy digitalizují. V samotné aplikaci eSpis lze dokument založit, sledovat jeho oběh, upravovat, vyhledávat jiné spisy, tisknout a mnoho dalšího. Každý spis by měl být v digitální spisové službě. Digitální spis má během své životnosti určitý cyklus (oběh, vyřizování, elektronické podepisování…). Tento spis se poté dále předává do příruční digitální spisovny a následně se předává po určité době do digitálního archivu, kde probíhá digitální archivace. Digitální archiv je umístěný mimo budovu Úřadu městské části – je to samostatná budova archivu, kde spisy podléhají skartačnímu řádu a po určité době se vymazávají z digitálního archivu. Součástí vymazání je protokol o vyřazení spisu.



Obrázek 15 - Desktopové rozhraní aplikace eSpis, zdroj: aplikace ÚMČ

4.6 Deduplikace (komprese) dat

Deduplikaci lze označit jako strategickou technologii, která má za úkol odstraňovat duplicity při procesu zálohování (ukládání) dat. Pokud již jsou patřičná data jednou uložena, není třeba ukládat ta samá data znovu. Stačí uložit jen informaci, kde již jednou uložená (patřičná) data najdeme pro případnou nutnost obnovy. V rámci Úřadu MČ bylo zjištěno, že celý tento proces je využíván zejména z důvodu úspory diskových kapacit, jak již bylo zmíněno, magnetické pásky se na Úřadu MČ nevyužívají, při deduplikaci jsou využívány dynamické disky (pole RAID), protože disponují pokročilými metodami ochrany dat. Dalším důvodem pro využívání deduplikace dat jsou nižší náklady na energii spojenou s napájením zálohovacích systémů pro ukládání dat. Bylo shledáno, že celý proces deduplikace efektivně zvyšuje spolehlivost celého informačního systému Úřadu MČ a usnadňuje virtualizaci.

4.7 Bezpečnost a skladování dat

Bezpečnost dat je jedním z nejdůležitějších aspektů, na které je ve veřejné správě brán ohled. Právě tato data (aplikační prostředí, elektronické dokumenty, digitální spisy a jiné) jsou hlavním zdrojem výkonu činnosti práce každého státního zaměstnance. Proto již při samotném procesu zálohování je z bezpečnostních důvodů zajištěn šifrovaný přenos dat pomocí softwarového prostředí Veeam, který disponuje vysokou rychlostí obnovení pro případ havárie stanice (Instant VM Recovery). Tím lze také chránit data před ztrátou v úložišti. Zajímavá funkce, která je často využívána se nazývá SureBackup, která umožňuje automatické ověření možnosti obnovení každé zálohy při každém zálohování.

Skladování dat ve veřejné správě na Úřadu městské části je koncipováno na dvou principech. Prvním, je datové centrum v rámci budovy, která podléhá protipožárnímu a technickému zabezpečení, druhým principem je využití datového centra mimo budovu, které má za úkol tvořit tzv. kopii prvního datového centra a přenos dat do něj je zajištěn po optickém vlákně. V případě kolize datového centra v rámci budovy je možné všechna data obnovit v jiné budově s využitím záložního datového centra v řádu jednotek hodin.

4.8 Vyhodnocení analytické části

Úřad MČ v Praze používá pro zálohování a archivaci dat velmi dobré technologie a software k tomu určený. Využití pole RAID 1 při zálohování je velice efektivní. V případě technické havárie jednoho z disků může být obsah bez problémů obnoven ze zrcadleného disku bez jakékoliv ztráty dat. Sofware Veeam Backup byl shledán jako profesionální nástroj, který má mnoho užitečných funkcí a jeho základním úkolem je virtualizace prostředí. Digitalizace spisů je zajištěna pomocí aplikace eSpis, která má velmi široké spektrum funkcí při zakládání a manipulaci se spisy. Součástí zálohovacího procesu je deduplikační technologie, která má za úkol odstraňovat duplicitní data při procesu ukládání – velmi efektivní pro diskové kapacity. V rámci bezpečnosti je zajištěn šifrovaný přenos dat pomocí Veeam Backup. Skladování dat na ÚMČ je zajištěno jak v rámci budovy, tak v rámci záložní budovy – tento způsob skladování je co se týče bezpečnosti hojně využívaný.

Ve spolupráci s IT technikem při navrhování efektivnějšího řešení na Úřadu MČ bylo zjištěno, že všechny zálohovací rutiny a procesy běží automatizovaně a efektivně. Jediný proces, který by se mohl zlepšit, je snížit požadovanou dobu skladování elektronických dat, o kterých je již známo, že nebudou do budoucna využity (například záloha aplikačního prostředí z roku 2012 nebo staré neaktuální přírůstkové zálohy).

5 Závěr

V této bakalářské práci bylo cílem přiblížit a pojednat o problematice procesu zálohování a archivace dat. V dnešní době – době nejrychlejšího rozmachu technologií je zálohování dat naprostá nutnost. Je mnoho potenciálních rizik, která uživatele mohou postihnout při každodenní práci s počítačovou stanicí, přičemž každý uživatel má své právo rozhodnout, jak velké úsilí do ochrany před ztrátou dat vloží, jak, v jakých intervalech a na jaká média bude zálohovat.

Cílem teoretické části bylo seznámit čtenáře o pojmech zálohování, archivace, proč svá data zálohovat a co je předmětem zálohování. Dále pak přiblížit téma o strategiích zálohování, o technologii RAID a o nejdůležitějších kritériích zálohovacích médií. Jednou ze součástí bylo dále popsat a rozdělit datové nosiče a také webová a cloudová úložiště. Závěr teoretické části je věnován softwaru pro zálohování a šifrování, ochranu digitálních dokumentů a zpracování dat před archivací.

V analytické části bylo cílem analyzovat proces zálohování a archivace dat ve veřejné správě, na Úřadu MČ. Za odborné asistence IT technika byla přiblížena právní legislativa dané problematiky a dále celý proces zálohování včetně využití softwaru k činnosti zálohování a archivaci digitálních dokumentů a dat ve veřejné správě. Na základě zjištěných nedostatků bylo cílem navrhnout s IT technikem možné řešení pro zlepšení celého zálohovacího a archivačního procesu za pomoci získaných teoretických znalostí.

6 Seznam použitých zdrojů

6.1 Knižní publikace

1. **Leixner, Miroslav.** *PC - zálohování a archivace dat.* Praha : Grada, 1993. 80-854-2473-8.

2. **Minasi, Mark.** *Pevné disky od A do Z.* Praha : Grada, 1992. 80-856-2335-8.

3. **Horák, Jaroslav.** *Hardware: učebnice pro pokročilé.* Brno : CP Books, 2005. 80-251-0647-0.

4. **De Guise, Preston.** *Enterprise Systems Backup and Recovery: A Corporate Insurance Policy.* New York : CRC Press, 2008. 987-1-4200-7639-4.

5. **Cubr, Ladislav.** *Dlouhodobá ochrana digitálních dokumentů.* Praha : Národní knihovna České republiky, 2010. 978-80-7050-588-5.

6. **Humphries, Mark.** *Data warehousing – návrh a implementace.* Praha : Computer Press, 2001. ISBN 80-722-6560-1.

7. **Kříž, Libor.** *Komprimační a archivační programy.* Praha : Computer Press, 2002. 80-7226-757-4.

6.2 Články v seriálové publikaci

1. **Brechlerová, Dagmar.** Paměťová média: Informace, aneb kam s nimi. *Security magazín: časopis pro vaši bezpečnost.* 2012, Sv. 18, 2.

2. **Čandík, Marek.** Zálohování dat v informačních systémech. *Security magazín: časopis pro vaši bezpečnost.* 2004, Sv. 11, 2.

3. **Kysela, František.** Zálohování a archivace jako součást bezpečnosti ICT. *Technik.* 2013, 12.

4. **Schreiber, Manuel.** Data v cloudu. *CHIP*. Měsíční, 2011, 12.

6.3 Elektronické dokumenty

1. **Tišnovský, Pavel.** Magnetické paměti pro trvalý záznam dat. *ROOT.cz.* [Online] 24. Červenec 2008. [Citace: 8. Listopad 2014.] http://bit.ly/1DpJeyF.

2. Cache. *Wikipedia.* [Online] 5. Listopad 2014. [Citace: 8. Listopad 2014.] http://cs.wikipedia.org/wiki/Cache.

3. **Hort, Tomáš.** Technologie a zajímavosti z oblasti SSD disků II. *PCTuning.* [Online] 5. Leden 2012. [Citace: 2. Prosinec 2014.] http://bit.ly/1IojB5A.

4. **Tišnovský, Pavel.** Vývoj optických pamětí: od DVD k Blu-ray. *ROOT.cz.* [Online] 11. Září 2008. [Citace: 2. Prosinec 2014.] http://bit.ly/1Iorkk2.

5. **Kraus, Josef.** Sedm nejlepších programů pro zálohování dat. *živě.cz.* [Online] 8. Leden 2012. [Citace: 18. Prosinec 2014.] http://bit.ly/1unEBD4.

6. **Teichman, Miroslav.** Proč vlastně archivovat data? *Data v péči.* [Online] 16. Říjen 2009. [Citace: 19. Prosinec 2014.] http://bit.ly/1EO1G5y.

7. **Team, WMware.** WMware Virtualizace. *WMware.* [Online] [Citace: 7. Únor 2015.] http://vmw.re/1xIfRRl.

8. **Team, Veeam.** Veeam Backup & Replication. *Veeam.* [Online] [Citace: 7. Únor 2015.] http://bit.ly/16usyIX.

9. —. Ochrana před ztrátou dat. *Veeam.* [Online] [Citace: 7. Prosinec 2014.] http://bit.ly/18Y2NTf.

10. **GJSZlin.** Pevný disk – schéma plotny pevného disku. *GJSZlin.* [Online] [Citace: 8. Listopad 2014.] http://bit.ly/1LYrPkm.

11. **Fujitsu.** Fujitsu. *Fujitsu.* [Online] [Citace: 8. Listopad 2014.] http://bit.ly/1DMATmb.

12. **Alza.cz.** Pevné disky. *Alza.cz.* [Online] Alza.cz, 2014. [Citace: 8. Listopad 2014.] http://bit.ly/1A14r0N.

13. Linear Tape Open. *Wikipedia.* [Online] 17. Únor 2014. [Citace: 8. Listopad 2014.] http://cs.wikipedia.org/wiki/Linear_Tape_Open.

14. **Team, WMware.** Virtualization Basics - What is a virtual machine. *WMware.* [Online] [Citace: 15. Leden 2015.] http://vmw.re/1CEmoEd.

15. **Křížová.** Příručka pro archivaci dokumentů. *ESF ČR.* [Online] 2008. [Citace: 5. Leden 2015.] http://bit.ly/1AFApRo.

16. **Team, Veeam.** Veeam Help Center (Technical Documentation). *Veeam.* [Online] [Citace: 8. Listopad 2014.] http://bit.ly/1DPZpDo.

17. **Mlejnek, Miroslav.** Zálohování dat. *SWMAG.cz - softwarový magazín.* [Online] 18. Listopad 2007. [Citace: 15. Listopad 2014.] http://bit.ly/1CFBcTc.

18. **Novák, Jiří.** Zálohování dat: Základní strategie a možnosti. *ICT manažer.* [Online] 25. Květen 2012. [Citace: 3. Listopad 2014.] http://bit.ly/1IqvEiW.

19. **Acronis.** Diferenciální - rozdílová záloha. *Acronis.* [Online] [Citace: 13. Listopad 2014.] http://bit.ly/18Y4Dn1.

20. —. Domácnosti a kanceláře - Acronis True Image 2015 for PC. *Acronis.* [Online] 2014. [Citace: 6. Prosinec 2014.] http://bit.ly/1AYzEhf.

21. **Corporation, Microsoft.** OneDrive grows with you. *OneDrive.* [Online] 2014. [Citace: 10. Prosinec 2014.] http://1drv.ms/1kENaxy.

22. **Support, Apple.** iCloud storage pricing. *Apple.* [Online] 2014. [Citace: 10. Prosinec 2014.] http://bit.ly/1KBf7nS.

23. **Box.** Pricing - Select a plan. *Box.* [Online] 2014. [Citace: 10. Prosinec 2014.] http://bit.ly/1xKgQR1.

24. **Team, Dropbox.** Choose your plan. *Dropbox.* [Online] 2014. [Citace: 10. Prosinec 2014.] http://bit.ly/MODcAL.

25. **Team, Google.** Úložiště na Disku. *Google Drive.* [Online] 2014. [Citace: 10. Prosinec 2014.] http://bit.ly/SgwLte.

26. **Šulc, Tomáš.** VelociRaptory a RAID - když jeden nestačí. *PCtuning.* [Online] 10. Listopad 2008. [Citace: 4. Listopad 2014.] http://bit.ly/1EO0vDe.

27. **Blšťák, Oliver.** Archivace – jediný možný cíl vašich dat. *System Online.* [Online] Srpen 2002. [Citace: 6. Listopad 2014.] http://bit.ly/1LYAALa.

28. **Rybka, Michal.** Kdo a proč zabil TrueCrypt? *PCTuning.* [Online] 6. Červen 2014. [Citace: 7. Březen 2015.] http://bit.ly/1GsTvMO.

29. **Corporation, Microsoft.** Ochrana souborů nástrojem BitLocker Drive Encryption. [Online] 2013. [Citace: 7. Březen 2015.] http://bit.ly/1BdAniA.

7 Seznam obrázků

Obrázek 1 - Schéma plného zálohování, vlastní tvorba....10
Obrázek 2 - Schéma přírůstkového zálohování, vlastní tvorba....12
Obrázek 3 - Schéma rozdílového zálohování, vlastní tvorba....13
Obrázek 4 - Schéma zálohovací rutiny 1, vlastní tvorba....14
Obrázek 5 - Schéma zálohovací rutiny 2, vlastní tvorba....15
Obrázek 6 - Schéma pole RAID 1, vlastní tvorba....16
Obrázek 7 - Schéma pole RAID 5, vlastní tvorba....17
Obrázek 8 - Pásková knihovna....20
Obrázek 9 - Pevný disk....22
Obrázek 10 - Schéma plotny pevného disku....23
Obrázek 11 - Porovnání HDD a SDD....27
Obrázek 12 - Porovnání cloudových úložišť, vlastní tvorba....28
Obrázek 13 - Schéma Veeam Backup....40
Obrázek 14 - Porovnání klasické a virtuální architektury....42
Obrázek 15 - Desktopové rozhraní aplikace eSpis, zdroj: aplikace ÚMČ....43