Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií

Diplomová práce

Zálohování a archivace dat ve zvolené firmě

Bc. Pavel Beran

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Provozně ekonomická fakulta

ZADÁNÍ DIPLOMOVÉ PRÁCE

Pavel Beran

Systémové inženýrství a informatika
Informatika

Název práce

Zálohování a archivace dat ve zvolené firmě

Název anglicky

Data backup in a medium-sized company

Cíle práce

Hlavním cílem práce je návrh komplexní zálohovací strategie pro zvolenou firmu a její následná realizace.

Dílčími cíli jsou:
· rešerše problematiky zálohování.
· analýza dostupných řešení pro zálohování, s ohledem na potřeby středně velké firmy.
· identifikace zdrojů, které je potřeba zálohovat – servery, úložiště, koncové stanice, síťové prvky, stroje.
· otestování vybraných řešení z hlediska rychlosti obnovy dat ze zálohy.
· formulace závěrů práce.

Metodika

Metodika řešené problematiky diplomové práce je založena na studiu a analýze odborných informačních zdrojů. V první části práce je uveden vhled do problematiky zálohování a její současný stav.

Praktická část je zaměřena na návrh komplexní strategie zálohování v konkrétní firmě, realizaci tohoto návrhu a otestování funkčnosti. Na základě syntézy teoretických poznatků a výsledků praktické části budou formulovány závěry práce.

Doporučený rozsah práce

60-80s.

Klíčová slova

záloha, data, obnova, archivace, médium, server, RAID, firewall, cloud, úložiště

Doporučené zdroje informací

Acronis True Image 2021: USER GUIDE. Acronis [online]. Dostupné z: https://dl.acronis.com/u/pdf/ATI2021_userguide_en-US.pdf

CROCETTI, Paul. Create your data backup strategy: A comprehensive guide [online]. 15 Jul 2020. Dostupné z: https://searchdatabackup.techtarget.com/Create-your-data-backup-strategy-A-comprehensive-guide

How do I back up my data to a remote Synology NAS or file server using Hyper Backup? Synology: Knowledge Center [online]. Dostupné z: https://kb.synology.com/en-global/DSM/tutorial/How_to_back_up_your_data_to_a_remote_Synology_NAS_or_file_server_with_Hyper_

Veeam Help Center Technical Documentation: Veeam Backup & Replication. Veeam [online]. Dostupné z: https://www.veeam.com/documentation-guides-datasheets.html

---

Předběžný termín obhajoby

2021/22 LS – PEF

Vedoucí práce

Ing. Martin Havránek, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 9. 8. 2021

doc. Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 19. 10. 2021

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 02. 02. 2022

Čestné prohlášení

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

V Praze dne 30.03.2022 _____

Poděkování

Rád bych touto cestou poděkoval vedoucímu práce Ing. Martinu Havránkovi, Ph. D., za odborné vedení a cenné rady při tvorbě práce. Také bych rád poděkoval rodině za podporu během studia.

Zálohování a archivace dat ve zvolené firmě

Abstrakt

Diplomová práce se zabývá problematikou zálohování ve středně velké firmě. Teoretická část práce se nejprve zaměřuje na analýzu současných trendů a používaných technologií v oblasti zálohování. Následuje představení vybraných řešení a identifikace zdrojů, které je potřeba zálohovat.

V praktické části je navržena a následně implementována komplexní zálohovací strategie pro zvolenou firmu. Po realizaci jsou implementovaná řešení otestována jak z hlediska rychlosti zálohy, tak rychlosti obnovy dat.

Klíčová slova: záloha, data, obnova, archivace, médium, server, RAID, firewall, cloud, úložiště

Data backup in medium-sized company

Abstract

The thesis deals with the problematic of backup in a medium-sized company. The theoretical part of the thesis first focuses on the analysis of current trends and technologies used in the field of backup. This is followed by an introduction of the selected solutions and identification of the resources that need to be backed up.

In the practical part, a comprehensive backup strategy for the selected company is designed and then implemented. After implementation, the solutions are tested both in terms of backup speed and data recovery speed.

Keywords: backup, data, data recovery, archiving, medium, server, RAID, firewall, cloud, data storage

Obsah

Seznam obrázků

Seznam tabulek

1 Úvod

Diplomová práce se zabývá problematikou zálohování a archivace dat ve firemním prostředí. Proces zálohování je extrémně důležitá činnost, což ve firemním prostředí platí dvojnásob. Zálohováním se firma, ale i soukromá osoba chrání proti ztrátě dat. Nejčastějším problémem je chyba uživatele, kdy něco omylem smaže. Méně častým, ale o to závažnějším problémem je selhání softwaru, nebo hardwaru. Třetím obvyklým případem je napadení systému útočníkem, respektive počítačovým virem. Jak se říká, existují dva druhy lidí. Ti co zálohují a ti co o svá data ještě nepřišli. A přijít o data ve firmě téměř vždy znamená ztrátu času a tím pádem i finanční ztrátu.

Hlavní motivací pro zpracování této práce je tvorba komplexní zálohovací strategie pro firmu, ve které je autor zaměstnaný. Jedná se o středně velký podnik s relativně širokým spektrem zdrojů, které je potřeba zálohovat. Mezi hlavní zdroje patří servery, datová úložiště, koncové stanice, síťové prvky, ale i průmyslové automatizované stroje ve výrobních halách. Popud pro tvorbu této práce vzešel ze stavu zálohování ve firmě, který je již nedostatečný, protože firma prošla výraznou modernizací a rozšířením infrastruktury a je tak potřeba na tuto skutečnost reagovat i z pohledu zálohování.

Během diplomové práce dojde k analýze zdrojů, kde bude definováno, co je potřeba zálohovat a co má jakou prioritu. Na základě této analýzy budou vybrána vhodná hardwarová a softwarová řešení. Následně bude navržena zálohovací strategie. Tato strategie bude mimo jiné obsahovat co, kam a kdy se bude zálohovat. V praktické části diplomové práce dojde na samotnou realizaci strategie. To znamená instalaci a konfiguraci vybraných řešení. Následně bude provedeno testování rychlosti zálohování a především rychlosti obnovy dat, která je při potencionálních problémech v budoucnu klíčová. Praktická část práce bude zároveň představovat dokumentaci, která jasně popíše celou implementovanou strategii, tak aby například nově příchozí zaměstnanci měli usnadněný vhled do problematiky v dané firmě.

2 Cíl práce a metodika

2.1 Cíl práce

Diplomová práce je zaměřena na problematiku zálohování ve středně velké firmě. Hlavním cílem je navrhnout a následně realizovat komplexní zálohovací strategii ve vybrané firmě.

2.1.1 Dílčí cíle

- Rešerše problematiky zálohování.
- Analýza dostupných řešení pro zálohování s ohledem na potřeby středně velké firmy.
- Identifikace zdrojů, které je potřeba zálohovat – servery, datová úložiště, koncové stanice, sítové prvky, stroje.
- Otestování vybraných řešení z hlediska rychlosti zálohování a obnovy dat.

2.2 Metodika

Metodika řešené problematiky diplomové práce je založena na studiu a analýze odborných informačních zdrojů. V první části práce je uveden vhled do problematiky zálohování a její současný stav.

Praktická část je zaměřena na návrh komplexní strategie zálohování v konkrétní firmě, realizaci tohoto návrhu a otestování funkčnosti. Na základě syntézy teoretických poznatků a výsledků praktické části budou formulovány závěry práce.

3 Teoretická východiska

V této kapitole bude čtenář seznámen s teoretickým podkladem k samotné praktické části této práce. Postupně bude detailněji vyložena problematika zálohování od příčin ztráty dat přes základní strategie zálohování až po časté chyby při tomto procesu. Dále bude probráno téma úložišť z různých úhlů pohledu. Následně budou uvedeny metody zálohování a zálohovací software.

3.1 Problematika zálohování

Proč zálohovat

Proces zálohování chrání uživatele proti ztrátě či poškození dat a především v případě firemní sféry s tím související finanční újmě. Ke ztrátě dat dochází neustále z mnoha důvodů, které jsou identifikovány dále a je nemožné tomuto jevu stoprocentně zabránit. Nicméně díky zálohování lze toto riziko minimalizovat, respektive lze minimalizovat případné následky.

3.1.1 Příčiny ztráty či poškození dat

Ke ztrátě či poškození dat dochází z mnoha důvodů. Zde jsou vypsány nejčastější problémy.

Chyba uživatele

Jedná se o suverénně nejčastější důvod ztráty dat a následné obnovy ze zálohy. Situace, kdy uživatel omylem něco smaže je na denním pořádku. Lidská chyba stojí v čele statistik s 50 % [1].

Porucha hardwaru

Dalším relativně běžným problém je selhání hardwaru. Typicky jde o pevný disk v počítači. S nástupem SSD disků se tato hrozba sice snížila, ale stále existuje. Podle statistik společnosti Backblaze selhalo v roce 2020 0,93 % pevných disků. To se může zdát jako relativně malé procento, ale vzhledem k obrovskému množství disků není konečné číslo v žádném případě zanedbatelné [2].

Ztráta/krádež zařízení

Při ztrátě či krádeži zařízení, jsou uživatelé bez záloh kompletně bez svých dat, která byla v daných zařízeních.

Přírodní katastrofa

Vytopená kancelář nebo serverovna, požár a podobně může lehce zničit veškerá elektronická zařízení. Proto je potřeba vytvářet zálohy a pokud možno na více fyzicky oddělených míst.

Aktualizace

Velice častý problém. Aktualizace operačních systémů, specializovaných firemních aplikací či firmwaru může v případě chyby v aktualizaci paralyzovat práceschopnost jednotlivých uživatelů nebo dokonce celé firmy. Proto je nutné mít dostupné zálohy systémů pro co nejrychlejší obnovu do stavu před aktualizací.

Virus a další útoky

Velký a složitě řešitelný problém. Virus může poškodit data. Dnes populární ransomware zašifruje veškerá data, ke kterým se dostane. Typicky vše v dané firemní síti. Proto je vhodné mít zálohy mimo hlavní síť, nejlépe i na mediích, která nejsou neustále připojena k síti.

Selhání softwaru

Operační systémy, počítačové programy a další aplikace nejsou bezchybné. To znamená, že při selhání daného softwaru, může dojít k poškození či ztrátě dat.

Nárůst dat

Objem dat ve světě narůstá obrovským tempem. S větším množstvím informací se také zvyšuje pravděpodobnost možnosti ztráty či poškození dat. Dle statistik IDC byl objem dat na světe v roce 2018 40 zettabajtů a do roku 2025 se zvětší na 175 zettabajtů. Pro představu 1 zettabajt je 1000 exabajtů, 1 exabajt je 1000 petabajtů a 1 petabajt je 1000 terabajtů [3].

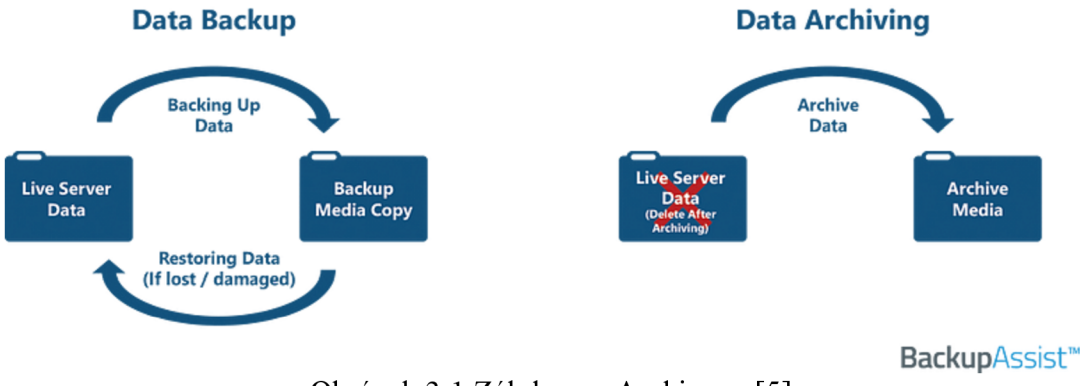
3.1.2 Základní pojmy zálohování

Záloha versus archivace

Pojmy záloha a archivace se relativně často zaměňují, či se považují za synonyma. To je však špatně. Obě činnosti jsou využívány v odlišných případech. V tomto ohledu mohou být vyzdvihnuty dva hlavní rozdíly mezi zálohou a archivací.

Zálohovány jsou aktivní data, která jsou každý nebo téměř každý den využívána. Jedná se například o výrobní data, aktuální objednávky apod. Po provedení zálohy zůstávají původní data na svém místě v produkčním systému. Druhým rozdílem je přístup k datům. Ten je v případě zálohování optimalizován na přenos velkých objemů dat tak, aby při problémech s produkčním zdrojem proběhla obnova co nejrychleji. Cílem zálohování je tedy mít k dispozici kopii dat a v co nejmenším možném časovém intervalu je použít v případě potíží s primárním zdrojem.

Oproti tomu archivace řeší uchování starších či neaktuálních dat, která nejsou využívána každý den. Velice často je uchování těchto dat vyžadováno zákonem. Typicky se archivují například staré emailové zprávy, smlouvy nebo dokumenty z účetnictví. V osobním prostředí lze za data vhodná k archivaci označit například staré rodinné fotografie. Oproti zálohování, které se provádí opakovaně v určitém časovém intervalu, se archivace provede jednou a následně jsou data uložena tak dlouho jak je potřeba. Ideálním scénářem je tak vhodně využívat obě techniky [4].



Obrázek 3.1 Záloha vs. Archivace [5]

Synchronizace

Se zálohováním se také často chybně zaměňuje technika jménem synchronizace. Důvod je jednoduchý. Pokud bude složka v počítači nebo třeba v mobilním telefonu, která se bude synchronizovat například do cloudu, nebo do NASu, tak se jakákoliv změna v jednom zařízení synchronizuje i do zařízení druhého. Pokud se nahrají chybná data, budou i na druhém zařízení. Pokud někdo omylem smaže soubor v počítači, soubor se smaže i na dalších zařízeních. Proto nelze proces synchronizace označit jako zálohování.

Replikace

Jedná se proces, při kterém se kopírují data z jednoho místa na druhé. Typickým příkladem je replikování virtuálních serverů z produkčního hosta na hosta jiného. Pokud nastane problém s primárním zdrojem, lze aktivovat spící repliku a zajistit tak chod systému.

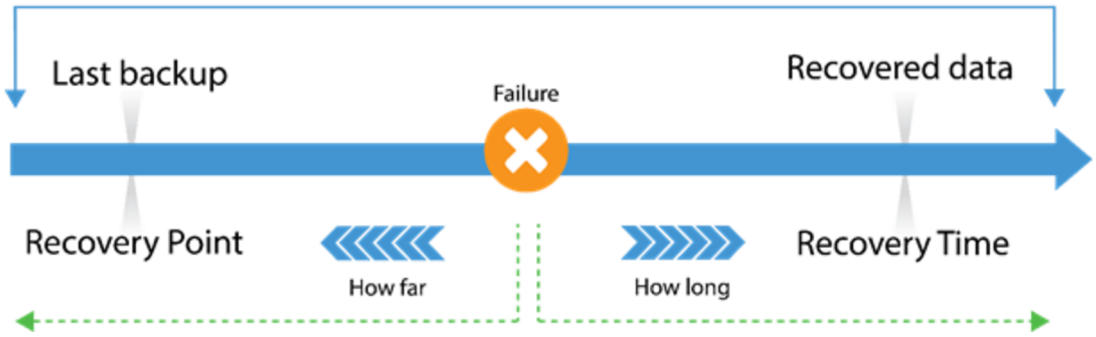
Obnova

Posledním bodem celého procesu zálohování je obnova dat v případě nestandardní situace. Stejně jako je potřeba vytvořit strategii vytváření záloh, tak je potřeba vymyslet plán pro obnovu. Samozřejmě velmi záleží na velikosti firmy a rozsáhlosti IT infrastruktury. Ale v každém případě pomůže vytvoření takového plánu minimálně k ujasnění priorit. Tento plán je označován jako Disaster Recovery Plan (DRP). Plán obsahuje postupy pro efektivní obnovu dat v případě potřeby.

V prvním kroku se provede tzv. business impact analysis (BIA), která určí závislosti firmy na IT infrastruktuře a dopady při jejím selhání. Jednotlivé služby, které poskytuje infrastruktura se rozdělí na kritické a nekritické a následně se pro každou službu vyhodnocují RPO a RTO.

RPO neboli Recovery point objective je vyjádření času do kterého je potřeba obnovit data a definuje maximální přípustné množství ztracených dat naměřených v čase od výskytu selhání do poslední platné zálohy.

RTO neboli Recovery time objective představuje, jak dlouho trvá obnovení z incidentu, dokud nebudou uživatelům dostupné běžné funkce systémů. Tento čas nelze dopředu úplně přesně určit, ale nastavením správných postupů lze tento čas zkrátit [6].



Obrázek 3.2 RPO a RTO [7]

Ve větších podnicích je také dopředu vhodné určit jednotlivé úlohy pro každého IT technika zabývajícího se případnou obnovou. Tento krok může výrazně urychlit vykonání daných plánů.

3.1.3 Druhy zálohovaných dat

Data, která je potřeba zálohovat lze rozdělit do dvou základních kategorií. Uživatelská data a systémová data. Uživatelská data jsou data, která byla přímo vytvořena uživatelem. Jedná se o fotografie, textové dokumenty, prezentace a další soubory. Naopak systémová data nebyla vytvořena přímo uživatelem jako takovým. Jedná se o operační systém, nainstalované programy a jejich data, konfigurační soubory, databáze a podobně.

3.1.4 Základní pravidla zálohování

Aby bylo zálohování smysluplné a účinné, je potřeba dodržovat základní pravidla, která pomohou zmenšit riziko ztráty dat z jakéhokoli důvodu.

Frekvence zálohování

Záloha dat by se měla provádět pravidelně. Četnost vždy záleží na konkrétních datech, jejich významu, nebo na tom, jak často se data mění. Někdy stačí provést zálohu jednou za měsíc, někdy jednou za týden. Naopak v určitých případech je potřeba provádět zálohu každý den, každou hodinu, nebo dokonce po každé změně dat, která se zálohují. Rozdíl je, pokud bude jednotlivec zálohovat fotografie z telefonu, kde bude stačit udělat zálohu například na cloud jednou týdně, nebo měsíčně (Podle toho, jak moc dotyčný fotí.), či pokud se bude zálohovat ve firmě účetní systém, kde je každý den vytvořeno x důležitých dokumentů.

Různé typy úložišť

Ukládat zálohy na stejné úložiště, kde jsou původní data, nedává smysl téměř v žádném ohledu. Toto řešení nepřináší v podstatě žádné výhody, pouze rizika. Pokud bude záloha dat uložena na stejném pevném disku jako původní data, tak se v podstatě provede pouze duplikace dat a pokud disk selže, dojde ke ztrátě originálních dat i záloh. Proto je vždy nutné zálohovat na externí média, která jsou oddělena od původních dat.

Dále je vhodné se nespoléhat pouze na jeden typ úložiště, ale rozvrstvit zálohy mezi více různých médií. Externí diskové pole, cloud, NAS, páska, nebo třeba optický disk. To vše výrazně sníží pravděpodobnost ztráty dat. Například kybernetický útok na jeden typ úložiště pravděpodobně nebude fungovat na všechna zařízení.

Také je velice vhodné mít zálohy na více geograficky oddělených místech. Pokud bude mít firma ve své vlastní serverovně veškeré produkční systémy a zároveň všechny technologie vypsané výše, kam bude ukládat zálohy, určitě to zvýší pravděpodobnost na ochranu dat. Nicméně pokud serverovna například vyhoří a nezvládne se nic zachránit, tak dotyčná firma opět přijde o všechny data. Proto je potřeba ukládat zálohy dat i na naprosto oddělené místo od originálu či dalších záloh [8].

Různá zálohovací prostředí

Nespoléhat se pouze na jeden software pro zálohování také zvyšuje šance vyhnout se problémům se ztrátou dat. Nicméně toto řešení je samozřejmě náročnější na implementaci i z pohledu financí. Z těchto důvodů není tento způsob příliš využíván.

Jiný operační systém

Je vhodné využít server, či úložiště, které nefunguje na operačním systému Windows. Většina útoků je vedena právě na prostředí Windows.

Jiné účty pro zálohovací systémy

Základní pravidlo. Zálohovací software potřebuje pro přístup k datům, která je potřeba zálohovat, oprávnění. Například v podobě uživatele. V žádném případě se nedoporučuje používat pro tyto účely administrátorský účet, který běžně používá správce firemní sítě. Pro tyto účely je tak vhodné vytvořit speciálního uživatele, který nebude používán na nic jiného. Samozřejmostí je velice silné heslo.

Neměnná úložiště

Takzvaná neměnná úložiště představují ochranu proti zašifrování ransomwarem. V tomto úložišti nemohou být vytvořené zálohy změněny nebo odstraněny po předem určený čas. Dokonce ani administrátor nemůže zálohy smazat či změnit.

Oddělené sítě

Zálohovací systém se doporučuje umístit tak, aby nebyl přímo dosažitelný přes hlavní produkční firemní síť. Zabrání se tím přímému útoku na zálohovací servery.

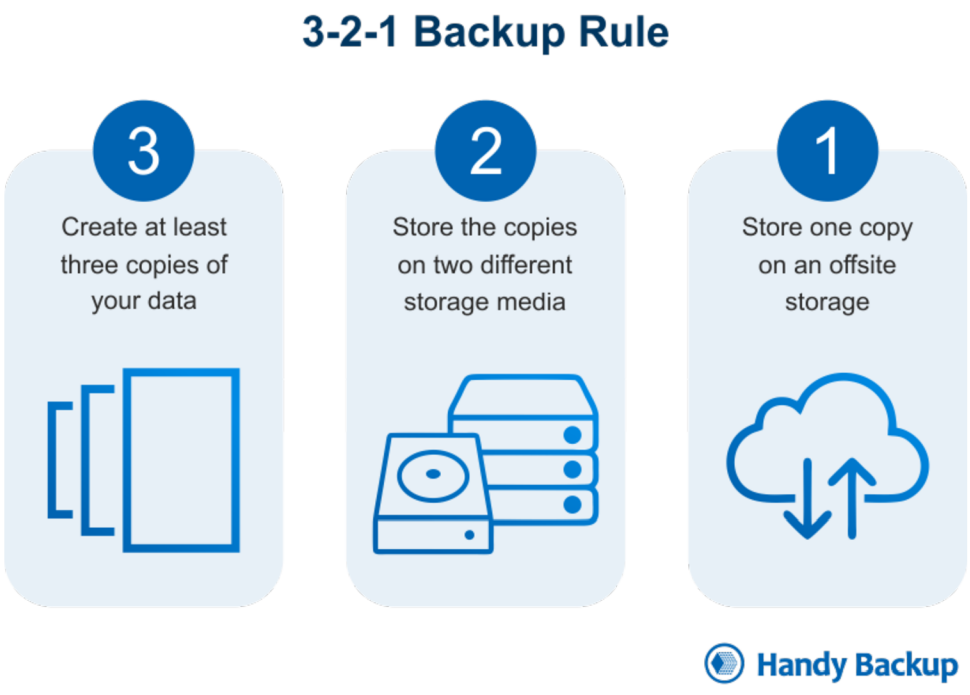
Čím důležitější data, tím více záloh

V případě extrémně důležitých a kritických dat je vhodné se držet jednoduchého pravidla, čím více tím lépe.

3.1.5 Zálohovací strategie

Pravidlo 3-2-1

Léty prověřený základní koncept pro zálohování. Jeho název vychází z jeho principu. Tedy tři kopie produkčních dat jsou uloženy na dvou fyzicky nezávislých úložištích, která jsou doplněna o jednu off-site zálohu. Jedná se o stále používaný přístup k zálohovaní i přesto, že obsahuje určitě nedostatky [9].



Obrázek 3.3 Pravidlo 3-2-1 [9]

Pravidlo 3-2-1-1 (3-2-2)

Vylepšení předchozího pravidla, které se snaží řešit problém s off-site zálohou. Tento problém spočívá v tom, že je často tato záloha použita v online podobě, například pomocí cloudového úložiště. Existuje tedy možnost, že se útočník k datům dostane. V případě tohoto pravidla je tak přidána ještě jedna záloha. Ta může být uložena například na pásce, nebo extérním disku kompletně mimo síť. Ovšem ani tato metoda není bezproblémová [10].

Pravidlo 3-2-1-1-0

Může být vytvořeno nepřeberné množství záloh, ale bez úspěšné obnovy jsou k ničemu. Toto pravidlo přidává kontrolu obnovy zálohy, tedy nula chyb. Nejen k tomu slouží proces DRP (Disaster Recovery Plan). Pomocí tohoto procesu lze prověřit všechny aspekty obnovy a objevit tím případné problémy, které lze poté řešit, či akceptovat riziko [11].

Obrázek 3.4 Pravidlo 3-2-1-1-0 [11]

Například společnost Veeam nabízí u svého řešení tzv. SureBackup. Jedná se o automatickou simulaci obnovy zálohy, kdy se zkontroluje, zda je daná záloha po obnově funkční.

Air Gap

Za Air Gap lze označit počítač, který není připojen k internetu ani lokální firemní síti a neobsahuje bezdrátové komunikační technologie. Takový stroj je vhodný k uložení citlivých dat, či ovládání kritických systémů. Doslova vzduchová mezera také komplikuje útoky hackerům. Samozřejmě možnosti, jak se dostat k takovému zařízení i tak existují, ale výrazně se tím snižuje riziko. Zde už musí útočník využít sociální inženýrství, případně podniknout přímý fyzický útok na daný stroj.

3.1.6 Aspekty pro výběr metody zálohování

Při volbě způsobu zálohování se musí zohlednit několik důležitých faktorů na základě kterých je poté zvolena vhodná strategie zálohování, vhodná média, iterace a podobně. Zde jsou vypsány ty hlavní.

Typ zálohovaných dat

Data se dají rozdělit na dva typy. Uživatelská data a systémová data. Tyto pojmy již byly vysvětleny v kapitole 3.1.3. Ke každé skupině dat se musí přistupovat trochu jiným způsobem a pro zálohování využít i odlišené nástroje.

Objem zálohovaných dat

Jedna z klíčových vlastností je velikost dat, která se plánují zálohovat. Na základě této informace je potřeba zvolit vhodná média, strukturu sítě a nástroje pro ukládání záloh. V případě běžných uživatelských dat na koncových stanicích ve firemním prostředí se většinou nebude jednat o velké objemy dat v řádech stovek gigabajtů až terabajtů, protože obvykle pracují s daty na síťových discích, vzdálených plochách atd. V osobním (domácím) počítači jsou naopak tyto větší objemy relativně normální.

Velké objemy dat pak budou ve firemním prostředí obsahovat systémová data. Samozřejmě záleží na dané firmě, její velikosti a na tom, jak rozvinutou má vlastní infrastrukturu. Příkladem mohou být virtuální servery, mailserver, datová úložiště, NAS zařízení a podobně.

Časová náročnost

Důležitý aspekt pro správné nastavení zálohování je trvání samotné zálohovací úlohy. V případě malých objemů dat, je možné provádět zálohy v podstatě kdykoliv. V opačném případě je potřeba pečlivě zvolit vhodnou dobu, kdy se má záloha provádět. Typicky se náročné zálohy provádějí večer a v noci.

Mimo jiné praktické důvody se tím lze vyhnout problémům s přetížením sítě či daných strojů v pracovní době. Samozřejmě ne vždy je toto možné. Pokud situace nedovoluje mít celý pracovní den bez zálohy, lze nastavit zálohu například na čas oběda, kdy se firemní prostory vyprázdní, a tak nevadí zatížení sítě, případně strojů. Vždy však záleží na konkrétním případu dané firmy.

Výkonová náročnost

Při volbě metod zálohování se musí brát zřetel i na výkonovou náročnost výsledného řešení. Obvyklý problém je špatně navržená firemní síť, která se může při zálohování stát kompletně nedostupná. Problémem může být i nedostatečně výkonný server, na kterém je provozován zálohovací software.

Výběr médií pro zálohu

Volba médií, na které se budou zálohy ukládat je velice důležitá. Médium je potřeba vybrat s ohledem na objem zamýšlených dat, která bude potřeba zálohovat, rychlost zálohy a následné obnovy dat, nebo spolehlivost média. Odlišná média se také zvolí pro archivaci nebo pro zálohu.

Výběr speciálního softwaru pro zálohování

V případě jednotlivce může stačit jednoduše zkopírovat důležitá data na záložní médium. Toto lze učinit v případě, kdy je potřeba zálohovat relativně malé množství dat, které se příliš často nemění. To je ve firemním prostředí nemyslitelné. Na základě potřeb dané firmy proto musí být vybrán vhodný software.

Šifrování citlivých dat

Citlivá osobní data nebo další kritická firemní data je potřeba šifrovat, nebo ukládat na šifrovaná úložiště.

Jak často bude záloha probíhat

Automatizované zálohy lze nastavit tak, aby se periodicky opakovaly. Jednou denně, jednou týdně, jednou měsíčně, jednou za hodinu atd. Vždy záleží na konkrétní situaci a důležitosti dat, podle které se provede nastavení.

Obnova dat ze zálohy

Kritická část zálohování. Přímo souvisí s výběrem médií a softwaru pro zálohování. Rychlost obnovy dat ze zálohy je naprosto zásadní. Čím déle obnova trvá, tím horší jsou ekonomické následky. Ovšem ani v případě rychlé obnovy nemusí být výsledek použitelný. Aby byla obnova úspěšná, musí záloha opravdu obsahovat funkční data. Může se stát, že záloha proběhne úspěšně a až po obnově je zjištěno, že záloha nic neobsahuje, nebo je

poškozena. Na ověření funkčnosti záloh existují automatizované nástroje. V potaz se také musí vzít zdroje nutné pro obnovu dat. Ať už hardwarové nebo lidské.

Ekonomická zátěž

Kompletní zálohovací řešení je nutné vybírat i s ohledem na finanční možnosti jednotlivce, respektive dané organizace. Toto se musí zohlednit především ve firemní sféře. V případě jednotlivce totiž velice často a velice dobře poslouží zdarma dostupná řešení, kdy jedinou placenou položkou bude médium pro ukládání záloh. V případě firem se mohou náklady za licence a hardware vyšplhat velmi vysoko. Ovšem ve výsledku bude cena za zálohování téměř vždy podstatně menší než finanční ztráta při případné velké ztrátě dat a nemožnosti data obnovit.

3.1.7 Časté chyby při zálohování

Špatně zvolený čas zálohy

Pokud proběhne pouze jedna záloha během pracovní doby, je více než pravděpodobné, že po provedení zálohy budou ukládána další důležitá data, která nebude daná záloha obsahovat. Pokud poté nastane problém, druhý den tato data nikdo neobnoví a budou ztracena.

Spuštěné aplikace během zálohy

S předchozím bodem částečně souvisí i tento bod. Pokud je prováděna záloha, nemělo by být spuštěno nic jiného. Většina zálohovacích softwarů přeskočí všechny spuštěné programy a tím se opět ztratí důležitá data.

Zálohy uloženy na stejném zařízení jako originální data

Tento problém se bude vyskytovat spíše u osobních domácích počítačů, kdy si uživatel s dobrým úmyslem udělá zálohu, ovšem zálohu uloží v daném počítači na stejném disku, kde je nainstalován systém a originální data. V tomto případě je záloha téměř bezcenná. Jediná situace, kdy lze takovou zálohu využít je, pokud nějaký program skončí chybou a poškodí data se kterými pracoval. Poté se může využít vytvořená záloha.

Neotestovaná záloha

Velice častý případ. Firma zálohuje a vše vypadá v pořádku, až do doby, kdy je nutné obnovit data ze zálohy. Poté zjistí, že systémy po obnově nefungují, či že je záloha dokonce úplně prázdná. Proto je nutné vytvořené zálohy testovat.

Záloha se neprovádí pravidelně

Dnes jsou k dispozici nástroje, které pravidelně provádějí zálohy, dle nastavení. Díky tomu se lze vyhnout řadě problémů. Není potřeba vyčlenit osobu na ruční zálohování, která lehce zapomene na nějaká data, nebo provede zálohu chybně. Při správném a ověřeném funkčním nastavení odpadá spousta zbytečné režie. Zálohovací periodu lze nastavit dle potřeb dané společnosti a daných systémů a je zde jistota, že se záloha vždy v daný čas provede. Pokud ne, nebo se provede s chybou, systém na to sám upozorní.

Špatně pojmenované zálohy

Jasně, věcně a logicky pojmenovat zálohy. Při krizové situaci může tato na první pohled drobnost, urychlit nebo naopak výrazně zkomplikovat celý proces obnovy. Při nutnosti obnovovat data opravdu není vhodné hledat, která záloha k čemu patří a co je potřeba obnovit.

Nekontrolování automatických procesů zálohy

Nastavením automatických záloh se ušetří spousta práce, nicméně se na toto řešení nelze stoprocentně spoléhat. Vždy je nutné pečlivě sledovat výsledky těchto procesů. Může se stát, že záloha na první pohled proběhne úspěšně, ale při bližším zkoumání se zjistí, že velikost zálohy je 0 bajtů.

Nenastavení posílání upozornění

S předchozím bodem souvisí i tento problém. Programy pro zálohování nabízejí možnost zasílat upozornění a hlášení o jednotlivých zálohovacích úlohách. Tato funkce je velice nápomocná při sledování a kontrole zálohovacích procesů. Pokud se záloha z nějakého důvodu neprovede, správce na to bude upozorněn. Ze samotného hlášení je také možné rovnou zjistit důvod, proč záloha neproběhla v pořádku. Výhodou je také fakt, že je toto emailové upozornění dostupné kdykoliv a kdekoliv.

Nedostatek místa v záložním úložišti

Jednoduchý princip. Pokud v cílovém úložišti není místo, není kam uložit zálohy. Některé zálohovací programy nabízejí možnosti, jak tento problém částečně řešit. Lze nastavit kolik záloh se má ukládat. Například deset verzí s tím, že při ukládání nové zálohy se nejstarší záloha z původních deseti smaže [12]. Více se tato problematika rozebírá v kapitole 3.3.5 Retenční politika.

3.2 Úložiště

3.2.1 Úložiště dle doby pro přístup

Úložiště lze rozdělit podle různých kritérií. Jedním z nejdůležitějších je doba pro přístup. V následující kapitole jsou použité výrazy online a offline. V tomto smyslu se nejedná o použití internetu, ale o to, jak se používá úložné médium. Online je přímo připojené k cílovému stroji, zatímco offline úložiště vyžaduje určitý fyzický zásah a montáž.

Online

Jedná se o úložiště, které je neustále připojené. Jeho výhoda je schopnost poskytnou požadovaná data v řádech milisekund. Na druhou stranu je také nejvíce náchylné na ztrátu či poškození dat. Důvodem může být napadení virem, smazání dat, nebo třeba špatná manipulace. Příkladem takového úložiště může být přímo připojené diskové pole.

Nearline

Úložiště, které se nachází mezi online a offline, ale principiálně stojí blíže k online úložišti. Nearline úložiště již vyžaduje určitý lidský zásah, ale většinou pouze pomocí počítače. Toto úložiště není tak rychlé jako online úložiště. Příkladem mohou být optická média nebo magnetická páska v automatické knihovně. Zde už je reálné zpoždění v řádech sekund, až minut [13].

Offline

Offline úložiště již není přímo dostupné, není přímo připojené a vyžaduje fyzický zásah pro jeho připojení. Většinou se jedná o levnější řešení. Příkladem může být zálohování na pásku, kde ale není pásková mechanika s vlastní knihovnou a je nutné pásky vlastnoručně vkládat do čtecí mechaniky. Zde se přístupová doba výrazně zvyšuje [14].

Externí úložiště

Ideální ochrana proti lokálním hrozbám jako je požár nebo krádež. Zde se ovšem přístupová doba dostává na úroveň hodin, někdy i dní. Taková úložiště představují například externí disky, nebo pásky uložené v bankovním trezoru. Podobné úložiště je vhodné pro ukládání dalších kopií záloh či archivaci, nikoliv primárních záloh, které budou v případě problému potřeba k okamžité obnově.

Cloud

V posledních letech zažil cloud extrémní rozmach. Jedná se o úložiště, které nabízí podstatné výhody, ale samozřejmě i některé nevýhody. Mezi hlavní přednosti tohoto řešení patří dostupnost dat v podstatě z celého světa. Dále nabízí teoreticky neomezenou kapacitu, podporu velkého množství operačních systémů, uložení dat mimo polohu původních dat nebo následné zálohování o které se stará poskytovatel služby. Nevýhodou je poté rychlost obnovy dat, která je závislá na rychlosti internetového připojení. Určitou nevýhodou pro někoho také může být svěřování dat do rukou třetí strany [15].

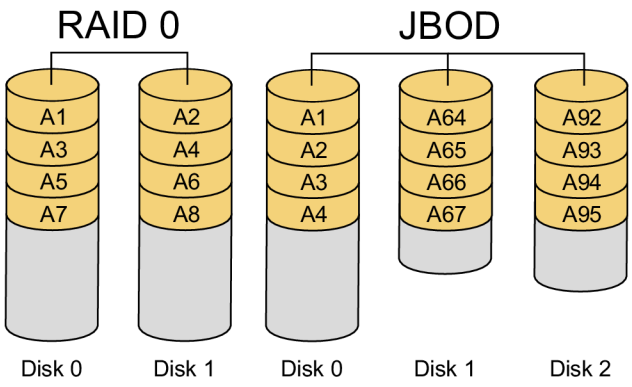
3.2.2 Pole pevných disků RAID

RAID je zkratka pro Redundat Array of Independet disks. Jde tedy o pole pevných disků uspořádaných určitým způsobem. Jedná se o metodu zabezpečení dat proti selhání pevného disku. Ono zabezpečení je tvořeno ukládáním dat na více nezávislých disků, díky čemuž jsou data uchována i při selhání některého z disků v poli. Úroveň zabezpečení závisí na zvoleném typu RAID. Nutno podotknout, že RAID v žádném případě nenahrazuje zálohování.

RAID 0

Tento typ ve skutečnosti není pravý RAID, protože neobsahuje redundantní data. V tomto případě se jedná pouze o spojení jednotlivých disků do jednoho logického celku. Toto spojení může být založeno na dvou principech. Zřetězení (neboli JBOD) nebo prokládání. V případě zřetězení jsou data postupně ukládána na několik disků. Když se naplní první disk, přejde se na druhý atd. Výhodou je snadné zvětšení kapacity pole přidáním dalšího disku. U prokládání se data ukládají na disky střídavě. Pole je rozděleno na stejně velké úseky a zápis nebo čtení delšího úseku dat tedy probíhá z více disků. To s sebou přináší problém při poruše disku, kde poté není příliš pravděpodobné, že by některý soubor zůstal

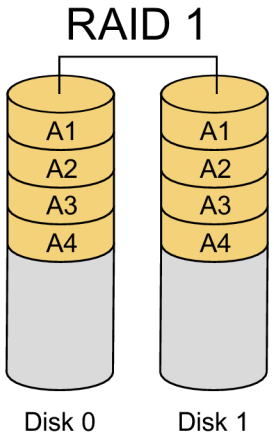
nepoškozen. Výhodou je rychlost čtení a zápis větších bloků dat, protože je možné v jeden moment číst data z jednoho disku a následující blok dat z jiného disku.



Obrázek 3.5 RAID 0 a JBOD [18]

RAID 1

Základní provedení, ale relativně efektivní ochrana dat. Provádí se takzvané zrcadlení obsahu disků. Data jsou současně zapisována na dva různé disky. Pokud jeden z disků selže, je okamžitě k dispozici druhý disk. Tuto metodu lze ještě vylepšit o použití dvou řadičů. Tím se vyřeší i případný výpadek řadiče a teoreticky se tím zvýší rychlost. Nevýhodou tohoto řešení je náročnost na velikost úložiště, které musí být dvojnásobné, protože dochází k duplikaci dat v poměru 1:1.

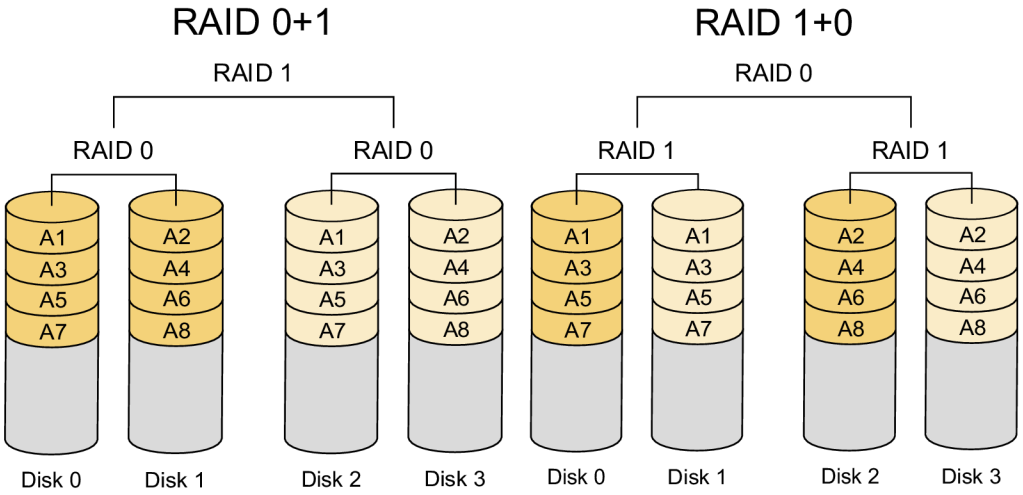


Obrázek 3.6 RAID 1 [18]

RAID 01 a 10

Kombinace RAIDu 0 a 1. Je tvořen pomocí alespoň dvou dvojic disků. Mezi těmito dvojicemi probíhá zrcadlení pomocí RAIDu 1. Samotná dvojice je poté řešena pomocí RAIDu 0. Tento RAID je tedy odolný proti selhání všech disků jednoho podpole.

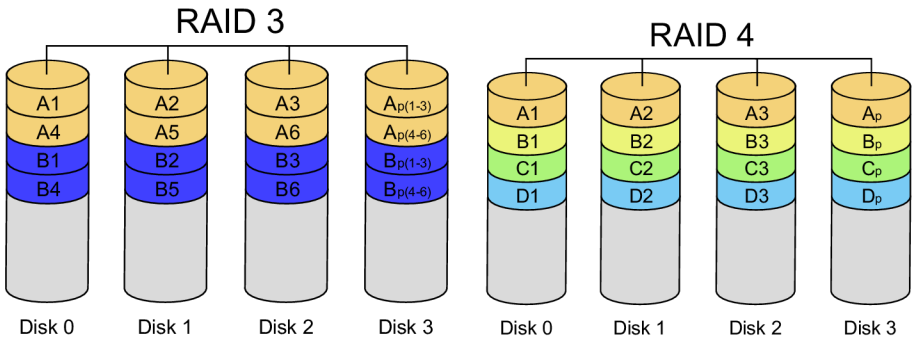
RAID 10 představuje opačný princip než RAID 01. V dané dvojici disků probíhá zrcadlení RAID 1 a mezi dvojicemi je RAID 0. Tento způsob je odolný proti výpadku jednoho disku v každém podpoli.



Obrázek 3.7 RAID 01 a 10 [18]

RAID 3 a 4

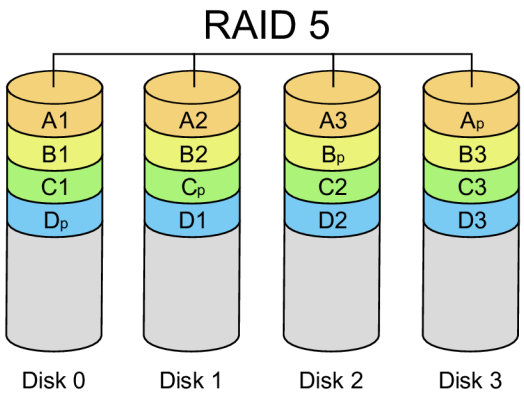
Je použito *x* stejných disků. Na *x-1* disků jsou ukládána data. Na poslední disk je uložena parita těchto dat. Pokud selže některý z disků s daty, lze pomocí ostatních disků a parity data obnovit. Pokud selže paritní disk, jsou data stále zachována. V případě RAIDu 4 jsou poté disky stripovány pro blocích nikoliv po bitech.



Obrázek 3.8 RAID 3 a 4 [18]

RAID 5

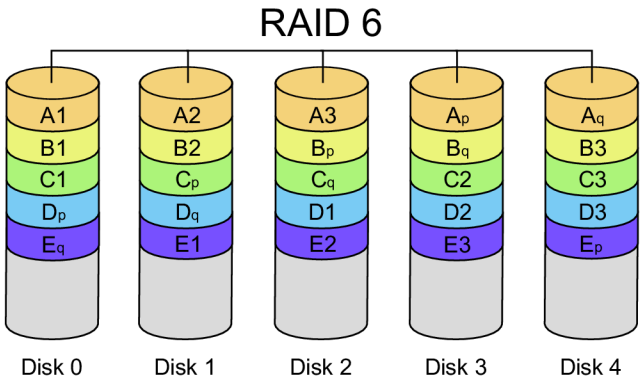
V roce 2021 jeden z nejvíce používaných systémů. Jedná se o vylepšení verzí 3 a 4. Rozdílem je, že parity jsou ukládány střídavě na všech discích a ne pouze na jednom. Celková kapacita je menší o kapacitu jednoho disku. Výhodou je vyšší rychlost čtení. Tento způsob je odolný proti výpadku jednoho disku.



Obrázek 3.9 RAID 5 [18]

RAID 6

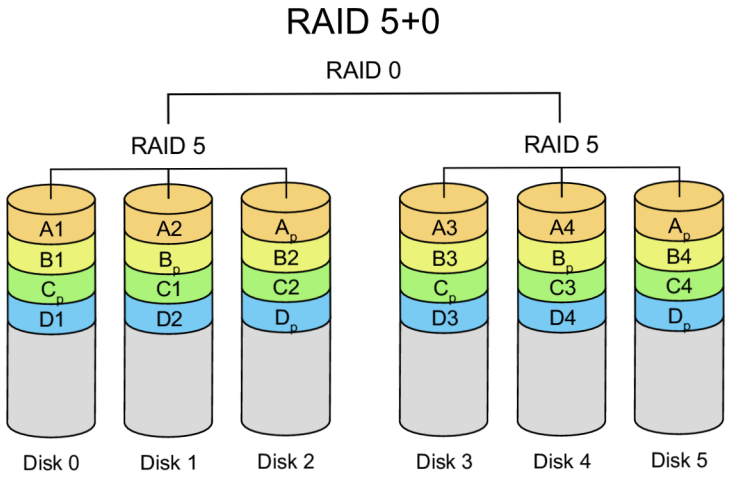
Rozšířený RAID 5. Zde jsou dva paritní bloky, na které se informace ukládají různými způsoby. Paritní data jsou také uložena střídavě na všech discích. Výhodou je odolnost proti výpadku dvou disků. Nevýhodou je nižší rychlost zápisu než u RAIDu 5, právě kvůli nutnosti výpočtu dvou sad paritních informací.



Obrázek 3.10 RAID 6 [18]

RAID 50

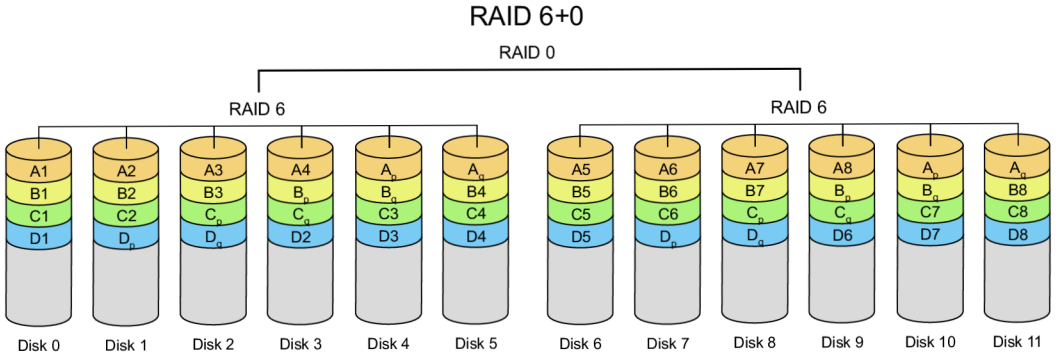
RAID 50 je dvouúrovňové pole, které se skládá z několika RAID 5 polí. Tyto podpole jsou spojeny RAIDem 0. V každém podpoli je potřeba jeden disk na paritní data. To znamená, že výsledné pole je odolné proti selhání jednoho disku v každém podpoli [16].



Obrázek 3.11 RAID 50 [18]

RAID 60

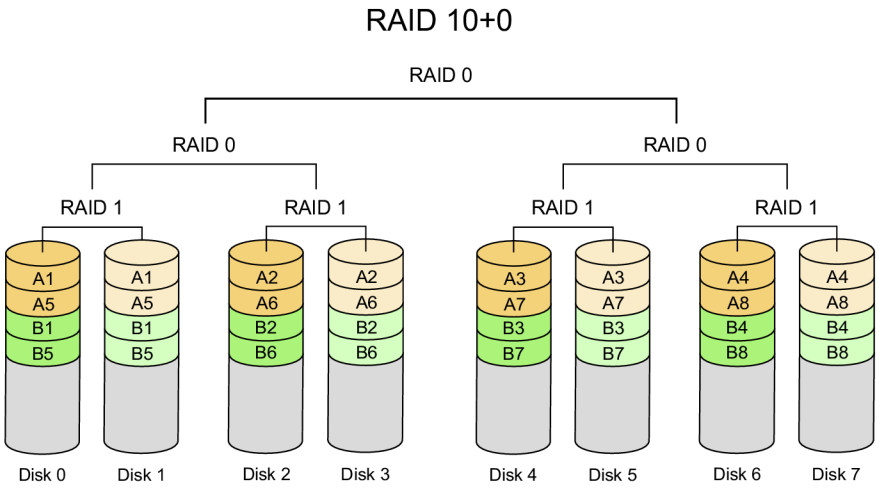
Opět se jedná o dvouúrovňové pole, složené tentokrát z několika polí typu RAID 6. Tím se dosáhne vyššího zabezpečení než v předchozím případě. Minimální počet disků je 8.



Obrázek 3.12 RAID 60 [18]

RAID 100

Tříúrovňové pole vytvořené dvouúrovňovým prokládáním dat na zrcadlené podpole. Prokládání přináší vyšší přenosové rychlosti. Takovéto pole je odolné proti výpadku jednoho disku v každém podpoli. Nevýhodou je možnost využít pouze 50% celkové kapacity pole [17].



Obrázek 3.13 RAID 100 [18]

3.2.3 Architektura úložiště

NAS

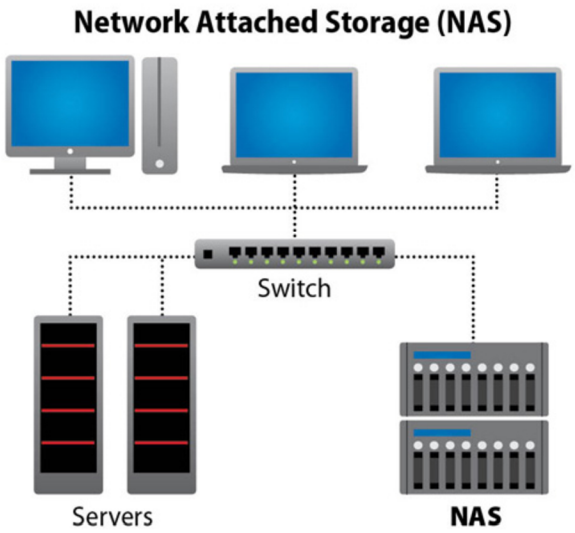
Pojem Network Attached Storage představuje datové úložiště připojené například do lokální firemní nebo domácí sítě. Skrz tuto síť je poté úložiště dostupné z řady zařízení. Pokud uživatel chce, může být NAS dostupný odkudkoliv na světě, pomocí připojení k internetu.

Výrobci zařízení NAS nabízejí široké spektrum produktů, vhodné jak pro malá domácí úložiště, tak pro profesionální firemní potřeby. To znamená malá kompaktní zařízení s jedním slotem na disk, až po velká disková pole v provedení určené pro rack, které je poté možné dále rozšiřovat pomocí expanzních jednotek.

Tento server obsahuje vlastní operační systém, z pravidla založený na bázi Linuxu, který nabízí velké množství užitečných funkcí. Samozřejmě záleží na výrobci a řadě vybraného zařízení. Například operační systém v zařízeních od společnosti Synology, lídra v této oblasti, nabízí nástroje pro správu uživatelů a jejich práv, zálohování, synchronizaci, RAID, funkce multimediálního centra a mnoho dalšího. Použití těchto zařízení obecně nabízí spoustu výhod.

Mezi výhody patří například:

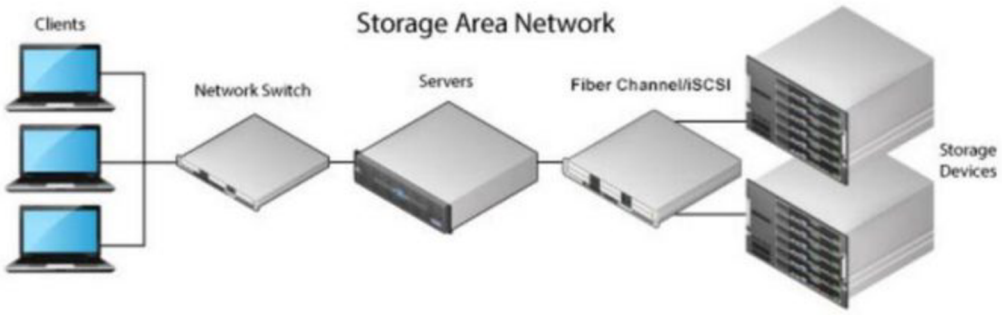
- Jednoduché sdílení dat mezi uživateli.
- Centralizace dat výrazně zjednodušuje proces zálohování.
- Integrace pro většinu zařízení a operačních systémů.
- Šifrování citlivých dat.
- Nesvěřování dat třetím stranám viz cloudová řešení.



Obrázek 3.14 NAS [20]

SAN

Storage Area Network je samostatná datová síť, pomocí které se připojují externí zařízení jako jsou disková pole, páskové knihovny a další úložná zařízení. Zařízení v síti jsou propojena optickými kabely a využívají protokol Fibre Channel (iSCSI). Toto řešení přináší podstatné výhody. Nezatěžuje primární firemní síť, protože všechny přenosy dat se odehrávají v oddělené síti. Dále přináší výbornou škálovatelnost, fyzické oddělení dat a serverů, vyšší propustnost, sdílení zdrojů mezi servery, možnost redundantních cest ke zdrojům, nebo podporu pro clusterová řešení. Nevýhodou je cena a komplikovanější správa. I když se rozdíly pomalu snižují, stále je SAN složitější a dražší řešení než použití NAS či DAS.

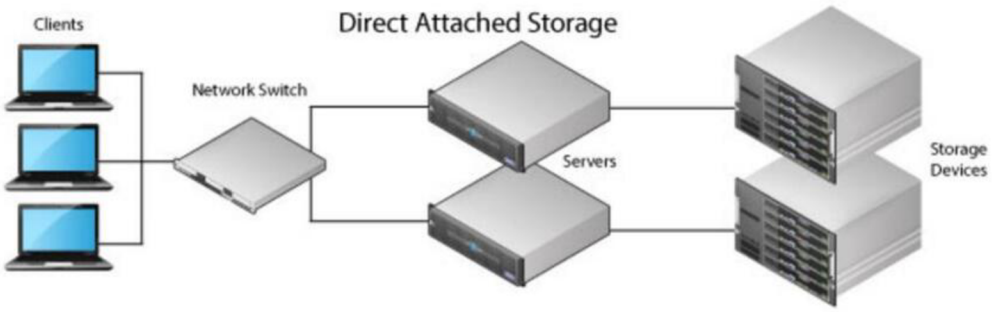


Obrázek 3.15 SAN [20]

DAS

Direct-attached storage je úložiště přímo připojené k serveru či počítači, který k němu přistupuje. Mezi taková úložiště patří například SSD a pevné disky, optické mechaniky, externí disky nebo disková pole. Zařízení je připojeno především pomocí rozhraní SATA, SCSI, SAS nebo Fiber Channel.

Mezi výhody patří snadná použitelnost, fakt že je tato technologie široce dostupná, vysoká datová propustnost, nízká přístupová doba a nižší cena. Nevýhodou je špatná škálovatelnost, nemožnost sdílení úložiště mezi více servery nebo nedostupnost dat při výpadku serveru [19].



Obrázek 3.16 DAS [20]

3.2.4 Úložná média

Existuje relativně velké množství různých druhů úložných zařízení. Opět je potřeba vybrat zařízení s ohledem na zamýšlené použití. Určitě lze říct, že zařízení vhodné k zálohování domácího PC rozhodně nebude vhodné pro zálohovací procesy ve firemním prostředí, kde se zpravidla o zálohování svých dat nestará každý uživatel samostatně.

Minimálně ve stejném schématu použití. Níže jsou uvedena datová média z obou konců spektra.

HDD a SSD

Druhý pevný disk nebo SSD disk v počítači může být použit pro nejjednodušší formu zálohy dat. Nutno podotknout, že takovéto zálohování není příliš efektivní. V podstatě poskytuje ochranu pouze proti selhání primárního disku. V dalších případech je tato záloha zbytečná. Mezi tyto případy například patří ztráta/krádež zařízení, napadení virem nebo mechanické poškození stroje.

Jiná situace je u externích disků. Zde již je pro použití zálohování více racionálních důvodů. Hlavní výhoda je jasná. Tedy že disk není neustále připojen k počítači. Naopak se může disk uložit na bezpečné místo fyzicky oddělené od zdroje originálních dat. Oblíbenosti tohoto řešení nahrávají i přívětivé ceny v přepočtu na jednotku dat a snadné použití. Běžná kapacita disků se pohybuje v jednotkách terabajtů.

Flash disk

Skladné, relativně rychlé a odolné úložné zařízení. Kapacita se pohybuje v řádech desítek až stovek GB. Pro časté zálohování však vhodné příliš není. Oproti externím diskům je zde vyšší cena za jednotku. Doporučit se dají spíše pro akutní zálohu menšího množství dat.

Optická média

DVD nebo Blu-ray je stále částečně používaný způsob pro ukládání dat. Výhodou je, že lze vytvořit mnoho kopií a ty různě rozmístit, čímž se lze pojistit například proti přírodní katastrofě. Nicméně je tento způsob vhodný spíše k účelům archivace než zálohování, což vychází z jejich vlastností. Optická média jsou náchylná na poškození, protože každý škrábanec může znamenat ztrátu dat, takže častá manipulace není žádoucí. Také je nutné uskladnit média v tmavém a suchém prostředí. Z druhé strany je ale problémem také krátká životnost média. Běžně lze narazit na případ, kdy jsou data po pár letech nečitelná. Existují i odolnější optická média, která používají například chemické ošetření nebo metalo-keramickou záznamovou vrstvu, která má vydržet 100 a více let. Zde už je ale problém vyšší cena.

Cloud

V posledních letech došlo v této oblasti k obrovskému rozvoji a dnes tak existuje velké množství poskytovatelů nabízející cloudové služby. Toto řešení je z mnoha důvodů velice populární. Dostupnost dat odkudkoliv na světě, ochrana dat, která se fyzicky nacházejí mimo počítač uživatele, není potřeba řešit selhání disků či další zálohy. O vše se stará poskytovatel služby. Dále jsou to možnosti automatické synchronizace mezi zařízeními, sdílení dat mezi uživateli, teoreticky neomezená kapacita úložiště, verzování souborů a mnoho dalších funkcí. V neposlední řadě zde roli hraje i příznivá cena, kdy v roce 2021 stálo například 2TB úložiště Google Drive 3000 Kč za rok.

Na druhou stranu, fakt, že o vše se stará poskytovatel může být i nevýhoda. Problémem je svěřování dat třetí straně, kdy musíte vybrané službě důvěřovat, že Vaše data nezneužije. Tuto hrozbu lze alespoň omezit šifrováním ještě před odesláním na cloud. Dalším problémem může být, že uživatel vůbec netuší, kde se fyzicky jeho data nachází. Primárně mohou být uloženy v datacentru někde v Evropě a zálohy mohou být rozesety téměř na všech kontinentech. Otázkou je, zda je vlastně informace o umístění dat přínosná, protože stejně nelze dosáhnout plné kontroly nad daty. Další potencionální nevýhoda může být v nutnosti disponovat stabilním a rychlým internetem.

Páska

Na první pohled stará technologie, která se ale stále relativně často využívá. Velkou výhodou ve srovnání s pevnými disky je výrazně lepší cena za uložení jednoho terabajtu. Dále nabízí výbornou spolehlivost a při správném skladování životnost až desítky let. Pásky nabízejí relativně velkou kapacitu (jednotky terabajtů) v poměru ke své velikosti, proto jsou snadno přenositelné, tudíž mohou být uloženy bezpečně mimo polohu originálních dat. Další podstatnou výhodou je ochrana proti hackerským útokům. Poté co jsou data uložena na pásku, je páska odpojena a útočníci k ní tak nemají přístup online. Nevýhodou je nižší rychlost. Vzhledem k těmto vlastnostem je páska vhodnější spíše k archivaci, případně k zálohování s menší retencí, jakožto další neprimární záložní řešení [21].

3.3 Metody zálohování

Ruční zálohování

Manuální zálohování lze relativně kvalitně provádět spíše na soukromé úrovni. Typicky se jedná například o zálohu rodinných fotografií, respektive dat, která se příliš často nemění nebo nepřibývají. Nicméně ani zde to není ideální přístup. Automatické zálohování je vhodné nastavit například v chytrém telefonu. Pořízené fotografie, nastavení nebo třeba kontakty se poté automaticky zálohují do cloudu a v případě ztráty či poškození zařízení jsou data stále dostupná.

Automatické zálohování

Největší problém u zálohování je lidský faktor. Pokud se z úvahy vynechají miniaturní firmy o pár zaměstnancích, které nepoužívají žádné složitější systémy, i zde je to ovšem velmi k zvážení, je ruční přístup k zálohování ve firemním prostředí v podstatě nereálný. Pokud by si každý zaměstnanec, případně v zastoupení IT technika, měl sám zálohovat svoje data, je naprosto jisté, že by dříve či později došlo k problému. Důvodem může být opomenutí uživatele, špatně provedená záloha, nekonzistentnost dat a velmi často naprostá neschopnost logického myšlení ze strany uživatelů.

V podnicích, které provozují vlastní síťovou a serverovou infrastrukturu jsou poté automatické zálohy absolutně nutné. Není reálné, aby IT oddělení zvládlo neustále a bezchybně ručně zálohovat třeba i desítky serverů, switchů, routerů, firewallů, stovky osobních počítačů a telefonů.

Softwary pro zálohování proto nabízejí poměrně široké možnosti nastavení automatického zálohování. Lze nastavit čas provedení zálohy, retenci, opakování zálohy v případě selhání, emailové upozornění na výsledek zálohy, kontrolu použitelnosti provedené zálohy atd. Samozřejmě ani zde není dobré se stoprocentně spoléhat na software a je vhodné provádění záloh pravidelně ručně kontrolovat, nicméně i tak automatické zálohování extrémně zefektivňuje tuto činnost.

3.3.1 Metody zpracování dat

Duplikace

Tvorba více stejných kopií dat uložených na různých úložištích v různých lokalitách. Tento proces snižuje šanci na ztrátu dat vlivem poškození jednoho z úložišť. Nevýhodou je náročnost na kapacitu úložišť, kdy stejná data zabírají násobně více prostoru.

Deduplikace

Tento proces je opakem duplikace. Princip je založen na předpokladu, že není nezbytné uchovávat více instancí stejných dat. Namísto kopie se tak ukládá pouze odkaz na původní umístění dat.

Příkladem využitelnosti může být společnost se 400 zaměstnanci, kde každý má osobní počítač s pevným diskem o velikosti v řádech stovek gigabajtů. Pokud bude nutné všechny počítače zálohovat, potřebná kapacita zálohovacího úložiště se bude pohybovat v řádech desítek terabajtů. V síti 100 Mbit bude úplné zálohování trvat dva až tři týdny. Toto množství dat by se dalo snížit deduplikací. Každý počítač má nainstalovaný stejný operační systém, stejné aplikace a často mnoho různých kopií stejných dat. Deduplikací, respektive přenášením pouze jedinečných dat, lze dosáhnout obrovských úspor jak v nárocích na úložiště, tak z pohledu času.

Během zálohování jsou data rozdělena do bloků a je kontrolována jedinečnost každého bloku pomocí kontrolních součtů uložených v databázi. Jedinečné bloky jsou odeslány do úložiště a duplikáty jsou přeskočeny.

Deduplikace tedy pomáhá snižovat využití úložného prostoru. Dále sníží zatížení sítě, protože se přenáší méně dat. Na druhou stranu deduplikované úložiště může vyžadovat více výpočetních prostředků jako je RAM nebo CPU [22].

Šifrování

Šifrování zajišťuje ochranu dat. Dá se říct, že každá firma, ale i soukromá osoba má data, která je potřeba šifrovat. Jedná se o citlivé informace, smlouvy, důležité výrobní postupy atd. Princip šifrování je v podstatě překlad originálních dat do nečitelného zakódovaného formátu. Klíč drží pouze osoba, která data zašifrovala. Pokud je potřeba data dešifrovat použije se právě onen klíč. Šifrování v této oblasti lze rozdělit na tři hlavní metody.

Šifrování celého disku zajistí ochranu celého prostoru, případně logické jednotky. Při této metodě je zašifrován celý disk, včetně prázdného místa. Pro přístup k datům je na disk zanesen pre-boot záznam pro zadání hesla, či nahrání certifikátu. Mezi výhody patří ochrana všech dat včetně SWAP souborů, nebo automatické šifrování všeho co uživatel uloží. Nevýhodou může být vyšší hardwarová náročnost díky šifrování a dešifrování velkého objemu dat, nebo časově náročné prvotní zašifrování. Hlavním rizikem je poškození bootovací části, což velmi pravděpodobně přinese ztrátu dat.

Souborové šifrování poskytuje ochranu pouze vybraných částí disku. Takových částí, které jsou pro uživatele důležité. To přináší menší nároky na hardwarové prostředky nebo snadné zálohování. Pro uživatele se na první pohled při práci nic nemění. Soubory může upravovat, kopírovat mazat atd., aniž by poznal, že jsou data šifrována. Při práci s daným souborem je tento soubor dešifrován, přenesen do operační paměti a po skončení práce je z ní zase odstraněn. Z jiného úhlu pohledu je však ponechání vlivu uživatele na šifrování ve firemním prostředí nevýhodou. Lehce může citlivá data uložit mimo zašifrované adresáře a tím pádem nejsou tato data chráněna.

Šifrování virtuálního disku nabízí kompromis mezi výše uvedenými metodami. Určité místo na fyzickém disku se vyhradí a vytvoří se virtuální disk, který se v operačním systému připojí jako fyzický disk. Uživatel má tak k dispozici zabezpečený prostor. Kapacita tohoto prostoru roste automaticky dle potřeby. Nároky jsou zde také menší než u šifrování celého disku [23].

Komprese

Komprese dat zmenšuje velikost vytvořených souborů při současném zachování informací obsažených v datech a tím pádem šetří místo na úložném médiu a snižuje náročnost na přenos v síti. Nevýhodou může být, že komprese ovlivňuje dobu trvání postupu zálohování. Proces komprese lze rozdělit na dvě hlavní formy. Ztrátová a bezeztrátová.

Bezeztrátová komprese vyhledá a odstraní statisticky nadbytečné části. V této technice nejsou ve skutečnosti odstraněny žádná data. To přináší výhodu možnosti komprimovaný soubor dekompresí rekonstruovat do původní podoby. Nevýhodou je menší kompresní poměr.

Ztrátová komprese přináší nenávratnou ztrátu některých dat a nelze je zpětně rekonstruovat. Tato metoda nabídne mnohem větší kompresní poměr za cenu možné zhoršení kvality výsledného souboru. Využití najde například u přenosu videa či zvuku, kdy

se počítá s tím, že lidský mozek si sám chybějící údaje domyslí, nebo si jich vůbec nevšimne [24].

3.3.2 Způsob zálohování

D2D

Metoda Disk to Disk je typ zálohy určený pro ukládání dat na pevné disky, případně na disková pole.

D2T

Záloha Disk to Tape se běžně provádí přes noc na magnetické pásky. Tyto pásky jsou poté odpojeny od sítě, případně jsou kompletně přemístěny na jiné fyzické místo.

D2D2T

Kombinace předchozích metod. Ke klasické záloze na disk, se přidává druhá vrstva v podobě pásky. Tento systém lze využít v klasické podobě jednoduché duplikace dat, případně v systému, kdy se na pásku vytváří základ plné zálohy a na disky se ukládají diferenciální nebo inkrementální zálohy.

D2D2C

Moderní způsob zálohování, který k záloze na disk přidává další vrstvu ve formě cloudu. Výhodou je na rozdíl od zálohy na pásku, přístupnost k datům z jakéhokoliv místa. Mezi další výhody patří nízká složitost dostat zálohy fyzicky mimo originální data, vysoká škálovatelnost a relativně příznivá cena. Nevýhodou je závislost na dostupnosti a rychlosti internetového připojení.

F2F2C

Flash to Flash to Cloud je nejnovější přístup, který přináší rychlejší čas obnovy virtuálních strojů a granulárních dat. Tzv. All-Flash úložiště díky použití SSD disků, případně dokonce NVME disků, přináší obrovský nárůst v rychlosti.

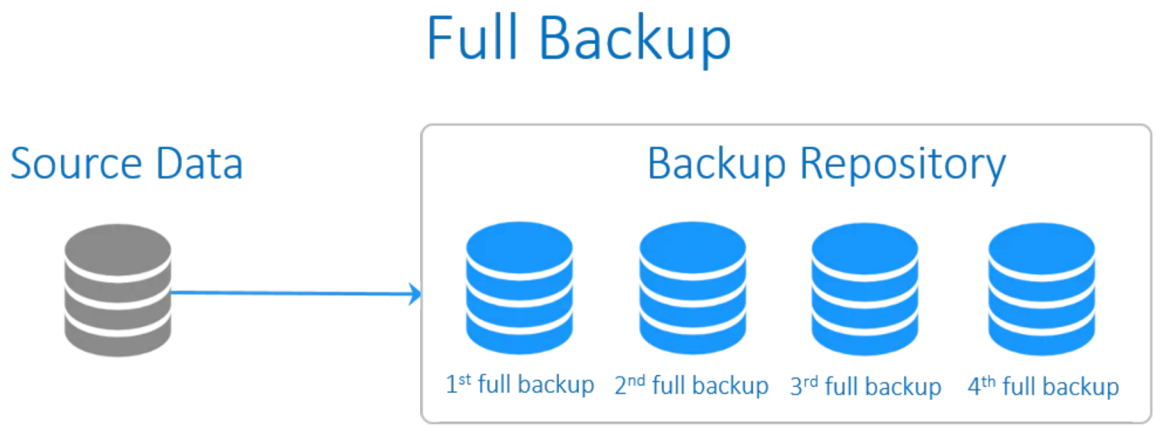
Nevýhodou je samozřejmě podstatně vyšší cena ve srovnání s HDD. Nicméně vyšší cena může být pro určité firemní subjekty naprosto zanedbatelná ve srovnání s možnými finančními ztrátami způsobené nefunkčními systémy, respektive časově delší obnovnou ze záloh po krizové situaci [25].

3.3.3 Druhy záloh

Jak již bylo uvedeno, při zálohování například osobního domácího počítače se lze v určitých scénářích relativně dobře spolehnout na ruční a naprosto jednoduché zálohování v podobě prostého kopírování vybraných dat na různá úložiště. Ve firemním prostředí je již tento přístup nemyslitelný. Jednak z důvodu mnohonásobně většího množství dat, ale také z principu samotných zdrojů dat. Zde je potřeba zálohovat celé systémy, servery, databáze atd. U těchto záloh je poté potřeba z jedné strany zajistit integritu dat, kontinuitu záloh nebo rychlou obnovu a z druhé strany minimalizovat náklady na úložiště a další zdroje infrastruktury. K tomu již slouží pokročilejší druhy záloh. Správný výběr záleží na mnoha faktorech a potřebách dané firmy a obecně nelze označit jeden druh záloh nebo strategii za špatnou nebo za dobrou.

Plná záloha

Jedná se o výchozí zálohu pro další typy záloh. Tato záloha obsahuje kompletně všechna data obsažená v originálním zdroji. To přináší výhodu rychlé a snadné obnovy, protože veškeré soubory jsou v jedné záložní sadě. Nevýhodou je naopak délka provedení samotné zálohy. Oproti dalším způsobům může být tato záloha i desetkrát pomalejší. Důvodem je, že každý soubor je zálohován při každé záloze znovu. Tím pádem je tato záloha také nejnáročnější, co se týče požadavků na velikost úložiště. Proto se tato záloha provádí typicky jednou týdně nebo měsíčně [36].



Obrázek 3.17 Plná záloha [26]

Zrcadlení

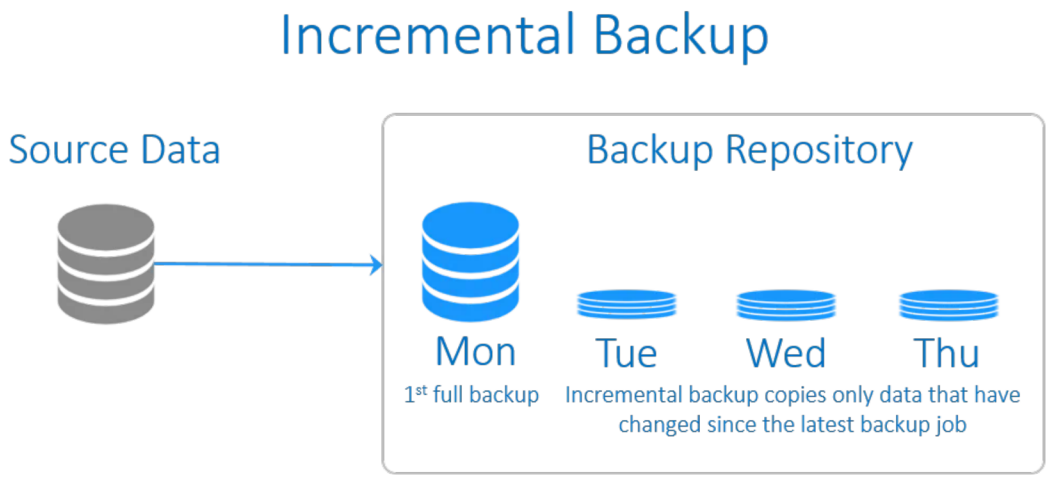
Princip je velmi podobný plné záloze. Mirror záloha však není z mnoha úhlů pohledu skutečná záloha. Pokud je ve zdroji smazán soubor, tak se stejný soubor smaže i v kopii. To znamená, že jakékoli úpravy dat například důsledkem lidské chyby, nehody nebo viru ve zdroji, mají stejný účinek i v kopii. Zrcadlení vytváří přesnou kopii zdrojových dat, ale v úložišti je uložena pouze nejnovější verze dat bez sledování různých verzí souborů.

Na rozdíl od ostatních záloh nejsou všechny soubory v kopii uloženy v jediném komprimovaném souboru, ale oddělené stejně jako ve zdroji. To umožňuje přímý přístup k záložním souborům bez provádění obnovy. Na druhou stranu tato metoda přináší nevýhody v podobě vysokých nároků na úložný prostor, vysoké riziko neoprávněného přístupu a poškození nebo zneužití dat.

Inkrementální záloha

Inkrementální neboli přírůstková záloha se skládá ze dvou částí. Na začátku je provedena plná záloha jakožto základ. Dále jsou prováděny inkrementální zálohy, které obsahují pouze nová, respektive změněná data ve srovnání s plnou zálohou a poté ve srovnání s poslední inkrementální zálohou.

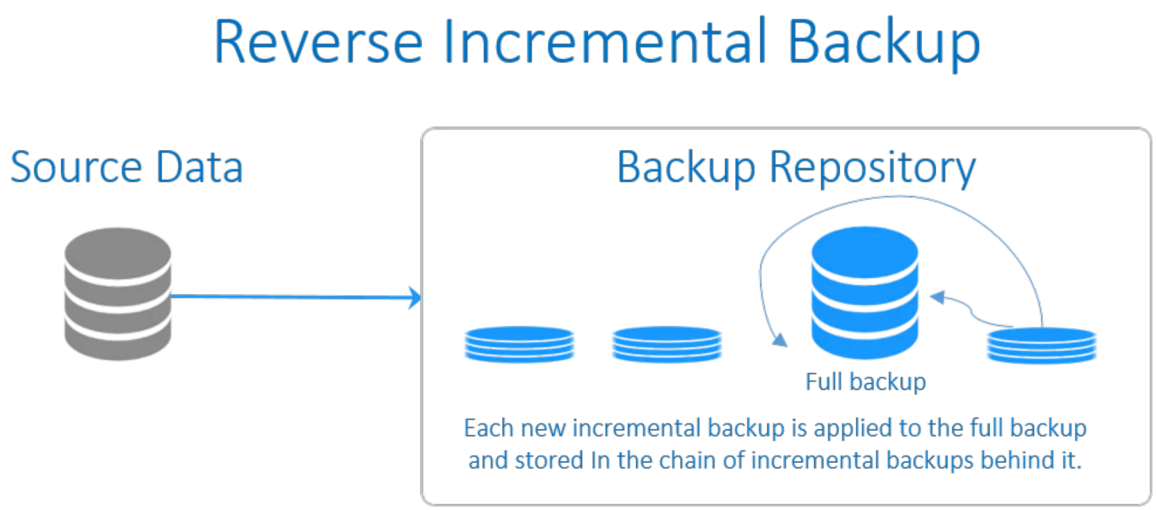
Tento přístup k zálohování přináší šetření úložného prostoru a rychlejší proces zálohování, ale obnova trvá naopak déle. A to z důvodu nutnosti obnovit plnou zálohu a následně všechny přírůstkové zálohy. Z tohoto důvodu pak použití v praxi vypadá tak, že se provádí plná záloha například jednou týdně a poté každý den inkrementální záloha. Dalším problémem je selhání jedné ze záloh v řetězci nutných k obnově. V tomto případě je celá záloha nefunkční [36].



Obrázek 3.18 Inkrementální záloha [26]

Reverzní inkrementální záloha

Upravená inkrementální záloha. Základ stojí opět na plné záloze. Následně se provádějí přírůstkové zálohy, které se následně doplní do původní plné zálohy. Vedle toho se přírůstkové zálohy ukládají do záložního řetězce. To přináší možnost obnovit plnou zálohu v situaci, kdy se obnovují starší verze dat. Výhodou tohoto řešení je rychlejší obnova.

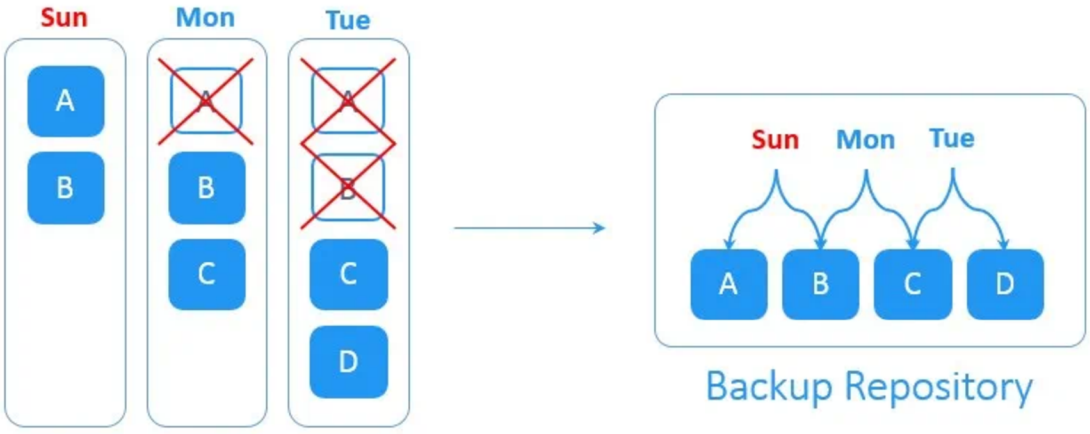


Obrázek 3.19 Reverzní inkrementální záloha [26]

Forever-Incremental záloha

Tento princip zálohování se od klasické inkrementální zálohy liší z hlediska organizace a zpracování zálohovaných dat. Jako u předchozích typů záloh začíná proces počáteční plnou zálohou jako referenčním bodem pro sledování změn. Následně jsou už prováděny pouze přírůstkové zálohy bez pravidelných plných záloh. Pro lepší ilustraci je uveden jednoduchý příklad. V sobotu byla provedena plná záloha. Od dalšího dne se denně provádějí přírůstkové zálohy. V neděli se vytvoří dva nové bloky dat A a B. V pondělí se blok A smaže a vytvoří se nový blok C. V úterý se smaže blok B a vytvoří se nový blok D. Systém sleduje všechny denní změny a odstraňuje duplicitní datové bloky, čímž se zmenší nároky na úložiště. Současně jsou k bodům obnovy přidány odkazy označující související datové bloky.

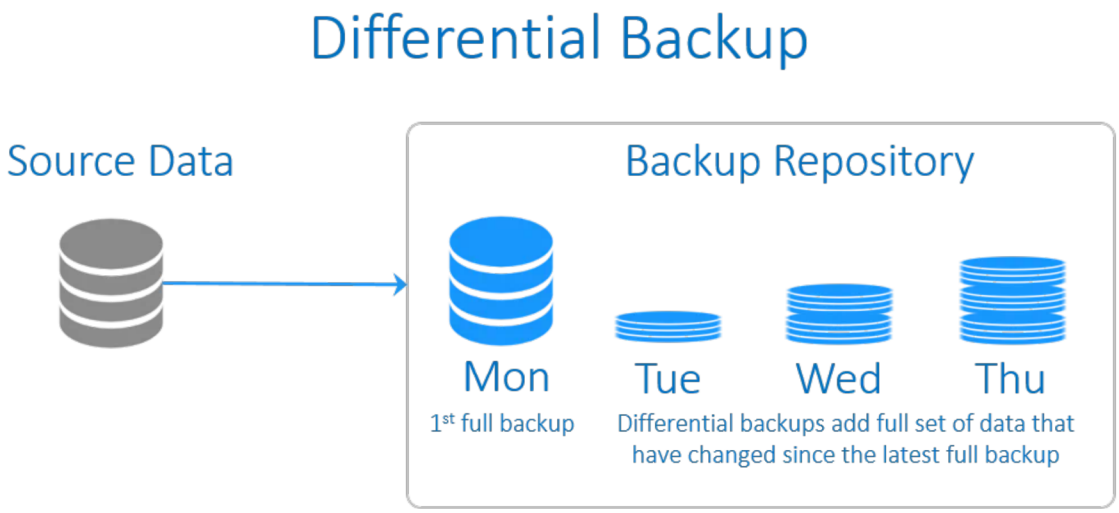
V závislosti na individuálním nastavení uchování záloh se po vytvoření řady přírůstkových záloh odstraní zastaralá data a body obnovy. Tím se uvolní úložný prostor. Všechna uložená data zálohy jsou organizována tak, že počáteční plná záloha i následující přírůstkové zálohy umožňují úplnou obnovu. Mezi výhody tedy patří rychlejší operace zálohy a obnovy, lepší správa úložiště a nízké zatížení sítě [26].



Obrázek 3.20 Forever-incremental záloha [26]

Diferenciální záloha

Stejně jako inkrementální tak i diferenciální neboli rozdílová záloha má vytvořen základ na plné záloze. Rozdíl oproti přírůstkové záloze spočívá v záloze dat, která se změnila od původní plné zálohy, nikoliv od předchozích záloh. Výhodou je rychlejší obnova než u přírůstkové metody, protože stačí původní plná záloha a poté některá rozdílová záloha, která se vztahuje k požadovanému bodu obnovy. Navíc případné poškození jedné z rozdílových záloh nemá vliv na ostatní zálohy jako v případě přírůstkové zálohy. Nevýhodou je postupné narůstání velikosti rozdílové zálohy a tím pádem větší nároky na úložiště a dobu trvání vytváření zálohy. Z tohoto důvodu je vhodné stejně jako u přírůstkové zálohy provádět například jednou týdně plnou zálohu [36].

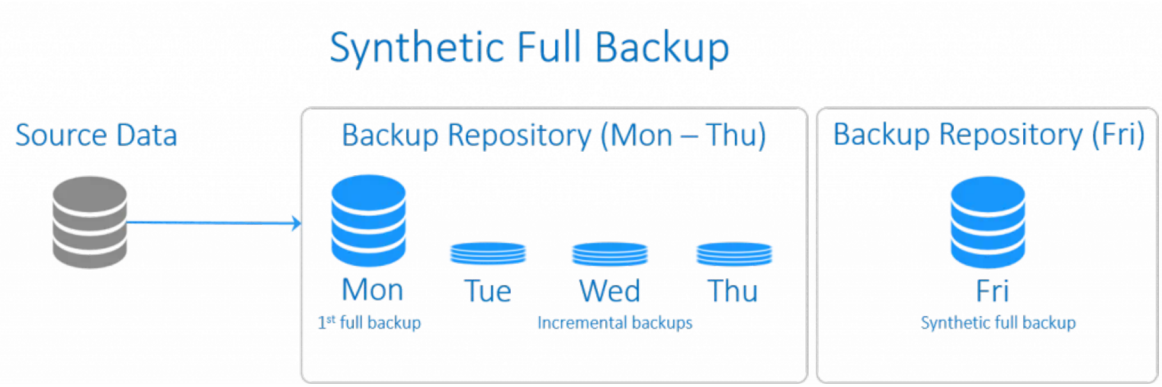


Obrázek 3.21 Diferenciální záloha [26]

Syntetická plná záloha

Klasická plná záloha zabírá spoustu času, je náročná na úložiště, způsobuje zatížení sítě a zdrojů dané infrastruktury. Z těchto důvodů byla vytvořena syntetická plná záloha.

Tato záloha má mnoho společného s reverzní inkrementální zálohou. Rozdíl lze najít ve způsobu manipulace s daty. Syntetická záloha začíná klasickou plnou zálohou, která se ale provede pouze jednou a poté už nikdy. Klasickou plnou zálohu následuje řada přírůstkových záloh. Zde přichází hlavní rozdíl. V určitém okamžiku jsou přírůstkové zálohy konsolidovány a aplikovány na existující plnou zálohu, aby se syntetizovala nejnovější záloha jako nový výchozí bod. Tento způsob zálohování tak přináší všechny výhody pravidelných úplných záloh, a přitom potřebuje méně času a spotřebovává méně úložného prostoru. Výsledkem je rychlejší záloha i obnova, lepší správa úložiště, menší požadavky na prostor a menší zatížení sítě [26].



Obrázek 3.22 Syntetická plná záloha [26]

3.3.4 Rotace záloh

Pod pojmem rotace záloh se skrývá strategie pro rotaci úložných medií. Typicky se jedná o magnetické pásky. Cílem je co nejvíce snížit opotřebení, respektive dosáhnout rovnoměrného opotřebení všech pásek nebo disků. Pokud se budou ukládat data pouze na jeden disk a ostatní se nevyužijí, je selhání onoho disku mnohem pravděpodobnější. Rotaci je vhodné využít při zálohování většího množství dat. Opakem rotace je kumulativní ukládání dat, kdy se pro zálohu používá jediné médium, které se po naplnění smaže a opětovně použije. Tato metoda je vhodná pro zálohování malého množství dat.

FIFO

Schéma First In, First Out je klasický princip fronty. Nové nebo upravené soubory se ukládají na média, která obsahují nejstarší zálohovaná data. Při každodenním zálohování na sadu 14 médií by tedy nejstarší záloha byla dva týdny stará. Nevýhodou tohoto schématu je situace, kdy zálohovaná data obsahují chybu, na kterou se přijde například až za 3 týdny.

GrandFather-Father-Son

Princip určený pro magnetické pásky. V tomto schématu se využívají typicky tři cykly záloh. Například měsíční (GrandFather), týdenní (Father) a denní (Son). Jednou měsíčně se provede plná záloha na pásku, která se poté vyjme z knihovny a nahradí se novou páskou. Mezí tím se každý týden provede týdenní plná záloha, ze které se celý následující týden provádí inkrementální denní zálohy. Denní záloha se provádí každý den a vždy se využije jiná páska. Následující týden jsou tyto pásky znovu přepisovány. Přehledný měsíční princip je vidět v tabulce.

Neděle	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota
Father	Son	Son	Son	Son	Son	Son
Father	Son	Son	Son	Son	Son	Son
Father	Son	Son	Son	Son	Son	Son
Father	Son	Son	Son	Son	Son	Grandfather

Tabulka 3.1 GrandFather-Father-Son

RoundRobin

Schéma založené na označení jednoho média pro každý den v týdenním cyklu. Jedná se o jeden z nejzákladnější typů rotace. Na každou pásku se ukládá aktuální denní záloha. Výhodou je rovnoměrné opotřebení pásek a jednoduchá implementace tohoto řešení. Nevýhodou je krátká historie záloh.

Týden	Páska 1	Páska 2	Páska 3	Páska 4	Páska 5	Páska 6	Páska 7
1	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota	Neděle
2	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota	Neděle
3	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota	Neděle
4	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota	Neděle

Tabulka 3.2 RoundRobin

Hanojská věž

Technika založená na matematické hádance, při které se využívá rekurzivní metoda pro optimalizaci cyklu zálohování. Nutnost častého zálohování přináší náklady na jejich uložení. Tento princip přináší kompromis. V příkladu s pěti sadami zálohovacích medií je postup následující. První sada médií se používá každý druhý den. Druhá sada se používá každý čtvrtý den. Třetí sada se používá každý osmý den. Každá další sada medií přidaná do rotace se používá pouze tehdy, když se nepoužívají předchozí sady. Čím déle je sada přidána, tím méně se používá a tím starší soubory ukládá. Princip je v následující tabulce, která ilustruje použití při třech média setech [27].

		Dny							
		1	2	3	4	5	6	7	8
Média set	1	A		A		A		A	
	2		B				B		
	3				C				C

Tabulka 3.3 Hanojská věž

3.3.5 Retenční politika

Definice retenční politiky zní takto. Jedná se o protokol, který definuje životní cyklus dat v organizaci. Správné nastavení retenční politiky pomáhá šetřit náklady v rámci zálohovací strategie a dodržovat zákony. Běžné chyby, které firmy dělají, je zálohování všeho co mají a uchování těchto záloh tak dlouho, jak je to možné. A to je v mnoha případech zbytečná chyba. Přináší to vysoké náklady za úložiště. Správně vytvořená politika řeší tento problém tím, že určí, co zachovat a co odstranit. Vedlejší, ale stejně tak důležitý produkt takové politiky je poté snadnější orientace v zálohách a rychlejší obnova. Životní cyklus obsahuje následující:

- Jak dlouho bude firma uchovávat určitou informaci.
- Jak budou tyto informace uloženy.
- Jaká data by měla být uložena a proč.
- Kdy smazat konkrétní data.

Pro tvorbu politiky jsou zásadní dva faktory. Externí a interní vlivy. Za prvé právní předpisy pro uchování určitých dat, jako jsou účetní dokumenty. Na data se mohou vztahovat různé zákony. Existují specifické nařízení pro jednotlivé státy, EU, nebo podle oboru podnikání. Za druhé jsou to potřeby dané organizace. Ty se v každé firmě mohou a pravděpodobně budou lišit. Obecně však lze dodržovat následující doporučené postupy.

Klasifikace dat dle typu a potřeb

Jak bylo řečeno, na zásady uchování dat, mají vliv externí a interní potřeby. Tímto rozdělením se určí cíl pro každou část dat. Některá data možná nebudou nikdy potřeba, ale musí zůstat uloženy na bezpečném místě. Z druhé strany jsou data, která budou potřeba v případě incidentu co nejrychleji obnovit. Při kategorizaci dat tímto způsobem zapadnou ostatní součásti politiky velice přirozeně. Navíc při tomto procesu dojde k určení RTO a RPO, které jsou středem strategie obnovy po havárii. Otázky, které je potřeba zodpovědět jsou:

- Jaká data jsou cenná z právního pohledu?
- Jaká data jsou cenná z vlastních potřeb firmy?
- Jaká data se týkají veřejných, chráněných nebo důvěrných informací?

Kategorizace dat dle životního cyklu

Životní cyklus zálohovaných dat je striktně definován potřebami dané společnosti. Z jednoduché logiky věci, by ne všechna dat měla být uchovávána stejnou dobu. U jednoho typu dat bude například vyžadována archivace po dobu deseti let a poté okamžité smazání. U jiných dat může být rozumnější a nákladově efektivnější je po nějaké době smazat, protože jsou například již příliš zastaralá a pro obnovu jsou už v podstatě nepoužitelná. Díky tomu, že budou tyto vlastnosti definovány, lze vytvořit plán zálohování pro každou specifickou datovou sadu v závislosti na době, po kterou by měla být uchovávána. Základní rozdělení lze provést následovně:

- Data, která je potřeba uchovávat měsíc.
- Data, která je potřeba uchovávat půl roku.
- Data, která je potřeba uchovávat rok.
- Data, která je potřeba uchovávat 3 roky.

Samozřejmě s ohledem na konkrétní firmu je potřeba definovat správné časové úseky a lze se pohybovat v menších i výrazně větších časových úsecích.

Určit co smazat a kdy

Včasné mazání dat je jedním z kritických pravidel, které firmy často opomíjejí, protože se na první pohled mohou zdát kontraproduktivní. Přesto je uchovávání dat, která by se měla smazat, relativně nákladným a hlavně zcela zbytečným procesem. Mezi typické důsledky patří:

- Nepřehlednost a složitá navigace v datech.
- Zbytečné vytěžování hardwaru.
- Prodlužování zálohy a obnovy.
- Riskování bezpečnosti.
- Vynakládání peněz na úložiště pro data, která nemají žádnou hodnotu.
- Případné stíhání při nedodržení zákonných postupů.

Například článek 5 obecného nařízení o ochraně osobních údajů (GDPR) uvádí, že jsou společnosti povinné zničit osobní údaje, u kterých se prokáže, že již nejsou potřebné pro obchodní nebo právní účely.

Definovat počet a typ verzí k uložení

Jednotlivé verze záloh jsou kopie originálních dat, které obsahují všechny změny, které byly se soubory provedeny. Parametry k definování:

- Dle počtu verzí k uložení.
    - Další (neaktivní) verze.
    - Poslední verze souborů, které byly smazány.
- Dle doby uložení.
    - Stávající data.
    - Smazaná data.

Rozhodnout o typu záloh a jejich frekvenci

Výběr správného typu zálohování a nastavení frekvence provádění záloh vždy záleží na konkrétní firmě, zdrojích a datech, která je potřeba zálohovat [28].

3.4 Software pro zálohování

Jak již bylo uvedeno, při zálohování ve firemním prostředí je nutná automatizace. Pro praktickou část práce jsou zde vybrány tři hlavní zálohovací softwary, které vyřeší zhruba 90 procent všech požadavků. Některé zbylé zálohovací úkoly budou vyřešeny integrovanými řešeními daných systémů.

Pro nezávislý výběr softwarového řešení není v této konkrétní implementaci příliš prostor. Zálohovací řešení, která zde budou představena jsou zvolena z následujících důvodů. V případě sofwaru Veeam má již firma zakoupenou licenci na tento software a tak bude dále využita. Na druhou stranu i v případě, pokud by licence zakoupená nebyla, tak řešení Veeam by i přesto bylo nejspíše použito. Mezi důvody patří výborná integrace s virtuální infrastrukturou založenou na platformě VMware. Dalším důvodem je velice kladná a naprosto bezproblémová zkušenost autora s tímto řešením. Software Synology HyperBackup je použit, protože se jedná o nativní řešení dodávané k použitému hardwaru a není zde s jednou výjimkou žádná další alternativa ve scénáři, ve kterém budou Synology zařízení použita v řešené firmě. Ona jedna výjimka navíc nabízí odlišné zaměření na zálohování než zmíněný HyperBackup.

3.4.1 Veeam Backup & Replication 11

Software pro zálohu a replikaci dat od společnosti Veeam je tradičním a jedním z nejrozšířenějších řešení na světě. Jejich řešení využívá přes 400 000 organizací ve více než 180 zemích. K dispozici je komplexní řešení ochrany dat a obnovy po havárii. Zálohovat lze virtuální, fyzické i cloudové stroje na úrovni bitové kopie a poté je obnovovat. Použité technologie optimalizují přenos dat a spotřebu zdrojů, což pomáhá minimalizovat náklady na úložiště a dobu obnovy v případě havárie. Další oceňovanou funkcionalitou je centralizovaná konzole pro správu operací zálohování, replikace a obnovy na všech podporovaných platformách. Navíc nabízí pro menší firmy licenci „Community Edition“, díky které lze zálohovat až 10 instancí zcela zdarma. Hlavní funkce jsou následující.

Zálohování

Tvorba záloh na úrovni bitové kopie virtuálních, fyzických a cloudových strojů.

Obnova

Obnovení ze záloh na původní nebo nové místo. Možnosti jako Instant Recovery, obnova na úrovni bitové kopie, na úrovni souborů, na úrovni aplikací atd.

Replikace

Vytvoření přesné kopie virtuálního počítače a udržování kopie v synchronizaci s původním virtuální počítačem.

Continuous Data Protection

Technologie replikace, která chrání kriticky důležité virtuální počítače a pomáhá dosáhnout RPO v řádech sekund.

Backup Copy

Kopírování záložních souborů do sekundárního úložiště.

Podpora úložných systémů

Zálohování a obnova virtuálních počítačů pomocí možností nativních snapshotů vytvořených v úložištích.

Podpora páskových zařízení

Ukládání kopií záloh na pásková zařízení.

Ověření obnovení

Testování záloh a replik virtuálních počítačů před ostrým obnovením.

Objekty k zálohování

Pomocí Veeam Backup & Replication lze zálohovat a obnovovat následující objekty:

- **Virtuální stroje:**
  - VMware vSpere
  - Virtuální počítače Microsoft Hyper-V
  - Nutanix AHV VM

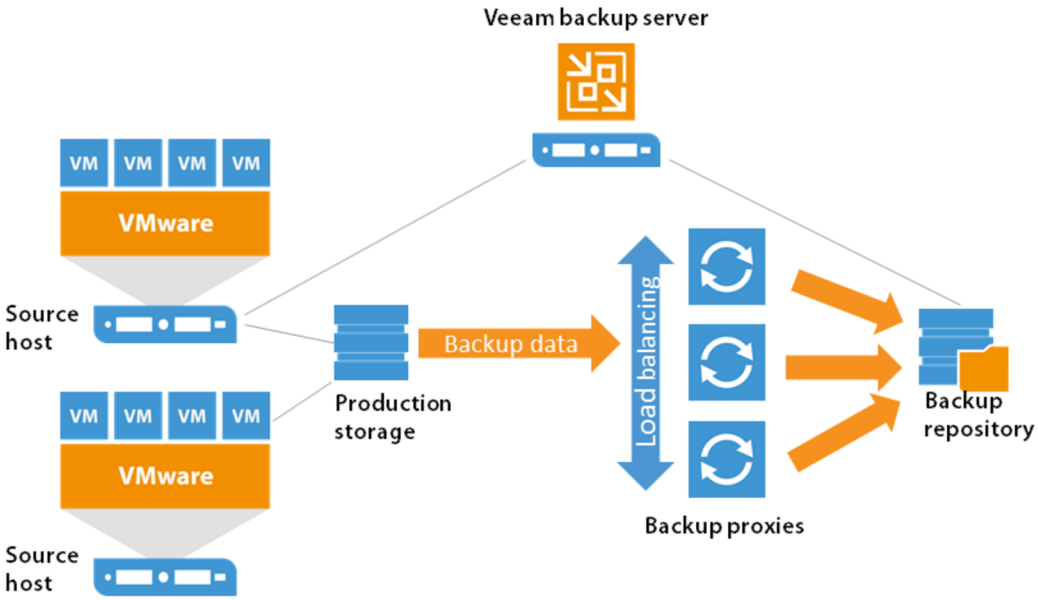
- **Cloudové virtuální počítače:**
  - Instance AWS EC2
  - Virtuální počítače Microsoft Azure
  - Google Cloud VM

- **Zálohování souborů NAS**
- **Fyzické stroje**
  - K zálohování fyzických počítačů s operačními systémy Windows, Linux nebo macOS se používají tzv. agenti, kteří jsou nainstalovaní na každém počítači. Veeam Back & Replication poté slouží jako centralizované řídící centrum pro nasazení a správu Veeam Agent.

Také jsou k dispozici nástroje jako *Veeam ONE*, který umožňuje monitorování v reálném čase a správu virtuální infrastruktury. Dále *Veeam Backup Enterprise Manager*, který slouží ke správě více instancí *Veeam Back & Replication* z jediné webové konzole. Neposledním nástrojem ve stáji Veeam je *Veeam Availability Orchestrator*. Jedná se o řešení, které řídí procesy obnovy po havárii v prostředích VMware vSphere, poskytuje obnovu jedním kliknutím, funkce pro dokumentaci nebo testování [29].

Architektura v prostředí VMware

Veeam Backup & Replication je modulární řešení, které umožňuje vybudovat škálovatelnou zálohovací infrastrukturu pro prostředí různých velikostí a konfigurací. Zálohovací infrastruktura Veeam se skládá ze sady komponent. Základní schéma nasazení je vidět na obrázku.



Obrázek 3.23 VMware + Veeam architektura [30]

Na obrázku jsou fyzické hosty (Source host), na kterých je nainstalován hypervisor VMware ESXi. Jedná se o samostatný operační systém pro spouštění virtuálních počítačů. K těmto hostům je připojen datastore. Přes záložní proxy se provádějí zálohy do záložního úložiště. Nad tím vším stojí Veeam backup server jako hlavní řídící prvek.

Veeam backup server je založený na Windows serveru a to na fyzickém nebo virtuálním stroji. Na tomto serveru je nainstalován Veeam Backup & Replication. Jedná se o základní komponentu v infrastruktuře zálohování, která plní roli konfiguračního a řídícího centra. Záložní server tak provádí všechny typy administrativních činností, čímž jsou:

- Koordinace úloh zálohování, replikace a obnovy.
- Řízení plánování úloh.
- Přidělování zdrojů.
- Využívá se k nastavení a správě komponent infrastruktury.

Kromě toho je server použit jako výchozí zálohovací proxy a také jako úložiště samotných záloh.

Veeam Backup & Replication konzole je samostatná komponenta na straně klienta, která poskytuje přístup k záložnímu serveru. Standardně se instaluje lokálně na záložní server, ale není to nutné. Konzoli lze nainstalovat i v samostatném režimu a k Veeam Backup & Replication se přistupuje vzdáleně.

Backup Proxy je komponenta postavená mezi ESXi serverem a ostatními komponenty v infrastruktuře. Zatímco záložní server spravuje jednotlivé úlohy, proxy zpracovává úlohy a zajišťuje provoz. Mezi hlavní úlohy proxy patří:

- Načítání dat virtuálního stroje z produkčního úložiště.
- Komprese dat.
- Deduplikace.
- Šifrování.
- Odesílání dat do úložiště záloh nebo jiného záložního proxy (V případě provádění replikace.).

Ve výchozím stavu je role proxy přiřazena samotnému zálohovacímu serveru, nicméně takové řešení stačí pouze pro malé instance s nízkou zátěží. U větších instancí se doporučuje použít samostatně vyhrazené proxy.

Backup repository je úložné místo, kde lze uchovávat záložní soubory, kopie virtuálních počítačů a metadata pro replikované virtuální stroje. Toto úložní místo mohou zastupovat následující systémy [30]:

- Microsoft Windows Server
- Linux Server
- CIFS (SMB) share
- Úložiště s rotačními disky
- Dell EMC Data Domain
- ExaGrid
- HPE StoreOnce

3.4.2 Acronis

Globální společnost založená v Singapuru v roce 2003. Ve svém portfoliu nabízí řešení v oblasti bezpečnosti, dostupnosti, soukromí, autenticity a zabezpečení. Jedním z hlavních produktů je Acronis Cyber Backup (dříve Acronis True Image). Toto řešení poskytuje zálohování pro:

- Fyzické servery
    - Windows, Linux
- Virtuální stroje
    - WMware, Hyper-V, KVM, RHEV, Oracle, Citrix
- Cloudové stroje
    - Azure, Amazon EC2
- Celé stroje, disky, svazky, soubory, adresáře
- Databáze Exchange, SQL, Active Directory, SharePoint, Oracle
- Officce 365, G Suit.
- Mobilní zařízení Android a iOS
- Koncové počítače s Windows a macOS.

Pro ukládání záloh lze využít:

- Místní disky
    - SATA, SCSI, IDE, RAID
- Síťová úložiště
    - SMB, NFS, iSCSI, FC (včetně NAS)
- Vyměnitelná média
    - ZIP, Rev, RDX, atd.
- Externí HDD a SSD
- Páskové jednotky a knihovny
- Acronis Cloud Storage

V tomto scénáři bude ale Acronis využit především pro jeho spolehlivé klonování a zálohování celých pevných disků, respektive fyzických strojů. K tomu slouží Acronis Disk Director. Klonování pevného disku znamená vytvoření přesné repliky disku. Naklonovaný disk bude obsahovat identická data od operačního systému, přes aplikace, po systémová

nastavení jako původní disk. Rozdíl od prostého překopírování je, že z naklonovaného disku lze korektně spouštět operační systém. To v případě obyčejného překopírování nelze. Tento postup se hodí v několika specifických situacích [31].

První situací, kdy je klonování disků vhodné, je záloha strojů ve výrobě. Typicky se jedná o počítače s buďto speciálním operačním systémem nebo s nějakou upravenou verzí operačního systému Windows. V mnoha případech ještě na verzích 2000, případně XP. Pokud u takového stroje dojde k selhání pevného disku, nebo k chybě softwaru v důsledku chybného nastavení, je následná oprava velice komplikovaná, zdlouhavá a nákladná. Jako nejideálnější forma ochrany se tak ukázalo klonování originálního disku. Vytvoření přesné kopie pevného disku, který obsahuje daný stroj. Při následném problému stačí jednoduše vyměnit poškozený disk za klon a řídící počítač, respektive celý stroj poté v naprosté většině případů bezproblémově funguje.

Druhým případem vhodným ke klonování je například výměna pevného disku za nový SSD disk. Pokud z nějakého důvodu nelze přistoupit k čisté instalaci operačního systému, lze provést klonování z původního HDD na nový SSD. Přitom nezáleží na tom, jestli jsou disky stejné velikosti nebo značky. Dokonce lze naklonovaný disk použít v úplně jiném počítači s odlišnou základní deskou, procesorem atd. To umožňuje Acronis Universal Restore, který nahradí ovladače pro původní hardware novými ovladači.

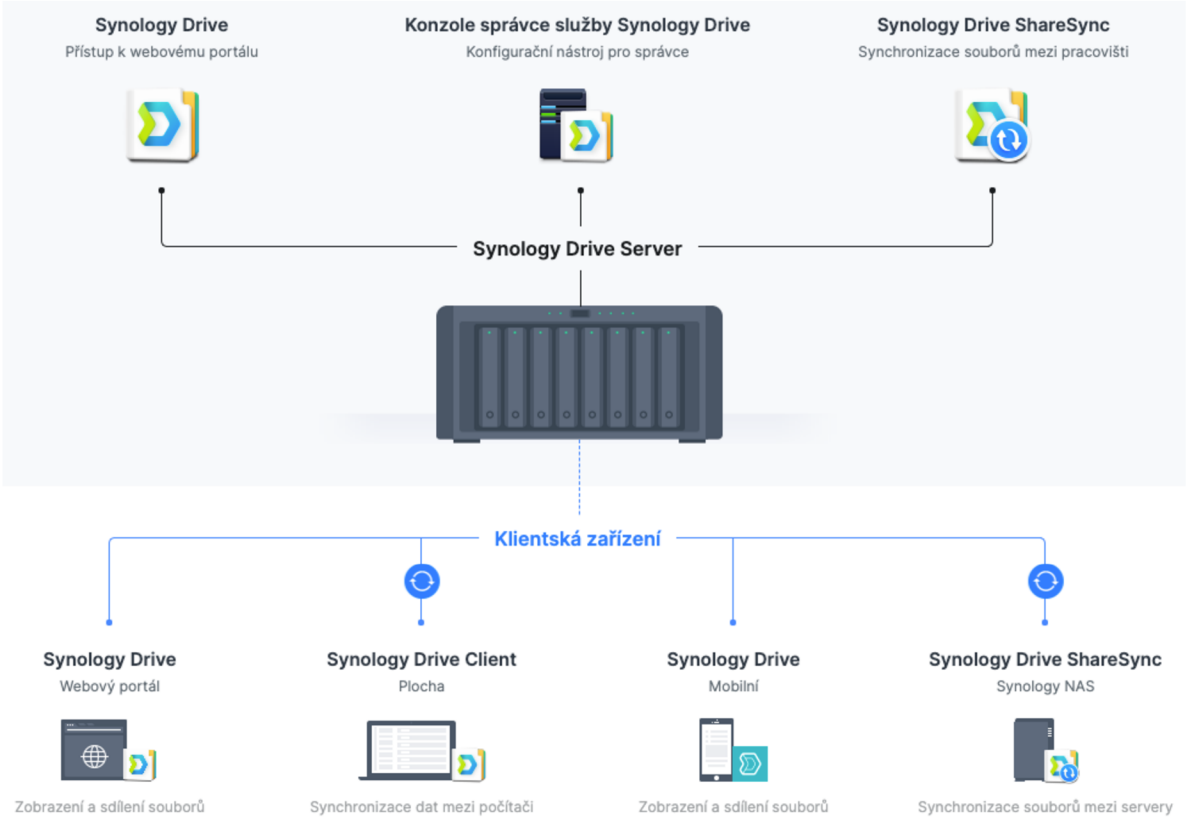
3.4.3 Synology Hyper Backup

Synology nabízí software Synology Hyper Backup. Toto zálohovací řešení je vybráno z důvodu využívání NAS úložišť od společnosti Synology. Jedná se o integrovaný zálohovací software určený pro tento hardware. Hyper Backup podporuje zálohování dat ze Synology NAS do:

- Synology
    - Lokální složky nebo USB zařízení
    - Vzdálené Synology NAS zařízení
    - Synology C2 úložiště
- File Server
    - Rsync
    - WebDAV
    - OpenStack Swift
- Cloud
    - Dropbox
    - Google Drive
    - MS Azure
    - HiDrive, JD Cloud, Rackspace, SFR NAS Backup, hubiC, hicloud S3, S3 Storage

Dále nabízí funkce jako deduplikace, šifrování, plánování, více verzí záloh, rotace záloh, přírůstkové zálohy, komprese nebo kontrolu integrity dat. Hyper Backup je řešení pro zálohu dat ze samotného NASu. Například v případě, kdy je vytvořena sdílená složka, kterou si uživatelé mapují jako normální síťový disk. Nejedná se však o zálohu dat z koncových stanic uživatelů, kteří budou NAS využívat.

K tomuto účelu slouží program Synology Drive. Jde o službu, která zajišťuje synchronizaci a zálohování dat mezi koncovými zařízeními a úložištěm NAS. Zároveň vytváří privátní cloud se stoprocentním vlastnictvím uložených dat. Služba je přístupná z Windows, macOS i mobilních zařízení. Dále nabízí například sdílení souborů i s lidmi, kteří žádné služby Synology nepoužívají. Bezpečnosti lze dosáhnout podrobným nastavením oprávnění a šifrováním SSL [32]. Princip celého použití je vidět na obrázku.



Obrázek 3.24 Synology Drive architektura [32]

Uživatel má na svém počítači nainstalován Synology Drive Client, kam si ukládá svoje soubory. V tomto klientu se poté nastaví synchronizace a zálohování na NAS. Svoje soubory má následně dostupné odkudkoliv jako v případě běžných cloudových řešení typu Google Drive.

3.4.4 Alternativy zálohovacích softwarů

Altaro VM Backup

Jedná se o pozitivně hodnocené zálohovací řešení pro virtuální infrastrukturu na bázi Hyper-V nebo VMware. Vyznačuje se relativně snadnou a intuitivní instalací a následnou konfigurací. Důležitý je také důraz na uživatelskou přívětivost. Samozřejmě nabízí všechny potřebné funkcionality jako je podpora pravidla 3-2-1, šifrování, komprese, instantní obnova, plánování, nastavení retence, full backup, replikace nebo automatické ověření vytvořené zálohy.

NAKIVO VM Backup and Replication

Velice rozšířené zálohovací řešení, které se silně zaměřuje na platformu VMware. Vytváří konzistentní „image-based“ zálohy virtuálních strojů. To znamená, že zálohuje celý virtuální stroj včetně celého disku, všech konfiguračních souborů a data aplikací jako jsou Active Directory nebo SQL Server. Dále nabízí inkrementální zálohu, instantní granulární obnovu souborů i celého virtuálního stroje, ověření vytvořené zálohy, kompresi, šifrování nebo monitoring pro VMware vSphere infrastruktury.

Další alternativy:

- Vembu BDR Suite
- Unitrends Backup
- Iperius Backup
- N-Able Backup

4 Vlastní práce

Tato část práce se zabývá vlastním praktickým zpracováním zálohovací strategie a její implementací ve firmě, kde je autor zaměstnán. V prvním kroku bude provedena analýza firemních zdrojů. To znamená identifikace dat, virtuálních i fyzických strojů, datových úložišť, koncových stanic, strojů ve výrobě a síťových prvků. Zde bude také vybráno nové řešení NAS úložišť. Následně bude představen aktuální stav zálohování a poté požadovaný stav. V další časti dojde na samotnou realizaci, tak aby se dosáhlo požadovaného stavu. V posledním kroku dojde k testování rychlosti zálohování i obnovy a funkčnosti záloh po úspěšné obnově.

4.1 Charakteristika společnosti

Název společnosti, ve které bude provedena praktická část práce, nebude po domluvě uveden. Nicméně se jedná o středně velkou firmu zabývající se výrobou produktů z plastu, dřeva a hliníku, od návrhu po jejich kompletní zhotovení. Na trhu je firma téměř 30 let a na konci roku 2021 měla přes 220 zaměstnanců. Firma působí na trzích v České republice, Rakousku, Německu a Itálii.

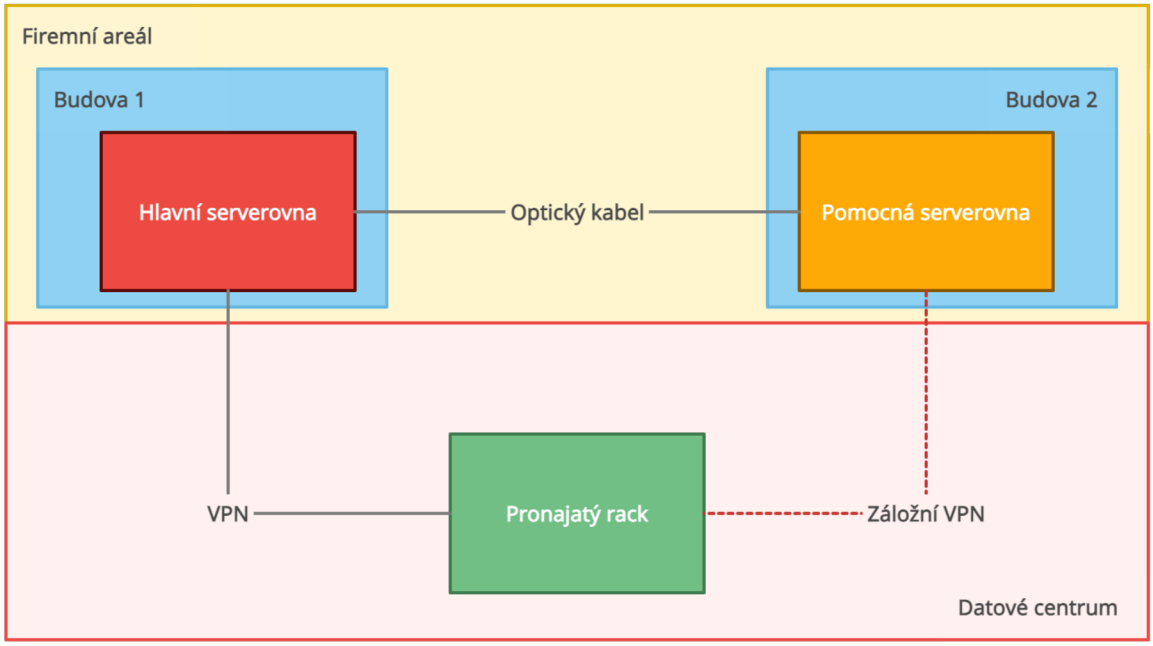
4.2 Analýza zdrojů

V této kapitole budou identifikovány firemní zdroje z pohledu dat, hardwaru od mobilních zařízení, přes servery až po stroje ve výrobě. Jednotlivé stroje budou charakterizovány z pohledu klíčových vlastností. Také bude uveden aktuální stav zálohování ve firmě a následně požadovaný stav po úpravách. Samotný návrh strategie a způsob realizace zálohování těchto zjištěných zdrojů bude obsažen v následující kapitole.

Stejně jako v případě výběru zálohovacího softwaru ani v případě použitého hardwaru není příliš prostor pro použití alternativ. Jak již bylo řečeno, celá infrastruktura byla v nedávné době výrazně modernizována, a tak ani není příliš potřeba dokupovat extra další hardware. Toto tvrzení má dvě výjimky. Jedná se o Veeam Backup Server v lokální infrastruktuře a dvě nová NAS zařízení. V případě Veeam serveru je již hardware nedostatečný a bude tak postaven nový server. Zbytek hardwaru bude využit tak jak je.

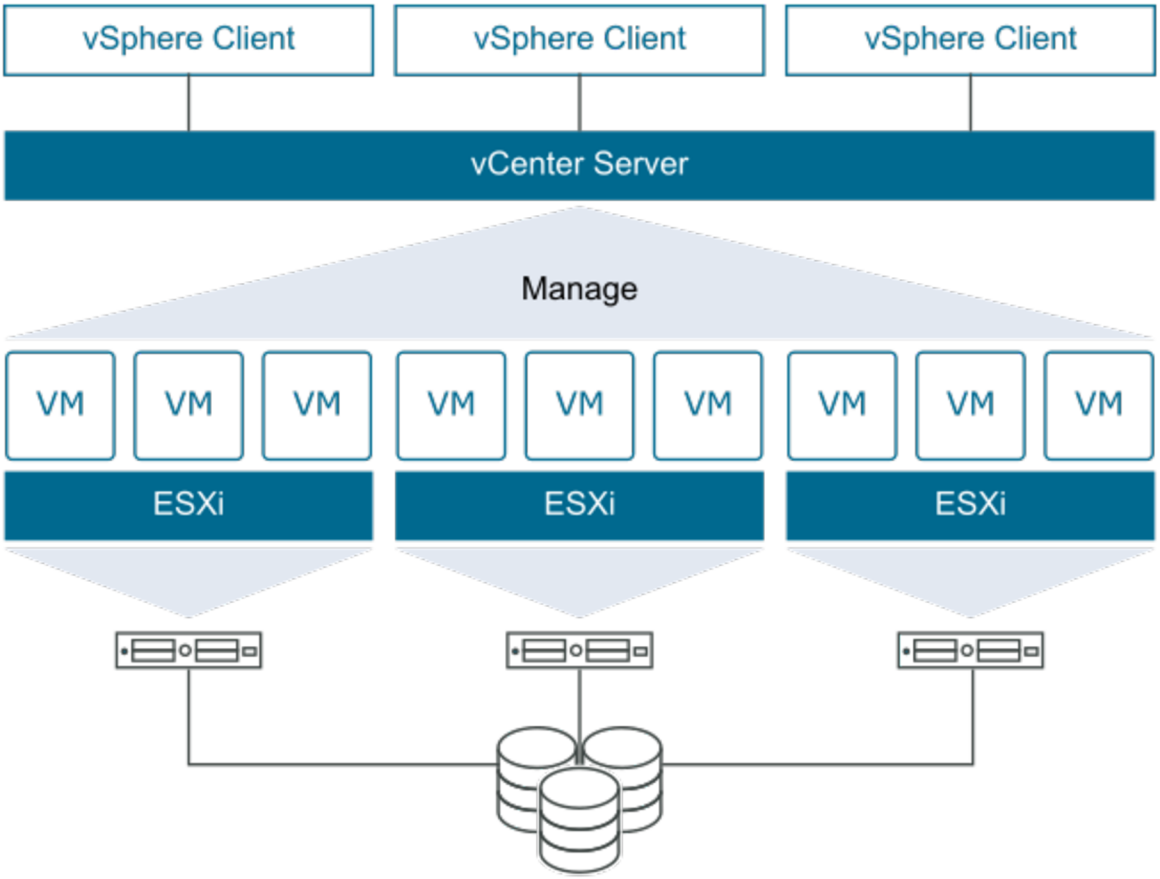
4.2.1 Infrastruktura

V této kapitole bude nastíněna firemní infrastruktura jak z pohledu fyzického rozmístění, tak z pohledu používaných technologií. Celá infrastruktura je rozdělena mezi dvě lokace. První je zázemí firmy a druhá je datové centrum, ve kterém si firma pronajímá rack, kde má svůj hardware. Obě lokace jsou propojeny pomocí VPN. To umožňuje využívat obě místa pro zálohování či replikaci virtuálních strojů. Lokální infrastruktura se poté ještě rozkládá mezi dvě oddělné budovy v areálu firmy, kde jsou serverové místnosti a které jsou propojeny optickou sítí. To nabízí efektivní možnosti zálohování. Následující analýza v dalších kapitolách tak bude rozdělena na lokální a vzdálenou část. Na obrázku níže je vidět schéma firemní infrastruktury.



Obrázek 4.1 Architektura firemní sítě – vlastní zpracování

Celá virtuální infrastruktura je postavena na řešení VMware. To znamená fyzické servery, na kterých je nainstalován hypervizor ESXi. Další vrstvou jsou již konkrétní virtuální stroje. Nad tím stojí vCenter Server, díky kterému lze spravovat větší počet připojených fyzických hostů. Poslední částí je vSphere Client. Jedná se o webovou aplikaci, která nabízí administraci s přístupem ke klíčovým funkcím bez nutnosti přistupovat k vShpere serveru napřímo. Přehledné schéma architektury je na následujícím obrázku.



Obrázek 4.2 VMware vSphere [33]

4.2.2 Servery

Virtuální stroje

Virtuální infrastruktura, postavená na té fyzické, je klíčovou částí pro chod firmy. Její zálohování je patrně ta nejdůležitější část celé zálohovací strategie. V následujících tabulkách jsou vidět virtuální stroje, jejich operační systém, zařazení, počet jader, paměť, velikost a priorita.

OS	Zařazení	CPU	RAM (GB)	Velikost (GB)	Priorita	Počet
Ubuntu 64-bit	Servis	2	4	32	Střední	1
MS Server 2019	Produkce	2	8	40	Vysoká	2
MS Server 2012	Produkce	4	16	120	Vysoká	1
Windows 10	Servis	2	4	60	Nízká	2
Windows 10	Produkce	2	8	100	Vysoká	1
Ubuntu 64-bit	Produkce	4	8	2000	Vysoká	1
MS Server 2016	Produkce	16	96	1024	Vysoká	1
Windows 10	Produkce	4	16	100	Vysoká	1
Windows 10	Produkce	4	8	80	Střední	1
MS Server 2012	Produkce	2	6	50	Vysoká	1
Windows 10	Produkce	2	4	60	Nízká	2
Windows 7	Servis	2	4	100	Střední	2
Windows 7	Produkce	2	4	60	Střední	2

Tabulka 4.1 Lokální infrastruktura – vlastní zpracování

OS	Zařazení	CPU	RAM (GB)	Velikost (GB)	Priorita	Počet
MS Server 2019	Produkce	2	8	40	Vysoká	1
Ubuntu 64-bit	Produkce	4	16	30	Vysoká	1
Ubuntu 64-bit	Produkce	4	8	90	Střední	1
MS Server 2019	Produkce	8	32	512	Vysoká	1
Windows 10	Servis	4	8	60	Střední	2
Ubuntu 64-bit	Servis	2	2	32	Střední	2

Tabulka 4.2 Datové centrum - vlastní zpracování

Fyzické stroje

Výše vypsané virtuální stroje stojí na fyzické infrastruktuře. Lokálním srdcem je serverový cluster složený ze tří serverů Dell PowerEdge R640 s připojenými datovými úložišti. K tomu se přidávají další servery, které slouží, nebo budou sloužit například jako Veeam backup server nebo jako fyzické Veeam backup proxy.

Lokální infrastruktura

Dell PowerEdge R640 3x

Jedná se o dvousocketovou platformu s procesory Intel Xeon Silver. Ideální pro vysokokapacitní softwarově definované úložiště nebo virtualizaci. Vhodné také pro použití ve výpočetním clusteru. Jak již bylo řečeno, tento cluster slouží jako zdroj virtuální infrastruktury. Parametry konkrétního serveru jsou vidět v tabulce.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	2	32	255	12x 2TB SSD	2x 10 GbE + 6x 1 GbE	1U

Tabulka 4.3 Dell PowerEdge R640 - vlastní zpracování

Dell PowerEdge R630

Oproti verzi R640 má tato řada několik omezení. R640 má dvě pozice na disky navíc v zadní části šasi. R640 má také větší možnosti v použití NVMe pamětí. Všech 10 předních pozic lze využít pro NVMe. U R630 lze bez dalších úprav využít pouze 4 pozice. R630 má také slabší procesor. Využit bude tento server pro nasazení virtuálních strojů, které nejsou kritické a nevyžadují neustálou dostupnost. Například pro testovací stroje nebo Veeam backup proxy. V krajní situaci lze na tento server přemigrovat nejdůležitější virtuální stroje z produkčního clusteru.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	2	32	96	10x 2TB SSD	2x 10 GbE + 6x 1 GbE	1U

Tabulka 4.4 Dell PowerEdge R630 - vlastní zpracování

Dell PowerEdge R430

Cenově dostupný dvou socketový server vhodný pro malé až střední firmy. Vybaven je procesory Intel Xeon E52600 v3 a dvanácti sloty DDR4 DIMM. Šasi nabízí možnost osazení čtyřmi 3,5" pevnými disky nebo až deseti 2,5" disky. V zadních pozicích má dva sloty PCIe Gen3. Server také disponuje dvěma zdroji. Primární použití tohoto serveru bude pro nasazení VMware vCenter Server.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	2	64	32	1x 500 GB SSD + 8x 2TB HDD	2x 10GbE + 4x 1 GbE	1U

Tabulka 4.5 Dell PowerEdge R430 - vlastní zpracování

HP ProLiant DL20

Kompaktní 1U server je cenově dostupný základní server pro menší firmy. Jedná se již o starší model, který byl uveden v roce 2015. Je osazen procesory Intel Xeon E3-1230 v5 a operační pamětí DDR4. Nabízí prostor pro dva 3,5" disky. Tento server je v současnosti využíván jako Veeam Backup Server.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	1	8	32	2x500 GB	2x 1 GbE	1U

Tabulka 4.6 HP ProLiant DL20 - vlastní zpracování

Datové centrum

Dell PowerEdge R610 2x

Podobně jako v případě R640 v lokální infrastruktuře tvoří R610 ve dvojici cluster v datovém centru. Na tomto clusteru opět stojí virtuální infrastruktura. Převážně se jedná o virtuální stroje, které jsou neustále dostupné pro firemní zákazníky, kteří si zde sami konstruují finální výrobky. Jelikož se jedná o službu zákazníkům, jsou tyto stroje umístěné právě v datovém centru, které zaručuje téměř stoprocentní dostupnost z pohledu sítě nebo elektrické energie. Ovšem za ostatní věci, jako je právě zálohování, již datové centrum nezodpovídá. Také zde se jedná o dvou socketový server s procesorem Intel Xeon. Servery jsou v následující konfiguraci.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	2	24	96	4x 2TB SSD	4x 1 GbE	1U

Tabulka 4.7 Dell PowerEdge R610 - vlastní zpracování

HP ProLiant DL580

Tento 4U vysoký server nabízí konfiguraci až ve čtyř socketovém provedení za použití procesorů Intel Xeon E7-4800/8800 v3 s podporou pro 18 jader nebo v4 s podporou pro 24 jader a 6 TB operační paměti. Díky své velikosti má také široké možnosti konfigurace úložiště. Například 10x SFF SATA SSD o velikosti maximálně 38,4 TB, nebo pět jednotek NVMe pro až 8 TB superrychlého úložiště. Dále nabízí v provedení v4 až 9 PCIe rozhraní nebo 25Gbe Ethernet s rozšířením více karet až na 100 GbE. Server je také použit pro virtualizaci pomocí řešení VMware. Konkrétní konfigurace serveru je v tabulce.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Vmware ESXi	2	72	768	5x 2TB SSD + 5x 2TB HDD	2x 40GbE + 8x 1 GbE	4U

Tabulka 4.8 HP ProLiant DL580 - vlastní zpracování

Veeam Backup Server

Veeam Backup Server je provozován na stroji, který byl postaven svépomocí z vybraných komponent. K vzhledem menšímu počtu cílů k zálohování je zde umístěn relativně slabší stroj. Ovšem hardwarové a softwarové požadavky udávající dokumentace pro Veeam Backup Server tento stroj splňuje bez problému. I přes tento fakt je ale v plánu v blízké budoucnosti upgrade tohoto serveru, protože například procesor Intel Core i3 generace Skylake je již relativně zastaralý. Server pak obsahuje SSD disk o velikosti 250 GB na běh systému a 8x 2TB HDD pro ukládání záloh.

OS	Socket	CPU	RAM (GB)	Úložiště	NIC	Rozměr
Windows 10 Pro	1	4	16	1x 250 GB SSD + 8x 2TB HDD	3x 1 GbE	4U

Tabulka 4.9 Veeam Backup Server - vlastní zpracování

4.2.3 Datová úložiště

Lokální infrastruktura

Dell Storage SCv3020

Nové modely SC v řadě 3000 byly vylepšeny o funkce dříve dostupné pouze u modelů vyšší třídy. Úložiště dostalo nové šestijádrové procesory Intel Xeon E5-2603v, dvakrát větší paměť a třikrát větší šířku pásma. Šasi nabízí prostor pro třicet 2,5" SSD disků nebo HDD disků s 15 000/ 10 000/ 7 200 otáčkami za minutu. Dostupné konfigurace pole jsou SSD, HDD nebo kombinace obojího. Další rozšiřující skříně SCv300/320/360 pak umožnují škálovat kapacitu až na 222 disků nebo 1 PB na jedno pole. Taková pole pak lze dále seskupovat do federovaných clusterů. Úložiště také disponuje duálními řadiči, 32 GB systémové paměti a síťovým připojením dvakrát 12Gb SAS, 10Gb iSCSI, 16Gb FC a redundantními zdroji 1485W. To vše ve skříni o velikosti 3U. Úložiště je připojeno k produkčnímu serverovému clusteru.

Synology RS4017xs+

Jedná se o větší a výkonnější diskové pole. Hlavní vlastnosti jsou vysoká spolehlivost, extrémní výkon a úložiště, poskytující akceleraci virtualizace a škálovatelnost pomocí expanzních jednotek až na 40 pevných disků. V základu se jedná o zařízení velikosti 3U a nabízí šestnáct slotů pro pevné disky. Vybaveno je procesorem Intel Xeon D-1541, pamětí DDR4 ECC UDIMM o velikosti 8 GB s možností rozšíření až na 64 GB. Obsahuje dva porty 10GBase-T a čtyři Gigabit porty. K tomu jsou k dispozici další dva sloty PCIe 3.0. Při použití šesti 10Gbe portů s link agregací nabízí sekvenční propustnost s rychlostí přes 4900 MB/s a 690 000 IOPS.

Zařízení je osazeno šestnácti disky, každým o velikosti 8 TB a je rozděleno na tři volumy. Toto úložiště již není přímo dostupné pro běžné uživatele v produkční síti. Volume 1 bude obsahovat úložiště pro Veeam server. Velikost je 35 TB. Volume 2 bude sloužit jako cílové místo pro zálohy různých síťových zařízení jako jsou firewally nebo routery. Dále pro zálohy nových dvou Synology. Velikost je také 35 TB. Volume 3 slouží jako úložiště připojené přes iSCSI, kde jsou uložené některé méně kritické virtuální stroje. Velikost je 14 TB a je plný z 50 %. Pro přehlednost bude tento NAS dále nazýván jako „Synology Backups“.

Dell PowerVault 124T

Jedná se o automatický zavaděč pásek o velikosti 2U pro menší a středně velké firmy. Podporuje až 24 TB LTO-5 kapacity s přenosovými rychlostmi 504 GB/h. Obsahuje 16 zásobníků pro kazety. Podporuje páskové jednotky LTO-3, LTO-4 a LTO-5. Dále nabízí funkce vzdálené správy pomocí webové aplikace, čištění, automatické zálohování a samozřejmě automatickou výměnu kazet, která snižuje riziko lidské chyby při ruční výměně. To vše při bezproblémové integraci do operačních systémů Windows nebo Linux a především pro integraci ve Veeam Backup & Replication.

Datové centrum

Dell PowerVault MD1200

Úložné pole o velikosti 2U, které nabízí 12 slotů pro SSD nebo HDD ve velikostech 2,5 nebo 3,5 palců. Maximální velikost úložného prostoru je 24 TB s podporou RAID 0, 1, 5, 6, 10, 50 a 60. Kapacitu lze navýšit propojením až osmi MD1200. Propustnost pole je 6Gb/s. Úložiště je připojeno v datovém centru jak k serverovému clusteru, tak k HP580.

Dell PowerVault MD3200

Podobné diskové pole jako výše zmiňované MD1200. Vyznačuje se výborným poměrem cena/výkon. Hlavním rozdílem jsou možnosti připojení a architektura úložiště. Řada MD3200 funguje jako SAN a nabízí osmkrát 6Gb/s k poskytování sdíleného úložiště a virtualizovaného prostřední. Zatímco MD1200 nabízí připojení pouze dvakrát 6Gb/s a jedná se o DAS úložiště, které lze připojit přímo k serveru Dell PowerEdge, nebo jako rozšiřovací jednotku k úložišti MD3200.

Dell PowerVault 124T

Stejně jako lokální infrastruktura je datové centrum vybaveno automatickým zavaděčem pásek.

4.2.4 Výběr nových NAS zařízení

V rámci této práce je nutné vybrat vhodné řešení v oblasti firemního NAS. Cílem je umístit jedno zařízení do datového centra a dvě do firemního zázemí. To umožní vytvořit poměrně kvalitní zálohovací strategii. K dispozici jsou především zařízení od výrobců

Synology a QNAP. Synology je lídrem na trhu, nicméně QNAP také nabízí zajímavé produkty, navíc je obecně o něco levnější než Synology. V případě této implementace existují teoreticky tři způsoby, jak se k výběru postavit. Jelikož firma již disponuje jedním Synology zařízením, stačí k dosažení cíle s třemi NAS, pořídit další dvě zařízení od Synology. Druhý způsob je kompletní přechod na technologie QNAP. Poslední možností je kombinace stávajícího Synology a dvou QNAP zařízení. Tato možnost byla zavržena, jelikož pro použití všech tří zařízení ke vzájemné záloze a pro plné využití poskytovaných funkcí, je vhodnější nasazení jednotné platformy.

V případě naprosto nového řešení, by QNAP jistě stál za úvahu. Nicméně v tomto případě, vzhledem k jednomu již používanému Synology, bylo rozhodnuto o doplnění dvěma dalšími úložišti od tohoto výrobce. I přes o něco levnější produkty od QNAP, by tři zařízení nebyla levnější než dvě nová zařízení od Synology. Důvodem je, že stávající Synology zařízení je datové úložiště z vyšší modelové řady s pořizovací cenou okolo 150 000 Kč bez daně. Dalším a podstatným důvodem je relativně velké množství bezpečnostních incidentů u zařízení QNAP. Speciálně v roce 2021 došlo k celé řadě masivních a úspěšných útoků na tato zařízení. Příkladem mohou být opakované problémy s ransomwary *DeadBolt*, *Qlocker* nebo *AgeLocker*. Problémy v takovém rozsahu se Synology zatím vyhýbají. Posledním důvodem je také bezproblémová zkušenost s dosud využívaným Synology RS4017xs+.

Synology RS820RP+ a Synology RX418

Jako nové zařízení bylo vybráno Synology RS820RP+ spolu s expanzní jednotkou RX418. Jedná se o poměrně výkonný NAS v provedení určeném do racku o velikosti 1U, vhodný pro středně velké firmy. Může pojmout čtyři 3,5" pevné disky. Verze RP+ má oproti standardní verzi RS820+ navíc druhý záložní systém napájení. Zařízení je vybaveno čtyřjádrovým procesorem Intel Atom C3538 s frekvencí 2,1 GHz. Operační paměť je v základu 2 GB, ale je zde prostor pro navýšení až na 18 GB. Díky PCIe 3.0 x8 slotu lze využít například m.2 kartu, nebo 10GbE kartu. Samozřejmostí je technologie hot-swap. Kapacita zařízení je při discích o velikosti 4TB a za použití RAID6 7.0 TB. K tomu je přidána expanzní jednotka RX418, která navyšuje místo pro další 4 disky 2,5"/3,5" SATA III HDD/SSD. Celková kapacita je tedy 24 TB. NAS lze také připojit do Windows Active Directory, LDAP a důvěryhodných domén.

Hlavním účelem tohoto úložiště bude využití jako privátního cloudu pro firemní zaměstnance. Každý uživatel bude mít vytvořenou tzv. home složku. Tuto složku má uživatel dostupnou ze všech platforem a odkudkoliv prostřednictvím nativních aplikací nebo pomocí webové aplikace. Jako další budou na tomto NASu vytvořeny sdílené složky pro jednotlivá oddělení a *Public* složka pro všechny uživatele. Tyto složky budou přístupné pouze v lokální síti a pro uživatele se budou tvářit jako běžné síťové disky, které mají připojené v počítači. V neposlední řadě bude sloužit toto i následující Synology zařízení pro ukládání záloh. Princip bude podrobně vyložen v dalších kapitolách. Pro přehlednost bude tento NAS dále nazýván jako „Synology Cloud".

Stejné zařízení bude využito v i datovém centru. Dále bude tento NAS nazýván jako „Synology Datastore". Využití najde především pro ukládání záloh virtuálních strojů a jako zdroj osobní složky pro firemní zákazníky.

Cena za jedno zařízení je přibližně 25 000 Kč bez daně. Rozšiřující jednotka pak stojí dalších 12 000 Kč bez daně. Celkově se tak jedná o 74 000 Kč bez daně. V případě použití podobně výkonného zařízení od QNAP, TS-451DeU, by se cena jedno zařízení pohybovala zhruba na hranici 17 000 Kč bez daně. Rozšiřující jednotka TR-004U stojí 7 000 bez daně.

4.2.5 Koncové stanice

Přesný počet koncových uživatelských stanic je relativně nestálý. Zaměstnanci přichází a odchází. V průměru se však jedná zhruba o 220 strojů. Jedná se o desktopové počítače, notebooky, tablety i mobilní telefony. Ale ne všechny zařízení potřebují kompletně zálohovat. Všichni uživatelé budou mít svůj Synology Cloud, kam si budou ukládat svoje data a ty se následně zálohují. Velká část uživatelů pak pracuje, respektive vytváří data kompletně na vzdálené ploše serveru a svůj pracovní počítač tak mají pouze jako prostředníka a pro mailového klienta, případně pro práci na internetu a podobně. Jejich počítače tak neobsahují žádný speciální software ani data, která by bylo nutné vybraným způsobem více zálohovat. Těmto uživatelům stačí zmíněný cloud. V případě selhání jejich stroje tak stačí vyměnit počítač a uživatel je znovu naprosto schopný práce. Jedná se zhruba o 80 % všech firemních uživatelů.

Další skupina uživatelů, respektive stanic již vyžaduje komplexnější přístup k zálohování. Typicky se jedná například o uživatele v oddělení konstrukce nebo v ekonomickém úseku, kteří již mají na svém pracovním počítači nainstalované složité

aplikace i s daty a jejich znovu zprovoznění bez kompletních záloh daných počítačů, by již nebyla otázka několika minut jako v případě předchozí skupiny uživatelů, ale spíše nižších jednotek hodin. Seznamy stanic s požadavky na zálohování jsou vypsány v tabulkách.

Typ zařízení	OS	Počet zařízení celkem	Komplexní záloha	Pouze Synology Cloud
Desktop	Windows	106	27	79
Desktop	macOS	8	4	4
Notebook	Windows	15	3	12
Notebook	macOS	3	0	3

Tabulka 4.10 Koncové stanice - vlastní zpracování

U firemních mobilních zařízení se využívají pro zálohování firemní účty Google nebo Apple a jejich funkcionality. Typicky záloha kontaktů, fotek nebo například souborů.

Typ zařízení	OS	Počet zařízení celkem
Telefon/tablet	Android	64
Telefon/tablet	iOS	13

Tabulka 4.11 Mobilní zařízení - vlastní zpracování

4.2.6 Stroje ve výrobě

Relativně specifické zařízení pro zálohování jsou výrobní stroje. Ovšem jedná se o naprosto kritickou část zálohování. Pokud z nějakého důvodu stroj nefunguje, velice pravděpodobně to znamená zpoždění v produkci zboží, a to je samozřejmě velký problém. Komplikovanosti nahrává fakt, že se využívají stroje poměrně širokého rozpětí data výroby a s tím solidní paleta různých operačních systémů. Jsou zde i 20 let staré stroje, ale i naprosté novinky instalované v roce 2022. Přehled strojů, které je nutné zálohovat je vidět v tabulce.

Stroj	OS	Výroba
Unicontrol	Windows 2000	Dřevo
Biesse	Windows XP	Dřevo
Univar	Windows 2000	Dřevo
Lis	Windows 2000	Dřevo
2x Rapid Alu	Windows XP	Plasty
Thornwest	Windows XP	Plasty
Svářečka	Windows 10	Plasty
Začišťovačka	Windows 10	Plasty
Shuttle	Windows 10	Plasty
Schirmer	Windows 10	Plasty

Tabulka 4.12 Stroje ve výrobě - vlastní zpracování

4.2.7 Síťové prvky

Síťové prvky jsou nedílnou součástí firemní infrastruktury. Některé složitější zařízení jako je firewall je nutné zálohovat. Také je potřeba záloha prvků jako switch a router, pokud obsahují komplikovanější konfiguraci. Podrobný seznam přináší tabulka.

Typ zařízení	Název	Počet
Firewall	Sophos XG210	2
Firewall	Kerio Control NG500	1
Switch	Netgear 16PT 10G	2
Switch	Cisco SG300-52	2
Switch	Zyxel GS1910-24	12
Switch	TP-link TL-SG3428X L2	5
Switch	Cisco SG200-18	6
Switch	Zyxel XGS-4526	2
Switch	Zyxel MGS-3712F	2
	Celkem:	34

Tabulka 4.13 Síťové prvky - vlastní zpracování

4.2.8 Data

Veškerá data, která se uchovávají by měla mít nějaký svůj význam. Ovšem jejich důležitost se bude u různých dat lišit. V případě bezpečnostního incidentu se musí při obnově postupovat od nejdůležitějších dat k těm nejméně důležitým. V tomto případě lze jako typický příklad extrémně důležitých dat označit například produkční virtuální servery, na kterých stojí celá firma od zaměstnanců v kancelářích po výrobu v halách. Naopak méně důležitá data budou například uživatelské adresáře v NAS Synology. Označení dat jako nekritická ale neznamená, že jsou nedůležitá. Nicméně, pokud budou tato nekritická data přiměřený čas nedostupná, nepřinese to žádné vážné následky. Naopak pokud budou nedostupná kritická data, tak jsou negativní následky v podstatě nevyhnutelné. Data budou identifikována na základě obecných kategorií a řazeny dle jejich priority. Vše bude opět rozděleno na lokální infrastrukturu a vzdálené datové centrum.

V samostatné kategorii poté stojí ukládané zálohy. Nejedná se o kritická data ve smyslu zde použitém, ale v případě bezpečnostního incidentu se tato data mohou stát kritickými velice rychle. Nicméně ukládané zálohy jsou až výsledek práce se zdroji, které jsou v této kapitole identifikovány a následně musí být zálohovány. Z tohoto důvodu nejsou tato data zahrnuta do následující seznamů.

Lokální infrastruktura

Kritická data

- Virtuální produkční stroje v serverovém clusteru – nejvyšší možná priorita.
- Sdílené složky(disky) ze Synology pro jednotlivá firemní oddělení.
- Data ze strojů ve výrobě.
- Virtuální produkční stroje na Dell R630.
- Virtuální servisní stroje v serverovém clusteru.
- Virtuální servisní stroje na Dell R630.
- Síťové prvky s komplikovanou konfigurací.
- Vybrané uživatelské počítače s komplikovanou konfigurací.

Nekritická data

- Uživatelské složky „home“ uložené na Synology NAS.
- Uživatelské počítače s jednoduchou konfigurací.
- Mobilní telefony a tablety.

Datové centrum

Jelikož se za pronajatý rack v datovém centru platí poplatky za řadu služeb (velikost racku, příkon, elektroměr, rychlost síťového připojení, servis, podpora, přístup k racku 24/7 a mnoho dalšího), je zde snaha zde držet opravdu pouze nutné zdroje a data. Datové centrum tak neobsahuje žádná nekritická data.

Kritická data

- Virtuální produkční stroje v serverovém clusteru – nejvyšší možná priorita.
- Virtuální produkční stroje na HP580.
- Virtuální servisní stroje v serverovém clusteru.
- Virtuální servisní stroje na HP580.
- Síťové prvky s komplikovanou konfigurací.

4.2.9 Současný a požadovaný stav

Aktuální situace v zálohování byla hlavní motivací pro zpracování této práce. Současný stav je velice nedostatečný. A to především díky podstatným změnám, které v nedávně době ve firmě proběhly. Došlo k relativně rozsáhlému rozšíření a modernizaci infrastruktury. Na to je nutné reagovat i z pohledu zálohování. V této kapitole bude představen stav zálohování dle jednotlivých kategorií, respektive zdrojů, které byly identifikovány v předchozí kapitole.

Lokální infrastruktura

Situace je taková, že se sice většina kritických i nekritických dat zálohuje, ale v tom, co se kam a jak zálohuje není zaveden jasný systém. Pokud by došlo k problému, tak by se nejspíše data dohledala, ale určitě by se při tom ztrácel drahocenný čas. Při instalaci nových strojů do infrastruktury bylo zálohování odsunuto do pozadí a nastavovaly se tak pouze vyloženě nutné zálohy produkčních virtuálních serverů. Zálohy jsou tak různě rozeseté mezi úložišti. Není splněno ani základní zálohovací pravidlo 3-2-1. Také bude potřeba nastavit nová Synology úložiště. Zde bude potřeba kompletně nastavit všechny procesy synchronizace i zálohování od nuly.

Dalším problémem je současný Veeam Backup Server. HP ProLiant DL20, na kterém je Veeam Backup Server, vykazuje problémy, musí se často restartovat a ani jeho výkon již není dostatečný. Proto bude v rámci realizace této práce postaven nový vlastní server, na

který bude nainstalován nový Veeam Backup Server. HP ProLiant bude místo toho využit pouze jako Veeam backup proxy.

Dále se vůbec nepoužívá pásková knihovna, která tak momentálně pouze zabírá místo v serverovně. Úkolem tak bude knihovnu zprovoznit a nastavit na ní zálohování. Dalším nevyužívaným nástrojem, kterým firma disponuje jsou blue-ray média a blue-ray vypalovačka. Ta bude využita pro archivaci dat, která se nemění a nepracuje se s nimi na denní bázi. Například archivy smluv, nebo data z účetního oddělení či správy majetku.

V případě strojů ve výrobě bylo zavedeno schéma zálohování, které není potřeba měnit. Stroje se zálohují ručně, a to přes Vánoční svátky a letní závodní dovolenou, kdy nejsou používány. Důvodem je především fakt, že kvůli jejich velmi nízkému výkonu trvá zálohování jednoho stroje klidně několik hodin. Druhým případem, kdy se zálohuje, je situace, při které se provedou na stroji nějaké úpravy nebo aktualizace. V případě strojů je tak nutné pouze správně určit úložiště, kam se budou zálohy následně ukládat, tak aby byla záloha vždy na více místech.

Zálohování síťových prvků probíhá většinou také ručně a opět pokud se provede nějaká změna v jejich konfiguraci. Výjimkou jsou firewally a vCenter Server, které se zálohují automaticky a pravidelně. V těchto případech tak bude opět především nutné správně nastavit místo pro ukládání těchto záloh.

Co se týče koncových stanic, tak většina vybraných uživatelských počítačů se momentálně nezálohuje. Všichni uživatelé budou mít dostupnou svoji „home“ složku na Synology, ale u vybraných počítačů bude potřeba nastavit komplexnější zálohování prostřednictvím Veeam Agent (dříve Veeam endpoint). Toto řešení je momentálně funkční pouze u 5 zařízení z 34.

Posledním úkolem, který je potřeba vyřešit, je přeorganizovat rozdělení serverů a úložišť mezi obě serverovny v lokální infrastruktuře. Momentálně je zbytečně vše v jedné serverovně. Pokud by došlo k nebezpečné události typu požár nebo vytopení místnosti, firma by přišla v podstatě o celou infrastrukturu.

Datové centrum

V datovém centru je situace podstatně jednodušší. Jednak z důvodu menšího počtu zdrojů, ale také z důvodu, že zde neproběhly žádné podstatné změny v poslední době. Hlavní úkoly tak budou dva. Správně nastavit zálohy a replikace produkčních a servisních virtuálních strojů na základě požadavků vycházejících z faktu, že zde provozované virtuální stroje za poplatky slouží firemním zákazníkům. Druhým úkolem bude konfigurace nového Synology zařízení.

4.3 Realizace

V této kapitole dojde na samotnou realizaci praktické části práce. V prvním kroku budou zvoleny strategie pro samotné zálohování a archivaci. Následně budou uvedeny jednotlivé kroky provedení a samotné konfigurace zálohovacích systémů.

4.3.1 Strategie

Předně je potřeba zvolit vhodnou strategii zálohování. V tomto případě půjde o tři různé strategie. První se bude zabývat zálohováním virtuálních strojů pomocí softwaru Veeam. Druhá bude řešit zálohování Synology zařízení. Třetí a poslední strategie se bude týkat archivace vybraných dat. V některých ohledech však budou všechny tři strategie provázané. První strategie bude rozdělena na lokální infrastrukturu a datové centrum. Opět ale budou obě lokality v určitém smyslu provázány. Nakonec budou zmíněné i plány na zálohování koncových stanic, výrobních strojů a síťových zařízení.

Základní princip lze postavit na obecných zálohovacích pravidlech. Konkrétní provedení v dané firmě je ale potřeba upravit dle požadavků a také dle možností jak z pohledu hardwaru, tak z pohledu financí.

Virtuální infrastruktura

Do této části spadá záloha virtuálních produkčních a servisních strojů. Protože se jedná o klíčovou část zálohování ve firmě, bude strategie založena s jedním dodatkem na pokročilém zálohovacím pravidlu **3-2-1-1-0**. Onen dodatek je, že offline záloha se bude vytvářet pouze u kritických virtuálních strojů a s delším intervalem mezi zálohami. Příklad zálohování jednoho virtuálního stroje tedy bude vypadat následovně.

Lokální infrastruktura:

- Primární záloha bude uložena na lokálním úložišti Veeam Backup Serveru.
- První kopie zálohy bude uložena na Synology Backups.
- Off-site kopie zálohy bude uložena na pásku.
- Offline záloha bude ukládána na externí HDD.
- Záloha bude ověřena pomocí Veeam SureBack.

Datové centrum:

- Primární záloha bude uložena na lokálním úložišti Veeam Backup Serveru.
- První kopie zálohy bude uložena na Synology Datastore.
- Off-site kopie zálohy bude uložena na pásku.
- Offline záloha bude ukládána na externí HDD.
- Záloha bude ověřena pomocí Veeam SureBack.

Nastavení zálohy

Zálohy se budou vykonávat jako inkrementální. Následně se provede synthetic full backup a také active full backup. Rozdíl spočívá v tvorbě zálohy. Active full backup vytváří zálohu z originálních dat virtuálního stroje. Synthetic backup vytváří zálohu z již provedených záloh. Výhodou je nevyužívání zdrojů v síti. Konkrétní časové nastavení se bude lišit u každé vytvořené úlohy, respektive virtuálního stroje a jeho potřeb. Nastavení jednotlivých úloh je vidět v tabulce. Retenční politika bude opět záviset na daném virtuálním stroji. Nastavit lze počet bodů k obnovení nebo počet dní. Inkrementální zálohy se budou provádět denně v určený čas. K tomu se přidají obě full backup zálohy. Konkrétní nastavení záloh pro jednotlivé virtuální stroje je vidět v tabulkách pro lokální infrastrukturu a pro datové centrum.

OS	Zařazení	Čas	Retence	Active full backup	Synthetic full backup
Ubuntu 64-bit	Servis	22:00	14 dnů	Jednou měsíčně	Jednou týdně
MS Server 2019	Produkce	23:00	7 dnů	1 za dva týdny	Jednou týdně
MS Server 2012	Produkce	23:30	7 dnů	1 za dva týdny	Jednou týdně
Windows 10	Servis	0:00	14 dnů	Jednou měsíčně	Jednou týdně
Windows 10	Produkce	0:30	14 dnů	Jednou měsíčně	Jednou týdně
Ubuntu 64-bit	Produkce	1:00	7 dnů	Jednou měsíčně	Jednou týdně
MS Server 2016	Produkce	1:30	21 dnů	Jednou týdně	Jednou týdně
Windows 10	Produkce	2:00	14 dnů	Jednou měsíčně	Jednou týdně
Windows 10	Produkce	2:30	14 dnů	1 za dva týdny	Jednou týdně
MS Server 2012	Produkce	3:00	7 dnů	1 za dva týdny	Jednou týdně
Windows 10	Produkce	3:30	21 dnů	1 za dva týdny	Jednou týdně
Windows 7	Servis	4:00	7 dnů	Jednou měsíčně	Jednou týdně
Windows 7	Produkce	4:30	7 dnů	Jednou měsíčně	Jednou týdně

Tabulka 4.14 Lokální infrastruktura - vlastní zpracování

OS	Zařazení	Čas	Retence	Active full backup	Synthetic full backup
MS Server 2019	Produkce	Každou hodinu	30 dnů	Jednou týdně	Jednou týdně
Ubuntu 64-bit	Produkce	0:00	21 dnů	Jednou týdně	Jednou týdně
Ubuntu 64-bit	Produkce	1:00	7 dnů	1 za dva týdny	Jednou týdně
MS Server 2019	Produkce	2:00	14 dnů	Jednou týdně	Jednou týdně
Windows 10	Servis	2:30	14 dnů	1 za dva týdny	Jednou týdně
Ubuntu 64-bit	Servis	3:00	21 dnů	1 za dva týdny	Jednou týdně

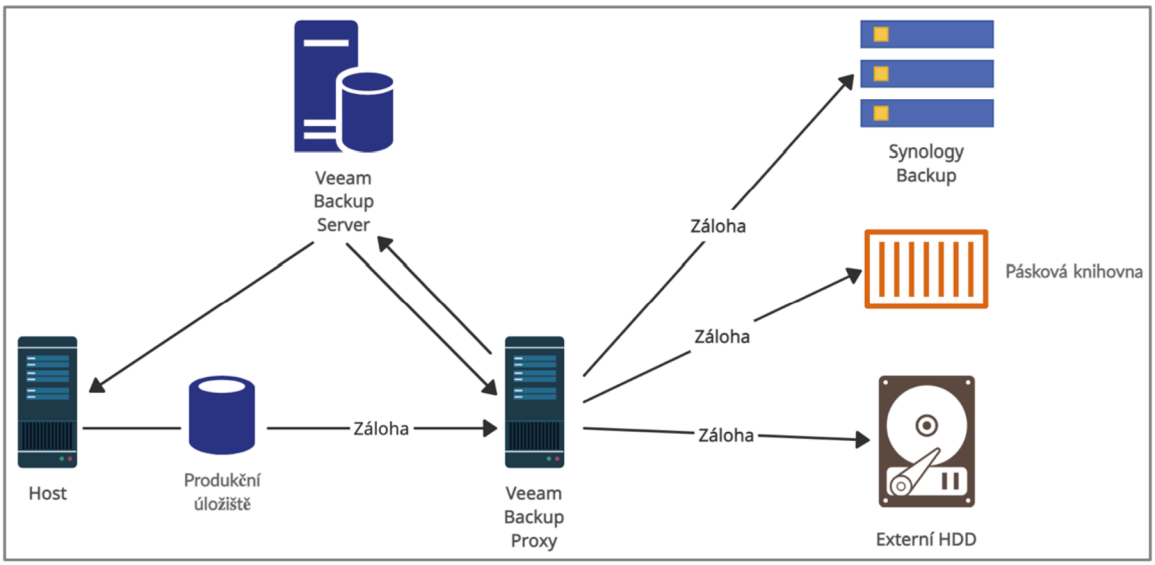
Tabulka 4.15 Datové centrum - vlastní zpracování

Pásková knihovna

Zálohy na páskovou knihovnu se budou provádět jednou týdně, tak aby v případě problému byla k dispozici původní data. I přes fakt, že se v tom případě může ztratit týden práce, pokud všechny ostatní zálohy selžou, stále je to lepší situace, než kdyby byly zašifrovány po případném útoku i zálohy na pásce.

Architektura zálohování virtuálních strojů

Struktura zálohování virtuální infrastruktury z celkového pohledu vypadá následovně.



Obrázek 4.3 Struktura zálohování virtuální infrastruktury - vlastní zpracování

Synology

Tato strategie se týká těch dat, která jsou primárně uložena na některém z těchto zařízení. Ostatní data, která se na tato zařízení ukládají již jako samotné zálohy, například zálohy virtuálních strojů z Veeamu, se zde neřeší. Výjimkou jsou zálohy síťových zařízení. U nich je většinou problém nastavit více cílů zálohy. Proto budou jejich primární zálohy dále zálohovány ze Synology. Celkově tak záloha Synology zařízení obsahuje následující data:

- Uživatelské „home“ složky.
- Sdílené složky pro jednotlivá firemní oddělení.
- Soukromé složky zákazníků.
- Zálohy sítových zařízení.

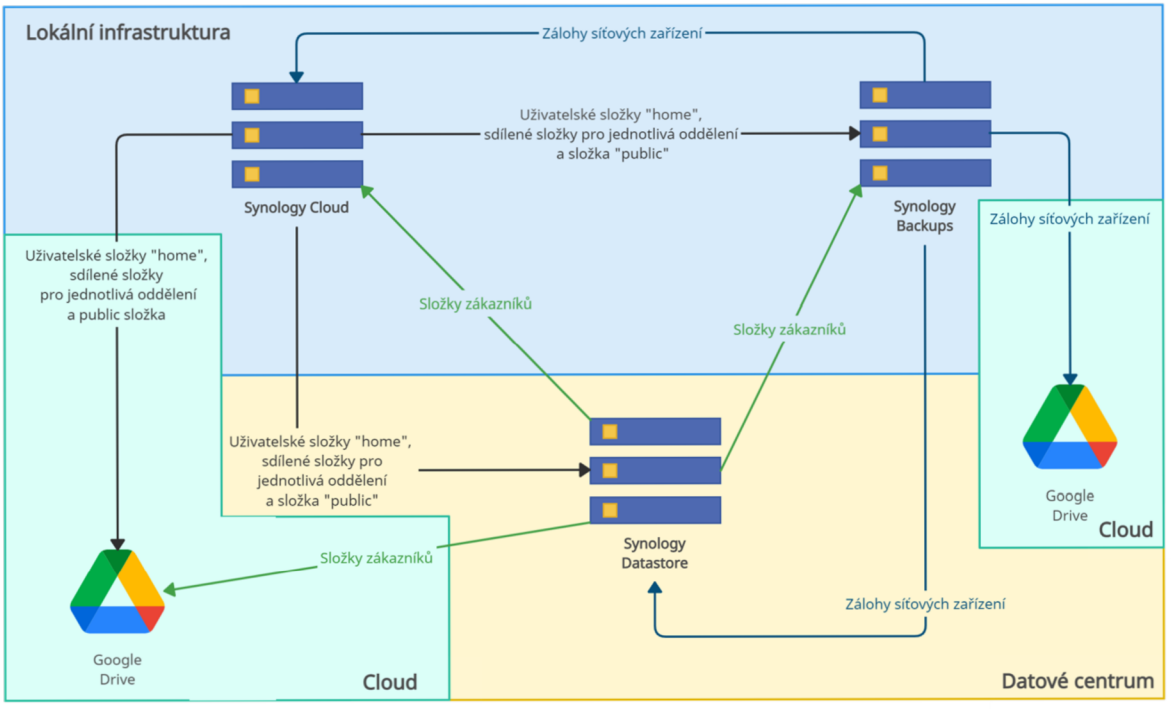
Strategie

Vzhledem k charakteru těchto úložišť a jejich možností bude použito pravidlo **3-2-1**. K dispozici jsou tři Synology úložiště. Princip je takový, že z každého Synology se budou zálohovat data na dvě zbývající. K tomu se provede šifrovaná záloha do cloudu na Google

Drive. Například složek zákazníků, uložených na Synology v datovém centru. Tím se docílí tohoto stavu:

- Originální data na prvním Synology.
- První záloha na druhém Synology.
- Druhá záloha na třetím Synology.
- Třetí záloha v cloudu Google Drive.

Struktura zálohování mezi zařízeními Synology je zobrazena na následujícím obrázku.



Obrázek 4.4 Struktura zálohování Synology zařízení - vlastní zpracování

Koncové stanice

Zálohy „složitějších“ počítačů budou prováděny jednou denně po pracovní době softwarem Veeam Agent. Cílem pro ukládání záloh budou Synology Cloud a Backups. Záloha mobilních zařízení je řešena automatickým zálohováním přes účty a služby Google a Apple.

Výrobní stroje

Zálohy výrobních strojů budou prováděny pomocí řešení Acronis. Zálohy se budou provádět během Vánočních svátků, během letní závodní dovolené nebo při změně konfigurace daného stroje. Cílem pro ukládání záloh budou opět obě Synology v lokální infrastruktuře a k tomu fyzický klon HDD/SSD daného stroje, uložený v trezoru.

Síťové prvky

U většiny síťových prvků probíhá záloha podobně jako u strojů. Tedy pokud dojde ke změně jejich konfigurace. Mezi výjimky patří firewally nebo vCenter Server, které se budou zálohovat automaticky a pravidelně. Z těchto strojů putují zálohy na všechny tři Synology zařízení.

Archivace

V této části je cílem návrh strategie pro archivaci následujících dat:

- Archiv smluv.
- Účetní data.
- HR.
- Správa budov.

Ve výše vypsaných případech se jedná o data, která se v podstatě nemění. V případě archivace je potřeba brát zřetel na podmínky dané zákonem. Normy určují, jak dlouho se které druhy dokumentů musí archivovat a také jak se mají archivovat. Tedy tak, aby bylo vše v souladu se směrnicí GDPR o ochraně osobních údajů. Dále musí být zaručena čitelnost a neporušenost dokumentů a také prokazatelnost osoby, která plnění uskutečňuje. Po uplynutí doby archivace lze data nenávratně zničit. Na tyto úkony a náležitosti je ve firmě určený specialista. Cílem této práce je tak vyloženě technické zajištění archivace.

Pro archivaci budou využity externí HDD disky a blu-ray média. V případě dokumentů bude použit uznávaný formát pro archivaci a to PDF/A. Média budou uložena na dvou odlišných místech v rámci firemního areálu. Samotný proces archivace bude probíhat vždy za jednotlivá čtvrtletí. Mezi tím se originální data, která ještě neprošla archivací, budou zálohovat.

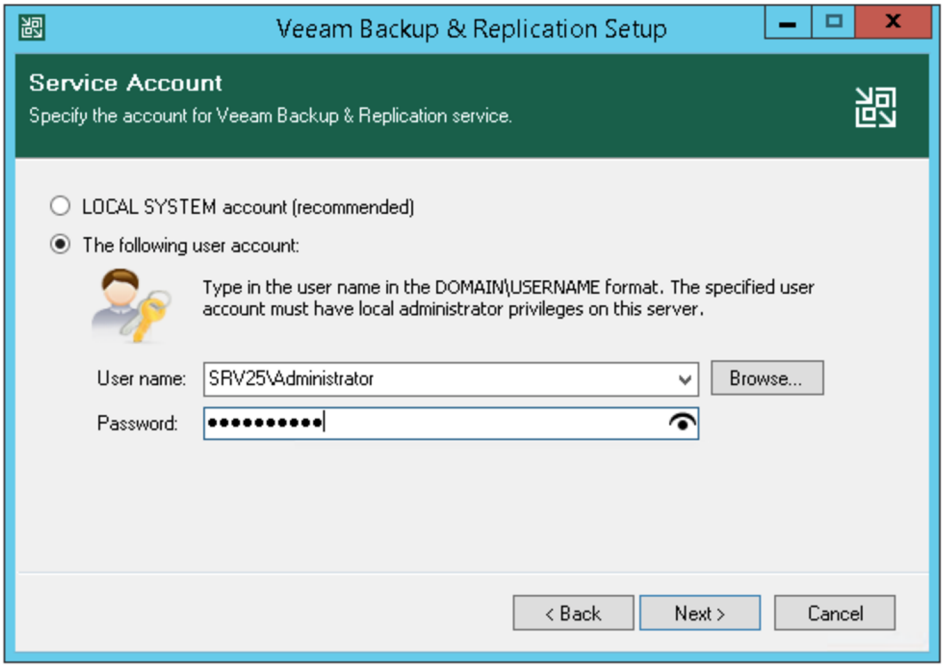
4.3.2 Konfigurace záloh

Veeam Backup & Replication

Stávající server, na kterém funguje Veeam Backup Server je nedostatečný z pohledu spolehlivosti a výkonu. Proto bylo rozhodnuto o sestavení nového serveru. Nepůjde o hotový rackový server, ale o svépomocí sestavený stroj. K sestavení budou využity takové komponenty, které splňují oficiální hardwarové požadavky výrobce. Čistě výkonem je však vysoce převyšují. Použité komponenty jsou následující:

- Základní deska – Gigabyte Z690 AORUS PRO
- Procesor – Intel Core i9-12900K 3.20 GHz
- Operační paměť – 32 GB
- SSD na operační systém - 250GB
- HDD na ukládání záloh – 4x 8 TB
- NIC – 4x Intel Gigabit CT + 1x TP-LINK TX401(10 Gbps)
- Case – Fractal Design Define R5 Black
- Operační systém – Windows Server 2019

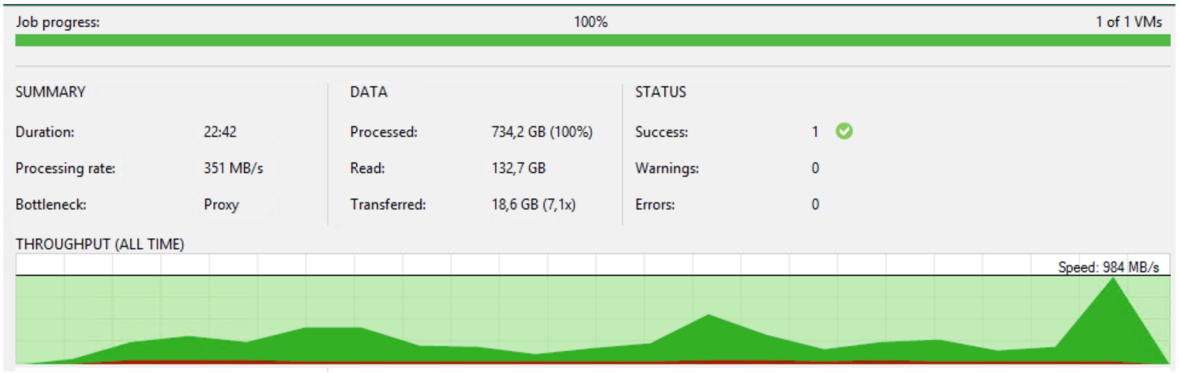
Po sestavení následuje instalace Windows Serveru. Jedná se o běžný instalační proces, který není nutné blíže rozvádět. Po instalaci a úspěšném zátěžovém testu je potřeba vytvořit pole RAID z HDD disků. Použit byl nástroj Intel Rapid Storage Technology. Čtyři disky jsou nastaveny na RAID5. Výsledná dostupná kapacita je tak 24 TB. Následně lze nainstalovat Veeam Backup & Replication a všechny potřebné komponenty. Opět se nejedná o nijak výjimečnou instalaci. Pozor se musí věnovat pouze u zvolení uživatele. Ve výchozím stavu je nastaven lokální uživatel. Pokud je stroj přidán do domény, potom se musí vybrat doménový uživatel, který musí být členem skupiny „Administrators“ a musí mít oprávnění pro úpravu databáze.



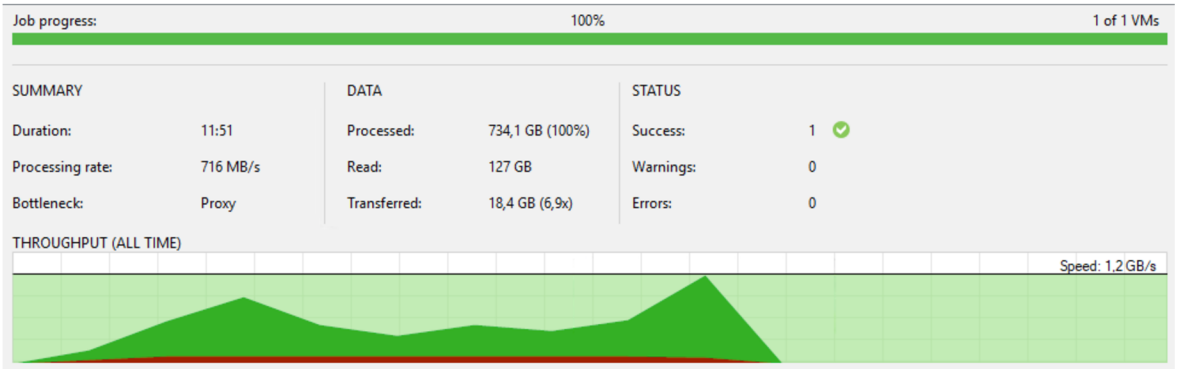
Obrázek 4.5 Instalace Veeam B&R [34]

Srovnání rychlosti zálohování se starým serverem

Po sestavení a instalaci softwaru byl proveden test, jenž měl za cíl ověřit smysl nasazení nového výkonného serveru. Test rychlosti zálohy byl proveden na produkčním serveru s OS Windows Server 2019 a velikosti 740 GB. Na původním Veeam Backup serveru trvala záloha 22:42 minut. Na nově postaveném stroji trvala záloha stejného serveru 11:51 minut. To je snížení času téměř o 50 %. Test byl proveden vícekrát, tak aby bylo zamezeno dosažení náhodného výsledku. Všechny výsledky byly téměř totožné. Srovnání testu je vidět na obrázcích. Na prvním je starý stroj a na druhém nový stroj.



Obrázek 4.6 Rychlost starého serveru - vlastní zpracování



Obrázek 4.7 Rychlost nového serveru - vlastní zpracování

Konfigurace

Po úspěšné instalaci je potřeba do Veeamu přidat infrastrukturu. To znamená:

- Backup proxy.
- Backup repository.
- VMware vCentre Server.
- Další jednotlivé hosty.
- Páskovou knihovnu.
- SMB shares.

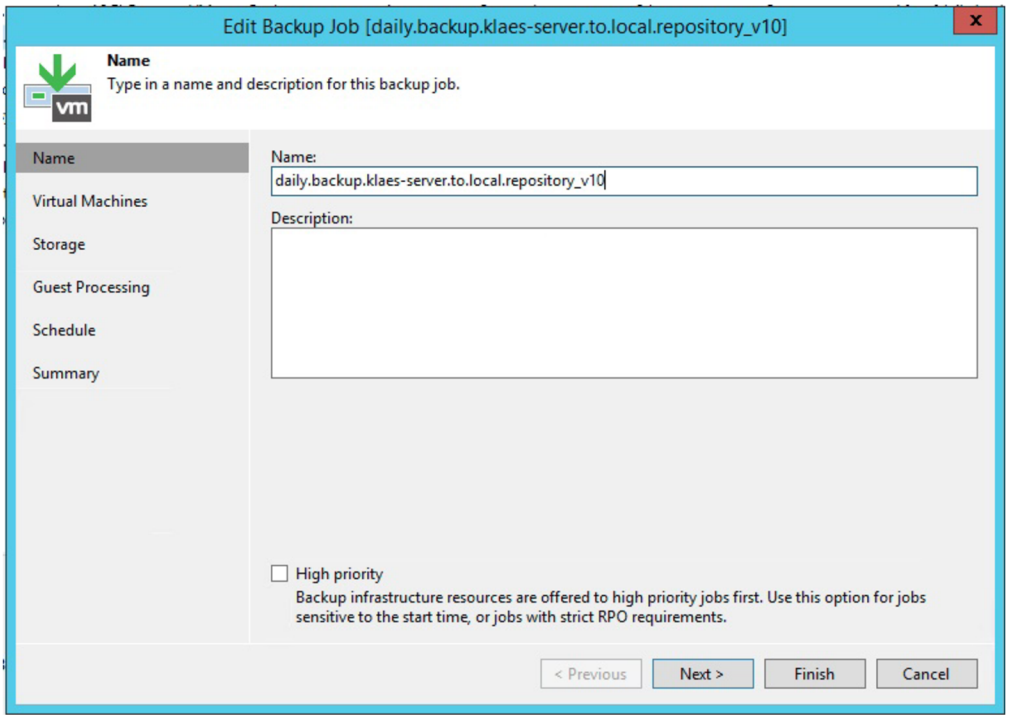
Samotné připojení jednotlivých komponent není nijak komplikované. Na vše je zde průvodce, kterým v podstatě stačí pouze projít a navolit konkrétní IP adresu, hostname nebo přihlašovací údaje.

Backup Jobs

Po přidání infrastruktury už lze začít vytvářet jednotlivé úlohy. Na následujících snímcích je vidět nastavení jedné konkrétní zálohovací úlohy.

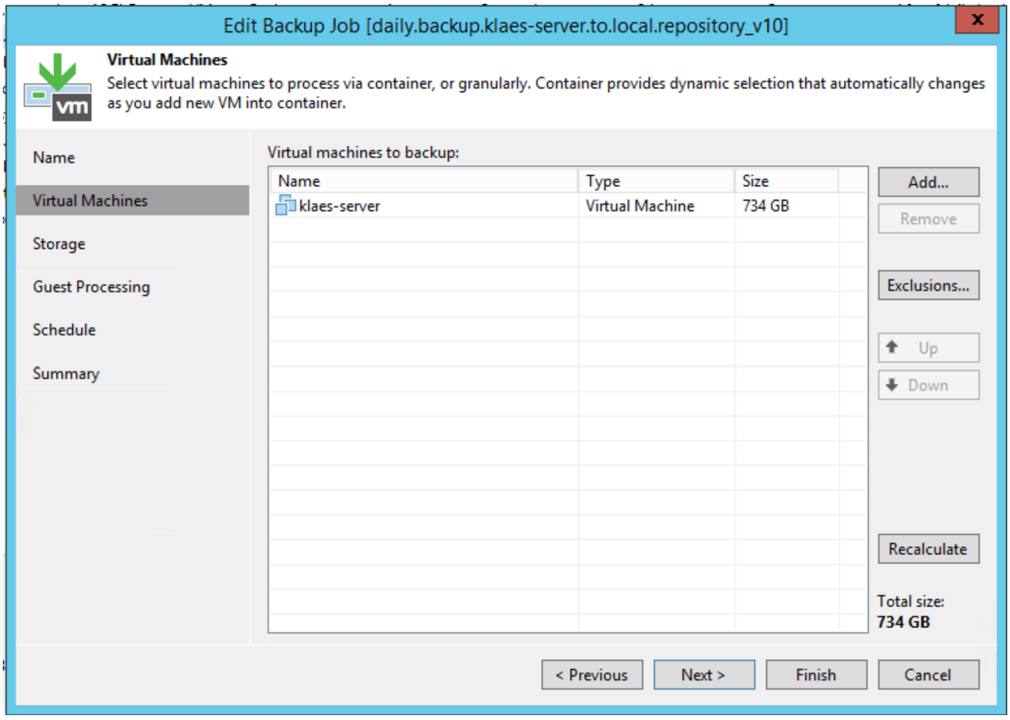
V první záložce je nutné zvolit název úlohy. Jedná se o relativně důležitou informaci pro pozdější práci. Název by měl být jednoduchý, ale na první pohled by měl ideálně obsahovat informace, které zasvěcené osobě pomohou identifikovat čas zálohy, typ, zdroj a cíl. Pod tímto názvem budou poté například přicházet informativní emaily s výsledky po dokončení úlohy. Také pro přehlednost v dalším nastavování je název důležitý. V případě desítek vytvořených úloh již může hrát tento detail podstatnou roli. Do pole „Description“ se sama vygeneruje časová značka vytvoření úlohy a uživatel, který ji vytvořil. Poslední částí je zaškrtávátko „High priority.“ Tím se nastaví této úloze vysoká priorita v přiřazování

zdrojů zálohovací infrastruktury. Tuto možnost je vhodno vybrat pro vysoce kritické virtuální stroje.



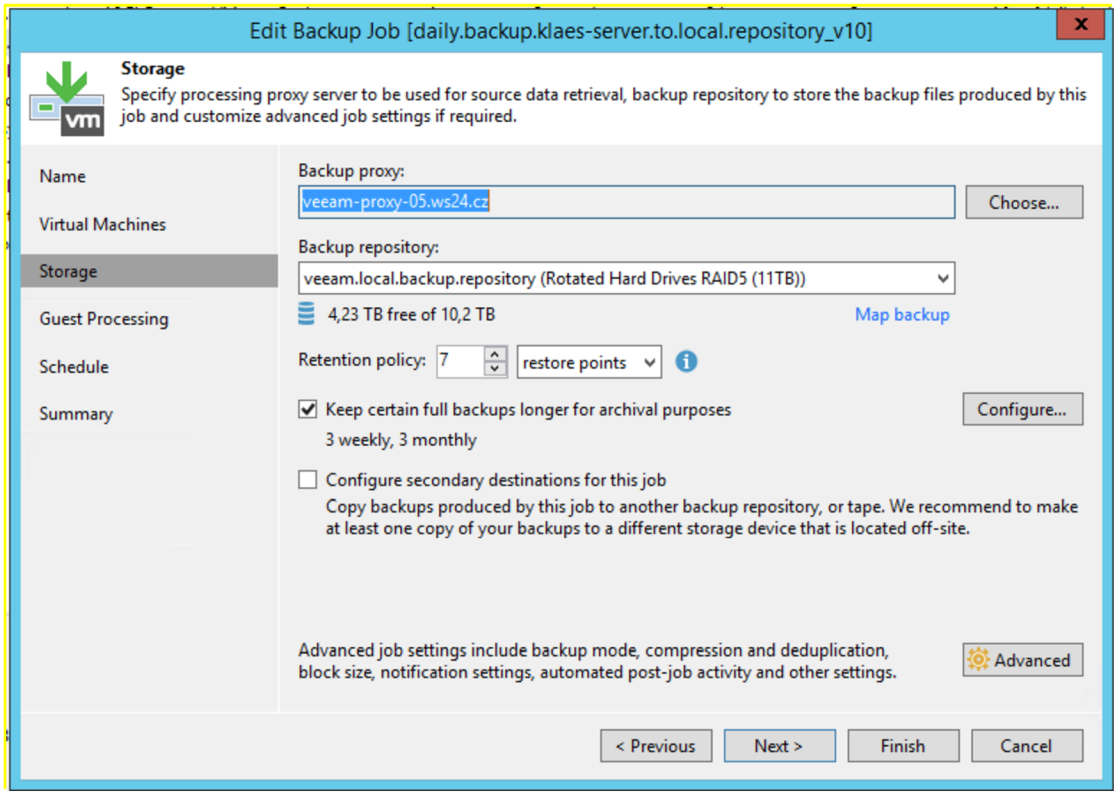
Obrázek 4.8 Veeam Backup job 1 - vlastní zpracování

Na další záložce je potřeba vybrat virtuální stroj, kterého se tento konkrétní job týká. Zde je možnost vybrat více strojů a zahrnou tak do jediné úlohy například určitou skupinu virtuálních počítačů, které spolu nějakým způsobem souvisí. V této úloze byl vybrán produkční server, který byl v předchozí částí použit k testování rychlosti zálohování.



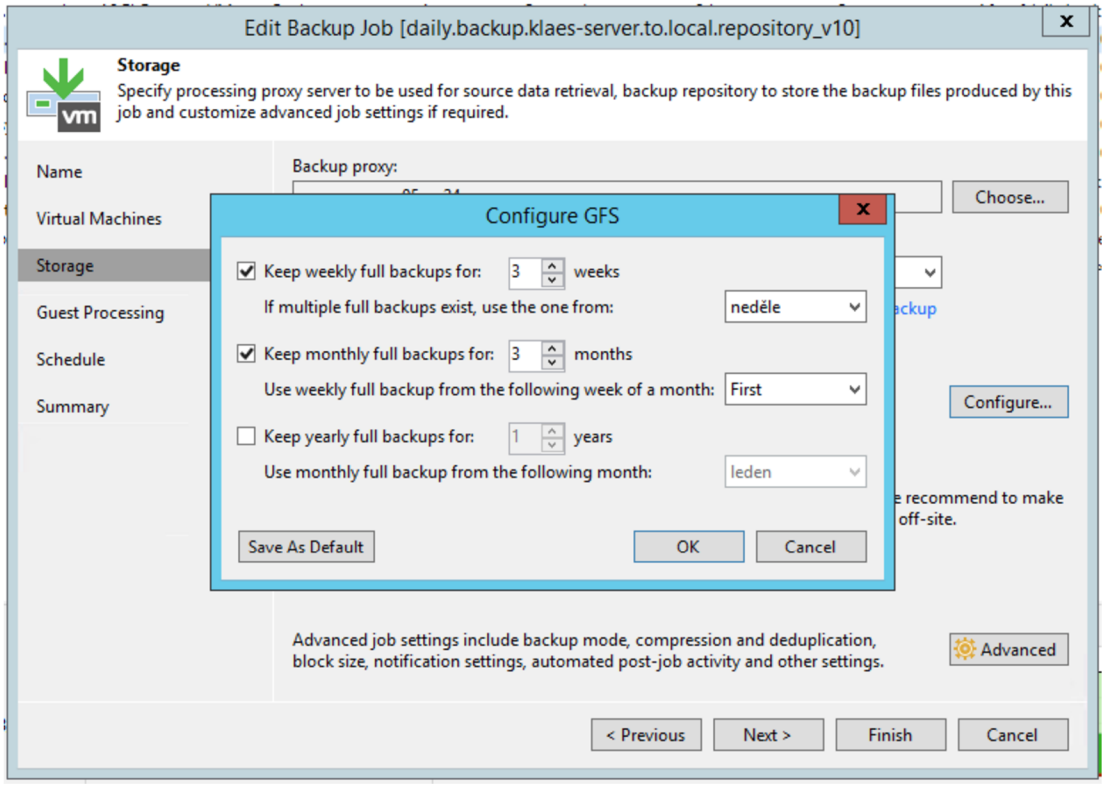
Obrázek 4.9 Veeam Backup job 2 - vlastní zpracování

V následujícím kroku se nastavuje vše, co se týká úložiště. V první řadě je nutné vybrat Veeam Backup proxy, přes které záloha proběhne. Následně se určí cíl pro uložení zálohy. Zde se také nastavuje retenční politika. V tomto případě se nastavila na 7 bodů.



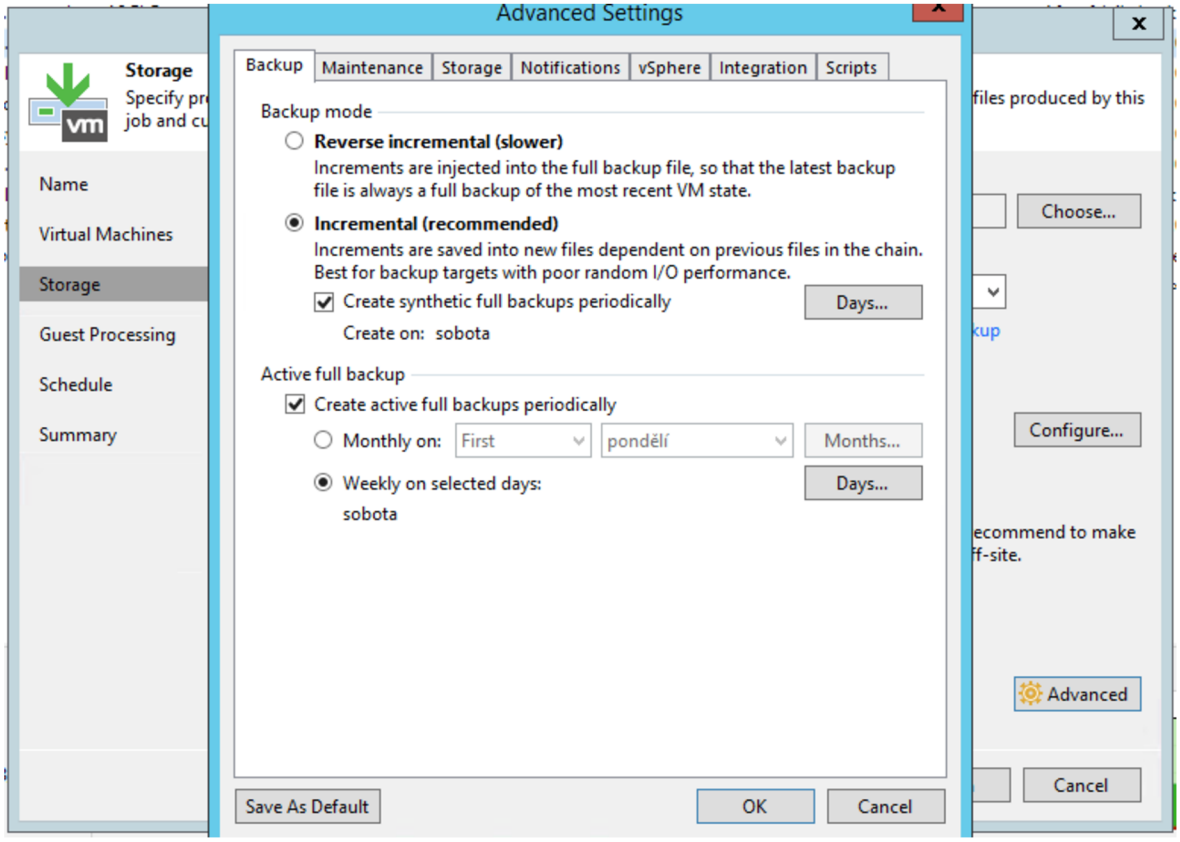
Obrázek 4.10 Veeam Backup job 3 - vlastní zpracování

Dále je zde možnost pro uchování plné zálohy na delší dobu a nastavení druhého úložiště. Týdenní plná záloha byla nastavena na 3 týdny. Měsíční plná záloha na 3 měsíce.



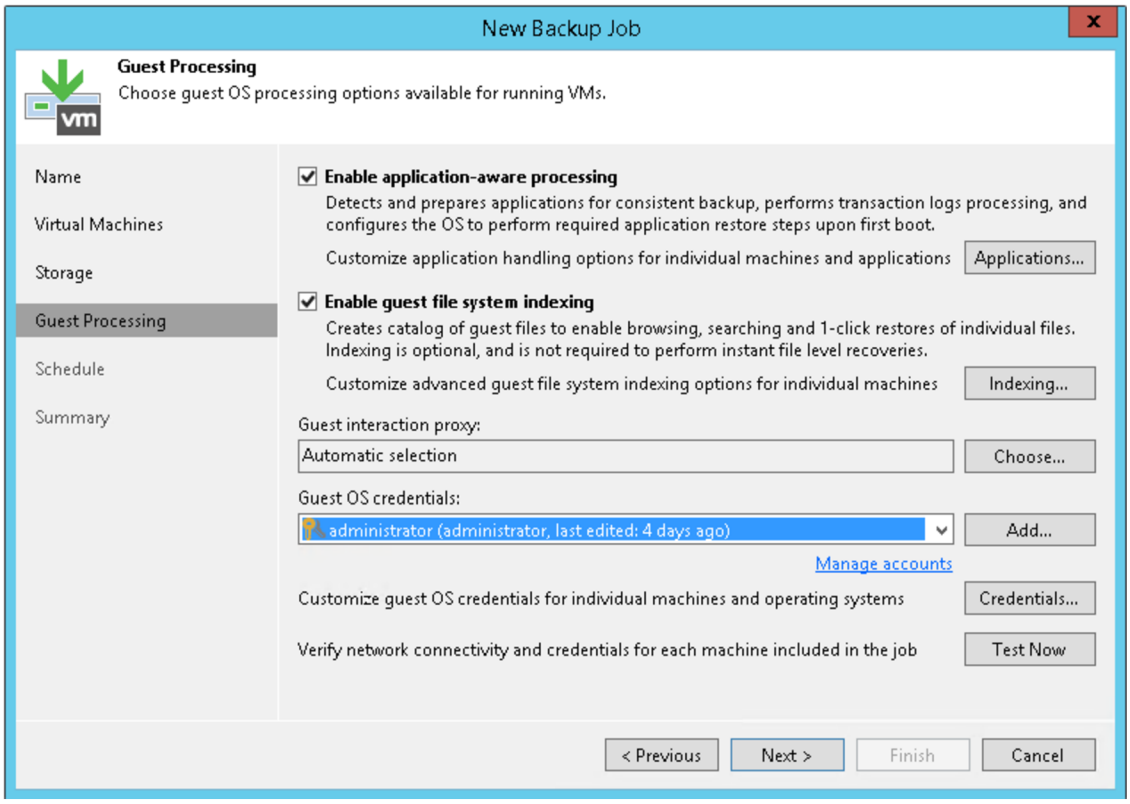
Obrázek 4.11 Veeam Backup job 4 - vlastní zpracování

Poslední možností je rozšířené nastavení zálohy. Zde se nastavuje typ zálohy či provádění syntetické plné zálohy, nebo aktivní plné zálohy. Zde byla vybrána možnost inkrementální zálohy s prováděním plné syntaktické zálohy každou sobotu, kdy není stroj vytížen. Ve stejný den proběhně pro jistotu také klasická plná záloha.



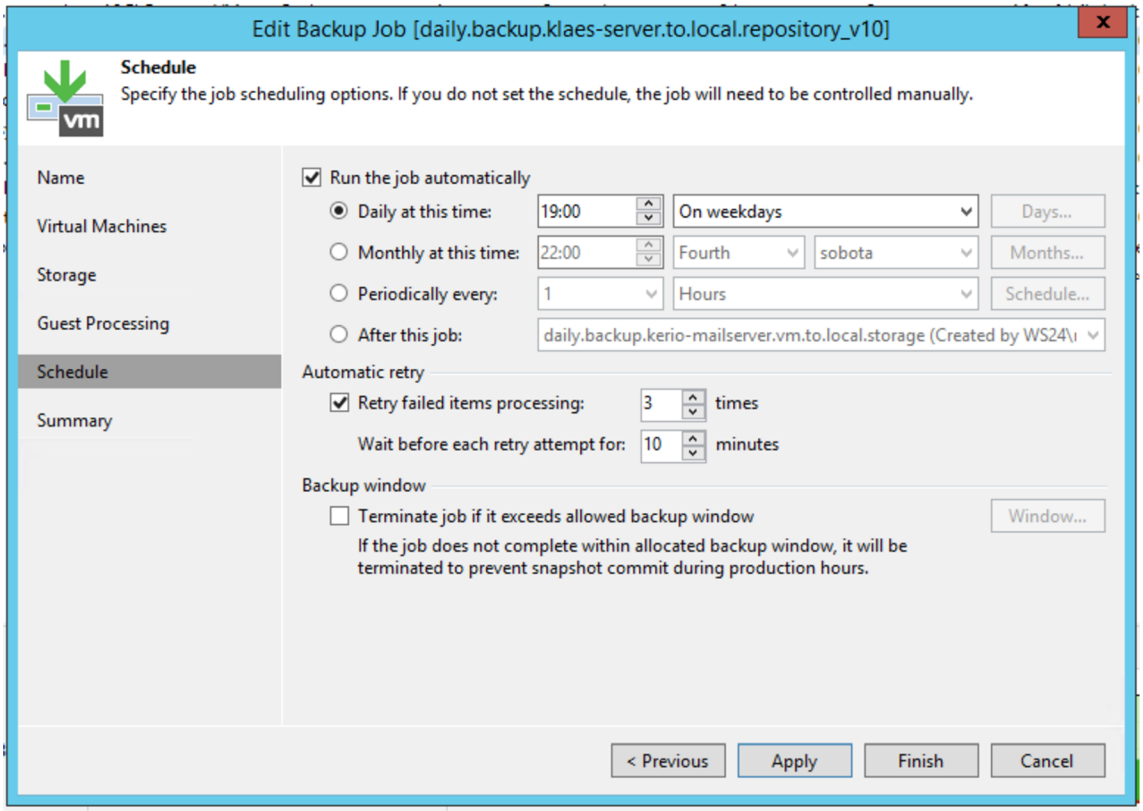
Obrázek 4.12 Veeam Backup job 5 - vlastní zpracování

V záložce „Guest Processing“ se povoluje indexace souborového systému a příprava aplikací pro konsistentní zálohu. Také se zde vybírá tzv. „Guest interaction proxy“. Toto proxy se právě stará o indexaci souborového systému a přípravu aplikací pro konsistentní zálohu. V posledním kroku je potřeba vybrat uživatele, pod kterým se záloha provede. Typicky to musí být účet s administrátorskými právy.



Obrázek 4.13 Veeam Backup job 6 - vlastní zpracování

V předposlední sekci se nastavuje automatické provádění zálohy. Zde se řeší opakování, čas provádění nebo automatické opakování v případě selhání úlohy. Tato úloha byla naplánována na každý den v 19:00. Opakování po neúspěšném pokusu bylo nastaveno na 3 pokusy. Poslední záložka „Summary“ slouží pouze ke kontrole toho, co bylo nastaveno v předchozích krocích.



Obrázek 4.14 Veeam Backup job 7 - vlastní zpracování

Replication Jobs

Nastavení replikace je podobné jako u zálohy, proto zde budou vyjasněna pouze nastavení, která se u klasické zálohy nenachází. První odlišností je výběr cíle pro vytvoření repliky. Zde je potřeba vybrat hosta, nebo cluster a datové úložiště, kde bude virtuální stroj uložen. V dalším kroku je potřeba vybrat uložiště pro metadata repliky. Typicky se bude jednat o backup repository. Na stejné záložce se také určuje sufix pro odlišení ostrého virtuálního stroje a repliky a následně se zde nastavuje počet bodů obnovení. Maximální počet je 28. Na další záložce se vybírají proxy. A to zdrojové a cílové. Tedy jedno proxy v tomto případě v lokální infrastruktuře a druhé proxy v datovém centru.

Další nastavení je velmi podobné klasické záloze. Nachází se zde například „Guest Processing“, časové nastavení a opakování úlohy nebo výběr zdrojového virtuálního stroje.

Záloha na pásku

Stejně jako v případě replikace, i zde budou pro přehlednost uvedena pouze specifická nastavení zálohy na pásku. První odlišností je výběr tzv. media poolu. Tento pool je nutné nejdříve vytvořit výběrem z pásek v automatické knihovně. Po připojení páskové knihovny do Veeamu se nastaví jednotlivé pooly pro následné zálohovací úlohy. Poté lze nastavit automatické vysunutí média z mechaniky po dokončení zálohy. Nastavení zálohy pomocí dostupné knihovny není příliš rozsáhlé, protože se jedná již o relativně starší zařízení. Nicméně jako doplňkové zálohovací zařízení nabízí vše, co je potřeba.

Veeam Agent

Pomocí tohoto softwaru se budou zálohovat koncové stanice. Přesněji řečeno osobní počítače. V první kroku je potřeba vytvořit „Recovery Media“. Pokud počítač z nějakého důvodu selže, lze z tohoto média provést bootování OS. Následně lze přistoupit k nastavení samotného zálohování.

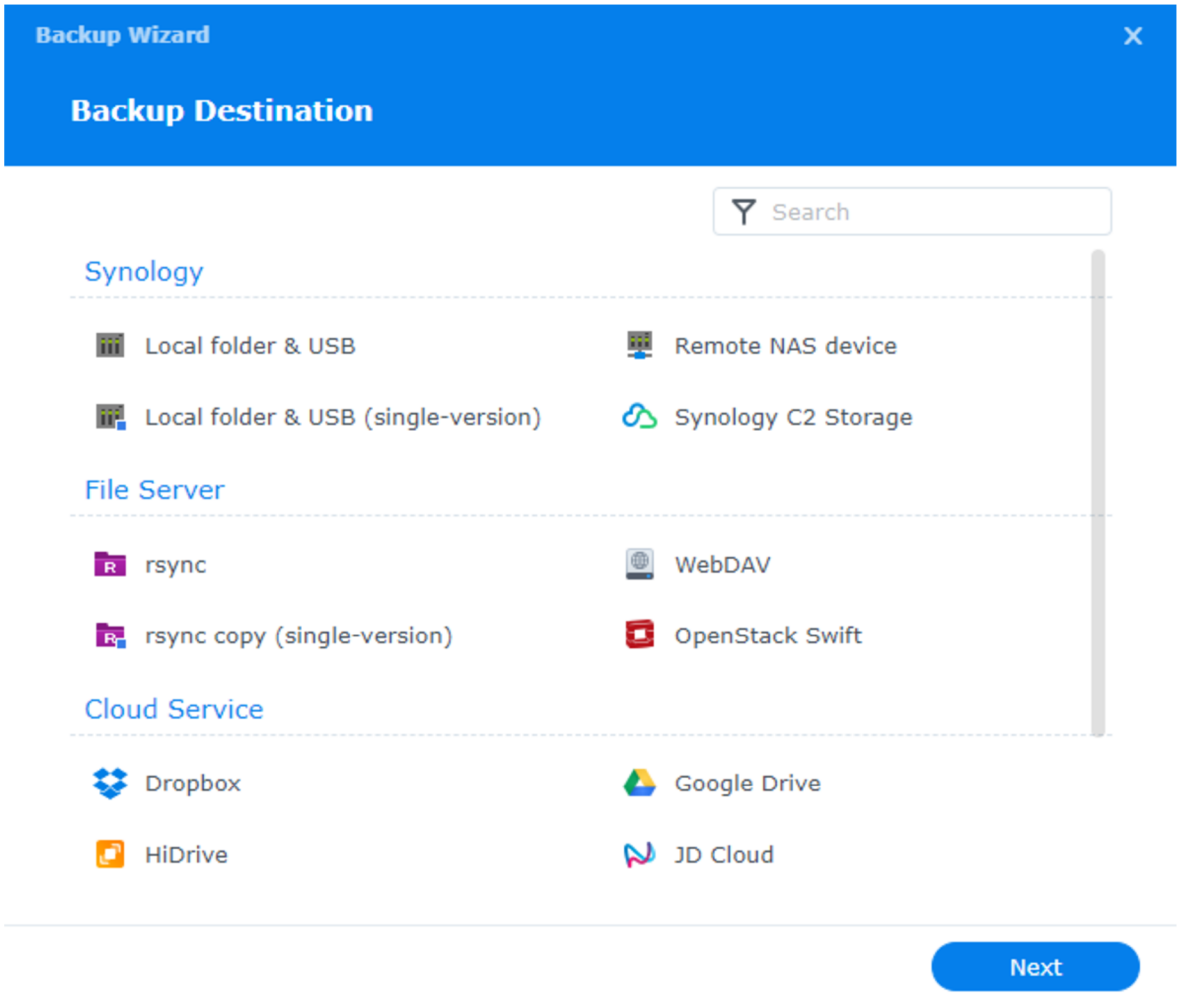
Na první záložce lze nastavit mód zálohy. Veeam nabízí zálohu celého stroje, jednotlivé volumy, nebo samostatné složky a soubory. V tomto případě bude vybrána záloha celého stroje.

V dalším kroku se nastavuje cíl pro zálohu. Veeam nabízí lokální úložiště, sdílenou složku, Veeam Backup repository nebo cloudové řešení. V tomto případě se bude zálohovat na Synology zařízení. K tomu je potřeba zadat uživatele s příslušnými právy.

Následně se nastavuje, kdy se má záloha provádět. Pracovní doba pro kancelářské zaměstnance končí v 16:15. Každodenní zálohy se tak budou provádět po pracovní době, tedy v 18:00. Na této kartě jsou také možnosti ohledně toho, co se má stát, pokud je počítač v době zálohy vypnutý, nebo co se má stát po dokončení zálohy. Pokud je vypnutý bude nastaveno, že se záloha provede hned po jeho zapnutí.

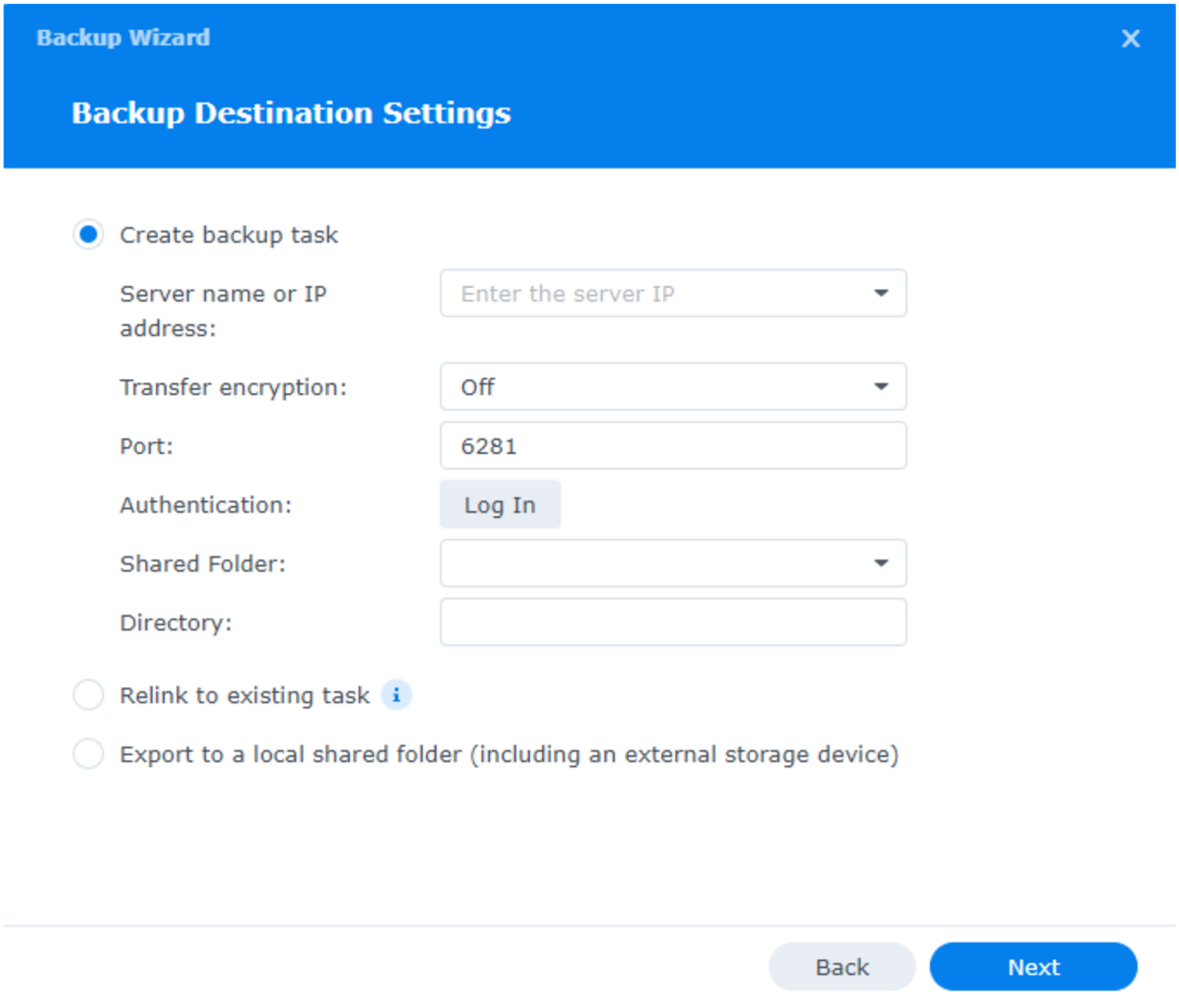
Synology Hyper Backup

V této části bude představen jeden konkrétní backup job na jednom ze Synology zařízení. V první řadě je potřeba aplikaci Hyper Backup nainstalovat na vybraný NAS. Aplikace je dostupná v „Package Centre“. Po instalaci lze přistoupit k samotnému zálohování. V prvním kroku se vybírá cíl pro zálohu. V tomto případě byla vybrána možnost „Vzdálené zařízení NAS“. To představuje další Synology zařízení.



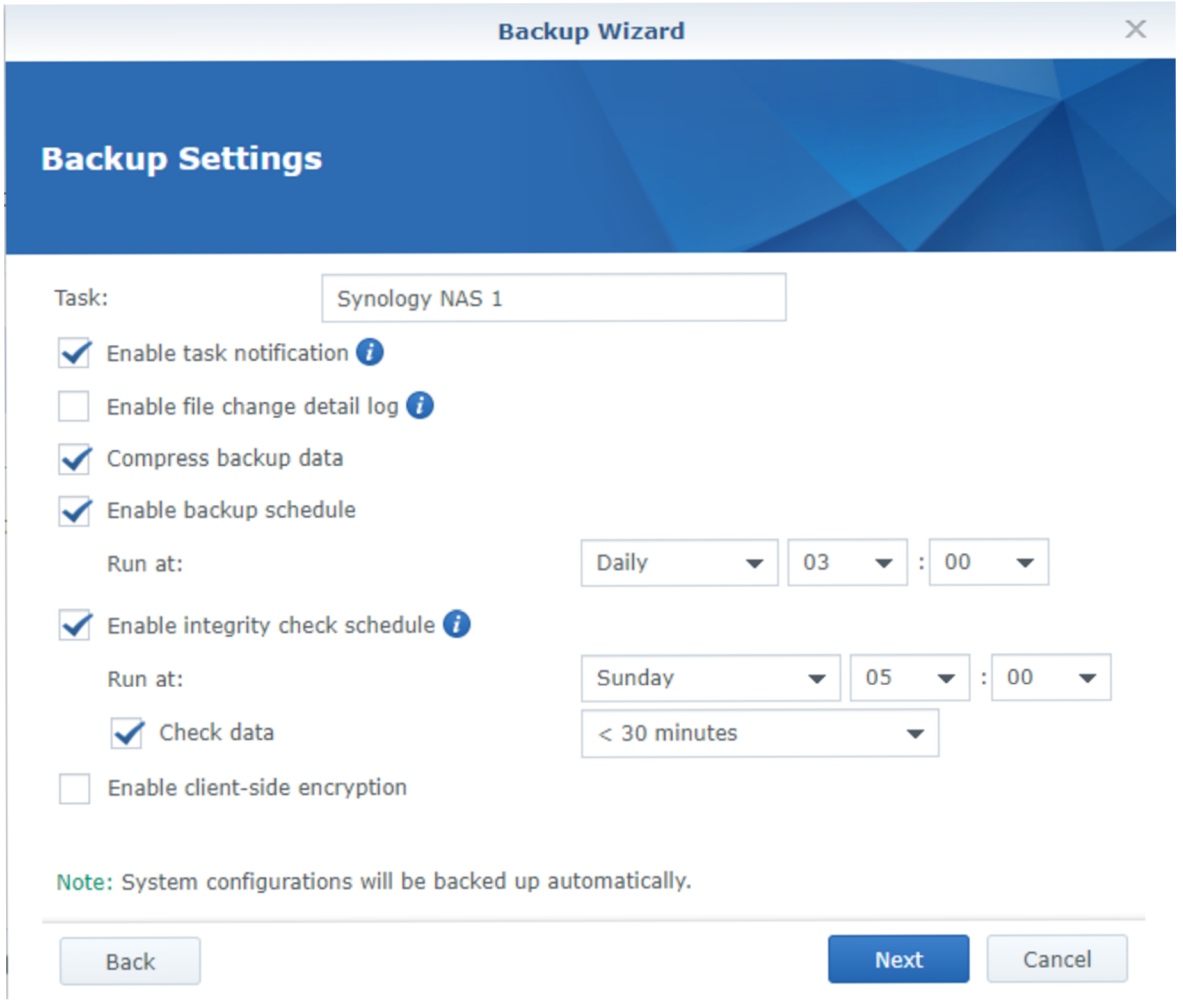
Obrázek 4.15 HyperBackup 1 - vlastní zpracování

V dalším kroku se nastaví IP adresa nebo hostname cílového Synology. Také je zde možnost zapnout šifrování. Tato možnost bude využita pro zálohování na Synology uloženém v datovém centru. Dále se nastaví uživatel, pod kterým se bude záloha provádět a vybere se cílová složka a adresář pro zálohu.



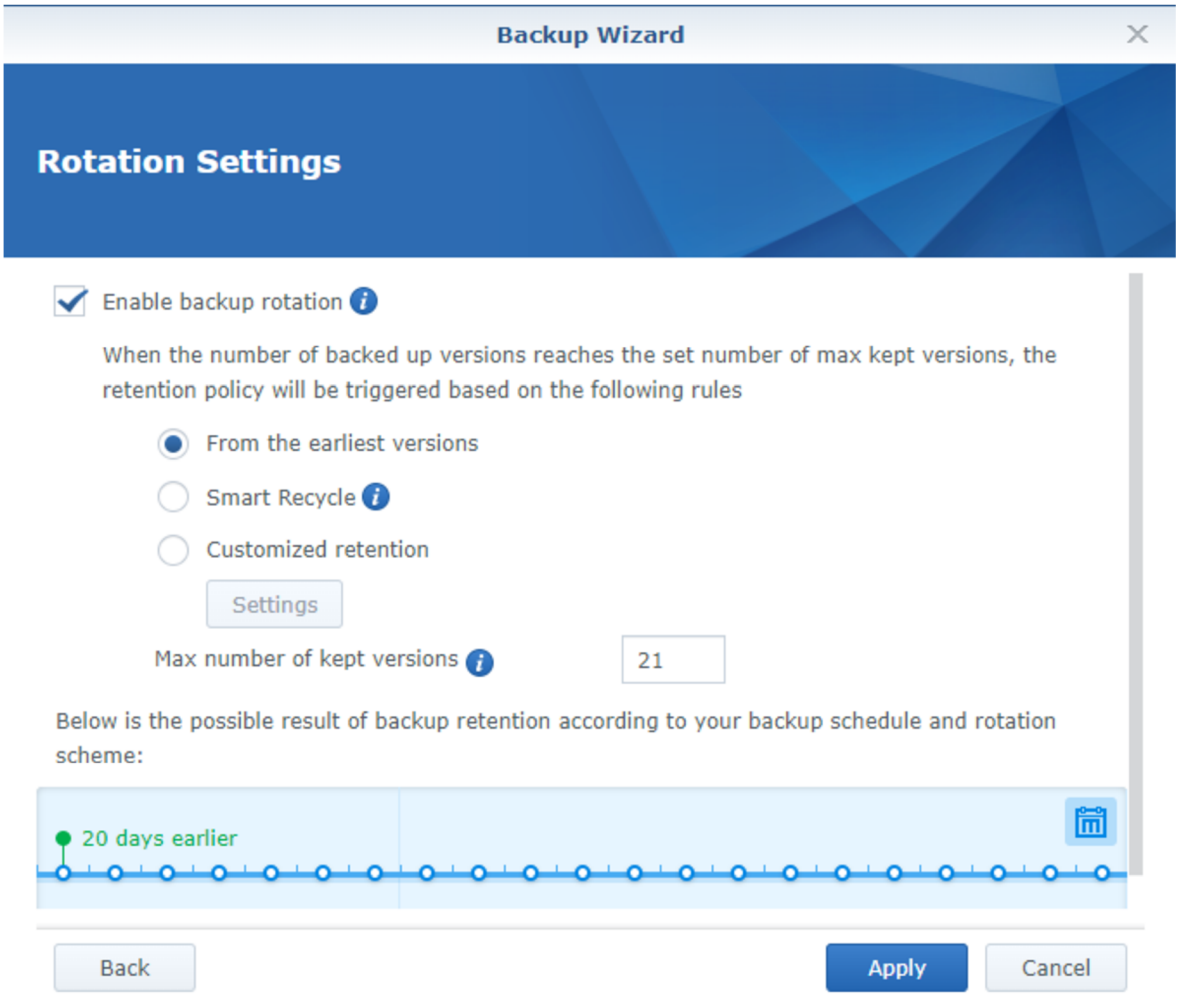
Obrázek 4.16 HyperBackup 2 - vlastní zpracování

Na další záložce se vybere zdroj pro zálohování. V následujícím kroku se nastaví automatické provádění zálohy. V tomto případě bude záloha nastavena na každý den v 03:00. K tomu se každou neděli provede kontrola integrity dat.



Obrázek 4.17 HyperBackup 3 - vlastní zpracování

V poslední části se nastavuje rotace zálohy. U této úlohy bylo nastaveno uchování verzí za 3 týdny zálohování. Poté se začnou nejstarší zálohy mazat a vždy se nahradí novou verzí. K dispozici tedy bude vždy 21 verzí, respektive dnů zálohy.

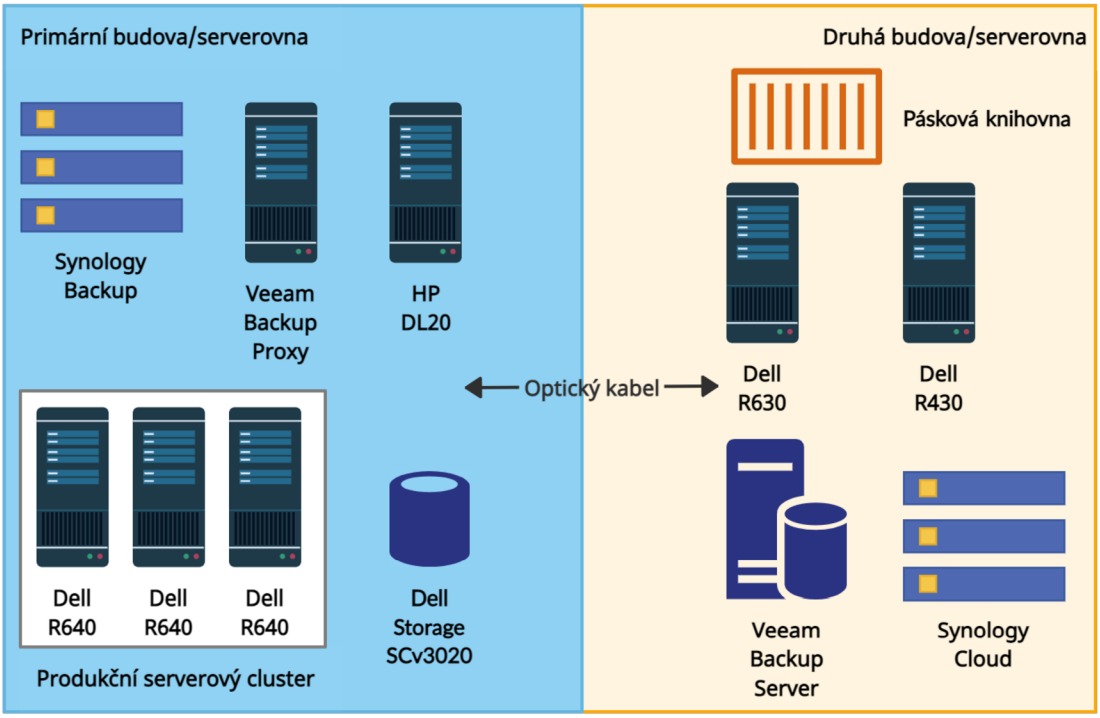


Obrázek 4.18 HyperBackup 4 - vlastní zpracování

4.3.3 Rozmístění zdrojů

Pro efektivní zálohovací strategii je potřeba ještě upravit rozmístění serverů a úložišť mezi obě serverovny v lokální infrastruktuře. Momentálně je vše v jedné hlavní serverovně. Mít k dispozici druhou a nevyužívat ji je zbytečný luxus. Toto rozdělení zvýší šance na uchování dat v případě nějakého opravdu závažného incidentu. Pokud první budova například vyhoří, stále bude k dispozici druhé stanoviště, kde budou zálohy a zdroje ze kterých půjde spustit minimálně kritická část virtuální infrastruktury v relativně přijatelném čase.

Na obrázku níže je vidět nové rozdělení infrastruktury. Nejsou zde například switche nebo UPS zařízení, kterých se nachází v obou serverovnách několik. Pro přehlednost se jedná čistě o infrastrukturu přímo týkající se zálohování.



Obrázek 4.19 Rozmístění zdrojů - vlastní zpracování

4.4 Testování

Cílem testování je především ověření funkčnosti záloh, rychlosti vytváření záloh a jejich následné obnovy. V rámci toho tak bude nakonfigurován Veeam SureBackup, který automaticky testuje vytvořené zálohy virtuálních strojů.

4.4.1 Záloha

V následující tabulkách jsou vidět výsledky s rychlostmi záloh jednotlivých dat. Rozdíl v rychlostech záloh virtuálních strojů mezi zálohami v lokální infrastruktuře a v datovém centru je dán především novým hardwarem použitým pro Veeam Backup Server v lokální infrastruktuře.

Virtuální infrastruktura

OS	Zařazení	Velikost (GB)	Full Backup	Inkrementální
Ubuntu 64-bit	Servis	32	0:03:45	0:00:55
MS Server 2019	Produkce	40	0:04:18	0:00:56
MS Server 2012	Produkce	120	0:12:08	0:03:39
Windows 10	Servis	60	0:07:02	0:01:21
Windows 10	Produkce	100	0:08:13	0:03:01
Ubuntu 64-bit	Produkce	2000	2:16:32	0:09:51
MS Server 2016	Produkce	1024	1:03:28	0:07:24
Windows 10	Produkce	100	0:07:06	0:03:09
Windows 10	Produkce	80	0:09:48	0:01:54
MS Server 2012	Produkce	50	0:06:14	0:01:11
Windows 10	Produkce	60	0:07:07	0:01:35
Windows 7	Servis	100	0:08:49	0:02:51
Windows 7	Produkce	60	0:07:26	0:01:42

Tabulka 4.16 Lokální infrastruktura - vlastní zpracování

OS	Zařazení	Velikost (GB)	Full Backup	Inkrementální
MS Server 2019	Produkce	40	0:09:58	0:01:27
Ubuntu 64-bit	Produkce	30	0:07:32	0:01:15
Ubuntu 64-bit	Produkce	90	0:12:51	0:03:54
MS Server 2019	Produkce	512	0:38:26	0:06:42
Windows 10	Servis	60	0:09:18	0:02:26
Ubuntu 64-bit	Servis	32	0:08:07	0:02:10

Tabulka 4.17 Datové centrum - vlastní zpracování

Synology

Rychlost jednotlivých záloh v zařízení Synology je velice různá. Vše záleží na počtu a velikosti změn zálohovaných dat. První záloha například všech uživatelských složek trvala 2 hodiny a 12 minut. Další dny už záloha trvala pouze 1 minutu. Ve dnech, ve kterých nastaly větší změny v souborech, se trvání zálohy dostalo maximálně na 22 minut. Stejné je to u ostatních úloh na všech třech zařízeních Synology.

Ostatní

Záloha síťových zařízení je otázka pár sekund. Naopak v případě strojů ve výrobě se záloha výrazně prodlužuje se stářím daného stroje. Od nejnovějších strojů, u kterých záloha proběhne do 30 minut až po 20 let stará zařízení, kde záloha probíhá i 3 hodiny. Problémem je jednoznačně nízký výkon počítačů v těchto zařízeních, protože velikost zálohy je menší než u nových strojů. Záloha celých uživatelských počítačů pomocí Veeam Agent s SSD diskem o velikosti 250 GB, což je typický stroj v řešené firmě, trvá průměrně 10 až 15 minut.

4.4.2 Obnova

Rychlost obnovy je velice důležitá informace. Snahou je samozřejmě docílit co možná nejkratší doby potřebné pro obnovu dat. Hlavní faktory ovlivňující rychlost obnovy jsou stejné jako u rychlosti zálohování. Tedy velikost obnovovaných dat a místa ze kterého a kam se data obnovují. To znamená lokální, sítové úložiště nebo cloud.

Virtuální infrastruktura

Veeam nabízí několik variant obnovy. Od obnovy na úrovni řádků v SQL databázi přes obnovu souborů a disků, po obnovu kompletního virtuálního stroje. V tomto případě je nejdůležitější čas obnovy celého virtuálního stroje. Přehled s časem obnovy jednotlivých virtuálních strojů se nachází v následujících tabulkách.

OS	Zařazení	Velikost (GB)	Obnova
Ubuntu 64-bit	Servis	32	0:00:45
MS Server 2019	Produkce	40	0:01:10
MS Server 2012	Produkce	120	0:02:26
Windows 10	Servis	60	0:01:35
Windows 10	Produkce	100	0:01:54
Ubuntu 64-bit	Produkce	2000	1:24:18
MS Server 2016	Produkce	1024	0:56:18
Windows 10	Produkce	100	0:02:04
Windows 10	Produkce	80	0:01:41
MS Server 2012	Produkce	50	0:01:23
Windows 10	Produkce	60	0:01:27
Windows 7	Servis	100	0:01:49
Windows 7	Produkce	60	0:01:18

Tabulka 4.18 Lokální infrastruktura - vlastní zpracování

OS	Zařazení	Velikost (GB)	Obnova
MS Server 2019	Produkce	40	0:02:45
Ubuntu 64-bit	Produkce	30	0:01:38
Ubuntu 64-bit	Produkce	90	0:06:31
MS Server 2019	Produkce	512	0:49:56
Windows 10	Servis	60	0:05:12
Ubuntu 64-bit	Servis	32	0:02:02

Tabulka 4.19 Datové centrum - vlastní zpracování

Synology

Rychlost obnovy uživatelských home složek se bude výrazně lišit u jednotlivých uživatelů. Důvodem jsou rozdílné velikosti adresářů a počet souborů. Pro testování byl použit adresář o velikosti 15 GB. Obnova tohoto adresáře ze Synology „Cloud“ na Synology „Backup“ trvala 8 minut a 26 vteřin.

Ostatní

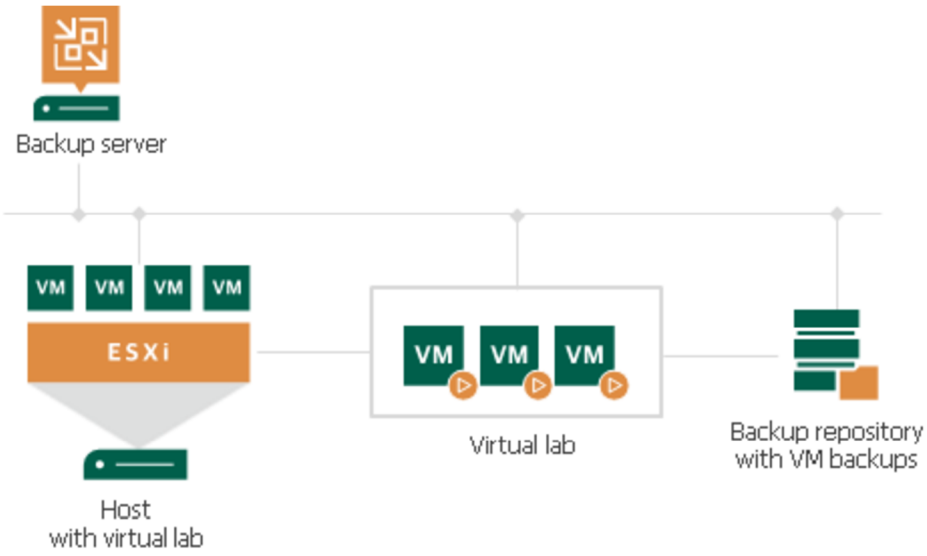
Stejně jako v případě zálohy i obnova je u síťových zařízení velice krátká. U strojů je rychlost již podstatný problém. Není možné zastavit výrobu a několik hodin čekat na obnovu. Proto je při zálohování provedena replika pevného disku daného stroje. V případě problému tak stačí za velice krátkou dobu vyměnit pevný disk v počítači nefunkčního stroje. Obnova uživatelských počítačů pomocí Veeam Agent trvá průměrně 5 minut.

4.4.3 Veeam SureBackup

Jedná se o úlohu, která ověřuje funkčnost vytvořených záloh. Pro tuto funkci musí být vytvořeno izolované prostředí nazvané virtual lab, ve kterém se vybrané zálohy obnovují. Dále je nutné vytvořit aplikační skupinu. Jedná se o skupinu komponent, které jsou nutné k ověření funkčnosti vybraného virtuálního stroje. Obvykle se jedná minimálně o doménový řadič, DNS a DHCP server. V závislosti na specifikách daného virtuálního stroje to mohou být jakékoliv další servery, jako jsou například databázový server, mail server a další.

Virtuální stroje se spouští přímo z komprimovaných a deduplikovaných záloh, které jsou umístěny v úložišti záloh. SureBackup zároveň nabízí možnost skenování k odhalení malwaru. Mimo to provede řadu testů samotného virtuálního stroje, čímž ověřuje jeho funkčnost. Během procesu ověřování zůstává záloha ve stavu pouze pro čtení. Veškeré změny, ke kterým dojde, když je virtuální stroj spuštěn, jsou po dokončení procesu

odstraněny. Proces se může ukončit automaticky, nebo lze nechat stroje spuštěné pro další ruční otestování funkčnosti [35].



Obrázek 4.20 Veeam SureBackup [35]

Konfigurace

Veeam jako u všech úloh i zde nabízí přehledného průvodce nastavení. To se skládá ze tří částí. Virtual Lab, Aplikační skupina a samotná úloha.

Virtual Lab

1. Nastaví se jméno. V tomto případě to bude „Sandbox Lab“.
2. Vybere se fyzický host, na kterém se budou virtuální stroje spouštět. V tomto případě to bude jeden z hostů v produkčním clusteru.
3. Určí se úložiště pro dočasné soubory ověřovaného virtuálního stroje a logy. Zde bude ponecháno lokální úložiště serveru.
4. Zvolí se proxy. To slouží jako brána, která poskytuje přístup z backup serveru k virtuálnímu stroji ve virtual labu.
5. V dalším kroku se zvolí síťový mód. Na výběr jsou tři módy. Single-host s automatickým nebo s manuálním nastavením a multi-host. Zde bylo zvoleno automatické nastavení pro single-host.
6. Poslední velmi důležitou částí je vytvoření izolované sítě.

Aplikační skupina

1. Nastaví se jméno. Zde to bude „Sandbox Group“
2. Následně byly přidány veškeré virtuální stroje, které jsou potřeba pro testování.
3. V dalším kroku se volí pro vybrané stroje, jakou mají zastávat roli. Například DNS server, doménový řadič atd.
4. Dále jsou k dispozici nastavení jako maximální povolený boot time, timeout nebo přidávání dalších testovacích skriptů. Zde byly nastaveny standardní hodnoty.

SureBackup job

Po vytvoření virtual labu a aplikační skupiny lze již vytvořit samotnou úlohu SureBackup. Tato úloha se v jejím vytváření zásadně neliší od klasických zálohovacích úloh.

1. Zvolí se jméno.
2. Vybere se virtual lab.
3. Určí se aplikační skupina.
4. Následuje výběr zálohovací úlohy, který se má testovat.
5. Nastaví se dodatečné možnosti ověření jako je test integrity, malware skenování a zasílání notifikací.

4.5 Výhody a nevýhody zvoleného řešení

Výsledné řešení by mělo přinést řadu výhod oproti původnímu řešení. Tou nejdůležitější změnou je zavedení strategie 3-2-1-1-0 u zálohování virtuální infrastruktury a strategie 3-2-1 u síťových úložišť. Výhoda v nasazení těchto strategií tkví již v prostém počtu míst s dostupnou zálohou, ale také v charakteru těchto kopií. Tím se snižuje riziko selhání hardwaru nebo softwaru na minimum. Je velmi nepravděpodobné, že v jeden okamžik dojde k selhání veškerých kopií zálohy. V dosavadním stavu byla záloha pouze na dvou místech. V tomto ohledu se tak jedná o podstatný posun k celkové bezpečnosti firemních dat. Díky nasazení těchto strategií se také snížilo riziko ztráty dat z důvodu napadení škodlivým kódem.

Další výhodou je také rychlost zálohování. V tomto směru došlo k výraznému snížení potřebného času k vykonání zálohy, a to díky nově postavenému serveru pro Veeam Backup.

Dalším pozitivním přínosem tohoto řešení je podstatně efektivnější využití firemních zdrojů v podobě hardwaru a celkově infrastruktury. Nevyužívaný hardware tak byl

zakomponován do celkové strategie. Rozložení infrastruktury mezi dvě serverové místnosti pak přináší benefit v podobě ochrany proti mechanickému poškození. Pokud jedna místnost například vyhoří, stále bude k dispozici druhá. V dosavadním stavu by firma přišla o veškerá data a zdroje jako jsou servery nebo úložiště.

Patrně největší nevýhodou či slabinou tohoto řešení jsou offline zálohy virtuálních strojů. Momentální způsob je relativně náročný z hlediska času a nutnosti zásahu administrátora. Na druhou stranu, pokud všechny ostatní zálohy budou například zašifrovány útočníkem, offline záloha by měla být vždy v pořádku. A i přesto, že bude obsahovat již podstatně zastaralejší data, stále se jedná o přijatelnější variantu než přijít o data kompletně.

Mezi nevýhody lze také zařadit relativně komplikovanější strukturu a časovou náročnost na implementaci tohoto řešení. Nicméně se jedná o řešení komplexního charakteru s vysokou mírou spolehlivosti, které v sobě zahrnuje veškeré potřeby zálohování ve vybrané firmě. Cenou za to je samozřejmě vyšší komplikovanost, než v případě základních požadavků menších firem či jednotlivců. Tento fakt se snažil autor zmírnit přesným představením strategie včetně nákresů, díky kterým by měli například nově příchozí zaměstnanci rychle získat vhled do problematiky.

4.6 Finanční zátěž

Finanční zátěž implementace tohoto řešení lze rozdělit do tří oblastí. Finance potřebné na software, hardware a plat administrátorů. Poslední část je v tomto případě poměrně složitě vyjádřitelná naprosto přesnými čísly. Jelikož realizace probíhala několik týdnů až měsíců spolu s další běžnou pracovní činností autora, která není složena pouze z tohoto jednoho úkolu, nelze přesně vyčíslit čas a tím pádem finance potřebné pouze k implementaci tohoto řešení. Odhad čisté časové náročnosti se pohybuje okolo 120 pracovních hodin. Průměrná hodinová sazba v ČR na pozici „systémový administrátor“ byla v březnu roku 2022 281 Kč [37]. Dle toho vychází cena implementace z hlediska obsluhy zhruba na 34 000 Kč.

V případě nákladů na hardware je situace jednodušší. První zařízení, které bylo placeno pouze kvůli této realizaci, byl nový server pro Veeam Backup. Náklady na sestavení serveru z jednotlivých komponent se vyšplhaly na 38 000 Kč bez daně. Další výdaje se týkají Synology zařízení v celkové hodnotě 74 000 Kč bez daně. Pro zbytek implementace byla využita stávající infrastruktura.

Náklady na software obsahují dvě části. Hlavní položkou je software Veeam. V době tvorby této práce měla firma již zakoupenou licenci, která vyprší v červnu 2022. V tu chvíli bude nutné licenci prodloužit. Cena za licenci na další rok je k vzhledem k počtu virtuálních strojů přibližně 55 000 Kč. Tato částka bude započtena do nákladů této práce, protože se již od počátku počítalo s dalším využitím služeb Veeam. Druhou položkou je software Acronis Disk Director. Zde se jedná o zanedbatelnou částku. K vzhledem k účelům, ke kterým bude tento software využíván, stačila běžná licence za jednorázových 1200 Kč. Celkové náklady nutné k realizaci řešení navrženého v této práci tak vycházejí na 202 200 Kč bez daně.

Náklady na příštích 5 let

Odhad plánované finanční zátěže na dalších 5 let v oblasti softwaru vychází především z předpokladu setrvání na zálohovací platformě Veeam. Za současných cen ročního poplatku bude pětileté období předplatného vyžadovat 275 000 Kč.

Obnova hardwaru bude obsahovat minimálně jednu položku. Tou bude nový Veeam server v datovém centru. Odhadovaná cena se bude podobat ceně nového serveru pro lokální infrastrukturu. Tedy 40 000 Kč bez daně. Další možnou investicí budou nové páskové knihovny. Stávající zařízení již začínají přesluhovat. Jelikož se ale nejedná o primární ani sekundární zálohovací řešení, lze s investicí do nového řešení vyčkat. Cena nové obdobně výkonné knihovny se pohybuje kolem 50 000 bez daně. Další investice v oblasti zálohování nejsou předpokládány. Změna může nastat v případě selhání některého hardwaru. Na základě aktuálních ceníků se tak celkový odhad investic na příštích 5 letech pohybuje mezi 315 000 až 415 000 Kč bez daně.

5 Diskuse

V bodě, kdy jsou veškeré zálohy nastavené, automaticky se provádějí a následně i kontrolují, by se mohlo zdát, že práce s celým zálohovacím systémem v podstatě skončila. To je ovšem milná představa. I s automatizovanými nástroji je nezbytná občasná ruční kontrola všech procesů. Mezitím je nutné pečlivě sledovat emailové zprávy s výsledky jednotlivých úloh. Zde je potřeba upozornit, že pokud nastane problém se zálohou, je velmi pravděpodobné, že problém se netýká přímo samotného zálohovacího softwaru, ale zdroje zálohy, nebo nějakého prvku v cestě, například virtuálního stroje. Údržba se tedy netýká pouze samotných zálohovacích softwarů a hardwarů, ale je úzce spojena s údržbou celé infrastruktury od fyzického hosta přes VMware až po jednotlivé části Veeam Backup. Dále je potřeba kontrolovat a instalovat aktualizace jak bezpečnostní tak funkcionální, a to u všech součástí infrastruktury.

Směr dalšího vývoje zálohování ve firmě je samozřejmě podmíněn jak potřebami firmy, tak ochotou vedení uvolňovat finanční zdroje na tuto problematiku. Je relativně běžné, že v této oblasti neorientovaní lidé, včetně hlavních představitelů firmy, naprosto bagatelizují význam a extrémní důležitost zálohování. Samozřejmě do doby, než nastane podstatný problém. Jedním z budoucích úkolů autora je tedy také komunikace s vedením společnosti a předkládání jasných argumentů založených na této práci, pro uchování životaschopnosti infrastruktury.

Konkrétní možnosti rozvoje se tak například týkají výběru zálohovacího softwaru pro virtuální infrastrukturu po vypršení dosavadního předplatného pro nástroj Veeam. Výběr bude proveden na základě sumarizace výhod a nevýhod všech vhodných možností. Jedním z možných směrů je využití bezplatných řešení. Při bližším pohledu je však zřejmé, že pro rozměry řešené společnosti není použití free verzí zálohovacích programů možné. Hlavní problém je ve většině případů počet virtuálních strojů, které lze v bezplatné verzi zálohovat. Za předpokladu setrvání provozování virtuální infrastruktury na platformě VMware je tak z dosavadních zkušeností a z momentálního pohledu řešení Veeam i do budoucna velice spolehlivé a vhodné.

Další vývoj se týká hardwarového vybavení. Co se týče datového centra, tak zde bude nutné také vytvořit nový výkonnější Veeam Backup Server, podobně jako k tomu došlo v rámci této práce pro lokální zázemí společnosti. V dnešním světě počítačových sítí je jednou z největších hrozeb pro firemní infrastrukturu ransomware. Tento vyděračský virus

jednoduše řečeno napadne síť a zašifruje vše k čemu se dostane. Časové a finanční náklady na nápravu jsou poté obrovské a řešení mnohdy stejně nefunkční. Jako jedna část vhodné ochrany se jeví použití neměnných záloh, respektive neměnných úložišť. Tyto technologie tak patří do plánu rozvoje firemního zálohování.

Výsledné řešení je navrženo na míru dané společnosti. Nicméně vychází z obecně platných principů a postupů, a tak je do jisté míry aplikovatelné i v cizích projektech.

6 Závěr

Výsledkem této diplomové práce je komplexní zálohovací strategie implementovaná v konkrétní středně velké firmě. Zdrojem zálohování jsou veškerá data, která se ve firmě vyskytují. To znamená vše od mobilních zařízení, přes počítače a servery až po automatizované stroje ve výrobě.

Práce je rozdělena na dvě hlavní části a následně na několik větších podkapitol. První část se zaměřuje na teoretické podklady, které jsou následně využity v druhé části. V té se autor zabývá konkrétním praktickým řešením v dané firmě.

Teoretická část práce je rozdělena do čtyř kapitol. První kapitola se zaobírá problematikou zálohování. V první podkapitole se čtenář postupně seznámil s nejběžnějšími příčiny ztráty či poškození dat. Následuje blok s vyjasněním základních pojmů zálohování. Další odstavce objasnily druhy zálohovaných dat, základní pravidla zálohování, zálohovací strategie a aspekty pro správný výběr metody zálohování. V poslední části byly vypsány časté a zbytečné chyby při zálohování.

Druhá kapitola je zaměřena na úložiště, a to z několika úhlů pohledu. V první podkapitole byla uvedena úložiště dle doby pro přístup k danému médiu. Další oddíl vysvětluje techniku RAID pro slučování pevných disků do polí. Následují části uvádějící síťovou architekturu úložišť a úložná média.

V třetí kapitole se autor zabýval metodami zálohování. To obsahuje témata jako jsou metody zpracování dat, způsoby zálohování, druhy záloh, rotace záloh nebo retenční politika.

Poslední kapitola teoretické části se věnuje představení softwaru pro zálohování všech částí infrastruktury. To znamená především řešení pro virtuální infrastrukturu v podobě softwaru od společnosti Veeam. Následně softwaru Acronis pro klonování a zálohu pevných disků například ze strojů ve výrobních halách. Posledním široce využívaným nástrojem je Synology Hyper Backup jakožto řešení pro úložiště Synology. Na konec byly představeny alternativy pro tyto nástroje.

Druhá hlavní část práce se zabývá samotnou realizací v dané firmě. V prvním kroku byla provedena analýza zdrojů, kterými firma disponuje a které nebo pomocí kterých bude zálohování prováděno. To znamená vše od charakteristiky infrastruktury, přes serverové vybavení, datová úložiště až po stroje ve výrobě nebo síťové prvky. Zde také došlo k výběru řešení NAS. Dále zde byla identifikována data, která byla rozdělena na kritická a nekritická.

V této části je také představen stav zálohování ve firmě před vypracováním této práce a následně požadovaný stav. Původní, a tedy nedostatečný stav byl hlavním zdrojem motivace pro tvorbu této práce.

Další blok je zaměřen na samotnou implementaci. Uvádí se zde strategie zálohování pro jednotlivé okruhy zdrojů ve firmě. Následuje konfigurace a ukázka nastavení u vybraných úloh v daném softwaru. Další kapitola řeší úpravu rozmístění zdrojů mezi dvě serverové místnosti v zázemí firmy. V posledních kapitolách práce je provedeno testování realizovaných řešení z pohledu záloh i následné obnovy. Vše je poté uzavřeno diskusí o dalším vývoji a nutné údržbě.

7 Seznam použitých zdrojů

[1] BROOKS, Ryan. Netwrix 2018 IT Risks Report: Summary and Key Takeaways. Netwrix [online]. 29. října 2018 [cit. 2022-02-01]. Dostupné z: https://blog.netwrix.com/2018/10/29/netwrix-2018-it-risks-report-summary-and-key-takeaways/

[2] KLEIN, Andy. Backblaze Hard Drive Stats for 2020. *Backblaze* [online]. 26. ledna 2021 [cit. 2022-02-01]. Dostupné z: https://www.backblaze.com/blog/backblaze-hard-drive-stats-for-2020/

[3] REINSEL, David, John GANTZ a John RYDNING. Data Age 2025: The Digitization of the World From Edge to Core. *Seagate* [online]. Listopad 2018 [cit. 2022-02-01]. Dostupné z: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

[4] REINSEL, David. Zálohování nebo archivace? *Acronis* [online]. 28. července 2016 [cit. 2022-02-01]. Dostupné z: https://www.acronis.cz/zalohovani-nebo-archivace/

[5] IPSEN, Adam. Archive vs Backup: What's the Difference? A Definition Guide. *BackupAssist* [online]. 28. března 2017 [cit. 2022-02-01]. Dostupné z: https://www.backupassist.com/blog/archive-vs-backup-whats-the-difference-a-definition-guide

[6] PALMER, Melissa. What is a disaster recovery plan?. *Veeam* [online]. 22. října 2020 [cit. 2022-02-01]. Dostupné z: https://www.veeam.com/blog/disaster-recovery-plan.html

[7] PURICICA, Cristian-Antonio. Demystifying Recovery Objectives. *Veeam* [online]. 9. října 2017 [cit. 2022-02-01]. Dostupné z: https://www.veeam.com/blog/rto-rpo-definitions-values-common-practice.html

[8] Zálohování dat. *Forensee* [online]. 30. října 2020 [cit. 2022-02-01]. Dostupné z: https://www.forensee.cz/2020/10/30/zalohovani-dat/

[9] 3-2-1 Backup Rule: The Three Steps of Planning Data Keeping. *HandyBackup* [online]. 30. října 2020 [cit. 2022-02-01]. Dostupné z: https://www.handybackup.net/backup_terms/backup_strategy.shtml

[10] Zálohování vždy až na prvním místě. *TotalService* [online]. [cit. 2022-02-01]. Dostupné z: https://www.totalservice.cz/novinky/zalohovani-vzdy-az-na-prvnim-miste-2021-04-06

[11] LOSSCHAERT, Nico. 3-2-1-1-0 Golden Backup Rule. *Veeam community* [online]. 18. ledna 2021 [cit. 2022-02-01]. Dostupné z: https://community.veeam.com/blogs-and-podcasts-57/3-2-1-1-0-golden-backup-rule-569

[12] LISCINSKY, Jakub. 10 Data Backup Mistakes You Can Easily Avoid. *Rescue magazine* [online]. 16. srpna 2021 [cit. 2022-02-01]. Dostupné z: https://intellope.com/10-data-backup-mistakes-you-can-easily-avoid/

[13] The Correct Use of the term Nearline. *IBM* [online]. [cit. 2022-02-01]. Dostupné z: https://www.ibm.com/support/pages/correct-use-term-nearline

[14] Digital Media Operations: Archiving. *BCE* [online]. [cit. 2022-02-01]. Dostupné z: https://www.bce.lu/archiving/

[15] Co je cloudové úložiště?. *Azure Microsoft* [online]. [cit. 2022-02-01]. Dostupné z: https://azure.microsoft.com/cs-cz/overview/what-is-cloud-storage/

[16] KOCMÁNEK, Vít. Přehled všech režimů RAID - rychlejší a bezpečnější ukládání dat. *Zive.cz* [online]. 2. dubna 2003 [cit. 2022-02-01]. Dostupné z: https://www.zive.cz/clanky/prehled-vsech-rezimu-raid---rychlejsi-a-bezpecnejsi-ukladani-dat/sc-3-a-111138/default.aspx

[17] MŮČKA, Jan. RAID disková pole: jaké jsou základní typy a v čem se liší?. *MasterDC* [online]. 12. srpna 2021 [cit. 2022-02-01]. Dostupné z: https://www.master.cz/blog/raid-diskova-pole-jake-jsou-zakladni-typy-a-v-cem-se-lisi/

[18] RAID. *Wikipedia* [online]. [cit. 2022-02-01]. Dostupné z: https://cs.wikipedia.org/wiki/RAID

[19] LEVENS, Skip. What's the Diff: NAS vs. SAN. *Backblaze* [online]. 14. ledna 2021 [cit. 2022-02-01]. Dostupné z: https://www.backblaze.com/blog/whats-the-diff-nas-vs-san/

[20] TECH, RS. Storage Architecture: NAS vs. SAN vs. DAS. *Router-switch.com* [online]. 22. března 2021 [cit. 2022-02-01]. Dostupné z: https://blog.router-switch.com/2021/03/storage-architecture-nas-vs-san-vs-das/

[21] VÁCLAVÍK, Lukáš. Archivovat data do cloudu, na HDD, SSD, DVD, nebo Blu-ray? Co je nejvýhodnější?. *Zive.cz* [online]. 11. ledna 2021 [cit. 2022-02-01]. Dostupné z: https://www.zive.cz/clanky/archivovat-data-do-cloudu-na-hdd-ssd-dvd-nebo-blu-ray-co-je-nejvyhodnejsi/sc-3-a-207888/default.aspx#part=3

[22] Backup Deduplication. *Acronis* [online]. [cit. 2022-02-01]. Dostupné z: https://www.acronis.com/en-us/articles/deduplication/

[23] LOUCKÝ, Milan. Metody šifrování dat. *Chip.cz* [online]. 3. dubna 2018 [cit. 2022-02-01]. Dostupné z: https://www.chip.cz/novinky/metody-sifrovani-dat/

[24] Data Compression. *Barracuda* [online]. [cit. 2022-02-01]. Dostupné z: https://www.barracuda.com/glossary/data-compression

[25] CARDIN, Jay. Flash-To-Flash-To-Cloud (F2F2C): What It Is And Why You Need It. *WEI IT* [online]. 15. října 2019 [cit. 2022-02-01]. Dostupné z: https://blog.wei.com/flash-to-flash-to-cloud-what-it-is-and-why-you-need-it

[26] MAYER, Alex. Backup Types Explained: Full, Incremental, Differential, Synthetic, and Forever-Incremental. *Nakivo* [online]. 6. listopad 2017 [cit. 2022-02-01]. Dostupné z: https://www.nakivo.com/blog/backup-types-explained-full-incremental-differential-synthetic-and-forever-incremental/

[27] Backup Rotation Scheme. *Networx security* [online]. [cit. 2022-02-01]. Dostupné z: https://www.networxsecurity.org/members-area/glossary/b/backup-rotation-scheme.html

[28] Backup Retention Policy: Best Practices for IT Admins and Business Owners. *Spin backup* [online]. 13. ledna 2021 [cit. 2022-02-01]. Dostupné z: https://spinbackup.com/blog/backup-retention-policy-best-practices/

[29] About Veeam Backup & Replication. *Veeam helpcenter* [online]. [cit. 2022-02-01]. Dostupné z: https://helpcenter.veeam.com/docs/backup/vsphere/overview.html?ver=110

[30] Veeam Backup & Replication. *VMware* [online]. 11. května 2020 [cit. 2022-02-01]. Dostupné z: https://kb.vmware.com/s/article/52533

[31] Klonování disku - jak naklonovat disk. *Acronis* [online]. [cit. 2022-02-01]. Dostupné z: https://www.acronis.cz/kb/klonovani-disku/

[32] Synology Drive. *Synology* [online]. [cit. 2022-02-01]. Dostupné z: https://www.synology.com/cs-cz/dsm/feature/drive

[33] VMware vSphere Documentation. *WMware Docs* [online]. [cit. 2022-02-01]. Dostupné z: https://docs.vmware.com/en/VMware-vSphere/index.html

[34] Installing Veeam Backup & Replication. *Veeam Help Center* [online]. [cit. 2022-02-01]. Dostupné z: https://helpcenter.veeam.com/docs/backup/hyperv/install_vbr.html?ver=110

[35] How SureBackup Works. *Veeam Help Center* [online]. [cit. 2022-02-01]. Dostupné z: https://helpcenter.veeam.com/docs/backup/vsphere/surebackup_hiw.html?ver=110

[36] DE GUISE, Preston. *Data protection: ensuring data availability* [online]. Boca Raton: Auerbach Publications, [2017] [cit. 2021-11-18]. ISBN 978-131-5169-620. Dostupné z: https://www-taylorfrancis-com.infozdroje.czu.cz/books/mono/10.1201/9781315169620/data-protection-preston-de-guise?context=ubx&refId=f5d75be4-f337-46ba-ac76-e2e70c2237b1

[37] *Průměrný plat na pozici Systémový administrátor* [online]. 2022 [cit. 2022-03-14]. Dostupné z: https://prumerneplaty.cz/pozice/systemovy-administrator