

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Bakalářská práce

Zálohování a archivace dat

Vojtěch Slepíčka

© 2023 ČZU v Praze

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Vojtěch Slepíčka

Informatika

Název práce

Zálohování a archivace dat

Název anglicky

Data backup and archivation

Cíle práce

Hlavním cílem práce je návrh zálohovací strategie pro malou firmu.

Dílním cílem bude analýza různých zálohovacích systémů a datových úložišť vhodných pro menší podniky s poukázáním na jejich rozdíly, výhody a nevýhody. V rámci srovnání zálohovacích systémů bude hodnocena jejich funkčnost, bezpečnost, přednosti a zápory.

Dále bude zpracován návrh efektivního řešení zálohování a archivování dat, v rámci kterého budou popsány různé možnosti a způsoby pro zálohu a archivaci dat, jejich rozdíly, pozitiva a negativa.

Metodika

Odborné poznatky a informace budou čerpany z literárních tištěných zdrojů a ověřených online zdrojů. Praktická část bude obsahovat návrh zálohovací strategie ve spolupráci s konkrétní firmou, na základě analýzy vhodných technologií pro zálohování.

Doporučený rozsah práce

35-45s.

Klíčová slova

Zálohování, archivace, data, server, úložiště, bezpečnost

Doporučené zdroje informací

BOTT, Ed, SIECHERT, Carl, STINSON, Craig. Mistrovství – Microsoft Windows 10. Computer Press, 2017.

ISBN: 978-80-251-4869-3

KOLOUCH, Jan, BAŠTA, Pavel a kol. CyberSecurity. CZ.NIC, z. s. p. o., 2019. ISBN: 978-80-88168-34-8

NELSON, Steven. Pro Data Backup and Recovery (Expert's Voice in Data Management). Apress, 2011. ISBN: 14-302-2662-5

PRESTON, W. Curtis. Backup and Recovery. U.S.A.: O'Reilly Media, Inc., 2007. ISBN: 9780596102463

SMEJKAL, Vladimír, SOKOL, Tomáš, KODL, Jindřich. Bezpečnost informačních systémů podle zákona o kybernetické bezpečnosti. Aleš Čeněk, s.r.o, 2019. ISBN: 978-80-7380-765-8

Předběžný termín obhajoby

2022/23 ZS – PEF

Vedoucí práce

Ing. Martin Havránek, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 17. 8. 2021

doc. Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 5. 10. 2021

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 13. 03. 2023

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 použitých zdrojů 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 15. 3. 2023

Poděkování

Rád bych touto cestou poděkoval Ing. Martinu Havránkovi, Ph.D., za jeho rady, připomínky a možnost zpracovávat bakalářskou práci pod jeho vedením.

Zároveň děkuji za podporu a pochopení své rodině.

Zálohování a archivace dat

Abstrakt

Bakalářská práce se zabývá možnostmi archivace a zálohování dat. Teoretická část je zaměřena na analýzu různých zálohovacích systémů a datových úložišť vhodných pro menší podniky s poukázáním na jejich rozdíly, výhody a nevýhody. V rámci srovnání zálohovacích systémů bude hodnocena jejich funkčnost, bezpečnost, přednosti a zápory. Dále bude zpracován návrh efektivního řešení zálohování a archivování dat, v rámci kterého budou popsány různé možnosti a způsoby pro zálohu a archivaci dat, jejich rozdíly, pozitiva a negativa. Praktická část se zabývá návrhem zálohovací strategie pro malou firmu.

Klíčová slova: zálohování, archivace, data, server, úložiště, bezpečnost

Data backup and archivation

Abstract

The bachelor thesis deals with the possibilities of data archiving and backup. The theoretical part is focused on the analysis of various backup systems and data storages suitable for smaller companies with reference to their differences, advantages, and disadvantages. As part of the comparison of backup systems, their functionality, security, advantages, and disadvantages will be evaluated. Furthermore, a proposal for an effective solution for data backup and archiving will be prepared, within which various options and methods for data backup and archiving, their differences, positives, and negatives will be described. The practical part deals with the design of a backup strategy for a small business.

Keywords: backup, archiving, data, server, storage, cybersecurity

Obsah

| | |
|---|-----------|
| 1 Úvod..... | 8 |
| 2 Cíl práce a metodika | 9 |
| 3 Teoretická východiska | 10 |
| 3.1 Seznámení se základními pojmy | 10 |
| 3.1.1 Zálohování dat | 10 |
| 3.1.2 Archivace dat | 10 |
| 3.1.3 Komprese dat | 10 |
| 3.1.4 Data | 10 |
| 3.1.5 Server | 11 |
| 3.2 Zálohování dat..... | 11 |
| 3.2.1 Nejčastější příčiny ztráty dat..... | 11 |
| 3.2.2 Skupiny zálohovaných dat | 13 |
| 3.2.3 Fyzická média pro zálohování dat | 14 |
| 3.2.4 Síťová úložiště NAS, SAN, DAS – jejich porovnání a rozdíly | 16 |
| 3.2.5 Online úložiště – cloud | 18 |
| 3.2.6 Ostatní služby pro zálohování dat..... | 19 |
| 3.2.7 RAID..... | 20 |
| 3.2.8 Úskalí použití hardware RAID v roce 2022 | 23 |
| 3.2.9 Zálohy a úrovně zálohování..... | 23 |
| 3.2.10 Porovnání elementárních způsobů zálohování | 24 |
| 3.2.11 Software pro zálohování dat | 24 |
| 3.3 Archivace dat | 26 |
| 3.3.1 Archivace dat versus zálohování | 26 |
| 3.4 Operace s daty | 28 |
| 3.4.1 Komprese | 28 |
| 3.4.2 Šifrování..... | 28 |
| 3.4.3 Deduplikace | 30 |
| 3.4.4 Verzování dat..... | 31 |
| 4 Vlastní práce..... | 32 |
| 4.1 Specifikace společnosti | 32 |
| 4.1.1 Geografické a základní údaje o společnosti..... | 32 |
| 4.1.2 Předmět činnosti společnosti | 32 |
| 4.1.3 Organizace společnosti | 33 |
| 4.1.4 Technické zázemí společnosti | 33 |
| 4.1.5 Aktuální stav zálohování a správy dat | 34 |
| 4.1.6 Jaká data společnost zpracovává a uchovává, jejich rozdělení..... | 35 |

| | | |
|----------|--|-----------|
| 4.1.7 | Možné příčiny a následky ztráty dat | 35 |
| 4.1.8 | Požadavky společnosti | 35 |
| 4.2 | Návrh zálohovacích strategií | 36 |
| 4.2.1 | 1. varianta návrhu – stanice NAS | 36 |
| 4.2.2 | 2. varianta – použití specializovaného softwaru pro zálohování | 41 |
| 4.2.3 | 3. varianta – kombinace NAS a cloud | 41 |
| 5 | Výsledky a zvolení nejlepší varianty návrhu zálohování | 43 |
| 6 | Závěr..... | 45 |
| 7 | Seznam použitých zdrojů | 46 |
| 8 | Seznam obrázků, tabulek, grafů a zkratk | 50 |
| 8.1 | Seznam obrázků | 50 |
| 8.2 | Seznam tabulek | 50 |

1 Úvod

Zálohování a archivace dat by neměly být v dnešní době pro nikoho neznámými pojmy. Jsou to jedny z nejdůležitějších procesů, které by měly uživatelé osobních počítačů a firmy, ať už malé či velké, znát a umět správně používat. V době, kdy probíhá světová digitalizace a veškeré citlivé dokumenty a údaje se transformují do elektronické podoby, by se bezpečnost dat neměla brát na lehkou váhu. V minulých letech unikly do prostředí internetu desítky miliard dat, a to ve většině případů jejich špatným zabezpečením a archivováním.

Více než 40 % českých uživatelů počítačů nepovažuje zálohování dat za podstatnou věc a svá data nezálohuje vůbec (Avast, 2020). Proč tomu tak je? Důvodů může být mnoho – neznalost, nedostatek času, lehkomyšlnost, či jiné faktory. Přitom mají na svých zařízeních mnohdy velmi citlivá data, která by se nikdy neměla dostat do rukou neoprávněných osob. Počítač nelze považovat za perfektní věc, bezvadné spotřební zboží, o které není nutné se starat a pečovat. V mnoha případech jde o velmi důležitou součást našeho osobního i pracovního života, a proto je důležité svá data chránit, zálohovat a archivovat, aby bylo možné předejít jejich nenávratným ztrátám.

Ve firemní sféře je zálohování a archivování dat nezbytná věc. Firmy uchovávají velké objemy dat, a to nejen o svých zaměstnancích, obchodech, zákaznících, ale mnoho dalších důležitých interních informací. Má-li firma „vyspělý“ systém zálohování a archivování, poukazuje to na její spolehlivost, důvěryhodnost, konkurenceschopnost, a především na to, že si svých zaměstnanců a zákazníků považuje. V dnešní digitální době by každá firma měla klást na tyto skutečnosti mimořádný důraz.

Bakalářská práce se bude v teoretické části snažit poukázat na to, jak správně nástroje pro archivaci a zálohování používat, čemu se vyvarovat, jaké nebezpečí hrozí při nesprávné manipulaci s daty. V praktické části se zaměří na samotný návrh strategie pro zálohování, archivování a zabezpečení dat v menší firmě. Přestože bezpečnost dat, zálohování a archivace dat tvoří podstatnou součást moderního digitálního světa, není k dispozici příliš spolehlivých zdrojů, ze kterých by běžní uživatelé, včetně malých firem, mohli čerpat potřebné informace a předejít tak problémům, které mohou ignorováním tohoto tématu nastat.

2 Cíl práce a metodika

Cíl práce

Hlavním cílem bakalářské práce je návrh zálohovací strategie pro malou firmu. Dílčím cílem bude analýza různých zálohovacích systémů a datových úložišť vhodných pro menší podniky s poukázáním na jejich rozdíly, výhody a nevýhody. V rámci srovnání zálohovacích systémů bude hodnocena jejich funkčnost, bezpečnost, přednosti a zápory. Dále bude zpracován návrh efektivního řešení zálohování a archivování dat, v rámci kterého budou popsány různé možnosti a způsoby pro zálohu a archivaci dat, jejich rozdíly, pozitiva a negativa.

Metodika

Odborné poznatky a informace budou čerpány z literárních tištěných zdrojů a ověřených online zdrojů. Praktická část bude obsahovat návrh zálohovací strategie ve spolupráci s konkrétní firmou na základě analýzy vhodných technologií pro zálohování. V praktické části bude použita k výběru vhodného hardwaru metoda vícekriteriální analýzy variant váženého součtu.

3 Teoretická východiska

3.1 Seznámení se základními pojmy

V oblasti výpočetní techniky se nachází mnoho specifických pojmů, které budou také použity v této bakalářské práci. Z toho důvodu jsou některé nejdůležitější a nejčastěji používané výrazy níže vysvětleny.

3.1.1 Zálohování dat

V případě zálohování dat jde o způsob ochrany dat před ztrátou nebo poškozením, které spočívá v pořízení duplicitní kopie originálních souborů. Soubory lze zálohovat buď přímo kopírováním na záložní médium (nejčastěji pevný disk HDD nebo SSD disk), nebo lze použít speciální zálohovací aplikace. Pro zálohování dat lze také využít kompresi dat (Diderot, 1999, s. 416).

3.1.2 Archivace dat

Archivace dat je ukládání jednotlivých údajů nebo celých datových souborů do informačního systému skladebného typu (např. databáze) se záměrem použít je v budoucnosti (Diderot, 1999, s. 205).

3.1.3 Kompresí dat

Kompresí dat dochází ke zkrácení délky souboru nebo zprávy pomocí zvláštního programu, kdy je současně uchován obsah zprávy nebo souboru. Nazývá se též „balení“ nebo „komprimace“ dat. Před vlastní prací s komprimovaným souborem je nutné tento soubor „rozbalit“, tj. uvést do původního stavu (Diderot, 1999, s. 151).

3.1.4 Data

Data jsou údaje, skutečnosti, resp. vhodným způsobem vyjádřená zpráva, která je srozumitelná příjemci a přizpůsobená k dalšímu zpracování. Data představují nejnižší prvek informačního systému (primární data), která procházejí dalším zpracováním (sekundární data). Uchovávají se na nosičích dat a zpracovávají se různými typy prostředků výpočetní techniky (Diderot, 1999, s. 186).

3.1.5 Server

V oblasti výpočetní techniky je serverem myšlen počítač nebo počítače, které řídí předávání dat po síti. Těmito daty mohou být webové stránky, program, soubory nebo informace. Servery naslouchají požadavkům od klientů, což jsou jiné počítače nebo zařízení. Když server slyší požadavek, přečte jej a odpoví odesláním dat zpět klientovi. Servery jsou postaveny za použití speciálního hardwaru – to umožňuje serverům zpracovat velké množství požadavků najednou a zůstat online nepřetržitě (Freecodecamp, 2021).

3.2 Zálohování dat

Nejdůležitějším a nejspolehlivějším způsobem, jak ochránit data před ztrátou nebo poškozením, je jejich pravidelné zálohování. Ke ztrátě dat může dojít nenadálým výpadkem elektrického proudu, selháním hardwaru nebo softwaru, lidskou chybou. Dat mohou být zálohována na různé druhy médií, např. pevný disk, CD/DVD nosiče, flash disky, NAS server, vzdálená online úložiště – cloudové systémy apod. (Windows Central, 2019).

Průzkum vypracovaný společností Spolková asociace pro informační technologie s názvem „Situace v oblasti bezpečnosti IT v Německu v roce 2021“ uvádí, že v roce 2019 bylo zachyceno 114 milionů nových variant malwaru, tzn. zhruba 320 000 denně. V roce 2021 to bylo již 144 milionů, tzn. 394 000 nových variant denně. Představuje to nárůst o 22 % během pouhých dvou let. O nutnosti zálohovat data není tudíž žádný pochyb (Synology, 2022).

3.2.1 Nejčastější příčiny ztráty dat

Pravidlem bývá, že způsob, jakým se provádí obnovení dat, určuje způsob, jakým se zálohování provádí. Je nutné si položit otázky: Před čím je nutné se chránit? Jsou všichni uživatelé zkušení uživatelé, kteří používají počítač inteligentně a nikdy nedělají hloupé chyby? Přišla by společnost o spoustu důležitých dat, kdyby byly soubory v počítačích uživatelů smazány? Kdyby hurikán zničil celou společnost, byla by schopna pokračovat v podnikání? Je nutné se ujistit, že jsou metody pro zálohování dat správné (Preston, 2007, s. 59).

Nejčastější příčiny ztráty dat:

- Chyba uživatele je zdaleka nejčastější příčinou ztráty dat. Ke ztrátě dat může vést např. náhodné vymazání nebo přepsání dat či nesprávné zálohování dat (Preston, 2007, s. 59).
- Chyba správce systému je méně častá než chyba běžného uživatele (pokud nemají uživatelé práva „roota“ nebo správce). Může se například stát, že správce systému odstraní složku s dokumenty uživatele (Preston, 2007, s. 59).
- Selhání systému, zde lze hovořit o dvou podobách selhání, a to selhání diskové jednotky a selhání celého systému (Preston, 2007, s. 60):
 - Ochrana systému před selháním diskové jednotky je poměrně jednoduchá. Je nutné pouze rozhodnout, jak moc chce být uživatel v bezpečí. Nejlepší ochranu nabízí zrcadlení, často označované jako RAID 1, které však zdvojnásobuje počáteční investice do hardwaru disků a řadiče. Většinou se však lze setkat s některou z dalších úrovní RAID, z nichž nejoblíbenější je RAID 5, přičemž RAID 6 získává stále větší popularitu. Svazky RAID 5 chrání před ztrátou jednoho disku tím, že na každém disku vypočítávají a ukládají paritní informace. RAID 6 přidává další ochranu tím, že ukládá paritu dvakrát, čímž umožňuje výpadek více než jedné jednotky (Preston, 2007, s. 60).
 - Většinu ochrany proti selhání celého systému zajišťují správné postupy správy systému. Je třeba reagovat na zprávy o špatných discích, řadičích, procesorech a paměti. Další metodou ochrany je použití tzv. žurnálovacího souborového systému. Žurnálování zachází se souborovým systémem podobně jako s databází a sleduje odevzdané a částečně odevzdané zápisy do souborového systému. Když se systém blíží k poruše, žurnálovací souborový systém může částečně odevzdané zápisy vrátit zpět, a tím odstranit korupci systému (Preston, 2007, s. 60).
- K vloupání do elektronických zařízení, vandalismu a krádežím v posledních letech dochází velmi často, o mnohých z nich se mluvilo po celém světě. Pokud dojde ke ztrátě dat v důsledku jedné z těchto příčin, je to velmi odlišné od jiných typů ztráty dat. Data sice jsou sice obnovitelná, ale nikdy není jisté, co se s nimi stalo v době, kdy jsme je neměli k dispozici. Proto je nutné dělat vše proto, aby se nic takového nestávalo. (Preston, 2007, s. 60).

- V případě přírodních katastrof v podobě hurikánu, tornáda, povodně, je nutné si klást otázky, zda je společnost na tyto možnosti připravena z pohledu ochrany dat. Pokud se používají externí úložiště, je toto zařízení v dosahu? Je připraveno zvládnout jakýkoliv typ přírodní katastrofy, který může nastat? Pokud se například kancelář nachází v záplavové zóně, ukládá společnost data v zařízení nacházejícím se ve vyšších patrech? Pokud ne, mohou být data ztracena při jednom velkém dešti. Obranou proti velké přírodní katastrofě by měla být například možnost ukládání dat v reálném čase mimo pracoviště na vzdáleném místě (Preston, 2007, s. 61).
- Kybernetické útoky jsou další častou příčinou ztráty dat. Hackeři mohou získat přístup do systému a ukrást či poškodit data různými způsoby. Aby se předešlo ztrátě dat v důsledku kybernetických útoků, je důležité zavést důkladná bezpečnostní opatření jako je používání antivirového softwaru, aktualizace softwaru a bezpečnostních protokolů a školení zaměstnanců, aby rozpoznali phishingové útoky a vyhnuli se jim. Je také dobré pravidelně zálohovat data, protože to umožní obnovu v případě jejich ztráty nebo poškození dat v důsledku kybernetického útoku (OpenAI, 2022).
 - Malware je škodlivý software, jehož cílem je infiltrovat nebo poškodit počítačový systém. Po instalaci může malware krást data, šifrovat soubory a požadovat za ně výkupné nebo mazat data (ESET, 2022).
 - Ransomware je typ malwaru, který zašifruje soubory oběti. Útočníci pak požadují od oběti výkupné za obnovení přístupu k souborům – nezaplacení výkupného obvykle znamená, že oběť nemůže zašifrované soubory obnovit (ESET, 2022).
 - Phishingové útoky jsou útoky, pomocí kterých mohou hackeři vylákat z lidí jejich přihlašovací údaje nebo jiné citlivé informace. To jim může pomoci získat přístup do systému a ukrást nebo poškodit data (ESET, 2022).
 - SQL injection je typ útoků, kdy útočníci mohou vložit do databáze škodlivý kód, který jim umožní ukrást nebo poškodit data (OWASP, 2022).

3.2.2 Skupiny zálohovaných dat

Systemovými daty se rozumí data, která se týkají provozu a údržby počítače nebo jiného zařízení nebo softwaru a systémů, které na něm běží. Systemová data mohou zahrnovat širokou škálu typů dat, včetně konfiguračních souborů, systémových protokolů,

chybových hlášení, uživatelských účtů a oprávnění, nastavení systému a systémových zdrojů, jako je využití paměti a úložiště. Ta jsou obvykle uložena v systémových souborech a adresářích počítače nebo zařízení a obvykle je spravuje operační systém nebo jiný systémový software. Používají se k řízení a monitorování provozu zařízení nebo systému a také k poskytování informací o výkonu a stavu systému. Systémová data je důležité chránit a udržovat, protože jsou nezbytná pro správné fungování počítače nebo zařízení. Jejich zálohování, pravidelná aktualizace a údržba systému mohou pomoci zabránit ztrátě dat a zajistit, aby zařízení nebo systém zůstaly stabilní a výkonné (OpenAI, 2022).

Personální data (osobní údaje) jsou veškeré informace, které lze použít k identifikaci konkrétní osoby. Mohou to být např. jméno, adresa, tel. číslo, e-mailová adresa, datum narození atd. Osobní údaje mohou zahrnovat také citlivé informace, jako jsou informace o zdravotním stavu, finanční informace a biometrické údaje. Jsou často shromažďovány, používány a sdíleny organizacemi pro různé účely, včetně marketingu, výzkumu a zákaznického servisu. Je důležité, aby si jednotlivci uvědomovali, jak jsou jejich osobní údaje shromažďovány, používány a sdíleny a aby podnikli kroky k ochraně svého soukromí. V mnoha zemích existují zákony, které shromažďování, používání a sdílení osobních údajů upravují, včetně obecného nařízení o ochraně osobních údajů (GDPR) v Evropské unii a kalifornského zákona o ochraně soukromí spotřebitelů (CCPA) ve Spojených státech (OpenAI, 2022).

3.2.3 Fyzická média pro zálohování dat

CD (Compact Disc) je optický disk, který je především užíván k ukládání a přehrávání digitálního zvuku a k ukládání různého typu dat. Disky CD jsou podobné diskům DVD (Digital Versatile Disc), ale disky CD jsou menší a pojmu pouze 700 MB dat. Disky CD jsou i dnes oblíbenou volbou pro zálohování a archivaci dat, protože jsou relativně levné, snadno použitelné a široce dostupné. Jsou také odolné a vydrží drobné poškrábání a otisky prstů, takže jsou vhodné pro dlouhodobé ukládání dat. Při použití disku CD pro zálohování dat, je potřeba mít vypalovačku, což je zařízení, které je schopno zapisovat data na disk CD. Pomocí softwaru pro vypalování disků je možné vytvořit disk CD z dat, která budou zálohována, a poté jej uložit na bezpečné místo pro možnost budoucího použití (OpenAI, 2022).

DVD je zkratka pro Digital Versatile Disc nebo Digital Video Disc a jedná se o další typ optického disku, který se používá k ukládání a přehrávání digitálního videa a dalších dat. Disky DVD jsou podobné kompaktním diskům CD, protože oba používají ke čtení a zápisu dat lasery, ale disky DVD mohou ukládat podstatně více dat než disky CD. Jednovrstvý disk DVD může obsahovat až 4,7 GB dat, zatímco dvouvrstvý disk DVD až 8,5 GB. Disky DVD se běžně používají k ukládání a přehrávání filmů, hudby a dalšího multimediálního obsahu a také k zálohování a archivaci dat. Používají se také v některých počítačových systémech jako prostředek pro instalaci softwaru nebo přenos dat mezi počítači (OpenAI, 2022).

Pevný disk (Hard Disk Drive) je typ paměťového zařízení, které se používá k ukládání a načítání digitálních dat. Skládá se z rotujícího disku neboli plotny, která je potažena magnetickým materiálem, a čtecí/zapisovací hlavičky, která může číst a zapisovat data na disk. Hlavička se při otáčení disku pohybuje a umožňuje přístup k různým částem disku. Pevné disky se používají v počítačích, serverech a dalších elektronických zařízeních k ukládání operačních systémů, aplikací a dalších typů dat. Jsou obvykle větší a pomalejší než jiné typy úložišť, například disky SSD, ale jsou také levnější a lze na ně uložit větší množství dat (OpenAI, 2022).

SSD (Solid-State Drive) je stejně jako HDD paměťové zařízení sloužící k ukládání dat. Používá paměť flash na bázi NAND. Říká se mu „solid-state“, protože nemá žádné pohyblivé části, na rozdíl od tradičního pevného disku, který ukládá data na rychle se otáčející disky (plotny). SSD se skládá z desky s plošnými spoji, řadiče a jednoho nebo více paměťových čipů, které ukládají data. Při přístupu k datům uloženým na disku SSD je řadič načítá přímo z paměťových čipů, místo aby musel čekat, až se roztočí disk a získá přístup k datům jako u pevného disku. Díky tomu jsou disky SSD mnohem rychlejší a spolehlivější než HDD, zejména pokud jde o přístup k velkým souborům nebo spuštění počítače. Jsou také odolnější, protože na ně nepůsobí fyzické otřesy nebo vibrace tak jako na HDD. Disky SSD jsou však obecně dražší (což se postupem času stává minulostí) a mají omezený počet cyklů zápisu, než se začnou opotřebovávat (OpenAI, 2022).

Flash disk, známý také jako USB disk, je přenosné paměťové zařízení, které k ukládání a přenosu dat používá flash paměť. Flash disky jsou malé a lehké, takže je lze snadno přenášet a používat na cestách. Jsou levné a dostupné, což z nich činí oblíbenou volbu pro přenos a zálohování dat. Pro použití tohoto typu disku k zálohování dat je třeba poměrně velká kapacita pro uložení dat, která chceme zálohovat. Ke zkopírování dat

z počítače pak lze použít operační systém počítače nebo samostatný zálohovací software. Po přenesení dat na jednotku flash by měl být disk uložen na bezpečné místo. USB disk je pohodlnou a přenosnou možností zálohování a přenosu malých až středně velkých objemů dat. Ve srovnání s jinými možnostmi zálohování, jakou jsou externí pevné disky nebo cloudová úložiště, však mají omezenou kapacitu (OpenAI, 2022).

3.2.4 Síťová úložiště NAS, SAN, DAS – jejich porovnání a rozdíly

Síťové úložiště je typ úložného systému, který je připojen k počítačové síti a umožňuje více uživatelům nebo zařízením přístup k datům a jejich sdílení prostřednictvím sítě. Síťová úložná zařízení mají obvykle vlastní vyhrazenou úložnou kapacitu a k síti jsou připojena pomocí ethernetových kabelů nebo bezdrátového připojení (OpenAI, 2022).

Úložiště jsou navržena tak, aby k nim mohlo přistupovat a používat je více uživatelů nebo zařízení současně, a lze je použít k ukládání a sdílení nejrůznějších typů dat, včetně dokumentů, fotografií, videí, hudby a dalších typů souborů. Chytrá síťová úložná zařízení lze také nakonfigurovat k provádění různých funkcí, jako je sdílení souborů, zálohování a obnova dat a streamování videí (OpenAI, 2022).

Tato zařízení jsou efektivní pro ukládání/sdílení dat v rámci sítě a často se používají v obchodním a podnikovém prostředí k usnadnění spolupráce a přístupu ke sdíleným zdrojům. Lze je také samozřejmě použít v domácích sítích (OpenAI, 2022).

NAS (Network-Attached Storage), SAN (Storage Area Network) a DAS (Direct-Attached Storage) jsou tři různé typy chytrých ukládacích systémů, které se používají pro ukládání a při správě dat v síťovém prostředí (OpenAI, 2022).

NAS (viz obrázek 1) je připojen k počítačové síti a umožňuje více uživatelům nebo zařízením přístup k datům a jejich sdílení v rámci sítě. Mají vlastní vyhrazenou kapacitu pro ukládání a jsou k síti připojena pomocí ethernet kabelů či bezdrátově. K systémům NAS se obvykle přistupuje pomocí protokolů založených na souborech, jako je SMB (Server Message Block) nebo NFS (Network File System). Stanice NAS patří mezi nejpoužívanější typ síťového úložiště a jeho popularita stále roste (OpenAI, 2022).

Obrázek 1 Sít'ové úložiště NAS



Zdroj: Synology.com (2022)

SAN (viz obrázek 2) slouží k poskytování vysoce výkonného úložiště a přístupu k datům serverům a dalším zařízením v síti. Systémy SAN používají k připojení k serverům a dalším zařízením specializované síťové protokoly, jako je Fibre Channel nebo iSCSI, a obvykle se používají pro ukládání a správu velkého množství dat v podnikovém prostředí. Dále se používají k podpoře kritických aplikací, které vyžadují rychlý a spolehlivý přístup k datům, jako jsou databáze, e-mailové systémy a virtualizační prostředí (OpenAI, 2022).

Obrázek 2 Sít'ové úložiště SAN



Zdroj: Lenovo.com (2022)

DAS (viz obrázek 3) je úložný systém, který je přímo připojen k jednomu zařízení nebo serveru, nikoli k síti. Připojen je fyzicky pomocí kabelu USB nebo SATA. DAS jsou přístupné pomocí blokových protokolů, jako je SCSI (Small Computer System Interface) nebo ATA (Advanced Technology Attachment), a obvykle se používají k lokální správě a ukládání dat v jednom zařízení či serveru (OpenAI, 2022).

Obrázek 3 Sít'ové úložiště DAS

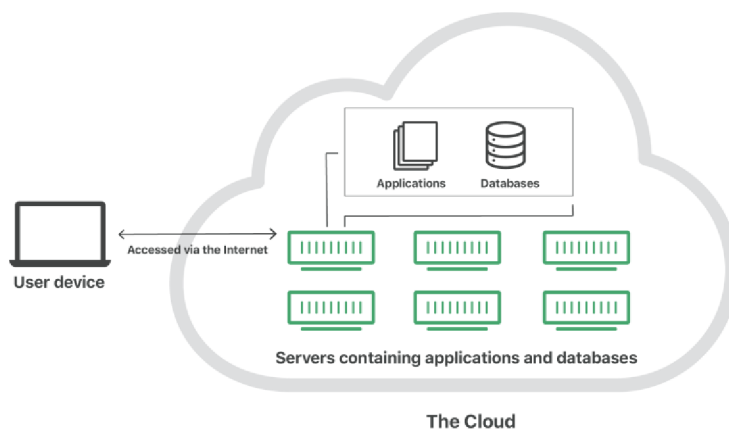


Zdroj: Petastor.com (2022)

3.2.5 Online úložiště – cloud

Pojmem cloud (viz obrázek 4) jsou označeny servery, které jsou přístupné přes internet, a software a databáze, které na těchto serverech běží. Cloudové servery jsou umístěny v datových centrech po celém světě. Díky využití cloud computingu nemusí uživatelé a společnosti sami spravovat fyzické servery ani provozovat softwarové aplikace na vlastních počítačích (Cloudflare, 2022).

Obrázek 4 Schéma cloudového úložiště



Zdroj: Cloudflare.com (2022)

Jako Cloud computing je označován „on-demand“ přístup prostřednictvím internetu k výpočetním zdrojům – aplikacím, serverům (fyzickým i virtuálním), datovým úložištím, vývojovým nástrojům, síťovým možnostem a dalším – umístěným ve vzdáleném datovém

centru spravovaném poskytovatelem cloudových služeb (CSP). CSP tyto zdroje zpřístupňuje za měsíční předplatné nebo je účtuje podle jejich využití (IBM, 2022).

V porovnání s tradičním místním IT a v závislosti na zvolených cloudových službách pomáhá cloud computing provádět následující činnosti (IBM, 2022):

- Snížit náklady na IT: Cloud umožní odložit část nebo většinu nákladů a úsilí spojených s nákupem, instalací, konfigurací a správou vlastní lokální infrastruktury (IBM, 2022).
- Zlepšit agilitu a dobu dosažení hodnoty: Díky cloudu může organizace začít používat podnikové aplikace během několika minut, místo aby čekala týdny nebo měsíce, než IT oddělení odpoví na žádost, zakoupí a nakonfiguruje podpůrný hardware a nainstaluje software. Cloud také umožní, aby si někteří uživatelé - konkrétně vývojáři a datoví vědci - mohli sami pomoci se softwarem a podpůrnou infrastrukturou (IBM, 2022).
- Snadněji a nákladově efektivněji škálovat: Cloud poskytuje elasticitu, jelikož místo nákupu nadbytečné kapacity, která zůstává nevyužitá v období útlumu, lze škálovat kapacitu nahoru a dolů v reakci na nárůsty a poklesy provozu. Možné je také využít globální síť poskytovatele cloudu a rozšířit své aplikace blíže k uživatelům po celém světě (IBM, 2022).

3.2.6 Ostatní služby pro zálohování dat

Služby zálohování dat jsou služby, které umožňují vytvářet kopie dat a ukládat je na odděleném místě k bezpečnému uložení (OpenAI, 2022).

- Lokální zálohování neboli zálohování na místě znamená proces zálohování systému, aplikací a dat na místní zařízení, jako je páska, disk, pevný disk, flash disk, CD, externí pevný disk nebo jiné médium, které je umístěno na místě v blízkosti zdroje dat. Ačkoli se doporučuje, aby organizace prováděla místní zálohování, je třeba vytvořit druhou zálohu na jiném zařízení (ideálně uloženém mimo pracoviště), aby byla data chráněna. Může být rychlejší než zálohování v cloudu, ale vyžaduje fyzický přístup k zálohovacímu zařízení a je náchylné ke stejným katastrofám, které mohou postihnout primární úložiště (Acronis, 2022).
- Hybridní zálohování kombinuje zálohování v cloudu a místní zálohování, což umožňuje ukládat část dat lokálně a část dat v cloudu. To může poskytnout pohodlí cloudového zálohování s rychlostí místního zálohování (OpenAI, 2022).

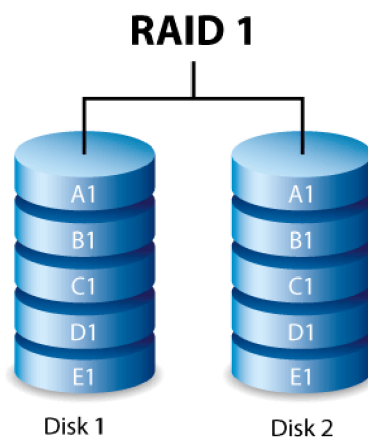
- Spravované zálohovací služby poskytují komplexnější podporu a správu pro potřeby zálohování. To může zahrnovat pomoc s nastavením, údržbou a obnovou, stejně jako pokročilejší funkce, jako je deduplikace, verzování a replikace (OpenAI, 2022).

3.2.7 RAID

Redundant Array of Independent/Inexpensive Disks (RAID) je technologie, která umožňuje ukládání dat na více pevných disků. Účelem RAID je dosáhnout redundance dat, aby se snížila ztráta dat a v mnoha případech také zlepšil výkon (Windows Cental, 2019).

- RAID 1 (viz obrázek 5) poskytuje vyšší zabezpečení dat tím, že všechna data jsou zapsána na každý disk ve svazku. Pokud dojde k selhání jednoho disku, data zůstávají k dispozici na ostatních discích ve svazku. Vzhledem k času potřebnému k vícenásobnému zápisu dat se však snižuje výkon. Kromě toho RAID 1 snižuje kapacitu disku o 50 % nebo více, protože každý bit dat je uložen na všech discích ve svazku (Seagate, 2022).

Obrázek 5 Schéma RAID 1

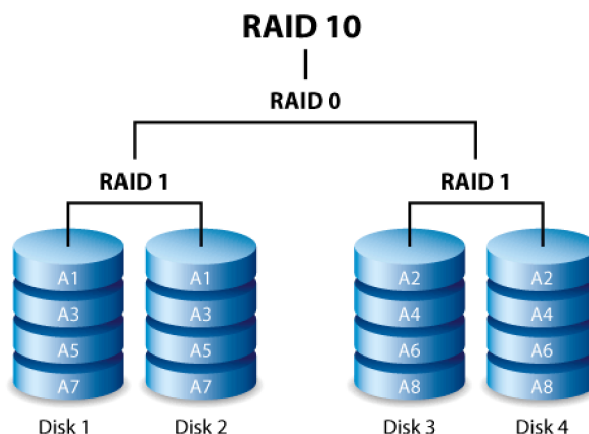


Zdroj: Seagate.com (2022)

- Seskupení disků do RAID 10 (viz obrázek 6) vezme sadu disků a vytvoří tzv. mirrored stripe. Disky v tomto typu pole RAID jsou konstruovány tím způsobem, že se vytvoří dvojice disků, které obsahují shodnou sadu bloků – říká se tomu zrcadlení. Tyto zrcadlené sady jsou pak seskupeny takovým způsobem, že pokud jsou na ně zapisována data, datové bloky se zapisují v proudech (streams) nebo prouzcích (stripes) napříč všemi zrcadlenými páry v sadě. Skupiny v RAID 10

se můžou skládat až ze čtyř disků a vždy musí mít sudý počet. Konfigurace RAID 10 toleruje selhání disku bez ztráty výkonu nebo dostupnosti. Důvodem je skutečnost, že pokud dojde k poruše disku v rámci sady, párový disk bude mít stejná data, která obsahoval vadný disk. Jakmile je disk vyměněn, data jsou zapsána zpět z párového disku na nový disk. RAID 10 má nejlepší výkonnostní charakteristiky ze všech typů úrovní RAID, přičemž vyvažuje výkon čtení i zápisu s nejvyšší úrovní ochrany. Nevýhodou jsou poměrně vysoké pořizovací náklady a složitější správa (Nelson, 2011, s. 63).

Obrázek 6 Schéma RAID 10

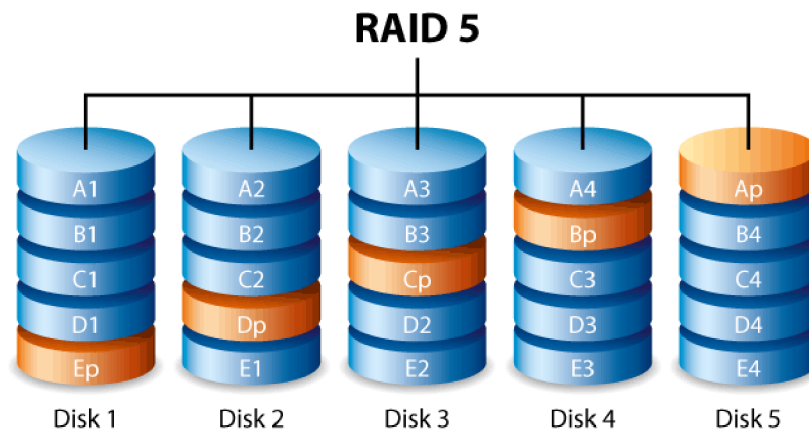


Zdroj: Seagate.com (2022)

- Pole RAID 5 (viz obrázek 7) má odlišný přístup k ochraně dat. Namísto „hrubé síly“, kdy se každý kus dat zapisuje dvakrát, metoda RAID 5 rozložila data na jednu sadu disků, ale pak vypočítá podpis datového proužku nazývaný parita, který se pak také запиše na sadu disků účastnících se pole RAID. Pomocí této metody parity může sada RAID 5 tolerovat přesně jedno selhání disku bez ztráty dat. Při poruše je také výrazně snížen výkon, stejně jako po výměně vadného disku. Když dojde k poruše disku v sadě RAID 5, nejsou již data v této části proužku dostupná a musí být regenerována „za letu“ skrz komplexní kalkulaci toho, jak by data měla vypadat, na základě informací obsažených v paritním pruhu. Tento přístup k výkonu počítače pokračuje ještě nějakou dobu po výměně disku, který selhal. Když je nový disk vložen zpět do pole RAID, data na disku musí být kompletně znovu vytvořena z paritních výpočtů. To má dopad nejen na výkon zápisu, ale také čtení. Navíc, jak se sada RAID 5 zvětšuje, prodlužuje se tak čas na

opětovné sestavení jakýchkoliv dat na konkrétním svazku, protože je třeba načíst a zpracovat větší paritu tak, aby bylo možné rekonstruovat datovou jednotku (Nelson, 2011, s. 63-64).

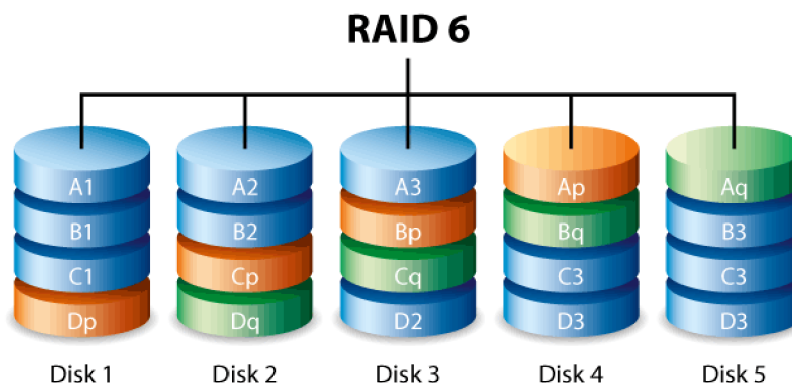
Obrázek 7 Schéma RAID 5



Zdroj: Seagate.com (2022)

- RAID 6 (viz obrázek 8) je relativně nová úroveň RAID pro zálohovací aplikace. V RAID 5 je ztráta více než jednoho disku katastrofální pro celé pole RAID, což má za následek úplnou ztrátu všech dat uložených v tomto poli, jak je popsáno výše. RAID 6 pomáhá tomuto problému předejít tím, že poskytuje model se dvěma paritami, ve kterém je parita pro jeden proužek uložena na dvou různých discích. To umožňuje selhání druhého disku během procesu obnovy prvního disku, který selhal, bez ztráty dat. Na pole RAID 6 se však stále vztahují stejné problémy jako na pole RAID 5, snížení výkonu zápisu v důsledku výpočtu druhé parity, který musí proběhnout během procesu zápisu dat. RAID 6 také snižuje množství použitelného úložiště oproti RAID 5, opět díky potřebnému prostoru pro výpočet sekundární parity (Nelson, 2011, s. 64-65).

Obrázek 8 Schéma RAID 6



Zdroj: Wikipedia.com (2022)

- JBOD je použití jednoho nebo více disků, které nejsou v konfiguraci RAID, ale jsou spravovány jako samostatné logické svazky (Western Digital, 2022).
- Spanning je kombinace jednotek lineárním způsobem k vytvoření jednoho velkého logického svazku. Výhodou použití tohoto režimu je, že se mohou přidat další disky, aniž by se musel přeformátovat systém (Western Digital, 2022).

3.2.8 Úskalí použití hardware RAID v roce 2022

Hardwarový RAID je v dnešní době v podstatě nepoužitelný pro high-endové, vysokorychlostní, 64+ jádrové systémy. Již dávno neexistuje žádná podpora na hardwarové úrovni pro systémy tohoto typu. RAID v podobě, v jaké byl znám, se vyvinul a stále se vyvíjí. Moderní hardwarové řadiče nekontrolují paritu, pokud není disk zničen nebo nehlásí chybu. Aby nedocházelo k problémům s postupným degradováním dat v důsledku nahromadění nekritických poruch v zařízení pro ukládání dat, je možné uvažovat o použití lepšího souborového systému jako BTFRS nebo ZFS (software RAID), protože ty kontrolují paritu správně. Jednou z velkých výhod software RAID je, že souborový systém ví, jaký prostor je aktuálně používán, proto se při prohledávání pole kvůli chybám, nemusí ztrácet čas s prohledáváním prázdného prostoru (Reddit - Level1Techs, 2022).

3.2.9 Zálohy a úrovně zálohování

Jako zálohy jsou nazvané kopie snímků pořízených v určitém časovém okamžiku, uložené v globálně běžném formátu a sledované po určitou dobu použitelnosti, přičemž

každá další kopie dat je udržována nezávisle na první. Lze vytvořit více úrovní záloh (Nelson, 2011, s. 16):

- Úplné zálohy (full backups) představují úplný snímek dat, která mají být chráněna. Plné zálohy jsou základem pro všechny ostatní úrovně zálohování (Nelson, 2011, s. 16).
- „Inkrementální záloha neboli přírůstková záloha je taková záloha, která obsahuje pouze data, která byla změněna od vytvoření poslední plné nebo přírůstkové zálohy. Je tedy menší než plná záloha a její tvorba zabere méně času. Protože však neobsahuje všechna data, je pro obnovení nutné mít všechny předchozí přírůstkové zálohy a původní plnou zálohu“ (Acronis, 2020).
- „Diferenciální záloha neboli rozdílová záloha, je taková záloha, která obsahuje jen ta data, která byla změněna od plné zálohy. Na rozdíl od přírůstkové zálohy, kde je každý vytvořený záložní soubor součástí „řetězce“, rozdílová záloha vytvoří nezávislý soubor obsahující všechny změny od vytvoření původní plné zálohy. Obecně by se měla rozdílová záloha obnovit rychleji než přírůstková, protože nemusí zpracovávat dlouhý řetězec předchozích záloh“ (Acronis, 2020).

3.2.10 Porovnání elementárních způsobů zálohování

Volba základní metody zálohování obecně závisí na potřebách organizace a na zdrojích, které jsou k dispozici pro zálohování a obnovu dat. Úplné zálohování poskytuje nejkompexnější ochranu před ztrátou dat, ale nemusí být praktické pro velké datové sady, které se často mění. Přírůstkové (inkrementální) a rozdílové (diferenciální) zálohování může být efektivnější, ale nemusí poskytovat úplnou kopii dat (OpenAI, 2022).

3.2.11 Software pro zálohování dat

Existuje řada softwarových programů, které umožňují provádět zálohování. Některé jsou specifické pro konkrétní výrobce, jiné jsou zcela běžně a volně dostupné. Každý z nich může mít v konkrétním prostředí určité výhody (Garfinkel, Spafford, Schwartz, 2003, s. 566).

Před samotným pořízením softwaru pro zálohování je nutné zjistit, zda řešení splňuje požadavky na tzv. RTO (cílový čas obnovy) a RPO (cílový bod obnovy). RTO udává, jak dlouhou dobu mohou být aplikace mimo provoz, aniž by došlo ke znatelným následkům spojeným s přerušením kontinuity podnikání. RPO říká, jak velké

množství dat lze ztratit, než dojde k dramatickému poškození podnikání. Tyto dva údaje určují četnost zálohování a dobu potřebnou k obnově dat. Řešení by proto mělo umožnit provádět zálohy tak často, jak je potřeba, a systém obnovy by měl umožnit obnovit data během několika minut nebo rychleji (ESET, 2022).

Cloudové řešení nebo on-promise - správným řešením zálohování by byla kombinace cloudového a lokálního řešení - pravidlo 3-2-1, kdy jsou tři kopie uloženy na dvou různých interních místech a alespoň jedna mimo firmu. Poté lze provádět běžné zálohování, které vyžaduje hodně místa v místním úložišti, a další zálohy, nebo jen nejdůležitější data, posílat do cloudu (ESET, 2022).

- Ruční a automatické zálohování - Jediný způsob pro vytvoření správné zálohy je použití specializovaného softwaru. Ze statistik vyplývá, že manuální zálohy jsou obvykle starší než půl roku. Pokud jde o obnovu, pak takové zálohy pravděpodobně nebudou obsahovat vše, co je potřeba. Velmi často se lze setkat se zálohováním dat pomocí skriptů, což se příliš nedoporučuje, protože poskytovatel softwaru nemá nad těmito skripty centralizovanou kontrolu. Rovněž optimální proces obnovy dat by měl fungovat téměř automaticky a měla by existovat snadná volba, jaká data chceme obnovit, a časový bod, od kterého je možné data obnovit (ESET, 2022).
- AMANDA (Advanced Maryland Automatic Network Disk Archiver) je open-source zálohovací software, který umožňuje správci sítě LAN nastavit jediný hlavní zálohovací server pro zálohování více hostů na jedinou velkokapacitní páskovou nebo diskovou jednotku. Amanda používá k zálohování nativní nástroje (například GNUtar, dump) a dokáže zálohovat velký počet pracovních stanic s více verzemi systémů Unix/Mac OS X/Linux/Windows (Amanda, 2017).
- IBM Spectrum Protect poskytuje škálovatelnou ochranu dat pro fyzické souborové servery, aplikace a virtuální prostředí. Tento software umožňuje organizacím spravovat až miliardy objektů na jeden zálohovací server. Je možné také snížit náklady na zálohovací infrastrukturu díky vestavěným funkcím pro efektivitu dat a ukládat data na pásky, lokální objektová úložiště i veřejné cloudové služby (IBM, 2021).
- Zálohování a obnovení je jedna z hlavních součástí zálohování systémů platformy Windows Vista, Windows 7 a novější. Dokáže vytvářet zálohy souborů a složek a také zálohy obrazů systému, které lze použít pro obnovu v případě poškození dat, selhání pevného disku nebo napadení škodlivým softwarem. Nahrazuje program

NTBackup, který byl součástí systému Windows od verze Windows NT 3.51. Na rozdíl od svého předchůdce podporuje jako zálohovací média CD, DVD a Blu-ray disky. Společnost Microsoft však v systému Windows 8 vyřadila funkci Zálohování a obnovení a doporučila používat aplikaci Historie souborů pro zálohování souborů a řešení třetí strany pro vytváření obrazů systému. Navzdory vyřazení funkce Zálohování a obnovení je však součástí všech následně vydaných verzí systému Windows (Wikipedia, 2022).

- Historie souborů nahrazuje nástroj Zálohování a obnovení jako hlavní zálohovací funkci od verze Windows 8. Automaticky vytváří přírůstkové (odkázat na téma) zálohy souborů uložených v knihovnách, včetně souborů uživatelů zapojených do skupiny HomeGroup, a složek zadaných uživatelem na jiné úložné zařízení (např. jiný interní či externí pevný disk (Wikipedia, 2022).
- Acronis Cyber Protect je softwarový balík, který má za cíl chránit systém před ransomwarem a umožňuje uživatelům zálohovat a obnovovat soubory nebo celé systémy ze záložního archivu, který byl dříve vytvořen pomocí tohoto softwaru (Acronis, 2022).

3.3 Archivace dat

Archivace dat se využívá k dlouhodobému uložení dat. Archivy jsou užitečné, pokud je nutné se vrátit k určité verzi souborů nebo pokud je nutné soubor z pracovní stanice odstranit a v případě potřeby jej později obnovit. Například je možné chtít uložit tabulky pro daňové účely, ale protože nejsou používány, není možné je ponechat na pracovní stanici (IBM, 2021).

Archivovaná data se obvykle již aktivně nepoužívají nebo nejsou potřebná pro každodenní provoz, ale stále jsou považována za cenná a je třeba je uchovat z právních, regulačních nebo obchodních důvodů (OpenAI, 2022).

Archivace dat je důležitou součástí správy dat a může organizacím pomoci zajistit přístup k důležitým historickým datům a splnit regulační požadavky (OpenAI, 2022).

3.3.1 Archivace dat versus zálohování

Velmi často dochází k záměně mezi archivací dat a zálohováním dat. Klasická zálohovací aplikace pořizuje pravidelné obrazy aktivních dat, aby poskytla metodu obnovy

záznamů, které byly smazány nebo zničeny. Většina záloh je uchovávána pouze několik dní nebo týdnů, protože pozdější záložní obrazy nahrazují předchozí verze (Computerworld, 2005).

Zálohování je v podstatě navrženo jako krátkodobá pojistka usnadňující obnovu po havárii, zatímco archiv je navržen tak, aby poskytoval průběžný rychlý přístup k podnikovým informacím trvajícím desítky let. Archivované záznamy mohou být dlouhou dobu umístěny mimo tradiční cyklus zálohování, zatímco operace zálohování chrání data, která se často mění (Computerworld, 2005).

Výkon je pro zálohování důležitým faktorem, ale protože většina operací zálohování zahrnuje velké soubory dat, je schopnost rychlého „streamování“ informací na zálohovací médium a z něj prvořadá. Rychlý náhodný přístup k malým datovým sadám při operacích obnovy je obvykle méně důležitý. Jako pojistku je také nutné minimalizovat náklady na zálohování snížením nákladů na každý uložený záznam. Médium volby pro aplikace zálohování a obnovy po havárii je tradičně pevný disk, protože splňuje kritéria výkonu a nákladů většiny organizací (Computerworld, 2005).

Požadavky na archivační úložiště jsou zcela odlišné od požadavků na zálohování (tabulka 1). V archivačních prostředích je mnohem důležitější dlouhá životnost médií a autenticita dat. Úložná média používaná v archivu by měla mít stabilní a dlouhou životnost, aby se zabránilo časté migraci dat v průběhu desítek let ukládání. Aby bylo možné dodržet firemní a vládní předpisy týkající se pravosti dat, je nezbytné, aby byly informace chráněny před modifikací. Pro datové archivy je rychlý náhodný přístup obvykle nejkritičtější výkonnostním hlediskem (Computerworld, 2005).

Tabulka 1 Rozdíly v požadavcích na zálohování a archivaci dat

| Požadavky na zálohování | Požadavky na archivaci |
|--|---|
| Vysoká kapacita médií | Autenticita dat |
| Vysoce výkonné čtení a zápis | Prodloužená životnost média |
| Nízké náklady na úložiště v přepočtu na GB | Vysoce výkonný náhodný přístup ke čtení |
| Geografické oddělení záloh | Nízké celkové náklady na vlastnictví |

Zdroj: vlastní zpracování, Computerworld (2005)

3.4 Operace s daty

3.4.1 Komprese

Komprese dat je proces kódování, restrukturalizace nebo jiné úpravy dat za účelem zmenšení jejich velikosti. V zásadě to zahrnuje překódování informací pomocí méně bitů než původní reprezentace (Barracuda, 2021).

Komprese se provádí programem, který využívá funkce nebo algoritmus k efektivnímu zjištění, jak zmenšit velikost dat. Algoritmus může například představovat řetězec bitů s menším řetězcem bitů pomocí „referenčního slovníku“ pro převod mezi nimi. Další příklad zahrnuje vzorec, který vloží odkaz nebo ukazatel na řetězec dat, který program již viděl. Dobrý příklad toho často nastává u komprese obrazu. Když je na obrázku nalezena sekvence barev, jako je „modrá, červená, červená, modrá“, vzorec může tento datový řetězec přeměnit na jeden bit, přičemž si stále zachová základní informace (Barracuda, 2021).

Komprese textu může být obvykle úspěšná, když jsou odstraněny všechny nepotřebné znaky, namísto toho vložením jediného znaku jako reference pro řetězec opakujících se znaků a následným nahrazením menšího bitového řetězce běžnějším bitovým řetězcem. Se správnými technikami může komprese dat efektivně zmenšit textový soubor o 50 % nebo více, čímž se výrazně sníží jeho celková velikost (Barracuda, 2021).

Pro přenos dat lze spustit kompresi obsahu nebo celého přenosu. Když jsou informace odesílány nebo přijímány přes internet, větší soubory, buď samotné, nebo s ostatními, nebo jako součást archivního souboru, mohou být přenášeny v jednom z mnoha komprimovaných formátů jako je ZIP, RAR, 7z nebo MP3 (Barracuda, 2021).

3.4.2 Šifrování

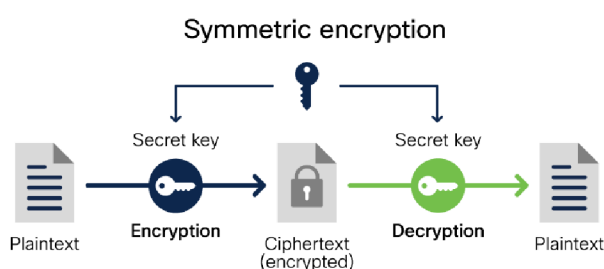
Šifrování v kybernetické bezpečnosti znamená převod dat z čitelného formátu do zakódovaného formátu. Šifrovaná data lze číst nebo zpracovávat až poté, co byla dešifrována (Kaspersky, 2021).

Šifrování je základním stavebním kamenem zabezpečení dat. Je to nejjednodušší a nejdůležitější způsob, jak zajistit, aby informace o počítačovém systému nemohly být odcizeny a čteny někým, kdo je chce použít ke škodlivým účelům (Kaspersky, 2021).

Šifrování zabezpečení dat je široce používáno jednotlivými uživateli a velkými společnostmi k ochraně uživatelských informací zasílaných mezi prohlížečem a serverem. Tyto informace mohou zahrnovat vše od platebních údajů po osobní údaje. Software pro šifrování dat, známý také jako šifrovací algoritmus nebo šifra, se používá k vývoji šifrovacího schématu, které lze teoreticky prolomit pouze s velkým množstvím výpočetního výkonu (Kaspersky, 2021).

- Symetrická šifra tento typ šifry odpovídá nejvíce tomu, co si většina lidí představí pod pojmem šifra. Jde o postup, který používá stejný klíč pro šifrování i dešifrování textu či dat obecně (viz obrázek 9). V kryptografii se ale pro data používá běžně výraz otevřený/šifrovaný text. Otevřený text je text určený k utajení, který se při procesu šifrování převede do šifrovaného textu. Dnes se běžně používá AES (Advanced Encryption Standard). Někdy je toto označení doplněno o délku klíče v bitech, tedy například AES-256. AES s 256bitovým true random (skutečně náhodným) klíčem je v současnosti pokládána za prakticky neprolomitelnou šifru, a to včetně využití kvantových počítačů při procesu lámání. AES se používá jako primární šifra pro ochranu dat obecně a používá ji například americká vláda pro utajení informací úrovně „Přísně tajné“ (ITnetwork, 2022).

Obrázek 9 Průběh symetrického šifrování

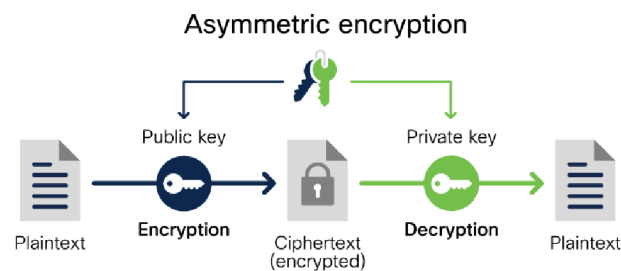


Zdroj: Cisco.com (2022)

- Asymetrická šifra využívá způsob šifrování, kde je nutný jeden klíč pro šifrování a zase druhý pro dešifrování. Je-li text zašifrován šifrovacím klíčem, je poté potřeba i dešifrovací klíč, aby bylo možné text zpět dešifrovat do otevřené podoby, přičemž z jednoho klíče nelze odvodit druhý. Proces (viz obrázek 10) asymetrického šifrování probíhá tak, že se na začátku vygeneruje z náhodných hodnot pár klíčů (key pair). Jeden klíč, tzv. tajný klíč, je pečlivě uschován do

nějakého souboru nebo na čipovou kartu tak, aby se k němu nikdo kromě vlastníka nemohl dostat. Druhý klíč, tzv. veřejný klíč, lze zveřejnit či někomu poslat otevřeným emailem. Pokud pak někdo bude chtít poslat zprávu, aniž by ji mohl číst kdokoliv jiný, stačí, když je zašifrována veřejným klíčem. Protože pouze vlastník má odpovídající tajný klíč, může ji zase dešifrovat (ITnetwork, 2022).

Obrázek 10 Průběh asymetrického šifrování



Zdroj: Cisco.com (2022)

3.4.3 Deduplikace

Deduplikace je proces analýzy dat na úrovni podsouborů a ukládání pouze těch prvků, které ještě nebyly uloženy na médiu (Nelson, 2011, s. 99).

U některých definic deduplikace je toto srovnání provedeno pouze na úrovni souboru, jako je například jeden dokument Word, který je jednotlivě uložen na více místech – všechny se zálohují. Nejedná se však o deduplikaci, ale ve skutečnosti jde o proces známý jako „úložiště jedné instance“ nebo SIS (Nelson, 2011, s. 99).

Při použití skutečné deduplikace se načtou nová data, rozdělí se na bloky, porovnají se s bloky, které jsou již uloženy, a uloží se pouze nové bloky. Bloky, které již existují, jsou vytvořeny jako „pointery“ na již uložená data. To snižuje množství dat tím, že se bloky ukládají pouze jednou pro každý následující přenos dat. Tato redukce dat se měří nikoli z hlediska množství uložených dat, ale jako poměr zpracovaných dat k uloženým datům (označovaným jako poměr deduplikace). Typicky dobrá deduplikace dosáhne poměru 10:1 a v mnoha případech bude dosahovat pravidelně přes 15:1 (Nelson, 2011, s. 100).

3.4.4 Verzování dat

Verzovací systém je systém, který zaznamenává změny souboru nebo sady souborů v průběhu času tak, aby bylo možné později vyvolat konkrétní verze. Používá se především v programátorské praxi, i když ve skutečnosti to lze dělat téměř s jakýmkoli typem souboru v počítači (Chacon, Straub, 2014).

Pokud chce grafik nebo webdesignér uchovávat každou verzi obrázku nebo rozvržení (což by zcela jistě chtěli), je systém řízení verzí (VCS) velmi rozumná věc. Umožňuje vrátit vybrané soubory do předchozího stavu, vrátit celý projekt do předchozího stavu, porovnávat změny v čase, zjistit, kdo naposledy upravoval něco, co by mohlo způsobovat problém, kdo a kdy zavedl problém a další. Použití systému VCS také obecně znamená, že pokud se něco pokazí nebo se ztratí soubory, můžou být snadno obnoveny. To vše navíc lze získat za velmi nízké režijní náklady (Chacon, Straub, 2014).

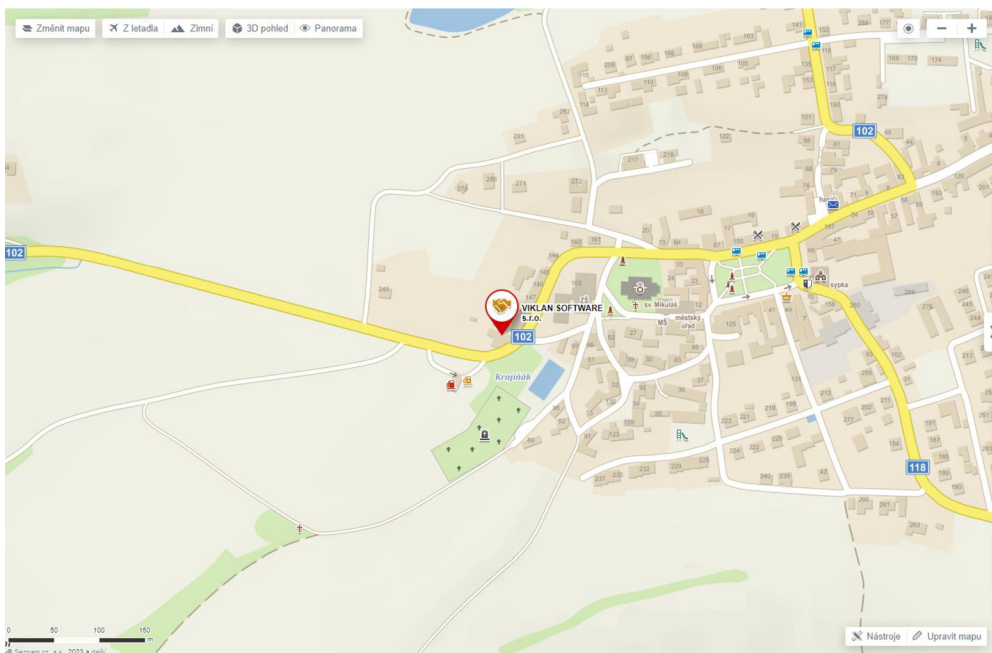
4 Vlastní práce

4.1 Specifikace společnosti

4.1.1 Geografické a základní údaje o společnosti

Firma, pro kterou se zadání zpracovává, se nachází v malém městě Krásná Hora nad Vltavou nacházející se v okrese Příbram, zhruba 15 km od města Sedlčany (viz obrázek 11). Název firmy je VIKLAN SOFTWARE s.r.o. (dále jen „firma“, „společnost“). Počet zaměstnanců se pohybuje do 10 lidí, včetně několika pracovníků, kteří pracují externě.

Obrázek 11 Sídlo společnosti VIKLAN SOFTWARE s.r.o.



Zdroj: Mapy.cz (2023)

4.1.2 Předmět činnosti společnosti

Společnost se zabývá implementací a vytvářením moderních systémů a aplikací, jako jsou databázové systémy, výpočetní programy pro technické účely, programy pro web, ERP systémy (účetnictví, IS podniku), nadstavby nad programy MS Office, či výpočetní programy v cloudu.

Aktuální hlavní náplní práce je vytváření knihoven na zakázku pro ERP systém Money S5, kdy firma vlastní oficiální licence a certifikace pro vytváření těchto knihoven. Dále převod historických programů do moderní podoby a správa a servis již zavedených aplikací a systémů.

4.1.3 Organizace společnosti

Společnost tvoří okolo 10 zaměstnanců, jak je uvedeno výše. Jedná se tedy o mikro podnik, kdy jeden zaměstnanec mnohdy zastává více rolí. Majitel firmy je zároveň senior-programátor a manažer. Ve firmě je zaměstnán ještě jeden senior-programátor a zároveň konzultant, několik dalších programátorů, dva junior-programátoři, jeden QA tester a brigádník. Společnost nemá žádné účetní či právní oddělení, tyto služby jsou dodávány externím subjektem.

4.1.4 Technické zázemí společnosti

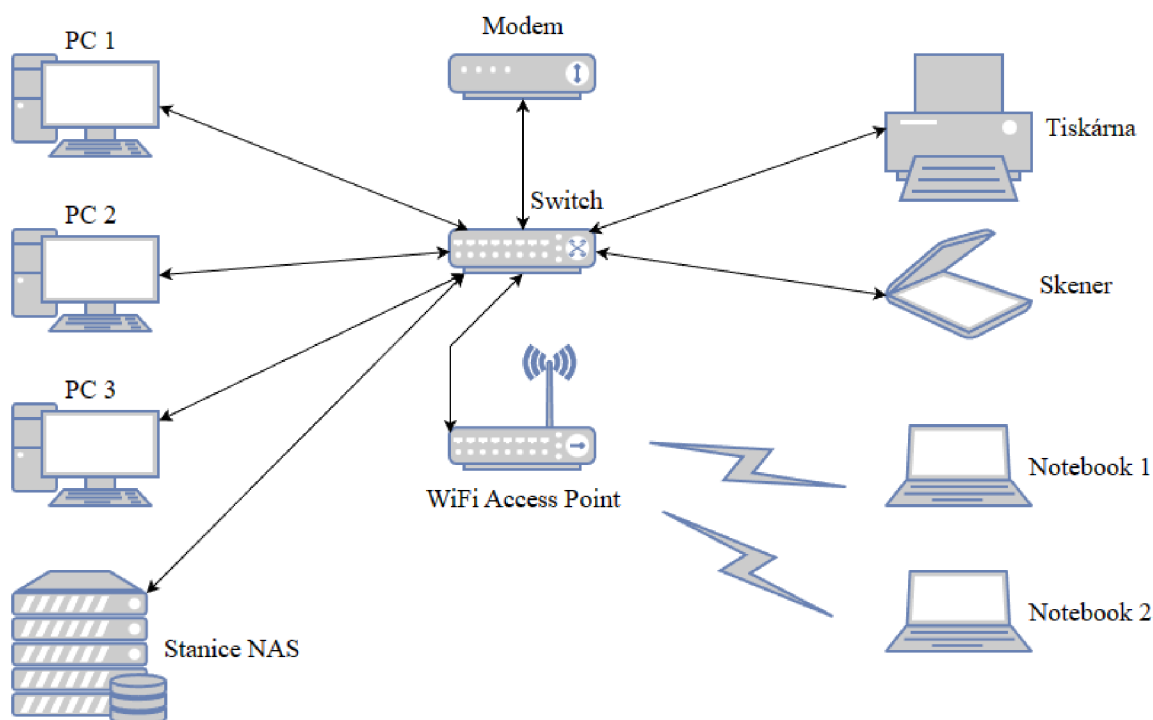
V době zpracovávání této práce probíhalo stěhování do nových prostor, které by měly poskytnout komfortnější zázemí a možnost lepší organizace. Dosud seděli programátoři a majitel firmy samostatně v oddělených místnostech. Hlavní místnost, kde se nacházejí programátoři, je poměrně malá, je zde velmi obtížná manipulace a je nutná kompletní reorganizace.

Technické zázemí firmy není příliš dobré. Hardwarové vybavení je spíše průměrné. Každý zaměstnanec má svojí pracovní stanici – někteří firemní, jiní využívají soukromé notebooky. Hardware v noteboocích a stolních počítačích je podobný – využívají procesory Intel Core i5, k dispozici je 8 GB operační paměti RAM, stolní PC obsahují většinou 1TB disk firmy Western Digital a notebooky jsou osazeny SSD disky o velikostech pohybující se od 256 do 512GB. K dispozici je sdílená tiskárna, skener a zastaralá NAS stanice osazena 1TB diskem, která plní funkci pro zálohování, verzování a sdílení souborů. Pracovní stanice a další hardwarové doplňky jsou k síti připojeny hvězdicovou topologií. Síťové prvky, které tvoří podnikovou síť, jsou switch, WiFi access point a modem od poskytovatele internetu (viz obrázek 12), centrální server není zaveden. Společnost je k internetu připojena bezdrátově, jelikož po městě není rozvedena optická síť.

Pracovní stanice mají nainstalovaný systém Windows 10, verze Pro nebo Home. Pro programování a vývoj je využíváno integrované vývojové prostředí Visual Studio

nejnovější verze, spolu s MS SQL serverem. Komunikace mezi zaměstnanci je vedena skrz pracovní i soukromé emaily. Verzování souborů se provádí manuálně, a to předáním vedoucím pracovníkům, kteří následně pomocí programu Total Commander soubory upravují a spojují. Ve společnosti není implementován žádný centrální informační systém.

Obrázek 12 Přibližné schéma síťové organizace společnosti



Zdroj: vlastní zpracování, použitý nástroj: app.diagrams.net

4.1.5 Aktuální stav zálohování a správy dat

Zálohování a správa dat jsou zpracovávány přes stanici NAS, která je již na hranici své životnosti. Stanice je osazena jedním 1TB diskem a její kapacita je využita zhruba ze 40 %. Zápis a čtení dat je velmi pomalé. Současný celkový stav celého procesu je nepřijatelný, potřebuje kompletní renovaci, a to především po stránce bezpečnostní a technické. Není jasně a jednoznačně stanoven žádný plán zálohování, ani správy dat. Zaměstnanci si zálohují data dle své potřeby na různá místa, nebo nezalohují vůbec. Verzování souborů, které je popsáno výše, není vyhovující a bezpečné. Z toho důvodu není vhodné ho tímto způsobem provozovat.

4.1.6 Jaká data společnost zpracovává a uchovává, jejich rozdělení

Firma pracuje především se soukromými daty společností, pro které zpracovává zakázky. Data tohoto typu tvoří největší podíl v procesu zpracovávání a uchovávání. Data nejsou pečlivě zabezpečena a jsou v rámci podniku sdílána nevhodným způsobem.

Dále uchovává data vlastní – především dokumenty, smlouvy, faktury, zdrojové kódy aplikací a systémů, fotografie, komunikace s klienty apod.

4.1.7 Možné příčiny a následky ztráty dat

Po analýze a charakterizování aktuálního stavu společnosti můžeme přejít k možným hrozbám, které mohly a mohou nastat, pokud podnik svojí celkovou správu a zabezpečení dat nezmění. Celý aktuální systém spoléhá na to, že uživatelé nebudou dělat chyby.

Pokud budeme vycházet z teoretické části této práce, hlavní tři faktory, které se jeví jako největší problém pro danou společnost, jsou: kybernetická hrozba, fyzické poškození či krádež dat a jako poslední – chyba zaviněná lidským faktorem.

Následné potíže, pokud by jeden z těchto problémů nastal, mohou mít pro firmu až likvidační dopad. Ztráta clientských dat nejenže poškodí dobré jméno firmy, ale může znamenat právní dohru v podobě žaloby ze strany klientů. Ztráta soukromých (interních) dat firmy naopak může vést ke ztrátě know-how společnosti, odcizení licencí a smluv, případně dalších soukromých informací o zaměstnancích, které navíc podléhají GDPR (Obecnému nařízení o ochraně osobních údajů).

4.1.8 Požadavky společnosti

Jelikož společnost, pro kterou je zpracováváno zadání, spadá do kategorie mikro podniku, bude kladen důraz především na efektivnost a co nejlepší poměr cena-výkon. Společnost by ráda celou svou infrastrukturu týkající se správy dat předělala, a to do takové podoby, aby splňovala aspoň minimální požadavky na bezpečnou práci s daty.

Hlavními požadavky jsou kompletní renovace zálohování a archivace dat, včetně technického vybavení, aby byla splněna obecná pravidla těchto úkonů. Tohoto cíle dosáhneme zvýšením bezpečnosti podnikové sítě a její reorganizací, migrací verzování z manuální do moderní podoby a v neposlední řadě sjednocením a bezpečným centralizováním pracovních procesů v rámci společnosti, u kterých to bude možné.

4.2 Návrh zálohovacích strategií

Při návrhu zálohovací strategie bylo nejdříve potřeba reorganizovat samotné technické zázemí. Počítače a veškerá technika byla přesunuta do nových prostor, kde se následně zřídilo zcela nové zapojení a zprovoznění firemní sítě. Firma upgradovala svůj síťový switch, sjednotily se ethernetové kabely (které byly i oštitkovány pro lepší orientaci) a také bylo přihlíženo k co nejlepšímu „cable managementu“ pro další přípravné úpravy, či nutnosti oprav. Při zapojení nové NAS stanice a její konfiguraci bylo již vše děláno v rámci nového zázemí. Firma navíc přešla na nový, moderní způsob verzování souborů pomocí nástroje GIT.

4.2.1 1. varianta návrhu – stanice NAS

Firma aktuálně využívá stanici NAS, která se pohybuje na hranici životnosti, navíc je nevhodně umístěna, nakonfigurována a prakticky se dá říct, že slouží spíše jako „externí disk“, přičemž není využit její plný potenciál. Řešení touto metodou není špatné, ale musí být provedeno správně.

Provede se srovnání nejnovějších stanic NAS, jejich vlastností, funkcí, konfigurace disků a osazení hardwaru. Požaduje se, aby mohla být stanice osazena nejméně dvěma, nejlépe čtyřma a více pevnými disky o velikosti 3,5" nebo 2,5". Musí být zajištěna podpora diskového pole RAID a pro vzdálenou správu je vhodná aplikace nutností. Výhodou je též možnost rozšíření operační paměti RAM (DDR4). Maximální cena stanice nesmí přesáhnout 20 000 Kč.

Stanice musí být osazena vhodnými disky. Pro NAS lze použít jak klasické pevné disky, tak speciální, určené právě pro použití se stanicemi NAS.

Porovnávání bude provedeno pomocí výběru kompromisní varianty metodou váženého součtu. Při kvantifikaci procesorů budou použity online benchmarky daných modelů.

- I. Výběr vhodných stanic NAS dle požadovaných vlastností na stránkách výrobců, případně prodejců (ceny jsou určeny jako průměr srovnání mezi několika e-shopy a v budoucnu se mohou výrazně lišit).
- II. Výpis modelů a jejich vlastností do tzv. matice variant (tabulka 2).
- III. Vlastnosti jednotlivých modelů jsou převedeny na váhy neboli preference.

- IV. Všechny vlastnosti (informace) se musejí kvantifikovat, použita je bodovací metoda a určuje se vlastnost kritérií (maximalizační, minimalizační).
- V. Kritéria jsou následně převedena na stejný druh → maximalizační a je určena ideální (maximální hodnoty) a bazální varianta (minimální hodnoty, opak ideální varianty).
- VI. V posledním kroku je vytvořena standardizovaná kritériální matice (tabulka 3) a užitek, ze kterého určíme pomocí pořadí nejvhodnější variantu.

Tabulka 2 Matice variant pro síťová úložiště

| MODEL | CENA | POČET SLOTŮ PRO DISKY | CPU | RAM | MAX. VELIKOST RAM |
|------------------|-----------|-----------------------|--|------|-------------------|
| Synology DS923+ | 16 200 Kč | 4 | AMD Ryzen R1600 3,1 GHz | 4 GB | 32 GB |
| Synology DS1522+ | 19 000 Kč | 5 | AMD Ryzen R1600 3,1 GHz | 8 GB | 32 GB |
| QNAP TS-433-4G | 11 900 Kč | 4 | ARM Cortex-A55 2,0GHz | 4 GB | 4 GB |
| QNAP TS-464-4G | 17 300 Kč | 4 | Intel Celeron N5095 2,9 GHz | 4 GB | 16 GB |
| QNAP TS-364-8G | 15 300 Kč | 3 | Intel Celeron Quad-Core N5105/N5095 2,9GHz | 8 GB | 8 GB |
| Synology DS723+ | 12 600 Kč | 2 | CPU AMD Ryzen R1600 2,6 GHz | 2 GB | 32 GB |

Zdroj: vlastní zpracování, Alza.cz (2023)

Tabulka 3 Kritériální matice pro síťová úložiště

| MODEL | CENA | POČET SLOTŮ PRO DISKY | CPU | RAM | MAX. VELIKOST RAM | UŽITEK | POŘADÍ |
|------------------|-------|-----------------------|-------|-------|-------------------|--------|--------|
| Synology DS923+ | 0,394 | 0,667 | 0,429 | 0,333 | 1,000 | 0,521 | 4. |
| Synology DS1522+ | 0,000 | 1,000 | 0,429 | 1,000 | 1,000 | 0,589 | 1. |
| QNAP TS-433-4G | 1,000 | 0,667 | 0,000 | 0,333 | 0,000 | 0,547 | 3. |
| QNAP TS-464-4G | 0,239 | 0,667 | 1,000 | 0,333 | 0,429 | 0,506 | 5. |
| QNAP TS-364-8G | 0,521 | 0,333 | 1,000 | 1,000 | 0,143 | 0,588 | 2. |
| Synology DS723+ | 0,901 | 0,000 | 0,429 | 0,000 | 1,000 | 0,457 | 6. |

Zdroj: vlastní zpracování

Po aplikování metody vícekritériální analýzy vychází jako nejvhodnější model Synology DS1522+. Toto zařízení je sice dražší, ale nabízí jako jediné 5 slotů pro disky, 8 GB paměti DDR4 RAM (rozšiřitelná až na 32 GB DDR4 RAM) a procesor AMD Ryzen R1600 3,1 GHz. Na stránkách výrobce je navíc uvedeno, že stanici lze velmi dobře modifikovat dle potřeb uživatele, pomocí přídatného modulu stanici rozšířit až o 1 % šachet pro další disky. Stanice má čtyři vestavěné 1GbE porty s možností rozšíření na 10GbE. Stanici obsahuje NVMe SSD disk jako mezipaměťové úložiště, což přináší mnoho výhod, hlavně pro zápis dat. Stanice je dodávána s vlastním softwarem, který otevírá možnosti nastavit další funkce NAS. S ohledem na vlastnosti a rozšiřitelnost, nebude firma nucena v budoucích letech stanici měnit.

Po výběru stanice je nutné vybrat vhodné disky. Synology DS1522+ podporuje disky 3,5" SATA HDD a 2,5" SATA SSD. Je nutné přihlídnout k tomu, že se firma v následujících letech nebude rozrůstat a nebude chtít konfiguraci NAS příliš měnit. Firma

nepotřebuje velkou kapacitu disků, nepracuje s příliš velkými daty, kapacita do 4TB/1 disk by měla být naprosto dostačující. Volíme tedy dva disky o velikosti 4 TB. Využito bude jednoduché diskové pole RAID 1, kdy se data ukládají na 2 disky zároveň a probíhá tzv. zrcadlení (viz kapitola RAID). Diskové pole tohoto typu sice nezrychlí proces čtení a zápisu, na druhou stranu je plně dostačující pro aktuální potřebu firmy a vzhledem k použití pouze 2 disků je to vhodné řešení. Pro přístup k datům, které se musí rychle předávat, bude osazeno navíc jedno SATA SSD o velikosti 500 GB. Na tomto disku nebudou data dlouho uložena a bude využíván čistě jako „tunel“ pro předávání dat.

Disky by měly být optimalizovány pro použití se stanicemi NAS. Disk by měl mít co nejvyšší životnost, být uzpůsoben pro nepřetržitý 24/7 provoz. Otáčky těchto disků se pohybují mezi 5400-7200 otáčkami za minutu. Připojení je zajištěno pomocí rozhraní SATA III. Vyrovňovací paměť cache větší velikosti je výhodou (256 MB+). Velikost je 3,5". Velmi důležitým parametrem je použitá technologie, resp. design disku. U disků určených pro použití ve stanicích NAS se nejčastěji setkáváme s technologiemi CMR a SMR, v našem případě budeme chtít disky s technologií CMR (viz níže).

SSD disk musí mít připojení pomocí SATA III, bude se hledět především na co nejlepší poměr cena/výkon a co nejvyšší rychlost čtení a zápisu. Udávaná životnost disku je dalším důležitým kritériem, ke kterému je potřeba přihlídnout.

CMR (Conventional Magnetic Recording) je konvenční magnetický záznam. Od roku 2000 se jedná o nejpoužívanější a standardní technologii ukládání dat na pevný disk. Zápis dat probíhá do magnetické stopy, které jsou vedle sebe, nepřekrývají se a operace zápisu na jednu stopu neovlivňují její sousedy.

SMR (Shingled Magnetic Recording) je novější typ zaznamenávání dat. Jedná se o tzv. šindelový magnetický záznam – šindelový, protože datové stopy si lze představit jako střešní šindele (částečně se překrývají). Díky tomu jsou výsledné stopy tenčí, takže se jich na danou plochu vejde více a dosáhne se lepší celkové hustoty dat.

Postup porovnání pevných disků HDD a SSD je obdobný jako u diskových stanic NAS a bude použita stejná metoda (tabulka 4), tedy kompromisní varianta váženého součtu.

Tabulka 4 Matice varianta pro HDD

| MODEL | CENA | OTÁČKY | RYCHLOST ČTENÍ/ZÁPISU | ZÁRUČNÍ DOBA | TECHNOLOGIE | CACHE | KAPACITA |
|------------------------------|----------|--------------|-----------------------|--------------|-------------|-------|----------|
| Seagate IronWolf 4TB CMR | 2 200 Kč | 5 400 ot/min | 202 MB/s | 3 roky | CMR | 256MB | 4 TB |
| Seagate IronWolf Pro 4TB CMR | 2 900 Kč | 7 200 ot/min | 214 MB/s | 5 let | CMR | 256MB | 4 TB |
| WD Red Plus 4TB | 2 400 Kč | 5 400 ot/min | 210 MB/s | 3 roky | CMR | 256MB | 4 TB |
| WD Red Pro 4TB | 3 200 Kč | 5 400 ot/min | 217 MB/s | 5 let | CMR | 256MB | 4 TB |

Zdroj: vlastní zpracování, Alza.cz (2023), Seagate.com (2023), Westerndigital.com (2023)

Tabulka 5 Kriteriační matice pro HDD

| MODEL | CENA | OTÁČKY | RYCHLOST ČTENÍ/ZÁPISU | ZÁRUČNÍ DOBA | UŽITEK | POŘADÍ |
|------------------------------|-------|--------|-----------------------|--------------|--------|--------|
| Seagate IronWolf 4TB CMR | 1,000 | 0,000 | 0,000 | 0,000 | 0,310 | 4. |
| Seagate IronWolf Pro 4TB CMR | 0,300 | 1,000 | 0,800 | 1,000 | 0,733 | 1. |
| WD Red Plus 4TB | 0,800 | 0,000 | 0,533 | 0,000 | 0,381 | 3. |
| WD Red Pro 4TB | 0,000 | 0,000 | 1,000 | 1,000 | 0,410 | 2. |

Zdroj: vlastní zpracování

Z porovnání nám vyplývá (tabulka 5), že nejlepší možný disk pro naše účely je Seagate IronWolf Pro 4TB CMR. Disky IronWolf od firmy Seagate jsou určeny přímo pro použití v NAS stanicích. Při zakoupení tohoto disku zároveň získáváme Seagate Rescue Service microsides, což nám umožňuje v případě poruchy disku zaslat disk přímo výrobci, který se pokusí zachránit data na něm uložená. Záruka 5 let, která je k disku dodávána, je skvělý bonus. Samotný disk stojí okolo 2 900 Kč, nákup páru bude stát 5 800 Kč.

Nejlepší SSD disk je Samsung 870 EVO, což nám opět potvrdila analýza (tabulka 6, 7). SSD disky od firmy Samsung jsou obecně velmi spolehlivé a jedny z nejpoužívanějších na trhu. Disk se může pochlubit skvělou rychlostí čtení a zápisu, a především dlouhou dobou životnosti 300 TBW (terabytes written).

Tabulka 6 Matice variant pro SSD

| MODEL | CENA | RYCHLOST ČTENÍ | RYCHLOST ZÁPISU | ŽIVOTNOST | ZÁRUČNÍ DOBA | KAPACITA |
|--------------------------|----------|----------------|-----------------|-----------|--------------|----------|
| Samsung 870 EVO | 1 300 Kč | 560 MB/s | 530 MB/s | 300 TBW | 5 let | 500 GB |
| WD Blue SA510 SATA | 1 000 Kč | 560 MB/s | 510 MB/s | 200 TBW | 5 let | 500 GB |
| ADATA Ultimate SU800 SSD | 1 400 Kč | 560 MB/s | 520 MB/s | 400 TBW | 3 roky | 512 GB |
| Crucial MX500 | 1 100 Kč | 560 MB/s | 510 MB/s | 180 TBW | 5 let | 500 GB |
| Kioxia Exceria | 1 000 Kč | 555 MB/s | 540 MB/s | 120 TBW | 3 roky | 480 GB |

Zdroj: vlastní zpracování, CZC.cz (2023)

Tabulka 7 Kriteriaální matice pro SSD

| MODEL | CENA | RYCHLOST ČTENÍ | RYCHLOST ZÁPISU | ŽIVOTNOST | ZÁRUČNÍ DOBA | UŽITEK | POŘADÍ |
|--------------------------|-------|----------------|-----------------|-----------|--------------|--------|--------|
| Samsung 870 EVO | 0,250 | 1,000 | 0,667 | 0,643 | 1,000 | 0,686 | 1. |
| WD Blue SA510 SATA | 1,000 | 1,000 | 0,000 | 0,286 | 1,000 | 0,659 | 2. |
| ADATA Ultimate SU800 SSD | 0,000 | 1,000 | 0,333 | 1,000 | 0,000 | 0,463 | 4. |
| Crucial MX500 | 0,750 | 1,000 | 0,000 | 0,214 | 1,000 | 0,586 | 3. |
| Kioxia Exceria | 1,000 | 0,000 | 1,000 | 0,000 | 0,000 | 0,460 | 4. |

Zdroj: vlastní zpracování

Samotné konfigurování NAS je velmi jednoduché – po osazení diskových šachet námi zvolenými disky, jsme připojili napájecí a síťový kabel. Nyní lze provést konfiguraci z jakéhokoli počítače, který se nachází ve stejné síti. Po spuštění webového prohlížeče a po zadání do adresového řádku: synologynas:5000, se spustí Web Assistant, který v síti vyhledá stanici NAS a pomůže ji nainstalovat.

Nainstaluje se webový operační systém DSM (Synology DiskStation Manager) ve verzi 7.1. Po dokončení instalace je potřeba založit administrátorský účet a zaregistrovat se. Následně je zobrazena plocha operačního systému stanice NAS. Pro konfiguraci disků následuje přesun do Správce úložiště, kde je zvoleno, pro jaké disky a jaký typ pole RAID (RAID 1) a souborový systém (Btrfs) se má vytvořit. V tuto chvíli lze NAS používat, ale konfigurace v tomto případě ještě není dokončena. Dále je nutné nastavit sdílení souborů dle instrukcí, případně přiloženého návodu.

K nastavení zálohování dat je nutné použít vyspělý a intuitivní balíček Active Backup Suite, který je součástí operačního systému DSM od Synology. Po nastavení tohoto balíčku lze skrz konzoli sledovat všechna zařízení a pomocí portálu pro obnovení kontrolovat a spravovat zálohovaná data. Tento balíček je velmi komplexní a lze zde nastavit opravdu cokoli, co je požadováno. Zálohovat lze např. osobní PC s Windows, fyzické servery s Windows a Linux, virtuální počítače, souborové servery, případně i další zařízení NAS.

Poslední krok, který je nutné nastavit pro možnost použití stanice NAS v rámci firmy, je vytvoření místních uživatelů a skupin a dále „zobrazit“ NAS na dalších počítačích v rámci sítě. Jednotliví uživatelé si musí nejprve založit svůj vlastní uživatelský účet, tento krok se provede v části Ovládací panel > Uživatel a skupina > Uživatel. Zvolí se vytvoření nového uživatele, musí se zadat jméno, email a heslo. Po přidání uživatele lze pokračovat nastavením přístupů, přidáním uživatele do skupin, oprávněním ke sdíleným složkám apod. Vytvoření skupiny pro uživatele probíhá obdobně, Ovládací panel > Uživatel a skupina > Skupina. Vytvoří se skupina a přiřadí se do ní členové, u kterých je to potřeba.

Skupině lze nastavovat oprávnění podobně jako pro konkrétní uživatele. Logický přístup je takový, že po vytvoření všech uživatelů a skupin přiřadíme každému uživateli soukromou složku, do které bude mít přístup pouze on. Složka bude využívána pro věci firemního charakteru, nikoli soukromé účely. Složka každého uživatele bude mít stejnou velikost.

Aby každý zaměstnanec/uživatel mohl vidět NAS na svém počítači, stačí ve složce Tento počítač toto zařízení přidat pomocí tlačítka Připojit síťovou jednotku.

Tento návod slouží pouze pro potřeby bakalářské práce, je tedy obecnějšího charakteru. Veškeré podrobnější informace jsou popsány v návodu přiloženém ke stanici NAS, nebo na stránkách výrobce (viz odkaz).

4.2.2 2. varianta – použití specializovaného softwaru pro zálohování

Softwarové zálohování dat je populární, avšak pro náš případ je to zcela nevhodné řešení, především z finančního hlediska a kvality softwaru.

Pokud by byl vybrán tento způsob řešení, byl by potřeba nákup licencí, a to buď jednorázových s trvanlivostí navždy, případně se licence platí na roční bázi. Licence jsou velmi nákladné (zhruba 2 000 Kč / jedna pracovní stanice).

V případě použití bezplatného, open-source řešení, nejsou k dispozici důležité nástroje, které nabízejí například diskové stanice či placený software a v prostředí firmy jsou velmi uplatnitelné a nedá se bez nich obejít.

4.2.3 3. varianta – kombinace NAS a cloud

Budeme vycházet z 1. varianty návrhu. NAS stanice s disky jsou vybrány a nyní dojde k výběru cloudové služby. Popularita cloudových služeb v posledních letech exponenciálně stoupá a v dnešním moderním online světě si bez nich již život představit nelze. Přidáním cloudové služby do našeho návrhu zálohování získáme možnost umístit naše zálohy, soubor apod. do online úložiště mimo naši firmu. Výhodou je tedy teoretická nemožnost odcizení našich dat a v případě celkového selhání systému možnost obnovy z místa, které je odděleno od firemní sítě a firemních systémů.

Největší výhoda cloudových služeb však spočívá v tom, že se dají škálovat dle potřeb uživatele.

Požadavky jsou, aby cloudová služba byla poskytována spolehlivou a ověřenou společností, která má v této oblasti zkušenosti. Velkou roli hraje samozřejmě cena

poskytované služby a možnost škálování. Premisa je taková, že se na cloudové úložiště budou jedenkrát týdně ukládat kritické dokumenty, soubory a také zálohy systémů.

Tabulka 8 Porovnání cloudových služeb

| Název | Cena | Velikost úložiště | Celková cena za měsíc |
|--------------------|------------------------------|-------------------|-----------------------|
| Microsoft OneDrive | 246 Kč/měsičně za uživatele* | 1 TB/uživatel | 10*246 = 2 460 Kč |
| DropBox | 280 Kč/měsičně za uživatele | celkově 5 TB | 10*280 = 2 800 Kč |
| Google Cloud | 282 Kč/měsičně za uživatele | 2 TB/uživatel | 10*282 = 2 820 Kč |
| iDrive Team | 45 Kč/měsičně za uživatele | celkově 10 TB | 10*45 = 450 Kč |

* včetně celého kancelářského balíčku Office

Zdroj: vlastní zpracování, Microsoft.com, Dropbox.com, Google.com, iDrive.com

Při porovnání cloudových služeb (tabulka 8) vyšlo jako nejlepší možné řešení iDrive Team od společnosti iDrive. Toto řešení jako jedno z mála navíc nabízí intuitivní zálohovací možnosti, které jsou zabudované přímo v poskytovaném softwaru. Navíc je zde integrované verzování souborů a dalších dat, což je další z kladů řešení od společnosti iDrive.

Alternativní řešení nabízí společnost Microsoft, která ke svému cloudu OneDrive navíc nabízí celý kancelářský balíček Office v ceně – Word, Excel, PowerPoint atd. 1 TB místa na uživatele není příliš, ale pro účely firmy je to naprosto dostačující. Vzhledem k tomu, že veškeré firemní pracovní stanice a notebooky běží na systému Windows, bude adaptace tohoto řešení jednoduchá.

5 Výsledky a zvolení nejlepší varianty návrhu zálohování

V této kapitole je vyhodnocena praktická část práce, je stanoveno výsledné a nejvhodnější řešení pro firmu, pro kterou byla praktická část zpracovávána. Dále jsou zde zobrazeny klady a zápory jednotlivých variant návrhu řešení (viz obrázek 13, 14).

Po přihlédnutí ke srovnání a analýze vypracovaných variant je nejvhodnější zvolit návrh č. 3 – stanice NAS a cloudové úložiště. Tento návrh jako jediný umožňuje splnit tzv. pravidlo 3-2-1. Celá praktická část klade nejvyšší důraz především na nejlepší poměr cena/výkon, podle toho byly zvoleny váhy jednotlivých vlastností analyzovaného a porovnávaného hardwaru.

Použití diskové NAS stanice Synology DS1522+ s dvěma HDD disky Seagate IronWolf Pro 4TB CMR a jedním SSD diskem Samsung 870 EVO zajistí bezproblémový provoz a dlouhou životnost, včetně možnosti dalšího škálování a konfigurování zálohování dat. Technologie diskového pole RAID 1 bude chránit firmu před možnou ztrátou dat, zároveň s použitím cloudových služeb zajistí, že společnost nebude nucena v následujících několika letech řešit tuto problematiku a bude se moct zaměřovat především na svoji hlavní činnost, tedy vývoj aplikací a systémů.

Vybrané cloudové řešení od společnosti iDrive splňuje všechny požadované vlastnosti a nabízí mnoho možností konfigurace, které samotné zálohování dokáže ještě více zautomatizovat a zjednodušit.

Obrázek 13 Klady, zápory, nové možnosti a hrozby návrhu zálohování č. 1

NÁVRH Č. 1

| | | | |
|----------------------|---|--|---------------|
| KLADY | <ul style="list-style-type: none">Ověřená technologieDlouhodobá životnostNízké náklady na provozKvalitní zpracováníProvoz 24/7Konfigurovatelnost | <ul style="list-style-type: none">HlučnostZávislost na dostupnosti elektrického prouduZávislost na dostupnosti internetuVyšší náklady na pořízeníNutnost aktualizací | ZÁPORY |
| NOVÉ MOŽNOSTI | <ul style="list-style-type: none">Rozvoj podnikáníDůvěra zákazníkůMožnost pracovat na více projektechZlepšení workflow | <ul style="list-style-type: none">Možnost hacknutí diskové stanice v případě špatné manipulaceSelhání obou diskůFyzická krádež | HROZBY |

Zdroj: vlastní zpracování

Obrázek 14 Klady, zápory, nové možnosti a hrozby návrhu zálohování č. 1

NÁVRH Č. 2

| | | | |
|-------|---|--------|---|
| KLADY | Ověřená technologie Dlouhodobá životnost Nízké náklady na provoz Kvalitní zpracování Provoz 24/7 Konfigurovatelnost +Škálovatelnost +Pravidlo 3-2-1 +Nemožnost odcizení dat | ZÁPORY | Hlučnost Závislost na dostupnosti elektrického proudu Závislost na dostupnosti internetu Vyšší náklady na zařízení Nutnost aktualizací -Bez internetu je cloud nepřístupný -V určitých ohledech může být cloud pomalý |
| | NOVÉ MOŽNOSTI | | Rozvoj podnikání Důvěra zákazníků Možnost pracovat na více projektech Zlepšení workflow +Přístupná data kdekoliv na světě |

Zdroj: vlastní zpracování

6 Závěr

Cílem bakalářské práce byl návrh zálohovací strategie pro malou firmu. Zálohování by nemělo být v dnešní době opomíjeno ani malými firmami, hlavně pokud pracují s citlivými daty svých zákazníků, což byl i případ firmy, pro kterou byla strategie zpracovávána. Významnou roli hrála finanční dostupnost řešení a z toho vyplývající co nejlepší poměr cena a výkon se zaměřím na použití co nejmodernějších součástí.

K dosažení tohoto cíle bylo použito studium literatury a materiálů zabývajících se problematikou zálohování a archivace dat.

První teoretická část práce slouží především jako výchozí podpora pro část druhou, respektive praktickou část. Jsou zde vysvětleny základní pojmy a principy, ale také srovnání nejrozličnějších hardwarových a softwarových komponent používaných pro zálohování a archivaci dat.

Ve vlastním zpracování byly navrženy tři strategie zálohování dat. Pomocí vícekritériální analýzy byl vybrán hardware dle požadovaných kritérií dané společnosti. Ze zpracovaných strategií byla vybrána ta nejvhodnější, a to použití síťové diskové stanice NAS v kombinaci s cloudovým úložištěm. Při výběru nejlepšího řešení se také dbalo na vhodný výběr disků. Výsledkem je jednoduchá a relativně finančně dostupná strategie zálohování. NAS stanice je k dispozici v celé firemní síti, což umožňuje zaměstnancům nenáročnou správu a manipulaci s daty. Díky vhodnému výběru je řešení dobře škálovatelné do budoucna a je možno jej rozšiřovat dle aktuálních potřeb firmy. Cloudové úložiště zajistí geografické oddělení dat a tím i vyšší bezpečnost a minimalizaci šance na ztrátu dat v případě jakéhokoliv typu havárie.

Porovnáním všech dostupných možností a kritérií bylo vybráno nejdostupnější řešení, co se týká ceny, výkonu a kvality, čímž bylo dosaženo cíle stanoveného v úvodu této bakalářské práce.

7 Seznam použitých zdrojů

Literární zdroje

GARFINKEL, S., SCHWARTZ, A., SPAFFORD, G. 2003. *Practical Unix & Internet Security*. 3. vydání. United States: O'Reilly & Associates, Inc. 984 s. ISBN 0-596-00323-4

NELSON, S. 2011. *Pro Data Backup and Recovery*. 1 vydání. USA: Apress. 290 s. ISBN 978-1-4302-2663-5.

PRESTON, W. 2006. *Backup and Recovery*. 1. vydání. USA: O'Reilly & Associates, Inc. 800 s. ISBN 9780596102463.

VŠEOBECNÁ ENCYKLOPEDIÉ DIDEROT. 8. svazků. 1999. 1. vydání. Praha: DIDEROT. ISBN 80-902555-2-3

Elektronické knihy

SCHACON, S., STRAUB, B. *Pro Git – Everything you need to know* [online]. 2. vydání. USA: Apress. 2022 [cit. 2022-12-20]. PDF format, 17.9 MB. Dostupné z: <https://git-scm.com/book/en/v2>.

Elektronické zdroje

ACRONIS.COM. *What is Cloud Backup? Cloud vs Local Backup Comparison* [online]. 2023-01-23. [cit. 2023-01-15]. Dostupné z: <https://www.acronis.com/en-us/blog/posts/cloud-vs-local-backup/>.

ACRONIS.COM. *Chraňte data před všemi hrozbami jedním uceleným řešením* [online]. 2023. [cit. 2023-02-04]. Dostupné z: <https://www.acronis.com/cs-cz/products/cyber-protect/>.

ACRONIS.COM. *Difference between Full, Incremental and Differential Backup* [online]. 2020-06-07. [cit. 2022-11-19]. Dostupné z: <https://kb.acronis.com/content/1536>.

AMANDA.ORG. *What is Amanda* [online]. 2017. [cit. 2023-02-04]. Dostupné z: <https://amanda.org/about.php>.

AVAST.COM. *Data a soubory nezálohuje 40 % Čechů* [online]. 2020-03-31. [cit. 2022-06-07]. Dostupné z: <https://blog.avast.com/cs/data-a-soubory-nezalohuje-40-procent-cechu>.

BARRACUDA.COM. *What is Data Compression?* [online]. 2022. [cit. 2022-11-19]. Dostupné z: <https://www.barracuda.com/support/glossary/data-compression>.

CLOUDFLARE.COM. *What is the cloud? Cloud definition* [online]. 2023. [cit. 2023-01-15]. Dostupné z: <https://www.cloudflare.com/learning/cloud/what-is-the-cloud/>.

COMPUTERWORLD.COM. *Backup vs archiving: It pays to know the difference* [online]. 2005-07-13. [cit. 2022-02-11]. Dostupné z: <https://www.computerworld.com/article/2556904/backup-vs--archiving--it-pays-to-know-the-difference.html>.

ESET.COM. *Jak najít správný software pro zálohování a obnovu?* [online]. 2022-08-08. [cit. 2023-02-04]. Dostupné z: <https://digitalsecurityguide.eset.com/cz/jak-najit-spravny-software-pro-zalohovani-a-obnovu>.

ESET.COM. *Malware* [online]. 2022. [cit. 2022-10-25]. Dostupné z: <https://www.eset.com/cz/malware/>.

ESET.COM. *Ransomware* [online]. 2022. [cit. 2022-10-25]. Dostupné z: <https://www.eset.com/cz/ransomware/>.

ESET.COM. *Phishing* [online]. 2022. [cit. 2022-10-25]. Dostupné z: <https://www.eset.com/cz/phishing/>.

FREECODECAMP.ORG. *Server Definition* [online]. 2021-03-18. [cit. 2022-05-20]. Dostupné z: <https://www.freecodecamp.org/news/server-definition/>.

IBM.COM. *What is cloud computing?* [online]. 2023. [cit. 2023-01-15]. Dostupné z: <https://www.ibm.com/topics/cloud-computing>.

IBM.COM. *IBM Spectrum Protect* [online]. 2021. [cit. 2023-02-04]. Dostupné z: <https://www.ibm.com/downloads/cas/L9MD4MEZ>.

IBM.COM. *When to back up and when to archive files* [online]. 2021-03-03. [cit. 2023-02-05]. Dostupné z: <https://www.ibm.com/docs/en/spectrum-protect/8.1.9?topic=data-when-back-up-when-archive-files>.

ITNETWORK.CZ. *Lekce 2 – Symetrická a asymetrická kryptografie* [online]. 2022. [cit. 2022-12-20]. Dostupné z: <https://www.itnetwork.cz/bezpecnost/symetricka-a-asymetricka-kryptografie>.

KASPERSKY.COM. *What is Data Encryption?* [online]. 2022. [cit. 2022-11-19]. Dostupné z: <https://www.kaspersky.com/resource-center/definitions/encryption>.

OPENAI.COM. *ChatGPT* [online]. 2023. Dostupné z: <https://openai.com/blog/chatgpt>.

OWASP.ORG. *SQL Injection* [online]. 2022. [cit. 2022-10-25]. Dostupné z: https://owasp.org/www-community/attacks/SQL_Injection.

REDDIT.COM. *Hardware Raid is Dead and is a Bad Idea in 2022* [online]. 2022. [cit. 2022-11-19]. Dostupné z: https://www.reddit.com/r/hardware/comments/tx4dnw/level1techs_hardware_raid_is_dead_and_is_a_bad/.

SEAGATE.COM. *RAID modes* [online]. 2023. [cit. 2022-11-20]. Dostupné z: <https://www.seagate.com/manuals/network-storage/business-storage-nas-os/raid-modes/>

SYNOLOGY.COM. *Co je to zálohování? Definice a podrobné vysvětlení* [online]. 2022-10-17. [cit. 2023-01-03]. Dostupné z: <https://blog.synology.com/cs-cz/co-je-to-zalohovani-definice-a-podrobne-vysvetleni/>.

WESTERNDIGITAL.COM. *What is RAID Storage?* [online]. 2023. [cit. 2022-11-16]. Dostupné z: <https://www.westerndigital.com/solutions/raid>.

WIKIPEDIA.COM. *Backup and Restore* [online]. 2023. [cit. 2023-02-04]. Dostupné z: https://en.wikipedia.org/wiki/Backup_and_Restore.

WIKIPEDIA.COM. *Features new to Windows 8* [online]. 2023. [cit. 2023-02-04]. Dostupné z: https://en.wikipedia.org/wiki/Features_new_to_Windows_8#File_History.

WINDOWSCENTRAL.COM. *What RAID means and why you might want one* [online]. 2019-05-19. [cit. 2022-11-20]. Dostupné z: <https://www.windowcentral.com/what-does-raid-mean>.

Obrázky

SYNOLOGY.COM. *DiskStation® DS420+* [online]. 2023. Dostupné z: <https://www.synology.com/img/products/detail/DS420plus/heading.png>.

LENOVO.COM. *Lenovo D1212 Direct Attached Storage* [online]. 2023. Dostupné z: <https://p1-ofp.static.pub/medias/bWFzdGVyfHJvb3R8MTQwNTQ3fGltYWdlL3BuZ3xoOTcvaDjLzExMDAyNzg1MzAwNTEwLnBuZ3xlOGRjY2Q1MDlmMTk4ZmRlMjcyMDRkOTk3OGZhN2VjMGNhMzNiMzYzMzRjOGMwNjliMzNkYWYzMTIyZmEwYmE2/lenovo-storage-d1212-hero.png>.

PETASTOR.COM. *IP SAN Storage* [online]. 2023. Dostupné z: <https://petastor.com/wp-content/uploads/PR424i.png>.

CISCO.COM. *What Is Encryption?* [online]. 2023. Dostupné z: <https://www.cisco.com/c/en/us/products/security/encryption-explained.html>

SYNOLOGY.COM. *DiskStation® DS1522+* [online]. 2023. Dostupné z: <https://www.synology.com/cs-cz/products/DS1522+>.

8 Seznam obrázků, tabulek, grafů a zkratk

8.1 Seznam obrázků

| | | |
|------------|--|----|
| Obrázek 1 | Síťové úložiště NAS | 17 |
| Obrázek 2 | Síťové úložiště SAN | 17 |
| Obrázek 3 | Síťové úložiště DAS | 18 |
| Obrázek 4 | Schéma cloudového úložiště..... | 18 |
| Obrázek 5 | Schéma RAID 1 | 20 |
| Obrázek 6 | Schéma RAID 10 | 21 |
| Obrázek 7 | Schéma RAID 5 | 22 |
| Obrázek 8 | Schéma RAID 6 | 23 |
| Obrázek 9 | Průběh symetrického šifrování | 29 |
| Obrázek 10 | Průběh asymetrického šifrování | 30 |
| Obrázek 11 | Sídlo společnosti VIKLAN SOFTWARE s.r.o..... | 32 |
| Obrázek 12 | Přibližné schéma síťové organizace společnosti..... | 34 |
| Obrázek 13 | Klady, zápory, nové možnosti a hrozby návrhu zálohování č. 1 | 43 |
| Obrázek 14 | Klady, zápory, nové možnosti a hrozby návrhu zálohování č. 1 | 44 |

8.2 Seznam tabulek

| | | |
|-----------|---|----|
| Tabulka 1 | Rozdíly v požadavcích na zálohování a archivaci dat | 27 |
| Tabulka 2 | Matice variant pro síťová úložiště | 37 |
| Tabulka 3 | Kriteriální matice pro síťová úložiště | 37 |
| Tabulka 4 | Matice varianta pro HDD | 39 |
| Tabulka 5 | Kriteriální matice pro HDD | 39 |
| Tabulka 6 | Matice variant pro SSD..... | 39 |
| Tabulka 7 | Kriteriální matice pro SSD | 40 |
| Tabulka 8 | Porovnání cloudových služeb | 42 |