

Analýza služeb datových úložišť

Jan Zálešák

Bakalářská práce
2016



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

akademický rok: 2015/2016

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: Jan Zálešák
Osobní číslo: A12255
Studijní program: B3902 Inženýrská informatika
Studijní obor: Informační technologie v administrativě
Forma studia: prezenční

Téma práce: Analýza služeb datových úložišť
Téma anglicky: An Analysis of File Storage and Synchronisation Services

Zásady pro vypracování:

1. Vypracujte obecnou literární rešerši na dané téma.
2. Provedte komplexní podrobnou analýzu, popis a identifikaci služeb jednotlivých datových úložišť.
3. Provedte porovnání nejběžnějších dostupných služeb, jejich možností nasazení v podnicích a podnikové administrativě.
4. Provedte analýzu licenčních podmínek, rizik a přínosů.

Rozsah bakalářské práce:

Rozsah příloh:

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

Seznam odborné literatury:

1. SODOMKA, Petr a Hana KLČOVÁ. Informační systémy v podnikové praxi: Petr Sodomka, Hana Klčová. 2., aktualiz. a rozš. vyd. Brno: Computer Press, 2010, 501 s. ISBN 978-80-251-2878-7.
2. BASL, Josef a Roman BLAŽÍČEK. Podnikové informační systémy: podnik v informační společnosti. 3., aktualiz. a dopl. vyd. Praha: Grada, 2012, 323 s. ISBN 978-80-247-4307-3.
3. POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. Business intelligence v podnikové praxi. 1. vyd. Praha: Professional Publishing, 2012, 276 s. ISBN 978-80-7431-065-2.
4. LABERGE, Robert. Datové sklady: agilní metody a business intelligence. 1. vyd. Brno: Computer Press, 2012, 350 s. ISBN 978-80-251-3729-1.
5. LACKO, Ľuboslav. Osobní cloud pro domácí podnikání a malé firmy. 1. vyd. Brno: Computer Press, 2012, 270 s. ISBN 978-80-251-3744-4.
6. TURBAN, Efraim. Decision support and business intelligence systems. 9th ed. Boston: Prentice Hall, c2011, xxii, 696 s. ISBN 978-0-13-610729-3.

Vedoucí bakalářské práce:

doc. Ing. Roman Šenkeřík, Ph.D.

Ústav informatiky a umělé inteligence

Datum zadání bakalářské práce:

5. února 2016

Termín odevzdání bakalářské práce:

1. června 2016

Ve Zlíně dne 5. února 2016

doc. Mgr. Milan Adámek, Ph.D.
děkan



Ing. Miroslav Matýsek, Ph.D.
ředitel ústavu

Prohlašuji, že

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

Prohlašuji,

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

Ve Zlíně, dne 30. 5. 2016

Zálesák
.....
podpis diplomanta

ABSTRAKT

Bakalářská práce je zaměřena na problematiku datových úložišť a jejich služeb. V teoretické části jsou popsány datová úložiště, disková pole a cloud computing a jeho modely nasazení, výhody a nevýhody, bezpečnost, rizika a cloud platformy. Poslední část teorie je věnována big datům. Praktická část je zaměřena na popis a porovnání osobních cloudů, jejich výhody a nevýhody obecně a zvlášť. Poslední část je analýza cloudových úložišť pomocí TOPSIS analýzy.

Klíčová slova: datová uložistiště, výhody, nevýhody, bezpečnost a rizika, služby uložistišť, cloud computing, Big Data, analýza.

ABSTRACT

This Bachelor thesis is focused on an issue of file storages and their services and it is divided into two parts. In Theory part, there are discribed file storages, disc arrays and cloud computing deployment models and cloud computing's advantages and disadvantages, including security and risks and cloud platforms. The last part of the theory is dedicated to Big data. The practical part is focused on description and comparison of private cloud services and their advantages and disadvtages in general and as an individual. The last part of the practical part is TOPSIS analysis of specific cloud services.

Keywords: file storage, advantages, disadvantages, security, risks, file storage services, cloud computing, big data, analysis

Tímto bych rád poděkoval svému vedoucímu práce panu doc. Ing. Romanovi Šenkeříkovi, Ph.D. za jeho ochotu, čas, rady a připomínky, které mi pomohly při zpracovávání bakalářské práce. Dále bych chtěl poděkovat mé rodině za jejich podporu během mého studia.

Prohlašuji, že odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

ÚVOD.....	9
I TEORETICKÁ ČÁST.....	10
1 DATOVÁ ÚLOŽIŠTĚ	11
2 DAS, NAS A SAN	12
2.1 ROZDÍL MEZI NAS A SAN.....	12
2.2 DAS – DIRECT-ATTACHED STORAGE	12
2.3 STORAGE.....	12
2.4 NAS – NETWORK-ATTACHED STORAGE	13
2.4.1 NAS v praxi.....	14
2.5 SAN - STORAGE AREA NETWORK	14
2.6 NAS NEBO SAN.....	15
3 RAID.....	17
3.1 RAID - BĚŽNÉ	17
3.2 RAID – SPECIFICKÉ.....	18
3.3 RAID – NESTANDARDNÍ TYPY	20
4 CLOUD COMPUTING	22
4.1 DEFINICE.....	22
4.2 MODEL NAsAZENÍ	23
4.2.1 Veřejný (Public cloud computing)	23
4.2.2 Soukromý (Private cloud computing)	23
4.2.3 Hybridní (Community cloud computing).....	23
4.2.4 Komunitní (Cominity cloud computing).....	23
4.3 DISTRIBUČNÍ MODEL	23
4.3.1 IaaS (Infrastructure as a service, infrastruktura jako služba)	23
4.3.2 PAAS (Platform as a Service, platforma jako služba)	24
4.3.3 SAAS (Software as a Service, software jako služba)	24
4.4 VÝHODY A NEVÝHODY CLOUD COMPUTINGU.....	25
4.4.1 Pohled z ekonomického hlediska	25
4.4.2 Pohled z technického hlediska	25
4.5 ZABEZPEČENÍ A RIZIKA	26
4.5.1 Identity	27
4.5.2 Informace	28
4.5.3 Infrastruktura.....	29
4.6 OSOBNÍ ÚDAJE.....	29
4.7 RIZIKA.....	29
4.7.1 Odcizení nebo zneužití dat	29
4.7.2 Ztráta dat	30
4.7.3 Hrozby.....	30
4.7.4 DoS a DDoS útoky.....	31
5 CLOUDOVÉ PLATFORMY	32

5.1	WINDOWS AZURE	32
5.2	GOOGLE CLOUD PLATFORM	33
5.3	OSTATNÍ.....	34
5.3.1	Amazon	34
5.3.2	Salesforce	34
6	BIG DATA	35
6.1	BIG DATA V PRAXI	35
6.2	ANALÝZA - 4V (5V).....	36
6.3	VIZUALIZACE A IR (INFORMATION RETRIEVAL)	37
6.4	NOSQL DATABÁZOVÉ KONCEPTY	38
6.5	APACHE HADOOP	40
II	PRAKTICKÁ ČÁST	41
7	CLOUDOVÁ ÚLOŽIŠTĚ	42
8	VÝHODY A NEVÝHODY CLOUDOVÝCH ÚLOŽIŠŤ	43
8.1	VÝHODY.....	43
8.2	NEVÝHODY	43
8.3	SYNCHRONIZACE DAT	44
9	PRIMÁRNĚ OSOBNÍ CLOUDY	45
9.1	GOOGLE DRIVE (DISK).....	46
9.2	MICROSOFT ONEDRIVE (DŘÍVE SKYDRIVE).....	47
9.3	APPLE ICLOUD	49
9.4	DROPBOX	50
9.5	MEGA.....	51
10	VÝBĚR ÚLOŽIŠTĚ – TOPSIS ANALÝZA.....	53
10.1	POSTUP.....	53
10.2	ZHODNOCENÍ.....	57
11	ZÁVĚR.....	58
	SEZNAM POUŽITÉ LITERATURY.....	59
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....	62
	SEZNAM OBRÁZKŮ	63
	SEZNAM TABULEK.....	64
	SEZNAM GRAFŮ	65

ÚVOD

Již z názvu bakalářské práce je jasné, že hlavní zaměření této bakalářské práce je na fenomén cloud computingu a konkrétně na jeho služby. Tyto služby jsou čím dál tím více využívány a to nejen ve sféře podniků, ale i mezi běžnými uživateli. Je to způsobeno rychlým rozvojem technologií a s tím spojených služeb. Jako příklad lze uvést růst mobilního průmyslu a s tím souvisí i růst sociálních sítí, které jsou vlastně jeden obrovský cloud.

Hlavně díky dostupnosti a jednoduchosti je cloud computing tak oblíbený. Běžný uživatel nemusí investovat čas ani finanční prostředky do budování své vlastní infrastruktury, pokud tedy nechce. Co se týče společností, většina z nich spoléhá na nějaké cloudové řešení. Ne všechny ale rozumí tomu, co vše takové nasazení obnáší, proto by si každá společnost měla předem zjistit výhody a nevýhody, jako jsou možná rizika spojená s bezpečností.

Bakalářská práce se skládá z části teoretické a části praktické. V první až třetí kapitole teoretické části se věnuje datovým úložištím, jako je DAS, NAS a SAN, a také různým typům RAID.

V další kapitole se nachází cloud computing, jeho modely nasazení, distribuční modely, výhody a nevýhody z ekonomického a technického hlediska a nakonec zabezpečení a rizika, jako je odcizení, zneužití nebo ztráta dat. Vše je uzavřeno pohledem na několik cloud platform.

Big data jsou poslední kapitolou praktické části, práce se soustřeďuje na jejich nasazení v praxi, hlavní problémy při zpracování dat a vizualizaci. Také jsou zmíněny NoSQL databáze a aplikace Apache Hadoop.

Praktická část pojednává o výhodách a nevýhodách cloudových úložišť. Je zde zmíněna synchronizace dat. Poslední kapitola bakalářské práce je popis a analýza nejběžnějších cloudových úložišť od velkých IT společností jako je Apple, Google a Microsoft.

I. TEORETICKÁ ČÁST

1 DATOVÁ ÚLOŽIŠTĚ

Objem dat se neustále zvyšuje, jak u fyzických osob, tak i u společností. Důvodů je více, ale asi ten největší je rozvoj technologií a to nejen informačních, ale všech a tudíž je nutné ukládat více dat. V minulosti se například dopisy nebo dokumenty zálohovaly do složek, a pak se někde uskladnily a zabíraly fyzické místo, dnes se čím dál tím více uskládají do datových úložišť a to nejen tyto soubory, ale i knihy, videa, fotky a hudba. Důležitou součástí dnešních úložišť je to, aby k nim měli lidé přístup pomalu odkudkoliv pomocí internetu a z mnoha jiných zařízení než jen z počítače, jako třeba z mobilních telefonů nebo tabletů. Tyto služby se nazývají cloud a všechny velké technologické společnosti, jako je Apple, Microsoft a Google mají cloudové služby. Jak již bylo zmíněno výše, dříve byly dokumenty a jiné soubory fyzicky skladovány a tudíž zabíraly hodně místa, díky úložištím je možné skladovat mnohem větší množství informací, což je velmi užitečné, a také jde vše velmi rychle dohledat a nemusíme se ručně přehrabovat v labyrintu fyzických dokumentů. Aby ale vše bylo přístupné, musíme si uvědomit, že jako vše, tak i úložiště trpí fyzickým opotřebením, proto je velmi důležité data zálohovat a to pokud možno na více než jednom místě. Také čas od času je nezbytné data v úložištích protřídit.

2 DAS, NAS A SAN

2.1 Rozdíl mezi NAS a SAN

Rozdíl mezi NAS a SAN je v tom, že NAS je v základu jeden disk, který umožňuje centrální sdílení dat (nikoliv databází), zato SAN představuje vysokorychlostní síť, se serverem komunikuje většinou pomocí optického kabelu. Vlastní síť má tvořenou jistým (podle potřeby) množstvím navzájem propojených úložných prostorů.

Pokud chceme porovnat tyto dvě technologie, zjistíme následující, NAS je levnější, mnohem lépe se s ním pracuje a instalaci i provoz by měl bez potíží zvládnout i běžný uživatel. Výhoda SAN je v lepší propustnosti, a také v tom, že se jedná o více zařízení, jež z hlediska bezpečnosti a zálohování jsou mnohem lepší řešení, ale také si za SAN hodně připlatíme, tudíž se vyskytuje většinou jen ve větších firmách.

2.2 DAS – Direct-Attached Storage

Jedná se o nejobyčejnější způsob jak ukládat data a to tak, že jsou uložena přímo na serveru, tedy na pevných discích v něm připojených. Jako vše to má své výhody a nevýhody. Za největší výhodu lze považovat rychlý přístup k datům. Nevýhoda je komplikovanější sdílení dat s dalšími servery. [1]

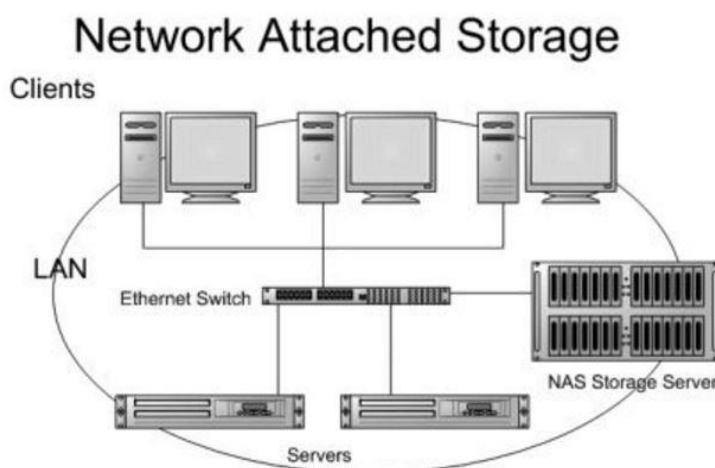
2.3 Storage

Pokud chceme vytvářet velká úložiště, budeme muset vytvořit cluster servery (skupina serverů, jež vykonávají stejné činnosti se stejnými daty, v případě poruchy se navzájem zastoupí a rozkládají si mezi sebou zátěž), a také je potřeba, aby potřebná data byla sdílena (například e-maily, zákaznické weby atd.). K tomuto slouží specifická zařízení, jež jsou k tomu, aby tato data uchovávala – storage nebo také disková pole. Diskové pole (storage) je zařízení s velkým množstvím disků, o obsah disků se stará a zpřístupňuje jejich obsah dalším. Data nejsou umístěna na jednotlivých serverech (např. webových), ale jsou umístěna na zmíněných diskových polích, odkud se s nimi pracuje, a tak dojde k oddělení logiky zpracování a uchovávání dat. Systém potom může mít jen malý pomalý disk, kde bude uložen operační systém a software a někde jinde storage s rychlými velkými disky, jež jsou i spolehlivější a jsou na nich uložena potřebná data. Storage je se systémem spojen pomocí nějaké počítačové sítě.

2.4 NAS – Network-Attached Storage

Srdcem NAS je co nejlepší datové úložiště, jež poskytne důležitým informacím bezproblémovou dostupnost (nikdy nelze garantovat 100%), a také dostatečnou ochranu. Velmi důležitou stránkou profesionálních NAS služeb je také nepřetržitý monitoring.

NAS blade/ NAS gateway, tak je většinou označován počítač, jehož operační systém je značně zjednodušen a zaměřen na nástroje pro sdílení, tedy pro síťové protokoly a ověřování atd. a slouží ke sdílení dat v rámci velké NAS sítě. Jako příklad takového operačního systému lze uvést třeba Microsoft Storage Server. Následně se tyto dvě části spojí (NAS gateway a NAS síť) a vytvoří nám jeden celek, a to komplexní NAS řešení. Výsledku je možné dosáhnout několika způsoby. Prvním je varianta „all in one“ – což znamená počítač, který má OS a rovnou i vlastní disky, na která se data ukládají (file server). Druhý způsob je NAS gateway natvrdo integrovaný do diskového pole. Je možné také počítač oddělit a připojit skrze iSAN/MultiPaths SAN. [1]



Obr. 1 NAS

Zdroj: www.svethardware.cz

Ať už je NAS řešený jakkoliv, je bezpodmínečně nutné, aby zvládal rychlé zálohování a obnovu dat, a také replikaci. Toto by měl být naprostý základ každého NAS. Každopádně mnohé NAS boxy toho zvládnou o mnoho více a nakonec se mohou stát velmi zajímavým doplňkem multimediálního vybavení každé domácnosti.

2.4.1 NAS v praxi

V poslední době se NAS boxy těší čím dál tím větší oblibě a tato storage technologie již není jen doménou sektoru firem, ale získává si stále větší popularitu mezi běžnými uživateli, kteří volí právě mezi NAS nebo externími harddisky.

Je zde i zlatá střední cesta v podobě „hybridů“ NAS a externích HDD¹, jako příklad lze uvést NAS External Hard Drive. Tyto disky disponují i síťovou konektivitou, a tak dokáží pracovat jako nezávislé jednotky na rozdíl od jiných úložišť. Je to vcelku dobrý kompromis, který je dostupný a je i relativně levný.

Oproti externím harddiskům nabízí NAS vcelku hodně výhod. Pro představu lze v jednom celku spojit zálohování a tiskový + file server. Dále WebStation (Hewlett Packard + MySQL), Multimedia Server, FTP² server, iTunes Server. Jako další velká výhoda je, že dokáží pomocí funkce Download Station stahovat data z internetu i při vypnutém PC. Pokud porovnáme ceny externího harddisku a NAS vyjde nám, že za trochu vyšší cenu dostaneme mnohem lepší úložiště pro naše data.

Zálohování je velmi jednoduché a poradit by si s tím měl každý. Pokud je potřeba zálohovat data, která se přímo nachází na tomto síťovém úložišti, stačí nainstalovaný program jako je Data Replicator a poté připojit třeba flash disk. Pokud máme rozsáhlejší síť, jde dle potřeby zálohovat rovnou z jednoho serveru na druhý dle potřeby, pokud máme USB X Mass Storage lze to zvládnout doslova stiskem jednoho tlačítka.

Nákladnější modely nabízí navíc vícenásobné stupně pro zálohování – jako je kombo Net-Bak Replicator (rychlá záloha nebo obnova PC) + Remote Replication (replicace mezi NAS za pomoci internetu) + Q-RAID 1 mirroring.

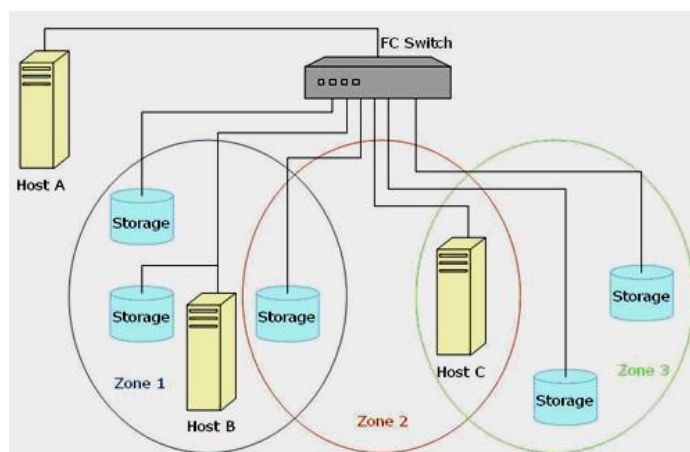
Samozřejmě je nezbytná i bezpečnost a o tu se starají velmi sofistikované utility, jež umožňují například HTTPS³, FTP skze SSL⁴ nebo TLS⁵ Encrypted Network Backup atp. [1]

2.5 SAN - Storage Area Network

Tato storage technologie je vcelku velmi nákladná a proto se s ní setkáme jen ve větších firmách a institucích (banky). Díky rychlému vývoji nicméně ceny klesají, proto si SAN v dnešní době mohou dovolit už i střední firmy. Problém je v tom, že hodně firem nevidí značné výhody k tomu, aby investovali do této technologie.

Asi největší výhodou SAN je podpora Disaster Recovery sites. V roce 1992 přišel share.org společně s firmou IBM⁶ s definicí tzv. „sedmi vrstev pro disaster recovery“. První vrstva je, když situace je úplně katastrofální, žádná záloha dat a nikdo ani neřeší plány na obnovu a tudíž není možné odhadnout čas pro případnou obnovu ztracených dat. Oproti tomu sedmá vrstva (Highly automated, business-integrated solution) znamená plnou automatizaci všech relativních procesů, a také skvěle propracovaný plán pro případ jakékoli jiné havárie nebo jiné události. Samozřejmě i zde platí, že čím je vrstva vyšší, tím vyšší jsou náklady nutné k jejímu dosažení. Jako další výhodu lze uvést podporu „no single point of failure“ architektury, což znamená, že zařízení s tímto označením (jako příklad síťové přepínače IP⁸) by mělo být teoreticky zabezpečeno vůči výpadkům, a také vůči útočníkům (poté lze anglický výraz brát doslovně jako „žádná skulina systému“).

Použitím optického kabelu (tloušťka 9 mikronů) lze data přenášet na velké vzdálenosti a tak se jedna síť SAN může rozkládat na mnoha kilometrech. Pokud tedy hledáme maximálně výkonnou dedikovanou síť, která je zejména určená pro čtení, zápis a ochranu dat je SAN to nejlepší co lze pořídit. [1]



Obr. 2 SAN

Zdroj: www.svethardware.cz

2.6 NAS nebo SAN

Toto rozhodnutí závisí na tom, kolik jsme do datového úložiště ochotni investovat, a také na tom, jak vážně to myslíme. Pokud se jedná o velké společnosti, je SAN samozřejmostí, ale

pokud se jedná o menší firmy nebo jen jednotlivce je výhodnější NAS. A když chce člověk jen něco, kde by uložil nepříliš důležitá data, tak klidně postačí jen externí disk, který nemá tolik výhod, ale v dnešní době jsou jeho pořizovací ceny vcelku nízké.

3 RAID

RAID (Redundand Array of Independent/Inexpensive Disks – Vícenásobné diskové pole nezávislých/levných disků). Díky snížení pořizovací ceny se disková pole dostala už i do sféry běžných uživatelů a nejsou již doménou pouze v oblasti serverů, co se týká základních konfigurací. Již v době, kdy byly řadiče Ultra ATA/66, bylo možné zapojit více disků do konfigurace, jež nám umožnily zvýšit spolehlivost, rychlost nebo rovnou obojí.

3.1 RAID - Běžné

RAID 0

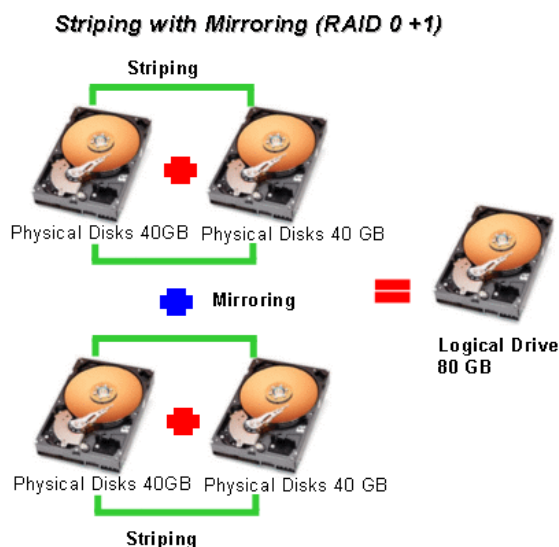
Raid 0 je první v číselném pořadí, jedná se o dva nebo více disků, jež jsou zapojeny do tzv. **strippingu**. U tohoto módu řadič rozděluje data na „proužky“, které se poté střídavě rozdělují mezi disky. Tudíž například 2GB⁷ soubor se střídavě zapíše na 2 disky, na každém bude 1024MB. Pomocí tohoto dělení je možné dosáhnout výrazně vyššího výkonu (zdaleka ne dvojnásobného). Při poruše bude dvakrát vyšší pravděpodobnost, že přijdeme o svá data, protože v případě havárie jednoho disku automaticky přijdeme i o data na druhém disku (pokud se nějak nepodaří data zachránit, což je ve většině případů nemožné). Další omezení tkví v tom, že je nejlepší použít stejné typy disků, nebo alespoň disky o stejné kapacitě, pokud tak neučiníme, bude velikost pole omezena na velikost dvojnásobku kapacity menšího z disků.

RAID 1

Další druh diskového pole, jež je označované jako **mirroring** – zrcadlení. Jak je již z názvu zřejmé, jeden ze dvou disků se použije pouze pro zálohu – to znamená, že na oba disky se zapisují identická data. RAID 1 se používá pro zvýšení bezpečnosti našich dat a je jen velmi nepravděpodobné, že se porouchají oba dva disky najednou. Omezení je stejné jako u RAID 0, a to, že je vhodné znovu použít disky o stejné rychlosti a kapacitě, aby se vzájemně neomezovaly.

RAID 0+1 (RAID 01)

Toto pole kombinuje výhody dvou předešlých polí, je tedy bezpečné a rychlé. Pro jeho sestavení je potřeba minimálně 4 pevných disků, poněvadž vzniklé pole RAID 0 se bude dále zrcadlit.



Obr. 3 RAID 01

Zdroj: www.svethardware.cz

RAID 10 (1+0)

Další typ RAID 10 je vcelku podobný předchozímu RAIDu 01, jen při sestavování se postupuje naopak. Jako první se budou obě dvojice disků zrcadlit a až poté se nad nimi vytvoří stripping RAID 0. Toto spojení poskytuje nejvyšší výkon u zabezpečených typů polí, a také je odolné vůči ztrátě až poloviny disků, jež jsou připojeny. [2]

JBOD (Spanning)

Mimo klasická RAID pole mohou řadič (nebo OS) podporovat tzv. **spanning**, jedná se o klasické spojení několika pevných disků do logického celku. Data jsou ukládána tak, že nejprve jsou uložena na první disk a až se tento první disk zaplní, data se budou ukládat na druhý disk, a tak dále. Spanning lze nazvat také pomocí zkratky **JBOD** (Just a Bunch Of Disks). Výhoda spočívá v tom, že pro vytvoření tohoto řešení není potřeba mít stejně velké disky, protože data nejsou ukládána prokládaně. Toto řešení se vyplatilo spíše v minulosti, protože umožňuje ukládat opravdu velké soubory, oproti tomu dnes při discích o kapacitách v řádu terabajtů, je to již poměrně zbytečné řešení. [2]

3.2 RAID – specifické

RAID 2

V podstatě RAID 0 s přidanou korekcí v podobě Hammingova kódu, jedná se o jednu z norm ECC⁸ (Error Checking and Correction). Nevýhoda je, že je nutná podpora ze strany disku a je vcelku nákladný na pořízení.

RAID 3

Také využívá stripping, dále také sofistikovaný způsob pro ukládání paritních informací, jež jsou vypočítávány za pomoci logické operace XOR⁹ (exkluzivní OR) neboli nonekvivalence. Ukládání těch informací, které při havárii jednoho disku umožňují rekonstruovat veškerá dříve uložená data, jež jsou poté realizována na další disk, který je k tomu vyčleněn. Výhodou je vyčlenění jen jednoho disku, za to u jiných typů polí to může být vyšší počet, ale zároveň se jedná i o největší nevýhodu daného řešení. Paritnímu disku je přezdíváno „Bottleneck“ (Hrdlo láhve), poněvadž je nejslabším místem celého spojení a také nejvytíženějším. Paritní informace se na něj musí zapisovat za jakýchkoliv okolností, a to znamená, že jeho životnost bude značně omezena v důsledku značného opotřebení. [2]

RAID 4

V mnoha ohledech se podobá předešlému RAID 3, rozdíl je ten, že data nejsou stripována po bitech, ale rovnou po celých blocích a paritní informace také tak, tím se dosahuje velkého výkonu při čtení velkého množství malých souborů. To se hodí zejména u databázových systémů, kde RAID 3 není nejvýhodnějším řešením.

RAID 5

Řeší problém RAID 4 s úzkým hrdlem láhve, zde již žádný disk navíc nenajdeme, paritní informace jsou ukládány střídavě na všechny disky po celém poli, ale vždy na jiné datové úložiště, než na kterém jsou originální data zapisována. Umožňuje to vyšší výkon spojení a tím i jeho velkou oblibu u většiny uživatelů. Nevýhoda je v tom, že diskové pole odolává jen poruše jednoho z připojených disků.

RAID 6

Tenhle typ přináší mnohem vyšší odolnost vůči výpadkům, protože parita je vypočítávána hned dvakrát (vždy jiným způsobem) a poté opět uložena mezi všechny disky, jež jsou připojeny. V případě výpadku není problém vše obnovit do stavu, v němž to bylo. Rychlost čtení je ekvivalentní RAID 5, ale u zápisu je dosažený výkon mnohem menší a to díky tomu,

že je potřeba dvakrát vypočítávat paritní informace. Asi největší nevýhodou pak je, že je nutné použít specializovaný a velmi často také drahý řadič. [2]

3.3 RAID – Nestandardní typy

Mimo již zmínění typy diskových polí existují i další, které většinou nejsou považovány za obecně přijatý standart a mohou být například jen registrovanou obchodní značkou společnosti, která tuto technologii uvedla na trh.

RAID 7

Za vznikem tohoto RAIDu stojí společnost Storage Computer Corporation. Princip vychází z RAID 3 a 4, plus přidává vyrovnávací paměť cache. To znamená, že systém je poté mnohem rychlejší a to jak při čtení, tak při zápisu. Bohužel všechny výhody jsou negovány tím, že pořizovací cena je příliš vysoká a to hlavně díky použití speciálního procesoru, který je určen pro složité výpočty v rámci celého pole v reálném čase.

RAID 30 (3+0)

Jedná se o další možnost, jak nestandardně spojit disky. Stripované větve tvoří zrcadla, ale jsou vzájemně propojeny pomocí RAID 3. Pokud zvolíme opačný postup, získáme RAID 03 (0+3).

RAID 100 a další

Existuje mnohem více různých kombinací, jako jsou například RAID 60, 53, 51 nebo 50. Zdaleka nejzajímavější z těchto řešení je poté RAID 100 (10+0), který si lze představit jako diskové pole RAID 10, nad nímž je vytvořeno další prokládané RAID 0 pole.

Hot-Spare

Jedná se o technologii, díky které lze minimalizovat časový interval, během něhož je pole „oslabeno“. Pokud dáme příklad třeba RAID 5, kterému se porouchá jeden z disků, nestane se v podstatě nic, činnost systému bude pokračovat dále a bezpečnost dat zůstane neovlivněna. Mohlo by se zdát, že tato situace je nereálná, ale opak je pravdou, stačí třeba vadný zdroj a přepětí se postará o zbytek.

RAID (5) ale ustojí jen výpadek jednoho z disků. Pokud se v době než dorazí technik k serveru, aby vše opravil, vypadne další disk, data budou ztracena. Hot-Spare ale umožňuje automaticky připojit k systému záložní pevný disk, kam budou následně chybějící data dopočítána. Tento záložní disk je rovněž možné sdílet mezi různými druhy polí, která jsou v daném počítači přítomná. [2]

4 CLOUD COMPUTING

Cloud computing je na internetu založený model vývoje a používání počítačových technologií. Lze ho také charakterizovat jako poskytování služeb či programů servery dostupnými z internetu s tím, že uživatelé k nim mohou přistupovat vzdáleně, např. pomocí webového prohlížeče nebo klienta elektronické pošty. Za předpokladu, že služba je placená, uživatelé neplatí za vlastní software, ale za jeho užití.

Principem u služeb a produktů v cloud computingu je to, že uživatelé propůjčují výpočetní výkon serverů. V mnoha případech se tak děje formou specializovaných aplikací, jejichž nabídka se pohybuje od kancelářských aplikací přes systémy pro distribuované výpočty až po operační systémy provozované v prohlížečích, jakými jsou např. eyeOS, Cloud či iCloud. [3]

Asi nejlepším příkladem pro běžného uživatele jsou e-mailové schránky (Gmail atd.). Pomocí webové aplikace k nim má člověk přístup odkudkoliv, kde je přístup k internetu. Další příklad jsou cloudová úložiště, jež jsou více rozebrány níže, jako jsou Google Drive, již zmíněný iCloud, OneDrive, Dropbox nebo Mega atd. a také obrázkové databanky, jako je například Flickr. Nakonec sem řadíme různé kancelářské balíky, online hry, a také sociální sítě (Facebook, Google+, Twitter, Instagram atp.)

4.1 Definice

Je mnoho definic co je to cloud computing, ale jako u všech definic, žádná není dokonalá a mnoho je jich špatných nebo nepřesných, Podle mě se nejvíce blíží ideálu definice NIST¹⁰ (National Institute of Standards and Technology), která zní: „ *cloud computing je model umožňující pohodlný, na požádání dostupný síťový přístup ke sdílenému fondu konfigurovatelných výpočetních zdrojů (jako jsou síť, servery, úložiště, aplikace a služby), které mohou být poskytnuty nebo uvolněny při minimálním úsilí pro správu nebo poskytovatele služeb. Tenhle cloud model podporuje dostupnost a skládá se z pěti základních charakteristik (samoobsluhou na požádání, širokým přístupem k síti, sdružováním prostředků, rapidní pružností a měřením služby), dále třemi modely služeb (SaaS – software jako služba, PaaS - platforma jako služba, IaaS - infrastruktura jako služba).* [4]

4.2 Model nasazení

4.2.1 Veřejný (Public cloud computing)

Označován jako klasický model Cloud Computingu. U tohoto modelu je široké veřejnosti nabídnuta a poskytnuta výpočetní služba. Za veřejný cloud lze považovat například Skype nebo Seznam.cz, protože je určen pro velký počet klientů nebo dokonce celou populaci a to znamená, že by měl pro všechny poskytovat stejné nebo velmi podobné funkce a služby.

4.2.2 Soukromý (Private cloud computing)

Někdy se nazývá též vnitřní oblak (internal cloud). Tento cloud je pouze pro organizace, a to pro organizaci samotnou, nebo pro organizace třetí strany. Výhoda tohoto modelu spočívá v jeho bezpečnosti a spolehlivosti. Kritika spočívá v menším ekonomickém prospěchu a ve využití možností cloud computingu, organizace si musí aplikace koupit a spravovat samy.

4.2.3 Hybridní (Community cloud computing)

Jak již název napovídá, jedná se o cloud, který kombinuje jak soukromé cloudy, tak i veřejné. I když ale vystupují navenek jako jeden jediný cloud, pravda je, že jsou propojeny pomocí standardizačních technologií.

4.2.4 Komunitní (Community cloud computing)

U tohoto modelu je infrastruktura cloudu sdílena mezi organizacemi nebo skupinou lidí, kteří ji využívají. Tyto organizace pak může spojovat množství věcí, jako je třeba bezpečnostní politika nebo stejný obor zájmu. [5]

4.3 Distribuční model

Zaměřuje se na to, co je v rámci služby nabízeno, ať už se jedná o software nebo hardware nebo jejich kombinace.

4.3.1 IaaS (Infrastructure as a service, infrastruktura jako služba)

V tomto případě se poskytovatel služeb zavazuje poskytnout infrastrukturu. Typicky se jedná o virtualizaci. Hlavní výhodou tohoto přístupu je to, že se o veškeré problémy s hardwarem stará poskytovatel. Na druhou stranu je někdy velice těžké toto akceptovat vzhledem k tomu, že hardware se bere jako něco, co vlastníme, na co můžeme sáhnout a jsme za to zodpovědní. IaaS je vhodné pro ty, kteří vlastní software (či jejich licence) a nechtějí se starat

o hardware. Příkladem IaaS jsou Amazon WS, Rackspace nebo Windows Azure. Zkratka IaaS může také znamenat integrace jako služba (z „Integration as a Service“). [5]

Výhody IaaS:

- Počáteční náklady jsou nízké
- Je možné si zvolit prostředí pro chod

4.3.2 PAAS (Platform as a Service, platforma jako služba)

Poskytovatel v modelu PaaS poskytuje kompletní prostředky pro podporu celého životního cyklu tvorby a poskytování webových aplikací a služeb plně k dispozici na internetu, bez možnosti stažení softwaru. To zahrnuje různé prostředky pro vývoj aplikace jako IDE¹¹ nebo API¹², ale také např. pro údržbu. Nevýhodou tohoto přístupu je proprietární uzamčení, kdy může každý poskytovatel používat např. jiný programovací jazyk. Příkladem poskytovatelů PaaS jsou Google App Engine nebo Force.com (Salesforce.com). [5]

Výhody PaaS:

- Výkon podle aktuální potřeby
- Neplatí se za extrém, ale za průměrnou spotřebu

4.3.3 SAAS (Software as a Service, software jako služba)

Aplikace je licencována jako služba, která je pronajímána uživateli. Uživatelé si tedy kupují přístup k aplikaci, ne aplikaci samotnou. SaaS je ideální pro ty, kteří potřebují jen běžný aplikační software a požadují přístup odkudkoliv a kdykoliv. Příkladem může být známá sada aplikací Google Apps nebo v logistice známý systém Cargopass.

Výhody SaaS:

- Počáteční náklady jsou nízké
- Vcelku časté aktualizace softwaru

4.4 Výhody a nevýhody Cloud computingu

4.4.1 Pohled z ekonomického hlediska

- **Velmi nízké investice** – není třeba pořizovat celou infrastrukturu. Co se týče provozních výdajů, je nutné si uvědomit, že se zvýší naše internetová konektivita (upload a download dat) a to díky tomu, že infrastruktura se nachází jinde.
- **Žádné skryté výdaje** – pokud máme vlastní cloudové řešení musíme se o něj starat, například dokoupit disk do diskového pole nebo přeinstalovat server po nečekaném pádu atd. Platíme za službu poskytovateli a ten se o vše musí postarat.
- **Nezávislost na IT zaměstnancích** – u menších firem, které nemají IT oddělení je to vcelku velká výhoda, protože se tak dá ušetřit. Nebo si najmout externí společnost.
- **Možnost dynamicky měnit kapacitu** – u velkých poskytovatelů je kapacita neomezená, ale je obrovská. Nevýhoda je, když chce uživatel snížit kapacitu, poněvadž některé firmy si za to účtují penále, nikoliv však u generačního navyšování kapacity. Je také možné dosáhnout na výhodnější nabídku. Některé společnosti (například e-shopy) využívají služeb Cloud computingu jen sezoně a to tak, že třeba v období kolem Vánoc (listopad, prosinec) si pronajmou vyšší kapacity než v průběhu roku.
- **Cena** – ta může být někdy vyšší než pokud bychom si Cloud computing realizovali sami a to hlavně co se týče dlouhodobého hlediska a to u infrastruktury, kde je možné dynamicky měnit její kapacitu, a také mít garantovanou dostupnost (většinou je to 99,9%), kdy pokud to není dodrženo, je zprostředkovatel sankcionován ve formě třeba slev (například měsíční nebo roční snížení ceny). [5]

4.4.2 Pohled z technického hlediska

- **Rychlost nasazení** – stačí vybrat poskytovatele, vyplnit registrační formulář nebo jen zaplatit a je hotovo. Pokud si jednotlivec nebo společnost řeší toto sám/i, je nutné nejdříve vybrat a nakoupit hardware, poté vybrat softwarovou platformu (to je stejně zdoluhavé, jako výběr hardware) a pak je nutné vše nainstalovat, nakonfigurovat, otestovat a až poté je možné začít vše využívat.
- **Odpadá péče o serverovou část infrastruktury** - nemusí se řešit výpadky, záplatování a ani hackerské útoky. Jediné co se řeší, je podpora vašich uživatelů, ale jen tehdy, pokud poskytovatel Cloud computingu toto nenabízí v potřebné formě sám.

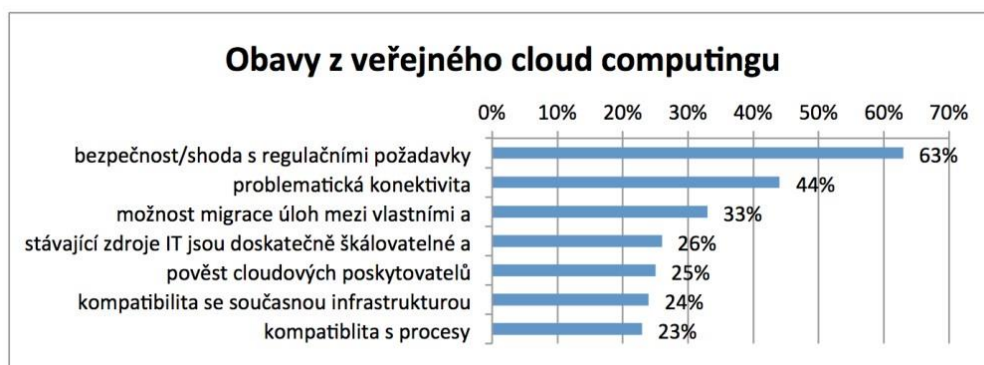
- **Přístup k novým funkcím (někdy i zdarma)** – zdarma jsou třeba aktualizace software, kde se nemusí nic zálohovat, instalovat, kontrolovat, migrovat a ani nic kupovat, to vše je v ceně služby a poskytovatel by se měl o vše postarat.
- **Garantovaná dostupnost služby** – anglicky Service Level Agreement je velká výhoda, protože ne mnoho firem má IT oddělení tak dobré, že by garantovalo dostupnost služeb. Je tedy nutné hledat společnost, která je schopna garanci dodržet a také tuto garanci dlouhodobě poskytnout.
- **Přístup k datům cizími lidmi** – stačí, aby se správce serveru podíval na účet, a vidí vše co tam je - jako jsou e-maily, různé soubory a tak dále. Je ale potřeba si uvědomit, že k těmto datům má přístup jen poskytovatel Cloud computingu a nikdo jiný a dále je vše šifrované, zabezpečené a to většinou mnohem lépe, než na serverech domácích. Pokud tedy nemáme data, která by mohla být velice lukrativní, není se potřeba obávat, že by je někdo zneužil.
- **Data putují po internetu** – pokud odejdeme od běžného využití, jako je hostování e-shopu, poštovního serveru nebo webových stránek, je nutné si uvědomit, že naše data putují po internetu, a to i když by nemusela. Toto se týká zejména kancelářských aplikací, ERP¹³ a CRP¹⁴ systémů nebo datových úložišť ve formě IaaS. Data jsou sice šifrována, ale pokud nastane nějaký velmi sofistikovaný útok, je více než pravděpodobné, že se k nim útočník dostane. To se ale může stát, i když máme data na HDD¹³ v počítači nebo na lokální síti připojené k internetu, jen útok bude vypadat jinak.
- **Limitovaná nabídka poskytovatelů** – jsme limitováni nabídkou různých poskytovatelů Cloud computingu, což může mít někdy vcelku negativní vliv pro uživatele, protože jsou zvyklí na software od jiné společnosti a někdy se jim těžce zvyká na nový. Jako u softwaru toto platí i u hardwaru, například tak, že jeden poskytovatel nabízí skvělé služby, ale špatný hardware a naopak. [5]

4.5 Zabezpečení a Rizika

Asi hlavním důvodem, proč není cloud computing ještě více rozšířen, než už je, je obava ze zabezpečení důležitých dat v cloudu, jak je vidět na obrázku níže od společnosti TechTarget.

[6] Poskytovatelé se samozřejmě snaží zabezpečit vše tak, aby k datům měli přístup jen

oprávněné osoby, a to tak, že nabízí mnoho služeb pro ochranu dat. Využívají různé pokročilé analytické nástroje, jak v privátním, tak hlavně ve firemním cloudu. Bohužel ne vždy je to úspěšné, jak dokazují nedávné úniky dat hollywoodských celebrit, které měli své fotky a videa na cloudu od společnosti Apple.



Zdroj: TechTarget, 2014

Obr. 4 Obavy z cloud computingu

4.5.1 Identity

Správa identit, autentizační služby a federovaná identita hrají v bezpečnosti cloudu významnou roli. Ochrana identit zajišťuje integritu a důvěrnost dat a aplikací a zároveň je zpřístupňuje autorizovaným uživatelům. Podpora zmíněných funkcí je jak pro uživatele, tak pro jednotlivé části infrastruktury, nedílnou součástí všech typů cloudu. Při řízení přístupu do cloudu je potřeba se zaměřit na následující:

Silná autentizace – pokud cloud slouží k provozu podnikových aplikací, je zapotřebí použít silnější autentizaci uživatelů než jen překonané uživatelské jméno a statické heslo. Doporučeno je použít silnou autentizaci (multifaktorová autentizace na bázi jednorázového hesla, federovanou (delegovanou) identitu k důvěryhodnému sdílení identit mezi různými subjekty a „risk-based“ autentizaci, která je založena na chování uživatele, kontextu a mnoha jiných faktorech. Vhodnou kombinací a vrstvením těchto autentizačních metod lze zajistit jak dodržení bezpečnostních SLA¹⁵, tak jednoduchost použití pro všechny typy uživatelů.

Granulární autorizace - zejména ve veřejném cloudu, je určena pro podnikové aplikace, kde je nutné velmi granulárně řídit jednotlivá přístupová práva uživatelů a jejich skupin, ideálně na bázi jejich rolí v podniku – „role-based authorization“. Hlavním důvodem je zde ochrana citlivých dat jednotlivých nájemců v rámci veřejného cloudu a s tím úzce související dodržení příslušných zákonů a předpisů. [7]

4.5.2 Informace

Bezpečnost v tradičních datových centrech je řešena nejen za pomoci IT prostředků, ale i fyzickým zabezpečením přístupu k hardwarové infrastruktuře. Tahle bariéra, ale s příchodem cloudu mizí a je nutné se soustředit na řízení bezpečnosti konkrétních informací než na stavbu jiných bariér. Po cloudu putující data i mimo něj mají vlastní zabezpečení, které je chrání po celou dobu. K dosažení tzv. „information-centric“ bezpečnosti je zapotřebí vyřešit následující:

Oddělení dat - Ve veřejném cloudu, kde se zpracovávají data mnoha zákazníků, musí být data izolována. Virtualizace, šifrování a granulární řízení přístupu výrazně pomohou v izolaci dat mezi nájemci, jednotlivými skupinami nebo uživateli.

Granulární bezpečnost dat - Se zvyšující se citlivostí informací a jejich počtem se musí prohloubit i klasifikace dat a důslednost ve vynucování jejich výhradně oprávněného použití. Dnešní podniková datová centra tuto oblast dost často opomíjela, neboť byla vždy pod kontrolou daného podniku. Bezpečnost dat v cloudu je tak kritická, že její granularitu musíme řešit už na úrovni souboru, tabulky či sloupce v databázi. To znamená inspekci dat (sledování obsahu), jejich tokenizace či šifrování a bezpečná správa šifrovacích klíčů po celý jejich životní cyklus.

Klasifikace dat - Nalezení ideálního poměr mezi uživatelským komfortem a požadavky na jeho zabezpečení není v cloudu jednoduché. Důležitými kroky k nalezení toho správného poměru je klasifikace dat a fungující procesy pro jejich vyhledávání, monitoring toku a použití. V této oblasti značně pomáhají systémy, které jsou určeny pro vyhledávání a ochranu citlivých dat, tzv. „data loss prevention“.

Information rights management - IRM rozšiřuje model „information-centric“ bezpečnosti o možnost procesního řízení přístupových práv přímo na úrovni dokumentu, a to i v případech, kdy opustí cloudové prostředí a je uložen mimo něj.

Monitoring a audit - Prostředí všech systémů, ve kterých se pracuje s citlivými či jinak zákonem chráněnými informacemi, musí být z pohledu bezpečnosti kompletně monitorováno a pravidelně auditováno, ideálně pomocí řešení pro sběr a analýzu logů z daných

systémů s reportingem do podnikového GRC (governance, risk, compliance) řešení pro řízení podnikových procesů a rizik. [7]

4.5.3 Infrastruktura

Celá infrastruktura by měla být v jádru bezpečná, nehledě na tom, zda stavíme privátní nebo veřejný cloud. To si vyžaduje:

- **Komplexní bezpečnost** – cloud by měl být již navržen jako bezpečný, postaven z bezpečných komponentů, implementován podle bezpečnostních know-how, bezpečně komunikující s okolím a podporující potřebná bezpečnostní SLA (označuje smlouvu sjednanou mezi poskytovatelem služby a jejím uživatelem).
- **Bezpečnou integraci** – v místě, kde dochází ke komunikaci mezi jednotlivými částmi cloudu, je nezbytné vynucovat dodržování bezpečnostních pravidel pro sdílení dat, pro zajištění jejich integrity a důvěrnosti. [7]

4.6 Osobní údaje

U cloudových služeb dochází ke zpracování osobních údajů a uživatel neví, kde jsou jeho data uložena a kdo k nim má přístup. Než se smluvně dohodneme s poskytovatelem cloudové služby je důležité si ujasnit několik věcí:

- Kdo a za jakých podmínek má k mým datům přístup,
- Je poskytovatel služeb věrohodný (má různé bezpečnostní certifikáty atd.),
- Opatření při výpadku cloud computingu, nebo při ukončení činnosti poskytovatele služby a zda navrátí nebo vymaže data,
- Monitorování cloud computingu,
- Kontrola místa, kde jsou data uložena,
- Používání šifrovaných dat a osobních údajů k jiným účelům, než je nezbytné,
- Technická a správní zodpovědnost za naše data a opatření při poruše dohody.

4.7 Rizika

4.7.1 Odcizení nebo zneužití dat

Asi největší obava společnosti, když se rozhodují pro cloud computing je právě odcizení nebo zneužití dat. Díky tomu, že cloudové technologie využívají jeden fyzický server pro vícero uživatelů je nezbytné data důkladně šifrovat a poskytovatelé tuhle službu nabízí

v dnešní době jako standard. Poskytovatelé ale nechtějí nést plnou zodpovědnost za naše data, která jsme jim „svěřili“ a proto své zákazníky nabádají ke sdílené zodpovědnosti za rizika. Jediné co garantují je fyzická bezpečnost. Proto velké množství uživatelů svá data šifruje, ještě než opustí jejich vlastní podnikovou síť a poté jsou data dešifrována až v další podnikové síti.

4.7.2 Ztráta dat

Na rozdíl od odcizení nebo zneužití, které jsou cílené invazivní útoky, je ztráta dat většinou neúmyslná, ať už se jedná o neúmyslné vymazání dat, selhání hardwaru nebo díky lidské chybě při údržbě atd. Patří sem ale i cílený útok za účelem vymazání nebo poškození dat určitého uživatele, nebo uživatelů daného poskytovatele služby. Například díky technické poruše může společnosti nastat nemalá škoda a odhadované finanční ztráty v důsledku ztráty dat se uvádějí v bilionech dolarů ročně. Asi nejlepší prevence před ztrátou dat, je nespolehat se výhradně na poskytovatele služby, ale mít data zálohována na nějakém lokálním úložišti. Samozřejmě i poskytovatelé mají data uložena na více geografických místech, ale na to se nejde vždy 100% spolehnout.

4.7.3 Hrozby

- **Zombie účty** – jsou to účty, které nejsou nikým používány, ale stále nebyly zrušeny nebo smazány. Jako příklad jde uvést, když pracovník odejde od společnosti a administrátor jeho účet uzavře, ale nikdy jej skutečně nevymaže. Nebezpečí, které plyne z tohoto chování, spočívá v tom, že pokud se někdo takového účtu zmocní, není nikdo, kdo by tyto účty sledoval, a poté si může takový uživatel dělat s daty, co chce, například ukrást nebo smazat data. Velké množství správců si ani neuvědomí, že zombie účty jsou stále v systému, protože například u aplikací SaaS se platí jen za aktivní účty, tudíž zablokované účty většinou nikde nejsou započteny.
 - **Ochrana:** Mazat všechny neaktivní účty.
- **Nepřející zaměstnanci** – lidská chyba je jedním z nejčastějších důvodů při ztrátě dat u jakékoli aplikace a je tomu tak i u cloudu. Je to způsobeno tím, že software nerozezná legitimní příkazy od nelegitimních. Například náhodné smazání, což tvoří až 2/3 všech ztrát dat. Ještě horší je to u zaměstnanců, kteří škodí cíleně. Tito nespokojení zaměstnanci mají přístup k datům na cloudu a mohou způsobit mnohem stejné ne-li větší škody než zombie účty, vzhledem k tomu, že vyhledávání a mazání dat dělají cíleně.

- **Ochrana:** Nevěřit nikomu, uživatelům přidělovat privilegia a nedávat nikomu větší přístup, než je zapotřebí.
- **Černá labuť** – označení události, kterou je pomalu nemožné předvídat nebo očekávat. Jako příklad lze uvést situaci, kdy hackeři objevili chybu naráz v systému Amazonu a platformě Applu, i když jsou chráněny obě služby zvlášť, kombinace chyb umožnila získat kontrolu nad všemi účty, což nikdo nemohl očekávat. Vzhledem ke komplexnosti SaaS aplikací, a také díky tomu, že cloudové služby jsou poměrně mladé a jejich bezpečnost ještě není dostatečně prověřena, je více než pravděpodobné, že se události jako černá labuť budou i nadále vyskytovat, a tak ohrožovat naše data. [8]
 - **Ochrana:** Zálohovat!

4.7.4 DoS a DDoS útoky

Jak již z anglického názvu (Denial of Service) vyplývá, u DoS útoku, útočníkův záměr spočívá v odepření služby. U tohoto případu to znamená odepření přístupu na internet, přerušení nějaké konkrétní služby nebo zahlcení celé vnitřní sítě. Dopad to nemusí mít jen finanční, ale i třeba prestižní. On-line bankovní dům se nejvíce obává nedostupnosti svých bankovních aplikací, třeba i v řádu minut, kdežto klasický úřad je schopný se obejít bez svého webu i týdny, ale stačí, aby byly zahlceny jeho vnitřní sítě a má to nedožrnné problémy.

Nesmí se zapomenout, že DoS může klidně sloužit jen jako zástěrka pro klasický hackerský útok, kdy útočník používá naše soustředění na obranu. Takový útok potom není ničivý, ale je zahlcující, aby hacker měl současně s napadenou stranou přístup k napadeným systémům a mohl třeba získat root hesla nebo jinak kompromitovat systém.

Zkratka DDoS (Distributed DoS) označuje poddruh DoS útoku, rozdíl je v tom, že je útok veden z velkého množství zdrojových počítačů. [9]

5 CLOUDOVÉ PLATFORMY

Jedná se o specializované cloud služby, které se hodí jen pro firemní sféru. Mezi ně patří Windows Azure od společnosti Microsoft a Google Cloud Platform od společnosti Google, dále Amazon Web Services od Amazonu atd. Hlavních rozdíl oproti osobním cloudům je ten, že u cloudových platform se očekává, že se budou hojně využívat, když se do nich investuje, proto jsou určeny jen firmám a to hlavně velkým firmám, protože poskytují velké množství specializovaných služeb, které běžný uživatel mimo firemní sféru nemůže v žádném případě využít.

5.1 Windows Azure

Jedná se o cloudovou platformu společnosti Microsoft. Využívána je k vytváření virtuálních serverů, spouštění aplikací, hostování a škálování webových aplikací pomocí datacenter Microsoftu. Nabízí tzv. Microsoft Online Services, což je balík online aplikací, které fungují jako služby (bez nutnosti instalace) k vyměňování dat, sdílení, online komunikaci (online konference) atd., vše je založeno na bázi cloudu. [10]

Windows Azure poskytuje širokou škálu služeb, jako jsou:

- **Virtual Machines** - jedná se o službu, která umožňuje provoz plnohodnotných virtuálních serverů v datových centrech Microsoftu. K dispozici je několik výkonných virtuálních serverů, které jsou konfigurovatelné, včetně instancí s vysokou kapacitou operační paměti. Jako operační systémy jsou k dispozici Windows Server, a to ve verzích 2008 (R2) a 2012 (R2) nebo také linuxové operační systémy. Dále je možné vytvořit virtuální server z image předinstalovanými aplikačními servery, jako jsou SQL Server nebo BizTalk Server, a také s předinstalovaným softwarem společnosti Oracle. Jedná se o skutečné virtuální servery, tudíž uživatel získá po vytvoření nad serverem úplnou kontrolu a může si v něm konfigurovat, jakékoli služby chce nebo instalovat libovolný software. [11]
- **SQL Database** – relační databázová služba. Vychází z technologie SQL Server. Zjednodušuje nasazení a správu velkého množství databází. Vývojáři nemusí nic instalovat, nastavovat, spravovat nebo aktualizovat. Velmi dobré jsou funkce jako vysoká odolnost proti chybám a skvělá dostupnost. [12]

- **Cloud Services** – služby pro zákazníky, kteří se chtějí věnovat svým aplikacím, a to jejich správě a rozvoji. Microsoft zajišťuje technickou podporu, správu, údržbu a aktualizace. Je možné také přejít na novější verze operačního (serverového) systému nebo databázového serveru. [12]
- **Web Sites** – jedná se o provoz webových stránek a webových aplikací.
- **Machine Learning** – je to cloudová služba, která umožňuje vytváření prediktivních analytických modelů. Zpřístupňuje metody pro strojové učení, a také výpočetní výkon platformy Azure. Je to technika, která využívá minulých zkušeností k předpovědi budoucího chování. [13] Využitím výpočetně náročných algoritmů a technik dokáže Azure Machine Learning, po dodání velkého vzorku pročištěných dat, nalézt mezi jednotlivými daty souvislosti a postupně upravovat nalezené souvislosti na základě měnících se dat – „učit se“.
 - je nezbytné definovat, co chceme predikovat jako výstup,
 - k problému je potřeba nashromáždit staticky reprezentativní sadu dat,
 - vytvořit prediktivní model,
 - model se musí natrénovat, to se provádí na základě předchozích zkušeností, musí se modelu nabídnout dostatečné množství dat, která popisují zkoumanou oblast,
 - analyzovaný případ se předhodí modelu a ten na základě statistické významnosti určí nejpravděpodobnější výsledek,
 - výsledky se musí interpretovat a vyvodit z nich konkrétní důsledky. [13]

5.2 Google Cloud Platform

Jedná se o skupinu cloudových produktů od společnosti Google, která se zaměřuje na zprostředkování IaaS a PaaS služeb. Společnosti se nemusí starat o samotnou infrastrukturu a mohou se soustředit na chod a vývoj vlastních aplikací. Nabízených produktů je velmi mnoho, ale rozhodl jsem se vyjmenovat tyto:

- **App Engine** – jedná se o cloud službu, díky které mohou vývojáři vytvářet nebo provozovat své vlastní aplikace na infrastruktuře od Googlu,
- **Compute Engine** – jedná se o nástroj pro náročné a rozsáhlé výpočty,
- **Cloud Storage** – ukládání velkého množství objemu dat online,
- **Cloud SQL** – pomocí tohoto nástroje, lze provozovat databázový server MySQL na cloudové infrastruktuře,

- **BigQuery** – nástroj, který využívá dotazy podobné SQL pro zpracování obrovských objemů dat,
- **Cloud Datastorage** – vysoce škálovatelné NoSQL databáze pro web a mobilní aplikace,
- **Cloud Machine Learning** – strojové učení z jakýkoliv dat a velikosti. [14]

5.3 Ostatní

5.3.1 Amazon

Poskytuje služby pod názvem Amazon Web Services. Mezi hlavní služby patří:

- **Amazon Elastic Compute Cloud** (Amazon EC2) – služba umožňující pronajmutí virtuálních počítačů.
- **Amazon Simple Storage Service** (Amazon S3) – úložiště od Amazonu
- **Amazon SimpleDB** – poskytuje databázové funkce [15]

5.3.2 Salesforce

Známa především díky automatizační aplikaci *Salesforce.com*. Skládá se ze:

- **The Sales Cloud** – aplikace cloud computingu
- **The Service Cloud** – platforma zákaznické služby
- **Your Cloud** – vývoj aplikací pomocí firemní platformy cloud computingu. [16]

6 BIG DATA

Pod tímto pojmem se rozumí analýza obrovského objemu dat, které by v minulosti vůbec nebylo možné zpracovat. Díky dnešním rychlejším počítačům, jež se stále zrychlují, a také stále levnějšímu úložnému prostoru a novým softwarovým nástrojům, je to možné. Poraden- ská firma Gartner definuje Big Data jako: „*soubory dat, jejichž velikost je mimo schopnosti zachycovat, zpracovat a spravovat data běžně používanými softwarovými prostředky v reál- ním čase*“. [17]

Jako nejlepší příkladem toho, kde se s Big Daty operuje, jsou závody Formule 1. V každém závodu může monopost vygenerovat gigabajty dat, a tyto data se vyhodnocují během závodu (v reálném čase). Díky těmto datům je možné lépe se přizpůsobit situacím, například, kdy zastavit v boxech o několik kol dříve, než bylo původně plánováno, nebo nasadit jiné pneu- matiky (mohou být více opotřeбенé, než se plánovalo) nebo pro změnu strategie. Tuto situaci komentuje ředitel McLaren Racing Jonathan Neale: „*Závodění je závislé na technologiích. Díky nim můžeme dělat během závodu informovaná rozhodnutí, která stojí na skutečných datech. Vlastně tak dostáváme šanci změnit původní rozhodnutí.*“

V roce 2013 podle analytické firmy IDC překročil celkový objem dat, která byla umístěna v databázích bilion terabajtů (1 TB = 1 099 511 627 776 bytů). Tyto data se v minulosti nevyužívala, protože nebyl dostatečný výpočetní výkon na jejich zpracování. Dnes díky do- statečnému výkonu a poklesu cen rychlých flash pamětí je to již možné. Pro ukládání vel- kých dat se tedy nepoužívají disková pole, ale ukládají se přímo do paměti serverů. Díky těmto tzv. in-memory databázím je možné zpracovat obrovské objemy dat v reálném čase a poté je okamžitě zobrazit na obrazovce. [18]

6.1 Big Data v praxi

Jako další příklad, kde se využívají Big Data je u automobilů a to konkrétně pomocí senzorů zabudovaných v autech, ale nejen v nich, ale třeba i v silnicích, semaforech a dalších do- pravních prostředcích. Pomocí GPS¹⁶ modulů v automobilech, spojených s GSM¹⁷ moduly lze poskytnout velké množství užitečných dat například pro sledování provozu na silnicích, a tedy pomoci řidičům přizpůsobit trasu cesty, aby se do cíle dostali co nejrychleji a s nej- nižšími náklady. Tyto informační data pomáhají i dopravním centrům, která pomocí nich mohou analyzovat a popřípadě upravit preference na jednotlivých semaforech pro mnohem plynulejší dopravu ve městě. Informace z celého města se mohou v reálném čase zobrazovat

v grafické podobě a obsluha v dopravním centru může okamžitě reagovat na nastalé situace (například odklonit dopravu při autonehodě).

Big Data jdou dále využít jako pomoc obchodníkům, kteří mohou pomocí nich lépe nabízet a analyzovat zboží. Kromě analýzy trhu lze pomocí Big Dat analyzovat i sportovní utkání, například v čem se má vylepšit hra atd. a nejen to, u sportovních událostí, jako jsou zápasy amerického fotbalu (NFL) týmů New York Jets nebo New York Football Giants mohou pořadatelé společně s mobilními operátory (americký Verizon) automaticky navigovat auta návštěvníků na volná parkovací místa nebo přizpůsobit kapacitu vstupních bran na stadiony, také je možné dostávat na chytrý mobil statistiky z utkání, jako například držení míče a to i jak dlouho jednotliví hráči drželi míč nebo jak se pohybovali po hřišti. [18]

6.2 Analýza - 4V (5V)

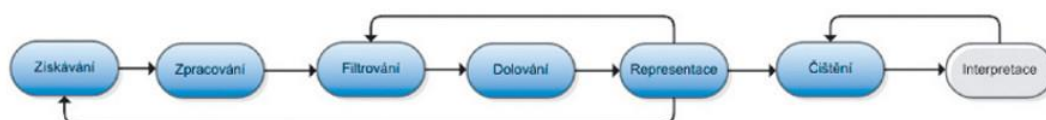
- **Objem dat** (Volume) – analýza je velmi obtížná, stačí se zamyslet nad množstvím e-mailů, fotek, videosouborů, senzorových dat, Twitter zpráv, které produkujeme a sdílíme každou sekundu. Nemluvíme zde o terabajtech, ale o petabajtech a do budoucna se nárůst objemu dat bude jen zvyšovat. Pokud si vezmeme všechna data generovaná od začátku naší civilizace až po rok 2008, stejné množství dat se dnes generuje každou minutu. Tudíž je jasné, že všechna tato data nemohou být uchována na jednom místě, ale na několika, a poté se pomocí softwaru spojí a analyzují.
- **Rychlost** (Velocity) – množství dat, které je potřeba analyzovat v reálném čase a která se hromadí, je obrovské. Například kamery na letišti, které musí informace o hrozcím nebezpečí vyhodnotit okamžitě, aby tyto informace měly význam. Nebo burzy, kolik dat se zde musí analyzovat a jak rychle, aby se mohlo rozhodnout, zda akcie koupit nebo ne. Technologie zpracování veledat nám umožňuje je analyzovat, zatímco jsou generována, aniž by kdy byla uvedena do databází.
- **Komplexnost dat** (Variety) – v minulosti byly počítače dobré pro zpracovávání číselných dat a statistických analýz, která se snadno analyzují jako databáze nebo se u nich daly počítat průměry. Naproti tomu data jako zprávy, konverzace na sociálních médiích, senzorová data, videa, přirozená řeč, obrazová data, analogová a nestrukturovaná data se analyzovala jen velmi těžce pomocí jednoduchých statistických metod. Pomocí Big Dat lze ale zpracovat obojí, jak strukturovaná, tak i nestrukturovaná data.

- **Věrohodnost** (Veracity) – vzhledem k tomu, s jakým množstvím dat se operuje, je jasné, že mnoho dat bude nedůvěryhodných nebo chaotických. Kvalita a přesnost dat jsou mnohem hůře kontrolovatelné než u běžných objemů dat (Twitter posty s hashtagy, různé zkratky, překlepy, hovorová řeč, ale také spolehlivost a přesnost obsahu). Díky novým analytickým softwarům lze již s těmito daty pracovat. Objem dat převáží kvalitu a přesnost dat.
- **Hodnota** (Value) – je sice pěkné mít obrovské množství dat k dispozici, ale pokud tato data nedokážeme nějak zhodnotit, jsou bezcenná. [19]

6.3 Vizualizace a IR (Information retrieval)

Vizualizace dat nám umožňuje transformovat množství dat tak, abychom mohli mnohem lépe získat potřebná data, a aby poté šla efektivním způsobem využívat. Zapotřebí jsou informační specialisté, aby mohli ze získaných dat získávat relativní informace, a to tak, že vědí, co v datech hledat a jakým způsobem lze tyto informace získat. Velkou roli zde hraje intuice, díky ní zkušený analytik dokáže zjistit, která data jsou užitečná a která nikoli. Z jednoho souboru dat lze získat rozdílné výsledky, díky tomu, že se na data díváme z různých úhlů. Podle toho jaké zvolíme jednotlivé filtry, způsob dolování dat nebo i jejich interpretace mohou vést ke zcela odlišným závěrům a poznatkům. [20]

U procesu vizualizace by mělo být mnoho odborníků z různých odvětví, ať už jde o matematiky a informatiky, díky nimž jde data přímo filtrovat a zpracovávat, dále sociology, kteří by formulovali nápady a hypotézy, až ke grafikům a designérům, kteří výsledky převedou do vizuálně přitažlivé a dobře pochopitelné podoby. Jako optimální prostředí pro tyto operace se tedy jeví různé univerzity a jejich fakulty.



Obr. 5 Diagram znázorňující jednotlivé fáze práce s velkými daty.

Zdroj: www.oldknihovna.nkp.cz

- **Získávání dat** – primární krok, za nímž se skrývají všechny dotazníky, kamerové systémy, testy atd. K vizualizaci Big Dat musíme mít podklady v dostatečné kvalitě a množství.
- **Zpracování** – jedná se o převedení konkrétní podoby dat na žádoucí a strojem snadno zpracovatelnou podobu. Je možné zahrnout filtrování obsažené informace (převod rozhovoru do značek) s ohledem na rychlost a snadnost dalšího zpracování (například algoritmického).
- **Filtrování** - navazuje na zpracování, rozděluje data podle určitého filtru na data, která chceme zpracovávat a na ta, která nepotřebujeme. Také lze použít vícestupňové filtrování pro výběr dat, která skutečně potřebujeme. Pokud bychom data nefiltrovali, mohlo by dojít k tzv. *Informačnímu přetížení* = záplavě informací, ve kterých se nelze orientovat.
- **Dolování** – fáze, která se zabývá získáváním dat z množiny všech předložených dat. Využívá matematickou statistiku, regulérní výrazy atp. Součástí je i filtrování šumu atd. Dalo by se říci, že se jedná o nejzajímavější část procesu, patří sem vytváření algoritmů, díky nimž lze identifikovat určité společné rysy dokumentů, a také snaha o automatické generování později analyzovaných struktur atp.
- **Reprezentace** - zabývá se zobrazením získaných dat do podoby, díky které budou vizuálně lehce pochopitelná. Pole, tabulky lze převést na grafy, mapy atd. Je nutné vidět důležité informace, trendy a umět je „vytáhnout“.
- **Čistění** – vylepšení již vytvořeného výsledku, rozhodnutí jak co graficky znázornit, aby to bylo přehledné atd. Pečuje se o vizuální stránku celého výstupu a znovu dochází k odstranění nepotřebných dat, informací, šumu, které nejsou potřebné.
- **Interpretace** – u této poslední fáze se hledají cesty, pro co možná nejlepší pochopení obsahu, například přidání vrstvy, díky níž uvidíme problém z jiného úhlu nebo animace v čase atd. Znovu se odkazuje k dolování dat. Během této fáze by se měl vždy účastnit odborník, kterého se téma týká a spadá do jeho kompetence. Porovnávají se hypotézy nebo vstupní modely s výsledky, vytváří se příběh a hledají se příčiny. [20]

6.4 NoSQL databázové koncepty

V minulosti se používaly hlavně klasické SQL¹⁸ databáze, dnes se čím dál tím více projektů nejruznějšího druhu spoléhá na řešení, které lze souhrnně označit jako NoSQL databáze. Jako základní vlastnosti SQL databáze lze vnímat:

- Data jsou uložena v tabulkách a každý sloupec je definován pro určitý datový typ a tak je možné za pomoci libovolného sloupce data filtrovat a vyhledávat v nich.
- Zásada ACID (Atomicity, Consistency, Isolation, Durability), což zajišťuje, že data jsou vždy konzistentní a že operace s databází nemají žádné nepředvídatelné řetězce.
- SQL databáze obsahují (standardně) jen znakové řetězce.

Tyto znaky nám říkají, že SQL databáze mají výborné vlastnosti, jako to, že nesou maximum zpracovatelných informací, ale mají také nevýhody. Ty jsou například, že obsah musí mít jasnou a pevnou strukturu, kterou nelze měnit, tabulky jsou velké a pomalé a nelze provádět více operací najednou atd. U databází vytvořených s Big Daty předpokládáme, že objem dat bude velký, máme zvýšené nároky na rychlost a většinou nepotřebujeme mít celou informaci, ale jen určité fragmenty nebo části výsledků. K tomu se NoSQL hodí mnohem lépe, mezi nejznámější řešení patří: [20]

Key-value – databáze jsou založeny na myšlence, že ke každému unikátnímu klíči je přiřazena určitá hodnota. Databázový systém svede jen poskytnout odpověď na otázku, která byla předem definována klíči a vrátit informaci o hodnotě, přidávat a upravovat záznamy atd. Vyhledávat lze jen pomocí klíče, který nelze nijak dynamicky měnit. Použití je u situací, kde je nutné zjistit rychle odpovědi, ale za cenu nižší možnosti databáze. Velmi často se používají ve spolupráci s klasickými a pomalejšími SQL databázemi, kde SQL databáze zajišťují pokročilejší a náročnější operace a rychlé Key-value databáze zajišťují základní dotazy. Toto řešení používá například LinkedIn (největší profesní síť na světě) nebo Facebook.

Grafově orientované databáze (Graph database) – odstraňují jeden z hlavních neduhů relačních databází, a to jejich nízkou škálovatelnost, která je dána pevně nastavenou tabulkou. Jedná se o mnohem více flexibilní řešení, kde každý uzel má kromě vlastních informací, také uloženy odkazy na sousední uzly. Toto řešení dnes využívá Facebook pro hledání informací od přátel.

Dokumentové databáze – snaží se řešit, jak lze prohledávat, analyzovat a odvozovat z dokumentů, které nejsou jen položkou v tabulce, ale obsahují souvislý text. [20]

6.5 Apache Hadoop

Apache Hadoop je open source projekt společnosti Apache Software Foundation, který se zabývá rozvojem platformy pro distribuované ukládání a zpracování velkých objemů dat v clusteru počítačů.



Obr. 6 Apache Hadoop

Zdroj: www.hortonworks.com

Distribuované uložení dat má na starost modul **HDFS** (Hadoop Distributed File System). Skládá se z řídicího uzlu (uzlů) namenode (s metadaty o samotném uložení dat) a uzlů data-node, kam se data fyzicky, s určitým stupněm redundance, ukládají. Základem distribuovaného zpracování jsou knihovny programovacího modelu **MapReduce**.

Fáze Map() zmapuje vstupní data a rozdělí práci dle uložení dat na příslušné uzly clusteru (což neznamená, že se výpočet provádí pouze na uzlech s daty), případně deleguje další podoperace Map na tyto uzly. Tato fáze poté provede operace, které lze odděleně provést na částech dat. Vedle toho provádí Fáze Map () např. i filtraci.

Fáze Reduce() zredukuje výsledky fáze Map() a provede operace závislé na celých datech, např. sumarizaci. Zároveň fáze Reduce() může probíhat na více uzlech. Trojici základních pilířů doplňuje modul YARN (Yet Another Resource Negotiator), který řídí přidělování zdrojů v momentě, tehdy pokud běží více úloh najednou.

Hadoop má celou řadu podprojektů, jejichž cílem je zjednodušení práce s touto platformou. Pro administraci např. existuje webové rozhraní Ambari. Pro programování výpočtů lze využít zjednodušeného jazyka modulu Apache Pig či SQL obdoby **HiveQL** modulu Hive pro datový sklad (jinak se při psaní úloh MapReduce používá Java). Velkou výhodou Hadoopu je, že do clusteru lze zapojit i běžné stolní počítače, že pád jednoho z uzlů clusteru je automaticky detekován (a při redundanci uzlů neohrozí uložená data ani probíhající výpočty) a že doba, kterou výpočet trvá, se paralelizací zkracuje. Hadoop využívají takové giganti jako Yahoo a Facebook. [21]

II. PRAKTICKÁ ČÁST

7 CLOUDOVÁ ÚLOŽIŠTĚ

V dnešní době jsou velkým hitem cloudová úložiště, jedná se o přesouvání dat z pevných disků a kompaktních disků na cloud. V praxi to funguje tak, že data jsou uložena v datacentrech poskytovatelů a odtud si uživatelé stahují (synchronizují) jejich kopie do svých zařízení. Tento trend není jen doménou firem, ale čím dál tím více i obyčejných domácích uživatelů. Je totiž velmi výhodné a pohodlné mít svá data uložena v cloudu, kde se k nim můžeme odkudkoliv na světě dostat, pokud máme přístup k internetovému připojení. Samozřejmě je možné si umístit server doma a mít vlastní datové úložiště, ale většina uživatelů se spíše spoléhá na různé internetové služby od velkých společností, jako je Google, Microsoft, Apple, Dropbox nebo na méně známe (Mega, Symplicity atd.). To ovšem neznamená, že by některé z nich výrazně zaostávali za těmito obry, jen nejsou tak rozšířeny v povědomí zákazníků.

Mnoho společností vám nabídne prostor zdarma ve svých datových úložištích, ale jen do určitého objemu dat. Poté pokud vám nestačí prostor, je možné si jej samozřejmě dokoupit. Bohužel ne všichni poskytovatelé cloudových služeb (platí spíše u služeb zdarma) garantují zachování dat.

Velké společnosti, jako ty, jež byly již vyjmenovány, se snaží nalákat zákazníky nejen na samotná datová úložiště, ale i na služby s nimi spojené, jako je třeba služba Microsoftu pro zálohu Windows nebo Googlu pro online práci v dokumentech a Applu na propojení se službou Apple TV.



Obr. 7 Cloud

Zdroj: www.technicloud.co.nz

8 VÝHODY A NEVÝHODY CLOUDOVÝCH ÚLOŽIŠŤ

Než se dostaneme k jednotlivým úložištím, bylo by dobré si na začátku shrnout základní výhody a nevýhody cloudových služeb, a to obecně, ne specificky ke každému úložišti, což se provede až níže. Začít by se mělo od výhod a poté přejít k nevýhodám. Výhody značně převažují nad nevýhodami. Co se týče nevýhod, není jich mnoho, ale bohužel jako u všeho, se najdou.

8.1 Výhody

Přístup odkudkoliv – hlavní výhoda cloudu je právě dostupnost dat odkudkoliv, nejsme omezení tím, že musíme používat jeden počítač, ale můžeme použít jakýkoliv počítač s přístupem na internet nebo mobilní telefon, tablet, prostě jakékoliv mobilní zařízení.

Sdílení dat s jinými uživateli – velmi opomíjená vlastnost, která umožňuje, aby s daty pracovalo více lidí najednou. Tím také odpadají velké přílohy v mailech, stačí přidat k datům oprávnění i pro ostatní uživatele a do mailu můžeme dát už jen odkaz, jenž si uživatelé stáhnou nebo budou pracovat online.

Zabezpečení, záloha a dostupnost dat – mnoho zprostředkovatelů úložišť garantuje zabezpečení dat, jenž jsou díky dnešnímu šifrování dat skutečně v bezpečí, jejich pravidelné zálohování a již zmíněnou dostupnost.

Datový prostor (základní) zdarma - neplatí to sice vždy, ale ve většině případů dostaneme od provozovatele datových úložišť prostor zdarma jen za registraci. Poté je samozřejmé, že za určitý měsíční nebo roční poplatek si budeme moci prostor navýšit.

8.2 Nevýhody

Internetové připojení – konkrétně tedy problémy s připojením, kdy rychlost nahrávání souvisí s rychlostí a kvalitou internetového připojení. Pokud vypadne internet, znamená to, že se k našim datům nedostaneme. Tento problém lze částečně eliminovat pomocí synchronizace, kdy se naše data automaticky ukládají na cloud, my s nimi samozřejmě můžeme pracovat, ale osoba, která s daty pracuje mimo nás online už ne. Ale je možné nastavit aplikace tak, že až budeme mít připojení na internet, data se nám automaticky nahrají na cloud.

Podmínky poskytovatelů – tato ustanovení většinou dávají souhlas, že poskytovatel cloudových služeb má na naše nahraná data práva. Nejedná se o autorská práva, ale práva, že může naše data používat. Je to ochrana, pokud by uživatel chtěl provozovatele zažalovat za neoprávněné držení dat. Nejedná se tedy o absolutní nevýhodu, jen je dobré to mít na paměti.

8.3 Synchronizace dat

Díky synchronizaci se z cloudu stává velmi dobré datové úložiště, které umožňuje mnoho služeb, jako je nahrávání různých verzí souborů při zapnutí automatické archivace. To umožňuje vrátit se k předešlé verzi, pokud se nám některá část nelíbí nebo jsme ji omylem smazali atd. Samozřejmě se ale musí počítat s mnohem vyšší datovou obsazeností. Jde také nastavit časový interval, jak dlouho mají být starší verze souborů zachovány. [22]

Velmi užitečná vlastnost synchronizace spočívá v jejím využití na více zařízeních, kromě počítačů jde synchronizaci nastavit i na mobilních telefonech nebo tabletech. Existují 2 hlavní možnosti synchronizace dat mezi zařízeními.

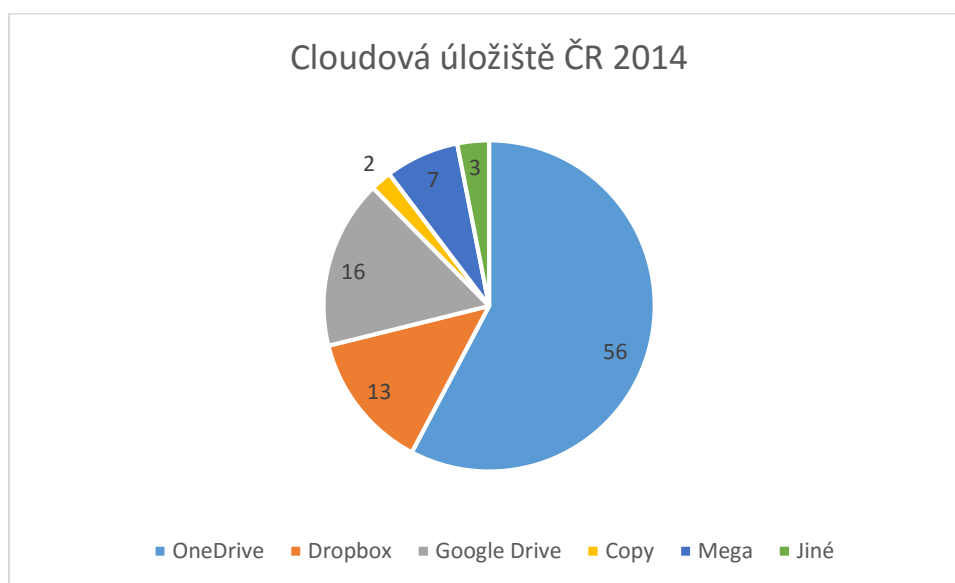
- **Centralizovaná synchronizace úložiště** – datově se jedná o nejnáročnější způsob jak data dostat do zařízení, protože synchronizuje celý obsah a ne jen vybrané soubory.
- **Decentralizovaná synchronizace úložiště** – u této synchronizace je možné nastavit, které soubory chceme synchronizovat

U synchronizace lze také nastavit co, kdy a jak se má synchronizovat. Například lze nastavit synchronizaci jen na popud uživatele nebo ji mít zapnutou stále na pozadí, dále také lze nastavit datum a čas synchronizace nebo podle toho zda je Wi-Fi síť zabezpečená, převádět soubory do různého formátu, synchronizovat nastavenou rychlostí, a také množství dat za určitý časový úsek.

9 PRIMÁRNĚ OSOBNÍ CLOUDY

Mezi cloudy pro primárně osobní použití by se automaticky měli zařadit cloudy 3 největších technologických gigantů, a to Googlu, Microsoftu a Applu, jelikož lidé používají cloudy hlavně ve spojení s ekosystémem, který mají a čím dál tím častěji používají pro zobrazení a práci svá mobilní zařízení (smartphony, tablety). Dále je nezbytné zmínit i další úložiště jako je Dropbox a Mega. Neznamená to ovšem, že by tato úložiště byla striktně pro osobní použití. Menší firmy jsou schopny tato úložiště rovněž využít, což je dáno i tím, že většina poskytovatelů má tarify pro společnosti.

Graf dole znázorňuje, jaké procentuální zastoupení měla jednotlivá cloudových úložišť v roce 2014 v České republice, hlasovalo se na portálu Živě.cz. Počet lidí, kteří hlasovali, byl 2604. [23]



Graf 1 Cloudová úložiště 2014

9.1 Google Drive (Disk)



Obr. 8 Google Drive logo

Zdroj: www.case.edu

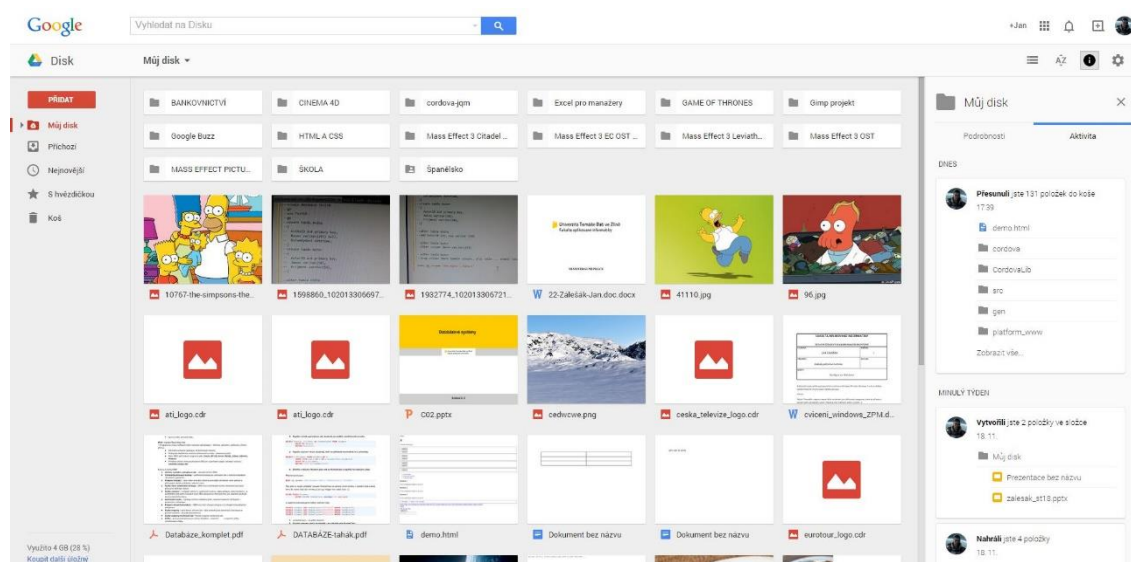
Datové úložiště od společnosti Google, jak již z názvu vyplývá. Jedná s o službu, která je velmi propojena s Gmailem a i ostatními službami od Googlu. V základu nabízí prostor o velikost 15 GB, který lze navýšit v podobě měsíčního poplatku.

- **15 GB – zdarma**
- **100 GB – 40 Kč/měsíčně (1,99\$)**
- **1 TB¹⁹ – 200 Kč/měsíčně (9,99\$)**
- **10 TB – 2000 Kč/měsíčně (99,99\$)**
- **20TB – 4000 Kč/měsíčně (199,99\$)**
- **30TB – 6000 Kč/měsíčně (299,99\$)**

Jak již bylo zmíněno velké plus spočívá v propojení velkého množství služeb od Googlu, jako jsou například Google Docs, s kterými jde pracovat online. Tuto funkci uměl velmi dlouho jen Google, ale nedávno ji začal nabízet i Microsoft a některá ostatní úložiště. Online jde na disku vytvořit textový dokument, tabulku, prezentaci, nákres nebo formulář (anketu) a na těchto dokumentech může pracovat více lidí online. Nesmí chybět ani synchronizace dat mezi PC a mobilním zařízením a tím pádem ani automatická záloha námi vyfocených fotek z telefonu nebo tabletu. Z mé osobní zkušenosti mohu potvrdit, že vše funguje bezchybně a synchronizace mezi počítačem a mobilním telefonem také.

Google Drive For Work – jedná se o cloudové řešení hlavně pro menší společnosti, lze zde provádět videokonference v reálném čase s více uživateli najednou, a také telekonference a chaty. Nechybí ani více specifické aplikace jako DocuSign pro elektronický podpis, Lucid-Chards pro vytváření modelů nebo Smartsheets pro projektový management. Kapacita je zde pomalu neomezená a cena je 10 dolarů na uživatele/měsíc.

V nedávné době prošel celý ekosystém Googlu vizuální přeměnou, která se týkala i Google Drive a podle všeho nás za chvíli čeká nový vzhled, jež bude sjednocovat všechny služby Googlu do Material designu. Podle mě je stávající design mnohem přehlednější než minulá verze. [24]



Obr. 9 Google Drive – nový vzhled

9.2 Microsoft OneDrive (dříve SkyDrive)



Obr. 10 Microsoft OneDrive

Zdroj: www.maketecheasier.com

Cloudové řešení od Microsoftu se jmenuje OneDrive, dříve SkyDrive, ale díky sporu muselo být přejmenováno. Velká výhoda pro uživatele s Windows 8 (8.1, 10) je záloha operačního systému na cloud a také u Office 2013 (2016) průběžně ukládání rozpracovaných dokumentů, aby o ně uživatel nepřišel. Speciální vlastností je, že si soubory můžeme jen nachystat a počkat se stažením souborů až budeme online nebo je mít přístupné i offline. OneDrive je také součástí systému Windows 8.1 a Windows 10.

V základu dostane uživatel již jen 5 GB (v minulosti to bylo 15 GB) prostoru pro svá data, který lze jako u většiny cloudů za poplatek rozšířit. Pro firmy je k dispozici velikost prostoru až 1 TB. Ještě donedávna bylo možné získat kapacitu 100 GB za 960 Kč/ročně (40 Kč/měsíčně) nebo 200 GB za 1920 Kč/ročně (80 Kč/měsíčně), to je již bohužel zrušeno a jsou zde nové tarify, důvod proč byly tyto změny provedeny, je vcelku jednoduchý, a to ten, že někteří uživatelé využívali neomezený tarif tak, že na cloud nahráli celou svou sbírku hudby a filmů (maximum uváděné Microsoftem je 75 TB), proto se Microsoft rozhodl snížit kapacity. Takové chování od uživatelů zvyšovalo datovou náročnost, požadavky na infrastrukturu, a také podle Microsoftu se tím snižovala „user experience“ pro ostatní uživatele. Tato změna proběhla na začátku roku 2016 a uživatelé mají 12 měsíců, aby data odstranili nebo je přemístili.

- **5 GB – zdarma (OneDrive)**
- **50 GB – 600 Kč/ročně, 49,99 Kč/měsíčně, \$1.99 (OneDrive)**
- **1TB pro 1 uživatele – 189,99 Kč/měsíčně (OneDrive + Office 365)**
- **1TB pro každého z 5 uživatelů (domácnost) – 269,99 Kč/měsíčně (OneDrive + Office 365)**

One Drive pro firmy – jako Google tak i Microsoft nabízí edici pro firmy, která stojí 4,2 euro/měsíčně. Největším rozdílem oproti klasickému OneDrive je, že se nachází na platformě **SharePoint**, která umožňuje mnohem širší možnosti správy než obyčejný OneDrive. Další rozdíl je v tom, že do služby se přihlašuje pomocí firemního nebo školního účtu Microsoftu, zato u klasického OneDrive přes osobní účet, tudíž administrátor firmy (školy) může například zakázat sdílení dokumentů atd. [25]

FUNKCIONALITA	ONEDRIVE	ONEDRIVE PRO FIRMY
Prostor	15 GB	1 TB
Cena	bezplatné	placené
Přihlášení	osobní účet Microsoft	firemní nebo školní účet Microsoft
Doporučené ukládané dokumenty	osobní	firemní
Pohledy a filtrování dokumentů	ne	ano
Metadata	přednastavená	možnost správy
Typy obsahu	ne	ano
Koš (na odstraněné soubory)	1	2
Verzování	základní	pokročilé
Nastavení upozornění	základní	pokročilé
Sledování dokumentů	ne	ano
Nastavení práv	základní	pokročilé
Rezervace sdíleného dokumentu	ne	ano

Obr. 11 Microsoft OneDrive – rozdíly (poznámka – prostor u OneDrive je nyní jen 5GB)

Zdroj: www.onedrive.live.com

9.3 Apple iCloud



Obr. 12 Apple iCloud

Zdroj: www.ceskymac.cz

Cloudové řešení od Applu nese název iCloud a na rozdíl od výše zmíněných úložišť je mnohem více omezeno, asi jako celý ekosystém Applu. To znamená, že je jen pro zařízení Apple. Pro přístup musí mít uživatel Apple ID a nově je možné data sdílet až se 6 rodinnými příslušníky. V základu nabízí 5 GB úložného prostoru. Úložiště postrádá jako jedno z mála historii verzí dat. Kapacitu jde jako vždy rozšířit a to až na 1 TB.

Velkou výhodou, jako u všech Apple zařízení, je jeho jednoduchost a snadné ovládání, kde vše funguje tak, jak má a nemusí se nic zdlouhavě nastavovat. Ať už se nacházíme na jakémkoliv zařízení, z každého je možné se dostat k dokumentům iWork, hudbě na iTunes nebo záložkám v Safari a také elektronické poště. Pokud uživatel ztratí zařízení Apple je možné ho pomocí iCloudu najít nebo jej lze zablokovat, aby ho nikdo nemohl zneužít. [26]

- **5 GB – zdarma**
- **50 GB – 0,99\$/měsíčně**
- **200 GB – 2,99\$/měsíčně**
- **1 TB - 9,99\$/měsíčně**

9.4 Dropbox



Obr. 13 Dropbox

Zdroj: www.commonswikimedia.com

Jedno z nejznámějších, nejprehlednějších a nejpraktičtějších úložišť ze všech nese jméno Dropbox, které se malinko změnilo, co se týče nabízeného množství prostoru pro data, zatímco v minulosti to bylo v řádu desítek GB, dnes jsou to pouhé 2 GB. V minulosti také po dalších nastaveních získal uživatel dalších 250 MB²⁰, a když si na Vaše doporučení někdo nový založil účet, dostali oba dalších 500 MB a to až do velikosti 16 GB, což znamenalo až

28 osob. Také šlo propojit Dropbox s Twitterem nebo Facebookem, což zaručovalo dalších 125 MB.

Na rozdíl od Googlu nebo Microsoftu nezakládáte při zakládání účtu také e-mailový účet, protože Dropbox je samostatnou službou, je tedy multiplatformní. V poslední době přinesl několik změn, jako například kancelářský balík Office online, který je nyní integrován přímo do webového prostředí, kde lze dokumenty vytvářet a upravovat. To samé platí i pro mobilní aplikaci, kde se základní editor také objevil. Pro propojení s počítačem nebo notebookem stačí nainstalovat jen aplikaci. Podobně jako Microsoft OneDrive je rovněž verze pro firmy s názvem Dropbox pro firmy (Dropbox for Teams).

- **2 GB + (16 GB) – zdarma**
- **1 TB – 270 Kč/měsíčně (2700 Kč/ročně)**

Dropbox for Teams – jedná se o verzi Dropboxu pro firmy, kde je neomezená kapacita a nabízí dvě úrovně. První se nazývá **Business** a platí se zde 12 euro na uživatele za měsíc. Druhá se nazývá **Enterprise** a u té je nutné kontaktovat Dropbox a společně se domluvit na poskytovaných službách, ale jak již z názvu vyplývá, jedná se o verzi výhradně určenou pro společnosti. [27]

9.5 Mega



Obr. 14 Mega

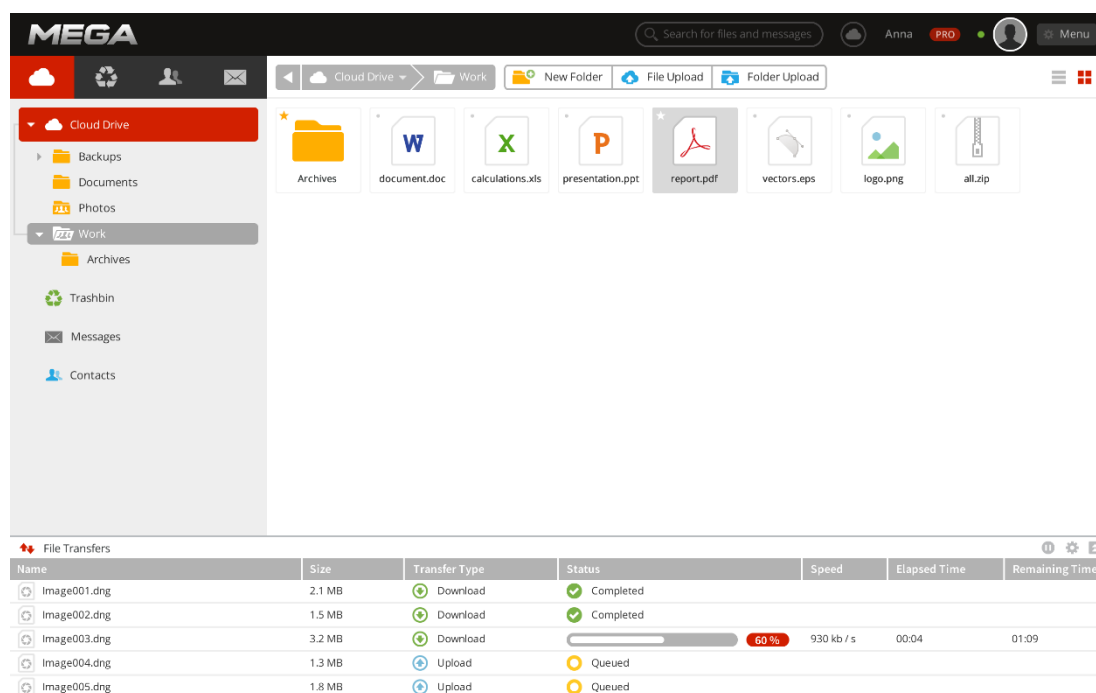
Zdroj: www.en.wikipedia.org

Mega je celkem nové úložiště, které je nástupcem Megauploadu, ale již vcelku známé. Každému uživateli nabízí zdarma 50 GB prostor. Kromě dostatku prostoru nabízí i přehledný a

líbivý design a široká nastavení jako je možnost nastavit, které složky se budou synchronizovat s počítačem, k tomu je ale zapotřebí nainstalovat aplikaci do počítače, která je jak pro Windows, tak i pro Mac.

Další velmi užitečná věc je nativní aplikace, která nejen že je vizuálně atraktivní, rovněž je chráněná heslem a také je možné v ní nastavit, že obsah se smaže, pokud 10 krát po sobě zadá špatné heslo. Asi největší nevýhoda tohoto úložiště spočívá v tom, že je zde nulová otevřenost pro cizí aplikace, tudíž nelze na Mega napojit žádnou další službu. Konkurence je na tom o poznání lépe a je to u nich běžnou vlastností. Webové rozhraní má také napojen šifrovaný chat. [28]

- **50 GB – zdarma**
- **200 GB – 135 Kč/měsíčně**
- **500 GB – 270 Kč/měsíčně**
- **2 TB – 540 Kč/měsíčně**
- **4 TB – 810 Kč/měsíčně**



Obr. 15 Mega – prostředí

Zdroj: www.cnews.cz

10 VÝBĚR ÚLOŽIŠTĚ – TOPSIS ANALÝZA

K výběru úložiště poslouží TOPSIS analýza, což je vícekritériální analýza variant.

TOPSIS (Technique for order preference by similarity to ideal solution) se snaží o nalezení takového řešení rozhodovacího problému, které se co nejvíce blíží optimální variantě, která je ale v reálu neuskutečnitelná a snaží se nastínit, jak by řešení vypadalo za optimálních podmínek.

10.1 Postup

- Jako první je třeba si zvolit, co se bude porovnávat, v našem případě se jedná o 5 hlavních úložišť, která jsou zmíněna v kapitole 9, a to jsou tato:
 - Google Drive
 - OneDrive
 - Apple iCloud
 - Dropbox
 - Mega
- Jako kritéria se zvolí základní kritéria, které by si zvolil běžný zákazník, a to ta, která jsou uvedena níže, a k tomu se přidají váhy ke kritériím, které každé odpovídá tomu, jak je dané kritérium důležité. Součet vah se musí rovnat 1.

KRITÉRIA	VÁHA
○ Kapacita	0,2
○ Cena	0,2
○ Ekosystém, přehlednost	0,3
○ Služby	0,2
○ Vzhled (vizuál)	0,1

- Dalším krokem je vytvoření tabulky s 5 zmíněnými úložišti a 5 zmíněnými kritérii. Hodnocení je od 1-10 (10 nejlepší) u každého úložiště a kritéria, tak, jak to znázorňuje tabulka níže.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	10	6	6	4	9
CENA	9	8	7	6	4
EKOSYSTÉM	10	8	8	5	4
SLUŽBY	8	8	9	6	5
VZHLED-VIZUÁL	8	7	9	6	8

Tabulka 1 Hodnoty

- Vytvoří se normalizovaná kritériální matice: $\mathbf{R} = (r_{ij})$, kde pro $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$;

$$r_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^m y_{ij}^2}}$$

- Poté se umocní každá hodnota v řádku zvlášť a výsledek se odmocní podle předchozí rovnice a vyjde nám tato tabulka.

KRITÉRIUM	mocnina	odmocnina
KAPACITA	269	16,40
CENA	246	15,68
EKOSYSTÉM	269	16,40
SLUŽBY	270	16,43
VZHLED-VIZUÁL	294	17,15

Tabulka 2 Mocnina a odmocnina

- Odmocnina se vydělí každým číslem v řádku (příklad – 10/16,40; 6/16,40...), aby se dostalo číslo mezi 0 a 1, což nám dá **přitažlivost**, která se poté zaokrouhlí na 2 desetinná místa a vznikne následující tabulka.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0,61	0,37	0,37	0,24	0,55
CENA	0,57	0,51	0,45	0,38	0,26
EKOSYSTÉM	0,61	0,49	0,49	0,30	0,24
SLUŽBY	0,49	0,49	0,55	0,37	0,30
VZHLED-VIZUÁL	0,47	0,41	0,52	0,35	0,47

Tabulka 3 Přitažlivost

- Konkrétní hodnoty kritérií, které jsou v matici (tabulce nahoře) se vynásobí vahou kritéria podle vzorce $\mathbf{z}_{ij} = \mathbf{w}_j \mathbf{g}_{ij}$ a pak vyjde tato tabulka.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0,12	0,07	0,07	0,05	0,11
CENA	0,11	0,10	0,09	0,08	0,05
EKOSYSTÉM	0,18	0,15	0,15	0,09	0,07
SLUŽBY	0,10	0,10	0,11	0,07	0,06
VZHLED-VIZUÁL	0,05	0,04	0,05	0,04	0,05

Tabulka 4 Váhy kritérií

- Vytvoří se **ideální** varianty (h_1, h_2, \dots) a **bazální** varianty (d_1, d_2, \dots)
 - $h_j = \max z_{ij}$
 - $d_j = \min z_{ij}$
- Dále se vypočítají vzdálenosti od ideální varianty (IH) a bazální varianty (BH).

$$d_i^+ = \sqrt{\sum_{j=1}^n (z_{ij} - h_j)^2}; i = 1, 2, \dots, m$$

$$d_i^- = \sqrt{\sum_{j=1}^n (z_{ij} - d_j)^2}; i = 1, 2, \dots, m$$

- Najdou se **ideální** hodnoty kritérií, což znamená, že se najdou maxima kritérií (řádků), která se zvýrazní červenou barvou.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0,122	0,074	0,074	0,048	0,110
CENA	0,114	0,102	0,090	0,076	0,052
EKOSYSTÉM	0,183	0,147	0,147	0,090	0,072
SLUŽBY	0,098	0,098	0,110	0,074	0,060
VZHLED-VIZUÁL	0,047	0,041	0,052	0,035	0,047

Tabulka 5 Maxima

- Od maxima v každém řádku se odečtou hodnoty z každého řádku a umocní se na druhou. Nakonec se sečtou všechny hodnoty ve sloupci a odmocní se. Dostaneme výsledek, který je zapsán dole ve žluté tabulce.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0	0,002304	0,002304	0,005476	0,000144
CENA	0	0,000144	0,000576	0,001444	0,003844
EKOSYSTÉM	0	0,001296	0,001296	0,008649	0,012321
SLUŽBY	0,000144	0,000144	0	0,001296	0,0025
VZHLED-VIZUÁL	0,000025	0,000121	0	0,000289	0,000025

0,013	0,063316664	0,064621978	0,13097328	0,137237021
-------	-------------	-------------	------------	-------------

Tabulka 6 Ideální hodnoty

- Najdou se **bazální** hodnoty kritérií, což znamená, že se najdou minima kritérií (řádků), které se zvýrazní červenou barvou.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0,122	0,074	0,074	0,048	0,110
CENA	0,114	0,102	0,090	0,076	0,052
EKOSYSTÉM	0,183	0,147	0,147	0,090	0,072
SLUŽBY	0,098	0,098	0,110	0,074	0,060
VZHLED-VIZUÁL	0,047	0,041	0,052	0,035	0,047

Tabulka 7 Minima

- Od minima v každém řádku se odečtou hodnoty z každého řádku a umocní se na druhou. Nakonec se sečtou všechny hodnoty ve sloupci a odmocní se. Dostaneme výsledek, který je zapsán dole ve žluté tabulce.

KRITÉRIUM	GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
KAPACITA	0,005476	0,000676	0,000676	0	0,003844
CENA	0,003844	0,0025	0,001444	0,000576	0
EKOSYSTÉM	0,012321	0,005625	0,005625	0,000324	0
SLUŽBY	0,001444	0,001444	0,0025	0,000196	0
VZHLED-VIZUÁL	0,000144	0,000036	0,000289	0	0,000144

0,15241063	0,101395266	0,102635277	0,033105891	0,063150614
------------	-------------	-------------	-------------	-------------

Tabulka 8 Bazální hodnoty

- Na závěr se vypočítají relativní ukazatele vzdálenosti od bazální varianty pomocí vzorce $UV = BH / (IH + BH)$. Tyto hodnoty se vypočítaly v předcházejících tabulkách, a to konkrétně ve žlutých. Pro přehlednost změním barvu BH na světle zelenou a IH zůstane žlutá. Třetí řádek je $IH + BH$.

GOOGLE DRIVE	ONEDRIVE	APPLE ICLOUD	DROPBOX	MEGA
0,013	0,063316664	0,064621978	0,13097328	0,137237021
0,15241063	0,101395266	0,102635277	0,033105891	0,063150614

0,16541063	0,16471193	0,167257255	0,164079171	0,200387635
------------	------------	-------------	-------------	-------------

0,921407711	0,615591512	0,61363722	0,201767786	0,315142269
VÝSLEDEK				

Tabulka 9 Výsledek

10.2 Zhodnocení

Podle metody TOPSIS se nejvíce blíží ideálu cloudové úložiště od společnosti Google s názvem Google Drive. Na druhém místě se umístil OneDrive od Microsoftu a jako třetí iCloud od Applu. Zbylé pozice obsadilo na 4. příčce úložiště Mega a poslední skončil Dropbox.

11 ZÁVĚR

Cílem bakalářské práce bylo popsat a na modelovém příkladu analyzovat většinu služeb datových úložišť. Cloud computing je relativně mladá služba, ale i tak se z ní stal světový fenomén. Nejedná se jen o firemní sektor, ale čím dál tím více i o běžné uživatele, kterým přináší mnoho výhod a jistou volnost při nakládání s jejich daty. Vzhledem k tomu, jakým tempem se vyvíjí nové technologie, je zcela zřejmé, že zlatá éra cloud computingu tak rychle neskonečí, naopak se bude ještě více rozvíjet. Jako hlavní problém, který by tento vývoj mohl pozastavit, vidím bezpečnost, protože právě ta je zatím asi největší překážkou, minimálně pro firmy, k většímu využívání této služby. Jako nejlogičtější varianta se zdá být ta, kde se využije hybridní model cloud computingu, který dovoluje využívat všechny výhody cloudů a zároveň mít data, která by mohla být zneužita v bezpečí svých datových úložišť.

Hlavním cílem praktické části bylo popsat datová úložiště, jejich formy a jaké mají výhody a nevýhody. Následně se zaměřit na cloud computing a jeho vlastnosti a z čeho se skládá, i zde byly zmíněny jeho výhody a nevýhody, a také jeho zabezpečení a možná rizika a cloudové platformy. Jako poslední část praktické části se nahlédlo na stále populárnější Big Data, která jdou až dnes řádně analyzovat, díky výkonnějším technologiím, jak hardwarovým, tak softwarovým.

Praktická část práce je věnována cloudovým úložištím, a to zhodnocení jejich výhod a nevýhod, synchronizace a poté vyjmenování nejběžnějších privátních cloudů. Poslední část práce je analýza těchto služeb za pomoci TOPSIS analýzy.

SEZNAM POUŽITÉ LITERATURY

- [1] DATACENTRUM WEDOS. Úložné systémy - NAS vs. SAN [online]. 2010 [cit. 2016-03-23]. Dostupné z: <https://datacentrum.wedos.com/a/78/ulozne-systemy-nas-vs-san.html>
- [2] VÍTEK, Jan a Petr STRÁNSKÝ. SVĚT HARDWARE. Funkčnost, rozhraní a technologie pevných disků: SMART, RAID a speciální typy disků [online]. 2009 [cit. 2016-03-15]. Dostupné z: <http://www.svethardware.cz/funkcnost-rozhrani-a-technologie-pevnych-disku/16088-4>
- [3] Cloud computing. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2016-03-11]. Dostupné z: https://cs.wikipedia.org/wiki/Cloud_computing
- [4] NIST. NIST Cloud Computing Program [online]. 2010 [cit. 2016-03-19]. Dostupné z: <http://www.nist.gov/itl/cloud/>
- [5] MÁCHA, Petr. DIMENSION DATA CZECH REPUBLIC. Cloud computing – historie a budoucnost: Co je to vlastně CLOUD computing? [online]. 2012 [cit. 2016-03-27]. Dostupné z: <http://www.ddconnect.cz/brezen-2012/datova-centra.html>
- [6] KŘÍŽ, Lukáš. IT BIZ. Cloudové kompendium pro SMB – parametry a benefity: Svatá trojice cloudových obav [online]. 2016 [cit. 2016-04-28]. Dostupné z: <http://www.itbiz.cz/zpravicky/cloudove-kompendium-pro-smb-parametry-a-benefity>
- [7] MATĚJŮ, David. Cloud.cz. ROLE BEZPEČNOSTI V DŮVĚRYHODNÉM „CLOUDU“ [online]. [cit. 2016-04-14]. Dostupné z: <http://www.cloud.cz/bezpenost/175-role-bezpecnosti-v-duveryhodnem-cloudu.html>
- [8] STANČÍK, Martin. COMPUTERWORLD. Tři největší hrozby pro vaše data na cloudu [online]. 2012 [cit. 2016-04-15]. Dostupné z: <http://computerworld.cz/securityworld/tri-nej-vetsi-hrozby-pro-vase-data-na-cloudu-49203>
- [9] HNĚVKOVSKÝ, Petr. Živě.cz. Už víte, co je to DoS. Tak teď se naučte bránit Více na: <http://www.zive.cz/clanky/uz-vite-co-je-to-dos-tak-ted-se-naucte-branit/sc-3-a-101772/default.aspx> [online]. 2001 [cit. 2016-04-5]. Dostupné z: <http://www.zive.cz/clanky/uz-vite-co-je-to-dos-tak-ted-se-naucte-branit/sc-3-a-101772/default.aspx>
- [10] Windows Azure Platform. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2016-04-28]. Dostupné z: http://cs.wikipedia.org/wiki/Windows_Azure_Platform

- [11] ŽALUD, Radek. Licence (nejen) ve Windows Azure Virtual Machines [online]. 2014 [cit. 2016-04-28]. Dostupné z: <http://www.zive.cz/clanky/licence-nejen-ve-windows-azure-virtual-machines/sc-3-a-172737/default.aspx>
- [12] Microsoft Azure: Moderní podnikový cloud [online]. 2014 [cit. 2016-04-28]. Dostupné z: <https://azure.microsoft.com/cs-cz/>
- [13] Azure Machine Learning: Jak Azure Machine Learning funguje? [online]. 2014 [cit. 2016-04-29]. Dostupné z: http://www.intecs.cz/azure-machine-learning-ml/?gclid=Cj0KEQjwmpW6BRCf5sXp59_U_ssBEiQAGCV9GvxfNn-VzRokXIGvMYAODRNQ5wsKAzaLmt_ssVo2v4egaAo3M8P8HAQ
- [14] Google Cloud Platform: Build What's Next Better software. Faster. [online]. 2016 [cit. 2016-05-12]. Dostupné z: <https://cloud.google.com/>
- [15] Amazon Web Services: PRODUCTS & SERVICES [online]. 2016 [cit. 2016-05-15]. Dostupné z: <https://aws.amazon.com/>
- [16] Salesforce: OUR PRODUCTS [online]. 2016 [cit. 2016-05-15]. Dostupné z: <https://www.salesforce.com/>
- [17] Big data. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2016-05-28]. Dostupné z: https://cs.wikipedia.org/wiki/Big_data
- [18] KUŽNÍK, Jan a Pavel KASÍK. Co umějí Big Data?: Třeba pošlou Formuli 1 do boxů nebo poradí fotbalistům [online]. 2014 [cit. 2016-04-15]. Dostupné z: http://technet.idnes.cz/big-data-c53-/veda.aspx?c=A141010_115149_veda_kuz
- [19] MARR, Bernard. Big Data: The 5 Vs Everyone Must Know [online]. 2014 [cit. 2016-04-16]. Dostupné z: <https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know>
- [20] ČERNÝ, Michal. Big data a jejich možnosti v kontextu knihoven [online]. 2013 [cit. 2016-05-22]. Dostupné z: <http://oldknihovna.nkp.cz/knihovna131/131104.htm>
- [21] CHLUP, Ondrej. Co mají Apache Hadoop a ArcGIS společného? [online]. 2014 [cit. 2016-05-20]. Dostupné z: <https://www.arcdata.cz/sluzby-a-podpora-zakazniku/podpora/clanek/co-maji-apache-hadoop-a-arccgis-spolecneho>
- [22] PAUL, Ian a David ČEPIČKA. Uvolněte místo na disku úpravou synchronizace s cloudem [online]. 2015 [cit. 2016-05-10]. Dostupné z: <http://pcworld.cz/internet/uvolnete-si-misto-na-pevnem-disku-upravou-synchronizace-s-ulozistem-na-cloudu-1-dil-47981>

- [23] KUPKA, Lukáš. Jaká jsou aktuálně nejlepší cloudová úložiště [online]. 2014 [cit. 2016-05-11]. Dostupné z: <http://www.zive.cz/clanky/jaka-jsou-aktualne-nejlepsi-cloudova-uloziste/sc-3-a-174542/default.aspx>
- [24] Google Drive [online]. 2016 [cit. 2016-03-11]. Dostupné z: https://www.google.com/intl/cs_cz/drive/
- [25] Microsoft OneDrive [online]. 2016 [cit. 2016-03-12]. Dostupné z: <https://onedrive.live.com/about/cs-cz/>
- [26] Apple iCloud [online]. 2016 [cit. 2016-03-12]. Dostupné z: <http://www.apple.com/cz/icloud/>
- [27] Dropbox [online]. 2016 [cit. 2016-03-13]. Dostupné z: <https://www.dropbox.com/>
- [28] Mega [online]. 2016 [cit. 2016-03-13]. Dostupné z: <https://mega.nz/>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

HDD	Hard Disk Drive
FTP	File Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
SSL	Secure Sockets Layer
TLS	Transport Layer Security
IBM	International Business Machines Corporation
GB	Gigabyte
ECC	Error Checking and Correcting
XOR	Exclusive OR
NIST	National Institute of Standards and Technology
IDE	Integrated Development Environment
API	Application Programming Interface
ERP	Enterprise Resource Planning
CRP	Capacity Requirements Planning
SLA	Service Level Agreement
GPS	Global Positioning System
GSM	Global System for Mobile Communications
SQL	Structured Query Language
TB	Terabyte
MB	Megabyte

SEZNAM OBRÁZKŮ

Obr. 1 NAS	13
Obr. 2 SAN	15
Obr. 3 RAID 01	18
Obr. 4 Obavy z cloud computingu	27
Obr. 5 Diagram znázorňující jednotlivé fáze práce s velkými daty.....	37
Obr. 6 Apache Hadoop	40
Obr. 7 Cloud	42
Obr. 8 Google Drive logo	46
Obr. 9 Google Drive – nový vzhled.....	47
Obr. 10 Microsoft OneDrive.....	47
Obr. 11 Microsoft OneDrive – rozdíly (poznámka – prostor u OneDrive je nyní jen 5GB)	49
Obr. 12 Apple iCloud.....	49
Obr. 13 Dropbox	50
Obr. 14 Mega	51
Obr. 15 Mega – prostředí	52

SEZNAM TABULEK

Tabulka 1 Hodnoty	54
Tabulka 2 Mocnina a odmocnina.....	54
Tabulka 3 Přitažlivost	55
Tabulka 4 Váhy kritérií.....	55
Tabulka 5 Maxima	56
Tabulka 6 Ideální hodnoty	56
Tabulka 7 Minima.....	56
Tabulka 8 Bazální hodnoty	57
Tabulka 9 Výsledek	57

SEZNAM GRAFŮ

Graf 1 Cloudová úložiště 2014	45
-------------------------------------	----