

Návrh a implementace datového serveru navázaného na DÚ Cesnet

Bc. Martin Sousedík

Diplomová práce
2018



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
akademický rok: 2017/2018

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Martin Sousedik**
Osobní číslo: **A16154**
Studijní program: **N3902 Inženýrská informatika**
Studijní obor: **Bezpečnostní technologie, systémy a management**
Forma studia: **prezenční**

Téma práce: **Návrh a implementace datového serveru navázaného na DÚ
Cesnet**

Téma anglicky: **The Design and Implementation of a Storage Server Linked to DS
Cesnet**

Zásady pro vypracování:

1. Zpracujte možnosti připojení k DÚ infrastruktury Cesnetu.
2. Definujte omezení využití DÚ pro různé aplikace používané na FAI.
3. Proveďte návrh implementace připojení naší infrastruktury k DÚ.
4. Implementujte řešení v testovací infrastruktuře.
5. Otestujte výkonnost a odezvu Vašeho řešení.
6. Stanovte limity počtu připojených zařízení k jednomu účtu DÚ.

Rozsah diplomové práce:

Rozsah příloh:

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

Seznam odborné literatury:

- MARINESCU, Dan C. Cloud computing theory and practise. Morgan Kaufmann; 1 edition. ISBN: 978-0124046276.
- NEMETH, Evi; HEIN, Trent R; SNYDER, Garth. Linux : kompletní příručka administrátora. Vyd. 1. Brno : Computer Press, 2004. 828 s. ISBN 8072269194
- LEIXNER, Miroslav. PC – zálohování a archivace dat : Nestůjte za dveřmi. Praha : Grada Publishing, a.s., 1993. 394 s. ISBN 8085424738.
- KASTNER, Aleš. Zálohování a archivace. Praha : GComp, 1997. 128 s. ISBN 8085649586.
- PECINOVSKÝ, Josef. Archivace a Komprimace Dat: Jak Zálohovat Data, Jak Komprimovat Soubory WinRAR, WinZip, WinAce, Windows a Nástroje Komprese Dat, Jak Archivovat Data Ve Windows. 2003.
- RYLICH, Jan. Cloudové služby: data i počítače v oblacích. Ikaros. 2012, ročník 16, číslo 9. ISSN 1212-5075. Dostupné z: <http://ikaros.cz/node/13965>

Vedoucí diplomové práce:

Ing. David Malaník, Ph.D.

Ústav informatiky a umělé inteligence

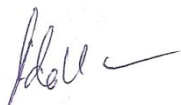
Datum zadání diplomové práce:

8. prosince 2017

Termín odevzdání diplomové práce:

28. května 2018

Ve Zlíně dne 8. prosince 2017



doc. Mgr. Milan Adámek, Ph.D.
děkan



doc. RNDr. Vojtěch Křesálek, CSc.
ředitel ústavu

Jméno, příjmení: Martin Sousedík

Název diplomové práce: Návrh a implementace datového serveru navázaného na DÚ
Cesnet

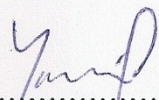
Prohlašuji, že

- 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 s tím, že vyrovnaní 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 23.5.2018


.....
podpis diplomanta

ABSTRAKT

Cílem práce je návrh a implementace připojení datových serverů na Univerzitu Tomáš Bati ve Zlíně, Fakultě aplikované informatiky a zálohování dat na Datové úložiště Cesnet. V teoretické části jsou popsány možnosti připojení, bezpečnost přenosu a datového úložiště a návrh propojení serverů s datovým úložištěm. Praktická část se zabývá implementací navrženého řešení, testování odezvy a výkonnosti řešení a stanovení limitů připojených zařízení.

Klíčová slova: Cesnet, zálohování, server, Debian, cloud, datové úložiště

ABSTRACT

The aim of the thesis is to design and implement data servers connection at Tomas Bata University in Zlín, Faculty of Applied Informatics and Data Backup to a Cesnet Data Storage. In the theoretical part are described the connection possibilities, the security of transmission and data storage and the design of connection of servers to data storage. The practical part deals with the implementation of the proposed solution, the testing of the response and the performance of the solutions and the determination of the limits of the connected devices.

Keywords: Cesnet, backup, archive, server, Debian, cloud, data storage

Tímto bych chtěl poděkovat svému vedoucímu, panu Ing. Davidu Malaníkovi, Ph.D., za vedení, užitečné rady a především zkušenosti, které mi předal během tvorby diplomové práce a také za čas, po který se mi věnoval.

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

OBSAH

ÚVOD	8
I TEORETICKÁ ČÁST	9
1 MOŽNOSTI PŘIPOJENÍ K DATOVÉMU ULOŽIŠTI	10
1.1 TOPOLOGIE SÍTĚ CESNET2	10
1.2 INFRASTRUKTURA DATOVÝCH ULOŽIŠŤ.....	11
1.2.1 Architektura datového centra	12
1.3 PROTOKOLY PRO PŘÍSTUP DO ÚLOŽIŠTĚ.....	13
1.3.1 Protokol NFSv4.....	13
1.3.2 Protokol Samba	13
1.3.3 Protokol FTPS.....	14
1.3.3.1 Protokol TLS.....	14
1.3.4 Protokol SCP/SFTP.....	14
1.3.5 Protokol Rsync	14
1.4 POROVNÁNÍ PROTOKOLŮ.....	14
2 OMEZENÍ VYUŽITÍ DATOVÉHO ULOŽIŠTĚ	16
2.1 BEZPEČNOST DŮ A ZACHÁZENÍ S DATY.....	17
3 MOŽNOSTI PŘIPOJENÍ INFRASTRUKTURY K DŮ	18
3.1 MOŽNOST Č. 1 – DVOUVRSTVÁ ARCHITETURA.....	18
3.2 MOŽNOST Č. 2 – TRÍVRSTVÁ ARCHITEKTURA	19
II PRAKTICKÁ ČÁST	20
4 IMPLEMENTACE ŘEŠENÍ	21
4.1 MONITOROVACÍ SYSTÉM MONITORIX	21
4.1.1 Instalace systému Monitorix [11].....	22
4.2 ZÁLOHOVACÍ APLIKACE BAREOS	24
4.2.1 Instalace zálohovací aplikace BareOS [12].....	24
4.2.2 Nastavení testovacích záloh hlavního serveru	25
4.2.3 Nastavení záloh prvního klienta (host1)	30
4.2.4 Nastavení záloh druhého klienta (host2).....	31
5 NASTAVENÍ A OTESTOVÁNÍ ODEZVY A VÝKONNOSTI	33
5.1 NASTAVENÍ PRO HLAVNÍ SERVER	33
5.2 NASTAVENÍ PRO KLIENTA HOST 1	34
5.3 NASTAVENÍ PRO KLIENTA HOST 2.....	35
5.4 OTESTOVÁNÍ VÝKONNOSTI A ODEZVY	35
5.4.1 Záloha hlavního serveru	36
5.4.2 Záloha klientů.....	37
6 STANOVENÍ LIMITŮ POČTU PŘIPOJENÝCH ZAŘÍZENÍ	41
ZÁVĚR	43
SEZNAM POUŽITÉ LITERATURY	44
SEZNAM OBRÁZKŮ	45
SEZNAM TABULEK	46

ÚVOD

Potřeba ochránit svá data před ztrátou nebo únikem citlivých dat je známa už od pradávna. Už i v dobách, kdy se psalo, či spíše rylo, do hliněných destiček tato potřeba uchování dat byla. S vynálezem papíru přišla úplně nová doba. Papír byl totiž mnohanásobně menší rozměry a je tedy dokumenty jednodušší uchovat z dosahu neoprávněných osob a hlavně pro uchování velkého objemu papírových dokumentů stačí mnohem menší prostor než pro takové destičky. Samozřejmě potřeba uchovávat data v papírové formě přetrvává mnohdy až dodnes. Ovšem vývoj je nezastavitelný a i když i dnes existují archivy pro uchování dokumentů a různé dokumentace, máme dnes již spoustu dalších možností, jak svá data chránit.

Žijeme v moderní době, která je plná elektrických zařízení, elektroniky a dalších novodobých vynálezů, které se od svého vynalezení za těch několik let proměnily k nepoznání. Žijí mezi námi lidé, kteří ještě zažili dobu bez všech dnešních elektronických vychytávek a jsou živoucím důkazem toho, jak moc rychlý je vývoj v této oblasti. Například doba, kdy existoval pouze stolní telefon. Kdo by si mohl tehdy myslet, že si stejným zařízením v budoucnu dokáže cokoli vyfotit? Nebo například začínající počítačová doba, kdy jeden počítač obsadil mnohdy i několik místností. Dnes si může každý dopřát svůj počítač a dokonce si jej může vzít všude s sebou.

Tyto moderní trendy tedy samozřejmě spějí k tomu, že spousta dat je v digitální podobě. Jde ale taková data uchovat před ztrátou či odcizením, když je uživatel nedokáže fyzicky držet v ruce? Je samozřejmě možné si data zálohovat na externí datová média, tedy flash disky, nebo externí disky a toto datové médium pak uschovat a zamknout, v případě, že jde o důležitá data. Dnes už ale existují tzv. Cloudová úložiště, která slouží majitelům dat pro uchovávání dat, o která nechtějí přijít. Někdo si může říct, že dát svá data někam na vzdálené úložiště nemůže být bezpečné. Pokud si uživatel přeje, aby mohl číst data pouze on sám, může si data zašifrovat a až pak teprve uložit. Co se týče životnosti datových médií, více nebezpečné je nechat si data na svém datovém médiu, protože společnosti, vlastníci datová úložiště, si dokáží ohlídat životnost svých datových médií a navíc mají mnohem lépe zajištěné uchování dat, i kdyby k poškození jednoho z jejich datového média došlo.

I. TEORETICKÁ ČÁST

1 MOŽNOSTI PŘIPOJENÍ K DATOVÉMU ULOŽIŠTI

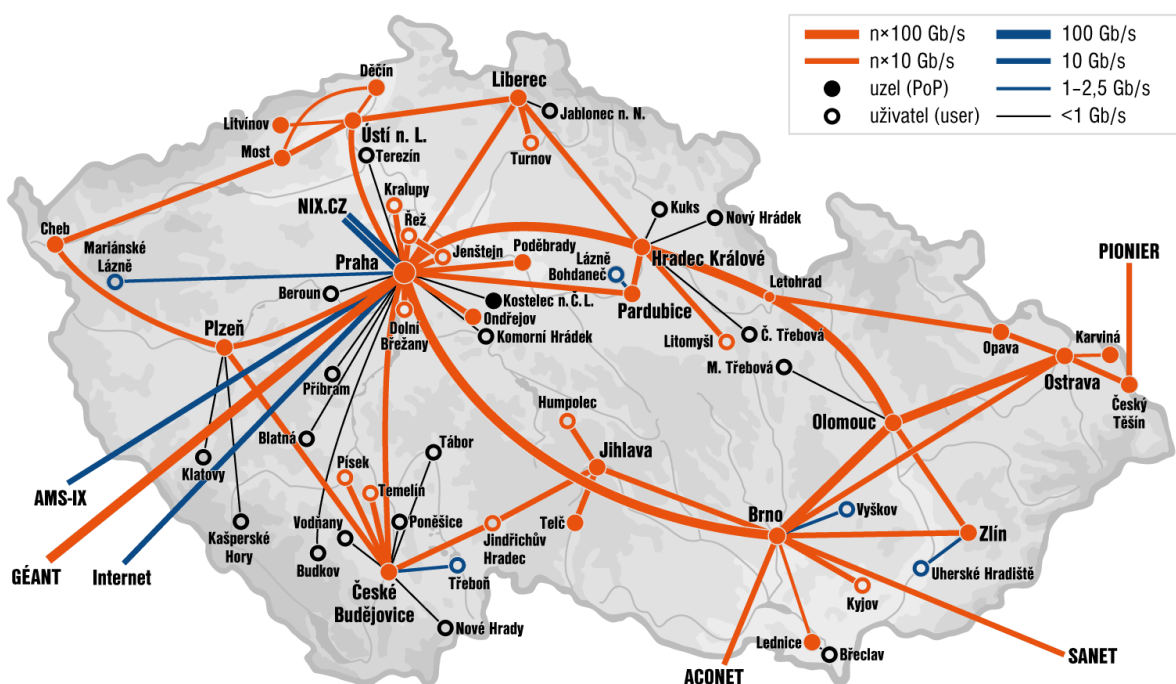
Poskytovatel datového úložiště nabízí zpravidla několik možností, kterými je možné se připojit k jeho infrastruktuře. Každá z možností má však své výhody a nevýhody, a je nutné vybrat vhodnou variantu z hlediska spolehlivosti, rychlosti a bezpečnosti. Vzhledem k tomu, že společnost Cesnet, z.s.p.o. má k dispozici více datových úložišť, je zapotřebí vybrat to nejvýhodnější.

1.1 Topologie sítě CESNET2

Síť CESNET2 je poměrně rozsáhlá a pokrývá téměř celou Českou Republiku. Hlavní páteřní síť je realizována pomocí optických kabelů. Ty jsou použity také pro připojení dalších vedlejších sítíových okruhů.

Pro co nejvíce ideální podmínky, z hlediska přenosové rychlosti, je potřeba zvolit vhodnou přenosovou trasu, tedy především jedno z nabízených úložišť v rámci sítě společnosti Cesnet, z.s.p.o.. Počítačová síť CESNET2 disponuje propojovacími trasami, které umožňují přenášet data s rychlostí 1 až 100 Gb/s. [1]

Výhodou je také možnost vytvoření VPN (Virtual Private Network – Virtuální privátní síť), ale především záruka vysoké spolehlivosti, které je dosaženo díky redundanci páteřního vedení, což znamená, že každý páteřní uzel má více nezávislých připojení.



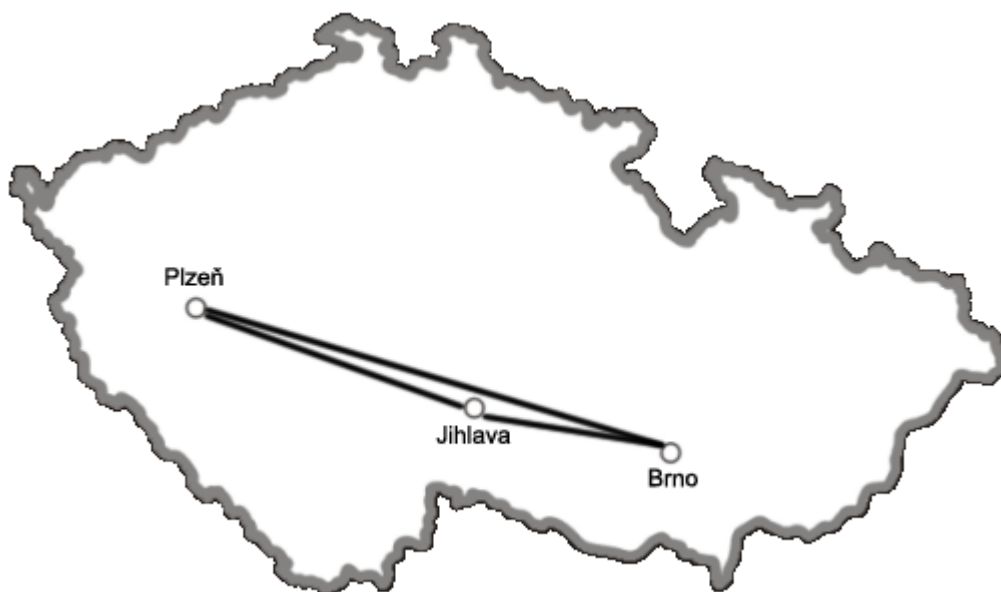
Obrázek 1: Topologie sítě CESNET2 [2]

Na obrázku je znázorněna topologie sítě CESNET2. Nás bude především zajímat, jak je propojené město Zlín se zbytkem sítě. Město Zlín je v síti propojeno se dvěma uzly (Olomouc a Brno) přes připojení o rychlosti 10 Gb/s, takže tato rychlost je nejvyšší, jakou lze v rámci připojení do sítě CESNET2 dosáhnout.

1.2 Infrastruktura datových úložišť

Pokrytí sítí CESNET2 je velké a disponuje množstvím propojovacích uzlů. Společnost Cesnet, z.s.p.o. má v České Republice datová úložiště ve třech městech (uzlech), konkrétně v Plzni, v Jihlavě a v Brně.

Připojení serverů nacházejících se na Fakultě aplikované informatiky ve Zlíně bude tedy nejvýhodnější použít datové úložiště v Brně, protože je nejbližší a vyšší rychlost než 10 Gb/s není možné v případě síťového uzlu (Zlín) dosáhnout.



Obrázek 2: Mapa datových úložišť v síti CESNET2 [3]

Úložiště využívají dva modely - distribuovaný model a hierarchický model.

Distribuovaný model je reprezentovaný geografickým oddělením jednotlivých úložišť. I když jsou úložiště vzájemně propojena, každé se nachází v jiné části České republiky. Díky tomu se snižuje pravděpodobnost ztráty dat především živelnou katastrofou, ale také poškozením médií nebo chybným úkonem uživatele či administrátora. Další výhodou je

také možnost výběru úložiště, které je co nejbližší k uživateli, což umožňuje dosáhnout co nejvyšší rychlosti přenosu dat, a případné následné řešení problémů je mnohem jednodušší. [4]

Hierarchický model funguje na principu přesouvání méně často používaných dat uložených na úložišti na levnější úložná média s nižší rychlostí, ale větší kapacitou, kterými jsou především páskové knihovny a MAID (Massive Array of Idle Disc) pole. Naproti tomu často používaná data jsou uložena na médiích s nižší kapacitou, zato větší rychlostí. Mezi tato úložná média patří paměti cache a rychlé disky. Výhodou hierarchického modelu je především to, že s nižšími náklady je možné dosáhnout mnohem vyšší kapacity, než kdyby bylo úložiště tvořené pouze disky. [4]

1.2.1 Architektura datového centra

Vzhledem k tomu, že k zálohování bude používáno datové úložiště v Brně, bude zde popsána pouze architektura tohoto úložiště. Jak už bylo zmíněno v předchozí kapitole, společnost Cesnet, z.s.p.o. využívá hierarchický model. Ten však není na každém úložišti stejný. Hierarchický model úložiště v Brně je na následujícím obrázku.



Obrázek 3: Hierarchický model datového úložiště v Brně [5]

Jednotlivé vrstvy (tier) úložiště jsou vzájemně propojené, aby mohlo docházet k přesunu dat mezi jednotlivými datovými médii. Celková kapacita úložiště 6134 TiB je rozdělena mezi jednotlivé vrstvy následovně:

- **Vrstva 1** – Dvě disková pole o kapacitě 498 TiB
- **Vrstva 2** – Tři pole MAID o kapacitě 2116 TiB

- **Vrstva 3** – Pásková knihovna o kapacitě 3520 TiB [4]

Díky této architektuře je možné dosáhnout velké kapacity datového úložiště a vzhledem k nízké ceně především páskových knihoven také lepší poměr ceny za 1 TiB.

Poznámka: Kapacita je udávána v jednotkách TiB (tebibyte). Jedná se o jednotku, která je používaná pro objem digitální informace. Předpony jsou odlišné, protože se jedná o binární předpony. Převod je následující: $1\text{TiB} = 1,1 \times 10^{12} \text{ B} = 1,1 \text{ TB}$.

1.3 Protokoly pro přístup do úložiště

Aby mohlo být úložiště využíváno, je zapotřebí se k němu připojit. K tomu slouží několik protokolů, skrze které je možné k danému úložišti přistupovat.

1.3.1 Protokol NFSv4

Protokol NFSv4 (Network File System verze 4, popsáný v dokumentu RFC 3530 [6]) byl vyvinut především pro uživatele, kteří vlastní zařízení s operačním systémem, jehož základem je operační systém Linux. Jednou z podmínek je však přístup s právy uživatele root.

V případě použití zařízení s operačním systémem Windows je nutné přistupovat přes komerčního klienta NFS Maestro.

V obou variantách operačních systémů je zapotřebí mít nainstalovanou podporu protokolu Kerberos, který zajišťuje spolehlivou autentizaci uživatelů. Protokol NFSv4 je sice multiplatformní, ale nelze jej použít pro operační systémy Mac OS z důvodu problému s mapováním uživatelů. Protokol NFS totiž místo UserID používá formát *uživatel@doména*. [7]

Díky protokolu NFSv4 lze připojit svazek tak, že jej uživatel vidí jako běžný lokální disk a lze na něj zapisovat data. Nevýhodou je složitost počátečního nastavení, proto je tento protokol vhodný spíše pro zkušenější uživatele. Velkou výhodou je především možnost využití šifrované komunikace. [8]

1.3.2 Protokol Samba

Samba je protokol, který byl vytvořen především pro uživatele zařízení s operačním systémem Windows. Velkou výhodou je bezesporu to, že nastavení je mnohem jednodušší než u protokolu NFSv4, tudíž protokol Samba mohou využívat zkušenější uživatelé, ale i ti

méně zkušenosti. Samba umožňuje připojit svazek stejně jako NFSv4, tedy uživatel jej vidí jako lokální disk. Nevýhodou je, že Samba neumožňuje šifrování. To jde ale vyřešit použitím některé z šifrovacích metod (například IPsec). [7]

1.3.3 Protokol FTPS

Protokol, který pro přihlášení, případně i pro data využívá šifrování pomocí protokolu TLS. Protokol FTPS je vhodný pro přenos velkých souborů, ale nezachovává vlastníka originálních souborů ani čas modifikace. Na tento protokol lze využít mnoho aplikací, například známý Total Commander a další aplikace pro správu souborů. [7]

1.3.3.1 Protokol TLS

Kryptografický protokol TLS (Transport Layer Security) je nástupcem protokolu SSL. Slouží pro zabezpečení internetové komunikace proti odposlechu nebo podstrčení falešných dat pomocí autentifikace (ověření pravosti). Používá se především pro autentizaci serverů, tedy koncoví uživatelé už nepotřebují autentizaci, protože je zaručená totožnost serveru. [7]

1.3.4 Protokol SCP/SFTP

Je vhodný především pro přenos velkých souborů, stejně jako FTPS, ale zachovává informace o souborech. Výhodou je, že přenos dat je šifrovaný a lze pro správu využít běžné správce souborů (FileZilla, apod.). [7]

1.3.5 Protokol Rsync

Jde o nástroj pro zálohování dat. Při zálohování nových souborů nástroj pracuje s metadaty. Kvůli tomu je rychlost přenosu maximálně 40 souborů za vteřinu, aniž by ale záleželo na velikosti. V případě kontroly existujících souborů je rychlost až 10 krát větší. U zazálohovaných souborů zachovává informace o vlastníkovi a jeho práva. [7]

1.4 Porovnání protokolů

Ze všech možných protokolů je potřeba zvolit takový, který nejlépe vyhovuje svými parametry dané aplikaci. Porovnávací kritéria jsou především možnost šifrované komunikace a rychlost přenosu dat. [7]

Název protokolu	Šifrování	Rychlost
NFSv4	ANO	50 MB/s
Samba	NE	10-12MB/s
FTPS	ANO	10 MB/s
SCP/SFTP	ANO	5-8 MB/s
Rsync	NE	40 souborů/s

Tabulka 1: Tabulka protokolů [Zdroj: vlastní]

V tabulce je rychlost dat vypsaná na základě reálných zkušeností uživatelů. Je jasné vidět, že parametry jednotlivých protokolů se liší. Ideální variantou je protokol, který splňuje požadavek možnosti šifrované komunikace a dokáže přenášet data s co největší přenosovou rychlostí. Vzhledem k těmto požadavkům vyhovuje nejlépe protokol NFSv4 a proto bude využit v praktické části při realizaci komunikace se serverem na síti CESNET2. Navíc protokol NFSv4 umožňuje datové úložiště připojit jako síťový disk, takže je možné k němu přistupovat odkudkoliv a kompatibilita mezi různými operačními systémy není problém, jako u ostatních protokolů. Další důležitou vlastností je, že protokol dokáže pomocí nástrojů (například tar) uchovat informaci o vlastníkovi a není pak problém s přístupem k souborům kvůli odlišným přístupovým právním.

2 OMEZENÍ VYUŽITÍ DATOVÉHO ÚLOŽIŠTĚ

Důležitou součástí ukládání dat je také bezpečnost vzdáleného úložiště z hlediska bezpečnosti uchování dat a nakládání s nimi, ale to skýtá různá omezení, ať už z hlediska přenosu či samotného uložení dat na datové úložiště.

Jedinou možnou variantou pro realizaci přenosu dat na datové úložiště je protokol NFSv4. Ten jediný splňuje požadavky na přenos z hlediska šifrování komunikace, rychlosti přenosu dat a dalších aspektů. Nevýhodou je sice to, že nedokáže sám o sobě zachovat vlastníka souborů, to lze ale vyřešit pomocí přídavných nástrojů. Dalším omezením je to, že protokol NFSv4 nedokáže pracovat s lokálními adresami serverů (např. 10.5.11.98) a vyžaduje tedy, aby byla serveru, který komunikuje skrze tento protokol, udělena veřejná IP adresa. Veřejnou IP adresu je možné, pokud má vlastník serverů veřejné IP adresy k dispozici. Dalším problémem je, že kromě veřejné IP adresy je potřeba zajistit resolvovatelné doménové jméno serveru, tedy ne ve tvaru IP adresy, ale adresy ve formě textu (doménového názvu), která se teprve přeloží na danou IP adresu.

Ze strany datového úložiště jsou samozřejmě také omezení, která se nedají přehlédnout. Pro zálohování dostane uživatel 2 TB místa na datovém úložišti, což je poměrně dostačující kapacita pro zálohy. Problémem je však především to, že maximální počet souborů, který je možné v rámci záloh uložit, je limitován počtem 100 000 souborů. Tento limit se může zdát jako dostatečně velký, ale například pro zálohování LMS (Learning Management System) Moodle není dostačující, protože počet souborů této aplikace maximální počet souborů překračuje téměř o jednonásobek limitu.

Informace ohledně šifrování dat na datovém úložišti Cesnet poskytl vedoucí oddělení datových úložišť, CESNET, z. s. p. o., pan RNDr. David Antoš, Ph.D. a vyjádřil se k tomuto problému následovně:

„My data nešifrujeme, protože šifrování na serveru nemá principiálně žádný smysl. Je sice pravda, že ušetří práci při případné likvidaci úložiště, protože pak na médiích zbude pouze „binární smetí“ a ne kusy souborů. My ale pro likvidaci používáme drtičku, což je spolehlivá metoda a není ani příliš drahá. Náš právní názor je, že libovolně citlivá data se dají učinit obyčejnými tím, že si je silně zašifruje uživatel na své straně a na úložiště si k nim nedá klíče, ale samozřejmě si nesmí klíče ztratit, protože by se k datům pak nedostal.

V každém případě jsou úložiště vybavena standardními přístupovými právy POSIXového souborového systému v případě služeb přístupu k souborovému systému.

Extra citlivá osobní data nebo medicínské dokumentace k nám na úložiště v otevřené podobě rozhodně nepatří, u dat tohoto typu je standardem, že neexistuje správce s přístupem ke všemu naráz (což u nás samozřejmě máme zavedeno).“

Z toho vyplývá fakt, že pokud chce mít uživatel svá data v bezpečí, musí si zajistit šifrování dat na serveru na své straně, což přináší zvýšené zatížení serveru a zpomalení záloh.

2.1 Bezpečnost DÚ a zacházení s daty

Datové úložiště umožňuje ukládat jak zálohy, tak i archivy. Pro archivy je potřeba si vymezit prostor, který je omezený jak objemově, tak i počtem souborů. Uložení záloh je omezené časově, přičemž maximální doba uložení jednotlivých souborů je 365 kalendářních dnů, ale omezení je také na počet souborů. Po překročení maximální doby uložení záloh je správce datového úložiště oprávněný data vymazat. Aby nedošlo k překročení limitů (doby uložení nebo objemu), systém zasílá upozornění dostatečně dopředu. [9, 10]

Uživatel DÚ má přístup pouze ke svým datům na první (nejvyšší) vrstvě, tedy k datům na diskových polích. Jak už bylo zmíněno výše, méně často používaná data se přesouvají na nižší vrstvy, což se děje automaticky. Uživatel má však možnost si detailně nastavit podmínky pro migraci dat mezi jednotlivými vrstvami. Migrací dat se zajišťuje právě ona spolehlivost a bezpečnost uložení dat. Pokud uživatel potřebuje přistoupit k datům, která jsou uložena na nižších vrstvách hierarchického modelu, pak se musí nejprve připravit. Tato operace je u souborů o velikosti desítek až stovek GB provedena v řádu sekund, maximálně jednotek minut, ale poskytovatel neuvádí konkrétní přístupovou dobu. V případě, že se uživatel potřebuje akutně dostat k datům, poskytovatel umožňuje nastavit této operaci vyšší prioritu a rychlost čtení tím zvýšit až na 200 MB/s. Integrita dat je zaručena vypočtením MD5 hashe, který je poté uložen do metadat souborů. [9, 10]

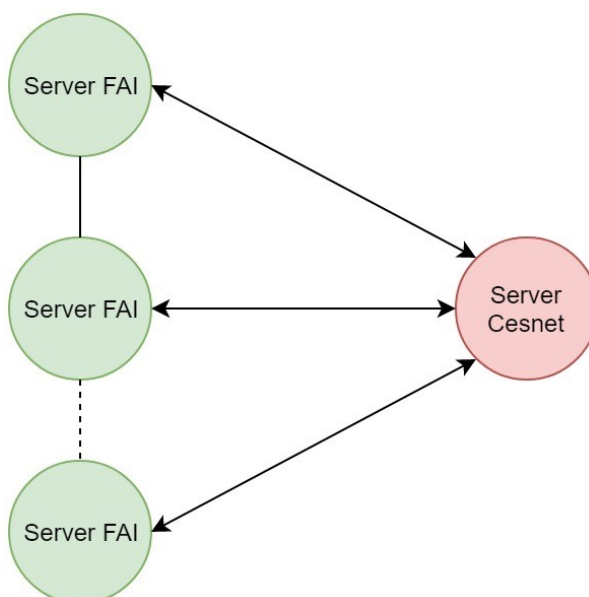
3 MOŽNOSTI PŘIHOJENÍ INFRASTRUKTURY K DŮ

Posledním krokem návrhu je návrh možností propojení v rámci infrastruktury na Fakultě aplikované informatiky a propojení této infrastruktury s úložištěm (serverem) firmy Cesnet, z.s.p.o..

Pro návrh se nabízí dvě varianty, z nichž každá má své výhody a nevýhody a je proto nutné zvážit, která z těchto dvou variant bude výhodnější a především jednodušší na správu z hlediska instalace a následné údržby a odstraňování poruch.

3.1 Možnost č. 1 – Dvouvrstvá architektura

První možností je architektura dvouvrstvá. To znamená, že všechny zálohované servery jsou spojeny přímo se zálohovacím datovým úložištěm (Server Cesnet).

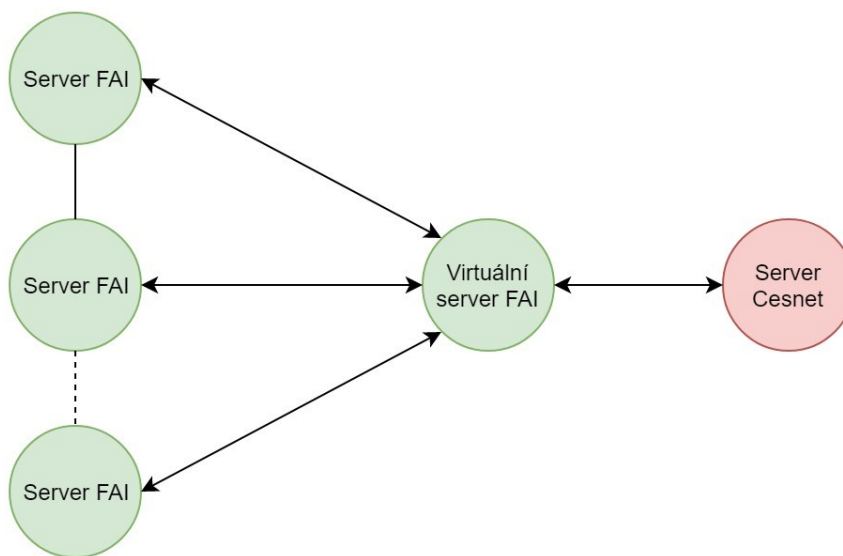


Obrázek 4: Možnost propojení č. 1 [Zdroj: vlastní]

Výhodou je především to, že každý server komunikuje s datovým úložištěm zvlášť a pokud se pokazí jeden ze serverů FAI, nemá to žádný vliv na ostatní servery. Poměrně velkou nevýhodou ale je složitá správa takové infrastruktury, například správa protokolů nebo fakt, že zálohovací software je zapotřebí nainstalovat na každý server zvlášť, což zabere poměrně hodně času a nelze dělat jednotné úpravy nastavení.

3.2 Možnost č. 2 – Třívrstvá architektura

Druhou možností je třívrstvá architektura. Třetí vrstva vzniká díky vřazení virtuálního serveru, který zajišťuje propojení serverů FAI ještě v rámci lokální sítě a teprve tento virtuální server komunikuje a posílá data na zálohovací server (datové úložiště) Cesnet.



Obrázek 5: Možnost propojení č. 2 [Zdroj: vlastní]

Tato architektura má jedinou nevýhodu, která je viditelná i z Obrázku 5. Pokud se pokazí virtuální server, tak není možné zálohovat. To je ale jediná nevýhoda, tohoto zapojení má více výhod. Mezi ně patří:

- všechny zálohované servery jsou propojené přímo v rámci sítě na FAI,
- veškeré problémy se dají řešit přímo na místě,
- instalace zálohovacího softwaru není nutná na každém serveru zvlášť, software stačí nainstalovat pouze na virtuální server, který komunikuje s datovým úložištěm,
- dají se snadno provádět změny, nemusí se složitě řešit přes Cesnet, z.s.p.o.,
- jednotlivé vrstvy jsou na sobě nezávislé.

Tato varianta je oproti předchozí mnohem lepší, spolehlivější a jednodušší na správu. Pokud bude virtuální server dostatečně spolehlivý, pak je tato architektura nejvhodnější pro tuto aplikaci a při realizaci bude využita právě tato struktura.

II. PRAKTICKÁ ČÁST

4 IMPLEMENTACE ŘEŠENÍ

V teoretické části jsou popsány možnosti, jakým způsobem lze komunikovat s datovým úložištěm, jakou architekturu (propojení v rámci sítě) je možné realizovat a také omezení v rámci ukládání dat. Praktická část je tedy zaměřena na výběr vhodné varianty, realizaci a instalaci potřebných nástrojů pro zálohování a sledování provozu na serveru.

4.1 Monitorovací systém Monitorix

Pro vytvoření statistik ohledně provozu na serveru je ideální variantou instalace vhodného programu, který takové statistiky dokáže vytvořit. Jako nejvhodnější se ukázal systém Monitorix. Velkou výhodou je, že se jedná o freewarový monitorovací systém a poskytuje nepřehledné množství statistik o systémových prostředcích a službách, běžících na serverech s operačním systémem Linux, nebo i jiných distribucích na bázi OS Linux. Veškeré statistiky jsou přehledně zobrazené pomocí grafů, které lze zobrazit pomocí webového rozhraní.

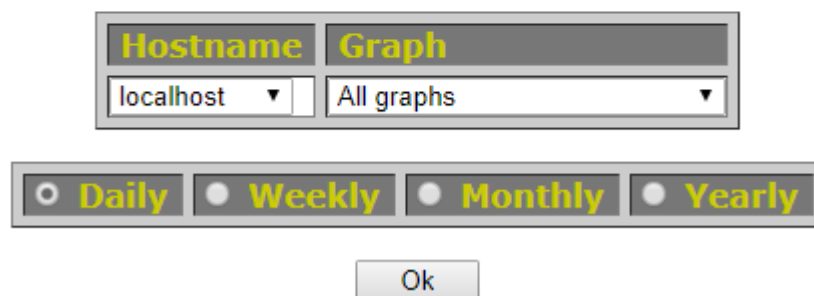
Mezi statistiky patří například:

- průměrné zatížení systému a jeho využití,
- využití jádra (kernelu),
- zobrazování teplot na jednotlivých senzorech,
- informace o zdraví disku,
- použití souborového systému a činnost vstupů a výstupů,
- použití adresářů,
- síťový provoz a jeho využití,
- síťový provoz na jednotlivých portech,
- statistiky serveru Apache,
- a mnoho dalších.

Jedná se o nenáročný systém, který zabere málo místa a jeho obsluha je jednoduchá. Program umožňuje zobrazit jednu konkrétní statistiku nebo všechny zároveň. Další možností je výběr výpisu statistik denních, týdenních, měsíčních nebo ročních.

Monitorix

v3.10.1



Obrázek 6: Úvodní obrazovka systému Monitorix [Zdroj: vlastní]

Pro instalaci systému Monitorix je použitý postup popsáný níže.

4.1.1 Instalace systému Monitorix [11]

První krok: Přidání úložiště Monitorixu

1. Vzhledem k tomu, že je potřeba změnit (doplnit repozitář Monitorixu) systémový soubor, pro jistotu jej zazálohujeme
`cp -pf /etc/apt/sources.list_backup`
2. Poté soubor otevřeme pomocí textového editoru pico
`Pico /etc/apt/sources.list`
3. Doplníme řádek s odkazem na repozitář na konec souboru
`deb http://apt.izzysoft.de/ubuntu generic universe`

Druhý krok: Přidání PGP klíče

4. Vzhledem k tomu, že je na serveru čistá instalace OS, doinstalujeme balíček wget
`apt-get -y install wget`
5. Přejdeme do složky, kde chceme klíč uložit a pomocí wget jej do této složky stáhneme

```
cd /tmp
```

```
wget http://apt.izzysoft.de/izzysoft.asc
```

6. Pomocí apt-key přidáme PGP klíč

```
apt-key add izzysoft.asc
```

7. Provedeme aktualizaci serveru, aby se načetly všechny změny

```
apt-get update
```

Třetí krok: Instalace balíčku Monitorix a konfigurace

8. Stáhneme balíček ze stránek vývojáře

```
wget http://www.monitorix.org/monitorix_3.10.1-izzy1_all.deb
```

9. Nainstalujeme balíček pomocí příkazu

```
apt-get -y install monitorix apache2-utils
```

10. Konfigurace systému Monitorix se provádí v konfiguračním souboru

```
nano /etc/monitorix/monitorix.conf
```

11. V konfiguračním souboru změníme následující údaje:

- V části httpd_builtin změníme port na 8888, aby nedošlo ke kolizím s jinými službami, jedná se o port, na kterém běží Monitorix,
- V části auth povolíme autentizaci, aby se k datům nedostal nikdo cizí

12. Zavřeme editor a restartujeme službu Monitorix

```
service monitorix restart
```

13. Posledním krokem je zadání hesla pro uživatele

```
htpasswd -d -c /var/lib/monitorix/htpasswd admin
```

Po tomto příkazu je nutné dvakrát zadat heslo a tím je vše nastaveno.

14. K Monitorixu je poté možné přistoupit přes webový prohlížeč na adrese

```
http://10.5.11.98:8888/monitorix/
```

4.2 Zálohovací aplikace BareOS

Pro zálohování dat je zvolena aplikace pro zálohování, archivaci, verifikaci a obnovování dat, BareOS (Backup Archiving REcovery Open Sourced), která je zdarma a bez omezení. Tato aplikace nemá problém s různými typy úložišť, jako jsou disky a páskové knihovny a umožňuje šifrování dat a šifrování přenosu dat. Doposud byla k tomuto účelu využívána aplikace BackupPC, která je také opensourcová a umožňuje jednoduché ovládání a konfiguraci skrze webové grafické rozhraní, ale aplikace BareOS má oproti BackupPC mnoho výhod, které rozhodly o použití k zálohování právě aplikací BareOS.

4.2.1 Instalace zálohovací aplikace BareOS [12]

Vzhledem k tomu, že byla zvolená architektura jednoho hlavního serveru (10.5.11.98) a dvou klientů (IP 10.5.11.91, 10.5.11.97), je sice potřeba nainstalovat aplikaci BareOS na všechny servery, ale každý server bude využívat jinou část nainstalovaného balíčku. Veškeré nastavení se pak provádí především na hlavním serveru.

Jednotlivé kroky instalace této aplikace jsou následující:

1. Definice proměnných pro odkaz (výběr verze OS a aplikace)

```
DIST=Debian_9.0
```

```
REL=release/17.2/
```

```
URL=http://download.bareos.org/bareos/$REL/$DIST
```

2. Připsání řádku s odkazem do souboru s odkazy (zdroji) a aktualizace těchto informací

```
printf "deb $URL \n" > /etc/apt/sources.list.d/bareos.list
```

```
apt-get update
```

3. Stáhnutí a přidání podpisového RSA klíče

```
wget -q $URL/Release.key -O- | apt-key add -
```

4. Samotná instalace balíčků potřebných pro správné fungování

```
apt-get install postgresql bareos bareos-database-postgresql bareos-webui
```

5. Nastavení katalogu a vytvoření databáze a databázových tabulek, včetně přidělení přístupových práv pro aplikaci bareos


```
su postgres -c /usr/lib/bareos/scripts/create_bareos_database
```

```
su postgres -c /usr/lib/bareos/scripts/make_bareos_tables
```

```
su postgres -c /usr/lib/bareos/scripts/grant_bareos_privileges
```

6. Vytvoření uživatele pro přístup přes webové rozhraní se provádí pomocí bconsole, což je součást aplikace BareOS, která slouží pro konfiguraci aplikace

```
bconsole
```

```
*configure add console name=admin password=admin profile=webui-admin
```

7. Spuštění služeb potřebných pro chod BareOS (Director, File daemon, Storage daemon)

```
service bareos-dir start
```

```
service bareos-sd start
```

```
service bareos-fd start
```

8. Přístup do webového rozhraní je možný přes webový prohlížeč na adrese:

```
10.5.11.98/bareos-webui
```

4.2.2 Nastavení testovacích záloh hlavního serveru

1. Definice klienta pro hlavní server

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/client

Soubor pro definici hlavního serveru (záloha sebe samého): **bareos-fd.conf**

```
Client {
```

```
Name = bareos-fd
```

```
Description = "Client resource of the Director itself."
```

```
Address = localhost
```

```
Password = "/4l9c2y4u4toT77Llpg7jGBQ7KBDeB2JqN7ixvrs3cOd" #
```

```
password for FileDaemon
```

```
}
```

Soubory ve složce client slouží pro definici klientů, tedy strojů (serverů), ze kterých se budou zálohovat data.

Vysvětlení jednotlivých řádků souboru:

- Client – jedná se o konfiguraci klienta
- Name (povinné) – název klienta
- Adress (povinné) – IP adresa serveru, ze kterého se zálohuje
- Password (povinné) – HASH otisk hesla pro přístup FileDaemona

2. Definice zálohovacích skupin

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/pool

Jedná se o konfigurační soubory skupin, které sdružují zálohy se stejným typem zálohování.

V této složce se nachází konfigurační soubory Differential.conf, Full.conf, Incremental.conf.

Vysvětlení jednotlivých souborů:

- Full – úplná záloha,
- Differetial – rozdílová záloha,
- Incremental – přírůstková záloha.

Pro každou skupinu se v rámci konfiguračního souboru definuje, zda se má jednat o zálohu nebo archiv, zda se mají data recyklovat, tedy přepisovat staré zálohy, jak dlouho se mají zálohy zachovávat (počet dní), maximální velikost záloh a maximální počet záloh.

3. Vytvoření konfiguračního souboru pro plánování akcí

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/schedule

Přidání konfiguračního souboru plánu, který každou půl hodinu provede zálohu

Soubor **HalfHour.conf**

```
Schedule {  
  Name = "HalfHour"  
  Run = Level=Incremental hourly at 0:00  
  Run = Level=Incremental hourly at 0:30  
}
```

Vysvětlení jednotlivých řádků souboru:

- Schedule – jedná se o konfiguraci plánu
- Name (povinné) – název plánu

- Run – definice kdy a jak se má spouštět tento plán
- Level = Incremental – provedení přírůstkové zálohy
- hourly at – každou hodinu v půl a v celou (tedy každé půl hodiny)

4. Definice souborů, které se mají zálohovat

Tento konfigurační soubor slouží k definici toho, co se má zálohovat, co ne a případně další možnosti, které umožňují zpracování souboru před zálohou.

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/fileset

Soubor **TestConfig.conf**

```
FileSet {  
    # name (required)  
    Name = "TestConfig"  
  
    Include {  
        File = /home/TestHlavni.txt  
    }  
}
```

Vysvětlení jednotlivých řádků souboru:

- FileSet – definice zálohovaných souborů/složek
- Name – název zdroje souborů
- Include – soubory/složky zahrnuté do zálohování
- File – cesta k souboru/složce

Volitelnou možností je přidání Exclude, což je definice souborů, které se zálohovat nemají.

Kromě těchto základních povinných položek je možné dále nastavit spoustu dalších možností, například kompresi dat, digitální podpis (MD5, SHA1, SHA256, SHA512) a jiné.

5. Konfigurace zálohovací akce

JobDefs slouží pro definování podrobného nastavení zálohovací akce.

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **Myjob.conf**

```
JobDefs {
  Name = "MyJob"
  Type = Backup
  Level = Incremental
  Client = bareos-fd
  FileSet = "TestConfig"          # selftest fileset          (#13)
  Schedule = "HalfHour"
  Storage = File
  Messages = Standard
  Pool = Incremental
  Priority = 10
  Write Bootstrap = "/var/lib/bareos/%c.bsr"
  Full Backup Pool = Full        # write Full Backups into "Full" Pool    (#05)
  Differential Backup Pool = Differential # write Diff Backups into "Differential"
  Pool (#08)
  Incremental Backup Pool = Incremental # write Incr Backups into "Incremental"
  Pool (#11)
}
```

Vysvětlení jednotlivých řádků souboru:

- JobDefs – jedná se o definici zálohovací akce
- Type – jedná se o zálohu
- Level – typ zálohy – přírůstková
- Client – klient, ze kterého se zálohuje
- Fileset – podle čeho se má zálohovat (který soubor/složka)
- Schedule – použitý plán – jak často se má zálohovat
- Storage – kam se bude záloha ukládat
- Messages – způsob vypisování zpráv o stav zálohy
- Pool – do jaké skupiny se má záloha zařadit
- Write Bootstrap - cesta, kam se má zapsat bootstrap

- Full/Incremental/Differential Backup Pool – který soubor se má použít pro konfiguraci jednotlivých typů záloh

6. Vytvoření zálohovací akce

Konfigurace Job slouží k definici akcí, které pak provádějí automatické zálohování. Jedná se v podstatě o propojení předchozích konfiguračních souborů.

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/job

Soubor **MyBackup.conf**

```
Job {  
  Name = "MyBackup"  
  JobDefs = "MyJob"  
  Client = "bareos-fd"  
}
```

Vysvětlení jednotlivých řádků souboru:

- Job – definice akce
- Name – název zálohovací akce
- JobDefs – konfigurační soubor, který má být použitý pro danou akci
- Client – ze kterého klienta se má zálohovat

7. Udělení přístupových práv pro nově vytvořené soubory pro aplikaci BareOS

```
cd /etc/bareos/bareos-dir.d  
chown -R bareos:bareos *
```

8. Definice host serverů

Aby centrální server věděl, jaké další servery (hosti) jsou k němu připojené a jaké jsou jejich IP adresy, musí se do souboru hosts přidat tyto klienti a jejich IP adresy.

Přidání klientských serverů do **/etc/hosts**

```
127.0.0.1    localhost  
::1         localhost ip6-localhost ip6-loopback  
ff02::1     ip6-allnodes  
ff02::2     ip6-allrouters
```

```
10.5.11.91 host1
10.5.11.97 host2
# --- BEGIN PVE ---
127.0.1.1 sousedikDP-central
# --- END PVE ---
```

9. Restart služeb aplikace BareOS

4.2.3 Nastavení záloh prvního klienta (host1)

Veškeré nastavení v této podkapitole se provádí na hlavním serveru (central). Postup je podobný jako v případě nastavení centrálního serveru, proto zde budou vypsány pouze jednotlivé parametry.

1. Přidání souboru s BareOS klientem

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/client

Přidání souboru **host1.conf**

Nastavení parametrů:

- Name = host1
- Description = "Host 1."
- Address = 10.5.11.91
- Password = "/4l9c2y4u4toT77Llpg7jGBQ7KBDeB2JqN7ixvrs3cOd"

2. Konfigurace zálohovací akce pro hosta 1

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **MyJobH1.conf**

Nastavení parametrů:

- Name = "MyJobH1"
- Type = Backup
- Level = Incremental
- Client = host1
- FileSet = "TestConfig"

- Schedule = "HalfHour"
- Storage = File
- Messages = Standard
- Pool = Incremental
- Priority = 10

3. Vytvoření zálohovací akce pro hosta 1

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/job

Soubor **MyBackupH1.conf**

Nastavení parametrů:

- Name = "MyBackupH1"
- JobDefs = "MyJobH1"
- Client = "host1"

4.2.4 Nastavení záloh druhého klienta (host2)

Veškerá konfigurace se opět provádí na hlavním serveru.

1. Přidání konfiguračního souboru klienta

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/client

Přidání souboru **host2.conf**

Nastavení parametrů:

- Name = host2
- Description = "Host 2."
- Address = 10.5.11.97
- Password = "/4l9c2y4u4toT77Llpg7jGBQ7KBDDeB2JqN7ixvrs3cOd"

4. Konfigurace zálohovací akce pro hosta 2

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **MyjobH2.conf**

Nastavení parametrů:

- Name = "MyJobH2"
- Type = Backup
- Level = Incremental
- Client = host2
- FileSet = "TestConfig"
- Schedule = "HalfHour"
- Storage = File
- Messages = Standard
- Pool = Incremental
- Priority = 10

5. Vytvoření zálohovací akce pro hosta 1

Konfigurace souborů ve složce: /etc/bareos/bareos-dir.d/job

Soubor **MyBackupH2.conf**

Nastavení parametrů:

- Name = "MyBackupH2"
- JobDefs = "MyJobH2"
- Client = "host2"

Pro testování funkčnosti byly použity pouze malé soubory, aby bylo možné rychle a jednoduše otestovat, zda daná konfigurace funguje správně a záloha proběhne úspěšně. Zálohy fungují, spouštějí se samy v daný čas, v případě, že už je vytvořena úplná záloha, pak se provede pouze přírůstková, a pokud není soubor změněný, pak má záloha nulovou velikost a tím se šetří místo na disku. Záloha (byť s nulovou velikostí) je pak vidět v provedených akcích, aby bylo jasné vidět, jestli zálohování funguje jak má.

5 NASTAVENÍ A OTESTOVÁNÍ ODEZVY A VÝKONNOSTI

Pro reálné otestování odezvy a rychlosti je potřeba vytvořit větší soubory, protože rychlost přenosu je velká a malé soubory se zazálohují velmi rychle. To má za následek to, že záloha souboru začne i skončí ve stejný čas a z grafů se nedá vyčíst dostatek informací. Z toho důvodu stáhneme soubor o velikosti 2,8 GB a z něj vytvoříme soubor o velikosti přibližně 8,8 GB (8810201088 B).

Tyto dvě operace se provádí následovně:

- `wget http://cdimage.kali.org/kali-2018.2/kali-linux-2018.2-amd64.iso`
- `cat kali-linux-2018.2-amd64.iso > test.iso && cat kali-linux-2018.2-amd64.iso >> test.iso && cat kali-linux-2018.2-amd64.iso >> test.iso`

Aby nedošlo k narušení funkčnosti kvůli přepisování souborů, je vhodné služby aplikace BareOS zastavit.

5.1 Nastavení pro hlavní server

1. Vytvoření nového zálohovacího plánu

Pro předchozí testování byl zálohovaný soubor malý a bylo možné jej zálohovat často. Vzhledem k tomu, že pro reálné otestování zálohování se bude zálohovat soubor o velikosti 8,8 GB, nastavíme prodloužený interval záloh, aby nedošlo k rychlému zaplnění disku na serveru.

Konfigurace souboru ve složce: `/etc/bareos/bareos-dir.d/schedule`

Přidání plánu, díky němuž se každý den o půlnoci provede záloha.

Soubor **Daily.conf**

Nastavení parametrů:

- `Name = "Daily"`
- `Run = Level=Incremental daily at 0:00`

2. Nastavení nové cesty k zálohovanému souboru

Konfigurace souboru ve složce: `/etc/bareos/bareos-dir.d/filesset`

Soubor **TestConfig.conf**

Nastavení parametrů:

- Name = "TestConfig"
- File = /home/test.iso

3. Konfigurace definice akce

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **Myjob.conf**

Vzhledem k tomu, že se změnila pouze cesta k souboru, není potřeba měnit název souboru s cestou k zálohovanému souboru.

Změní se pouze plán:

- Schedule = "Daily"

4. Nastavení zálohovací akce

Soubor **MyBackup.conf** zůstane beze změn, protože parametry není potřeba měnit.

5.2 Nastavení pro klienta host 1

1. Nastavení nové cesty k zálohovanému souboru

Vzhledem k tomu, že na každém serveru se jmenuje testovací soubor jinak, je potřeba vytvořit nový soubor pro definici cesty k souboru.

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/fileset

Soubor **TestConfigH1.conf**

Nastavení parametrů:

- Name = "TestConfigH1"
- File = /home/test1.iso

2. Konfigurace definice akce

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **MyjobH1.conf**

Změna nastavení parametrů:

- FileSet = "TestConfigH1"
- Schedule = "Daily"

5.3 Nastavení pro klienta host 2

1. Nastavení nové cesty k zálohovanému souboru

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/fileset

Soubor **TestConfigH2.conf**

Nastavení parametrů:

- Name = "TestConfigH2"
- File = /home/test2.iso

2. Konfigurace definice akce

Konfigurace souboru ve složce: /etc/bareos/bareos-dir.d/jobdefs

Soubor **MyjobH2.conf**

Změna nastavení parametrů:

- FileSet = "TestConfigH2"
- Schedule = "Daily"

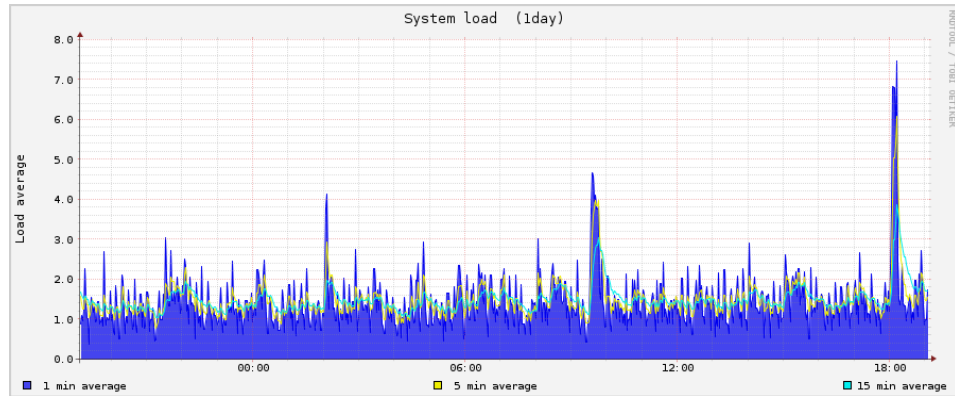
Tímto jsou nastavená veškerá potřebná nastavení pro testování výkonnosti a odezvy řešení a můžeme tedy přistoupit k samotnému testování.

5.4 Otestování výkonnosti a odezvy

Prvním krokem je otestování zálohy hlavního serveru, při kterém je sledováno zatížení systému, zatížení procesoru a jednotlivých jader a alokace vyrovnávací paměti.

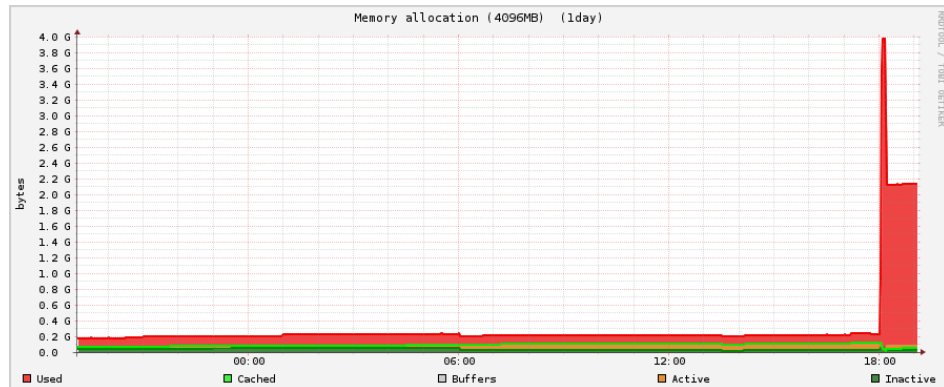
5.4.1 Záloha hlavního serveru

Plná záloha testovacího souboru na hlavním serveru trvala v průměru 330 sekund. To znamená, že záloha proběhla s rychlostí zápisu 26,7 MB/s.



Obrázek 7: Záloha hlavního serveru - zatížení systému [Zdroj: vlastní]

Při zálohování dat na serveru samozřejmě zatížení systému vzrostlo, ale zdaleka nedosáhlo 100%, špičková hodnota dosáhla maximálně 75%.



Obrázek 8: Záloha hlavního serveru – využití paměti [Zdroj: vlastní]

Co se týče vytížení vyrovnávací paměti, chvilkově sice došlo k plnému zatížení, ale to nemělo vliv na průběh zálohy, server běžel v pořádku a záloha proběhla úspěšně.

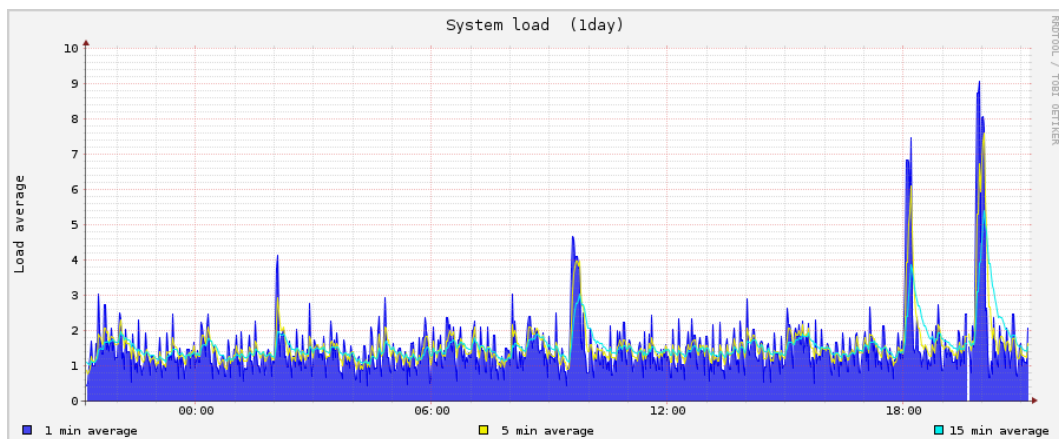


Obrázek 9: Záloha hlavního serveru – využití procesoru [Zdroj: vlastní]

Při záloze hlavního serveru neměl procesor žádný problém a maximálně byl vytížen na přibližně 66%.

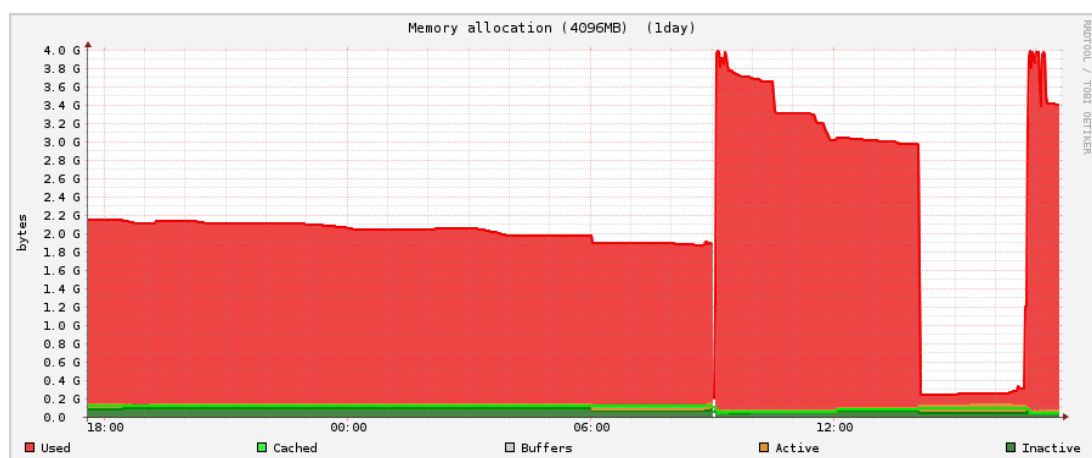
5.4.2 Záloha klientů

Pro otestování zálohy klientů byly použité soubory o stejné velikosti. Oproti záloze hlavního serveru se v tomto případě zaměříme i na využití gigabitového ethernetového portu, aby bylo jasně vidět, jak tento proces využívá síť.



Obrázek 10: Záloha klientů – zatížení systému [Zdroj: vlastní]

Oproti záloze hlavního serveru, kdy zálohování probíhá pouze v rámci samotného serveru, je zatížení systému vyšší, protože systém musí provádět navíc operace při komunikaci s klienty.



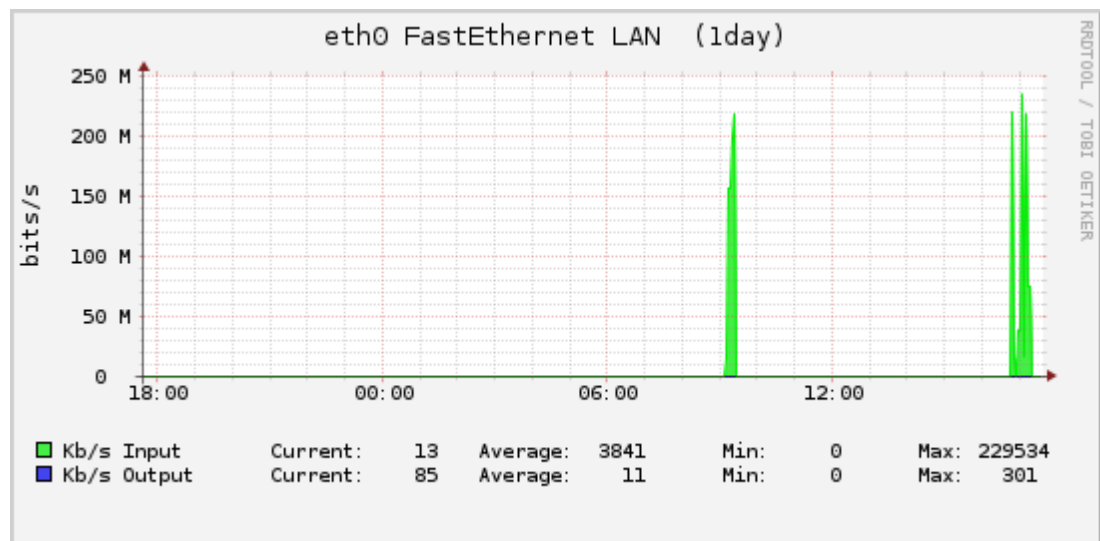
Obrázek 11: Záloha klientů – využití paměti [Zdroj: vlastní]

Z hlediska využití vyrovnávací paměti, ta byla při této záloze vystavena déle maximálním hodnotám kvůli komunikaci s klienty, a taky proto, že objem přenášených dat byl v podstatě dvojnásobný.



Obrázek 12: Záloha klientů – využití procesoru [Zdroj: vlastní]

Z hlediska využití procesoru není téměř žádný rozdíl oproti záloze hlavního serveru. Maximální využití v průběhu zálohy klientů bylo 70%, což je hodnota velmi blízká hodnotě využití procesoru u lokální zálohy hlavního serveru.



Obrázek 13: Záloha klientů – využití ethernetového portu [Zdroj: vlastní]

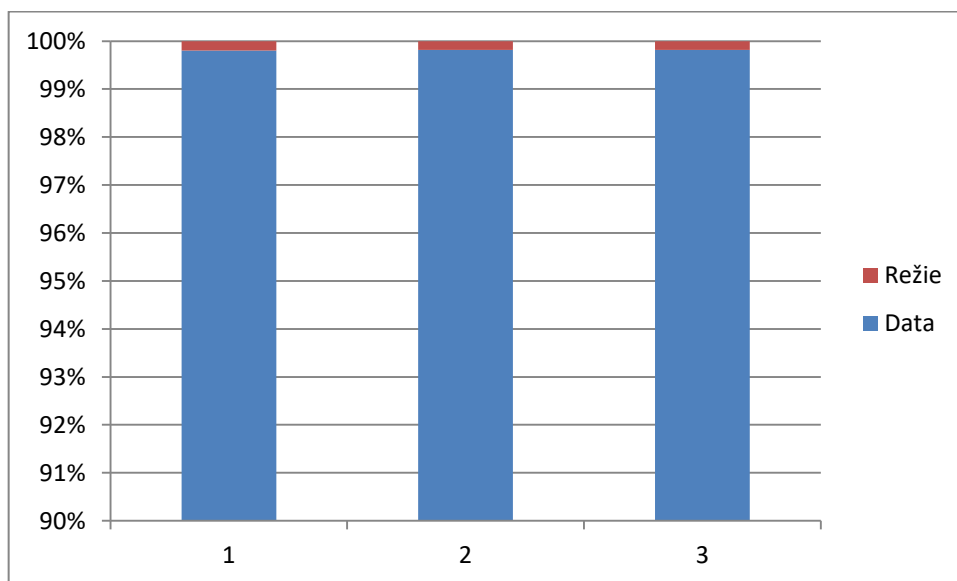
Při zálohování klientů je důležitý také přenos po síti. Špičková hodnota přenosové rychlosti byla 301 Mbit/s, což znamená, že ethernetový port nebyl vytížen na maximum.

6 STANOVENÍ LIMITŮ POČTU PŘIPOJENÝCH ZAŘÍZENÍ

Tato kapitola je věnovaná stanovení počtu zařízení (klientských serverů), která je možno připojit k datovému úložišti. Omezení může být dané několika aspekty. Maximální počet klientů ale udává především komunikační protokol.

Pro komunikační protokol NFSv4 není obecně stanovený maximální počet zařízení, která lze připojit k jednomu úložišti. Pokud ale využijeme oficiální stanovení limitů pro Linuxový systém Red Hat Enterprise Linux verze 6.3, ten pak stanovuje maximální počet 128 připojených zařízení.

Jedním z dalších omezení může být využití maximální kapacity ethernetového portu. Při přenosu dat od klientů do centrálního serveru totiž dochází mimo přenos samotných zálohovaných dat také k přenosu dat, která slouží pro komunikaci mezi centrálním serverem a klientským serverem. V našem případě jsou servery propojené pomocí gigabitového ethernetu (rychlost 1000 Mb/s), který by měl sloužit primárně pro přenos zálohovaných dat. Problém nastává totiž tehdy, kdy se takzvaná režijní data začnou svou velikostí přibližovat velikosti přenášených souborů. Režijní data totiž slouží pro komunikaci mezi klientským a centrálním serverem a při zvyšování počtu připojených zařízení roste samozřejmě také objem těchto dat.



Obrázek 14: Graf poměru přenesených dat a režijních dat [Zdroj: vlastní]

Na základě změřených dat byl vytvořen graf poměru přenosu právě těchto režijních dat a dat přenesených pro zálohu z klientů na centrální server. Při komunikaci s jedním klientem byl objem přenesených režijních dat přibližně 0,2% celkového objemu přenesených dat. Při dvou klientech byl tento objem dvojnásobný. Lze tedy říci, že zvyšování objemu režijních dat má lineární charakter a na základě toho byl určen také poměr přenesených dat v případě tří připojených klientů. Pokud si určíme hranici, kdy objem režijních dat může být maximálně 25% objemu celkově přenesených dat, pak maximální počet takto připojených klientů je 125. Tímto je tedy určen limit maximálního počtu klientů, protože je nižší než umožňuje samotný komunikační protokol.

ZÁVĚR

Cílem diplomové práce bylo navrhnout, realizovat a otestovat zálohování interních serverů na Fakultě aplikované informatiky, které slouží pro běh webových aplikací a stránek, které se nacházejí na subdoménách UTB.

Vzhledem k tomu, že se jedná o služby, se kterými pracují zaměstnanci a studenti a jsou na těchto serverech uložena data jejich, ale i samotné Univerzity Tomáše Bati, je potřeba tyto služby zálohovat.

První část se skládá z návrhu možností připojení k datovému úložišti. Pro komunikaci musí být využitý takový protokol, u kterého je zabezpečena bezpečnost přenosu pomocí šifrování, ale také další důležité parametry, například co nejvyšší přenosová rychlost nebo také zachování údajů o vlastníkovi souborů, protože bez tohoto údaje by mohly nastat problémy při obnově dat. Jako nejvíce vhodný protokol byl vybrán protokol NFSv4, který splňuje požadavky na šifrovanou komunikaci a dokonce dokáže přenášet data s velkou rychlostí.

I když se povede vybrat nejvhodnější protokol, neznamená to, že nemá žádné omezení. Omezení však neurčuje pouze protokol, ale také provozovatel datového úložiště, protože bez limitů by si mohl na úložiště každý ukládat jakýkoliv objem dat a tento uchovávat po jakoukoliv dobu. Z tohoto důvodu jsou na datovém úložišti limity na objem souborů, počet souborů a dobu uchování záloh, které jsou pro některé aplikace mnohdy nedostačující. Do návrhu patří také propojení serverů v rámci vnitřní infrastruktury, které je rozhodující pro správnou funkčnost, ale také pro jednodušší správu. Byla zvolena varianta jednoho hlavního serveru a dvou klientů, protože díky této architektuře je možné zálohy spravovat z jednoho místa, tedy z hlavního serveru.

Praktická část je realizací zvoleného návrhu řešení a toto řešení je zde také otestováno. Zálohy se podařilo nastavit a funguje i automatické zálohování hlavního serveru a klientů podle stanovených plánů. Testovací infrastruktura se osvědčila a byla spolehlivá.

Posledním bodem bylo stanovení limitů počtu připojených zařízení. V rámci testovací infrastruktury byly k dispozici dva servery a na základě měření a testování bylo možné určit, jaký maximální počet zařízení je možné připojit, aby nedošlo k zahlcení ethernetových spojů režijními daty, která slouží pro výměnu informací mezi hlavním serverem a klientem. Jde tedy o to, že režijní data musí mít mnohem menší objem než data zálohovaná.

Práce bude mít praktické využití pro zálohování některých lokálních serverů na FAI.

SEZNAM POUŽITÉ LITERATURY

- [1] *Topologie sítě CESNET2* [online]. [cit. 2018-05-21]. Dostupné z: <https://www.cesnet.cz/sluzby/pripojeni/topologie/>
- [2] *Mapa sítě CESNET2* [online]. [cit. 2018-05-21]. Dostupné z: <https://www.cesnet.cz/wp-content/uploads/2012/10/cesnet2-topo1.gif>
- [3] *Topologie úložišť* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/_media/cs/mapa-datacenter.png?cache=&w=508&h=311&tok=26252a
- [4] *Infrastruktura úložišť* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/cs/infrastruktura_ulozist/start
- [5] *Architektura DÚ v Brně* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/_media/cs/hsm_brno.png?cache=&w=601&h=278&tok=b91271
- [6] *RFC 3530* [online]. [cit. 2018-05-21]. Dostupné z: <https://tools.ietf.org/html/rfc3530>
- [7] *Přehled protokolů* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/cs/prehled_protokolu_a_sluzeb_s_jejich_doporucenimi/
- [8] *NFSv4* [online]. [cit. 2018-05-21]. Dostupné z: <https://www.gsp.com/cgi-bin/man.cgi?section=4&topic=nfsv4>
- [9] *Provozní pravidla* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/cs/provozni_pravidla/start
- [10] *Zacházení s daty* [online]. [cit. 2018-05-21]. Dostupné z: https://du.cesnet.cz/cs/navody/faq/csdata_management/start
- [11] *Performance monitoring with Monitorix* [online]. [cit. 2018-05-21]. Dostupné z: <https://www.howtforge.com/tutorial/performance-monitoring-with-monitorix-on-ubuntu-16-04/>
- [12] *BareOS ® manual* [online]. [cit. 2018-05-21]. Dostupné z: <http://doc.bareos.org/master/html/bareos-manual-main-reference.html>

SEZNAM OBRÁZKŮ

Obrázek 1: Topologie sítě CESNET2 [2]	10
Obrázek 2: Mapa datových úložišť v síti CESNET2 [3]	11
Obrázek 3: Hierarchický model datového úložiště v Brně [5]	12
Obrázek 4: Možnost propojení č. 1 [Zdroj: vlastní]	18
Obrázek 5: Možnost propojení č. 2 [Zdroj: vlastní]	19
Obrázek 6: Úvodní obrazovka systému Monitorix [Zdroj: vlastní]	22
Obrázek 7: Záloha hlavního serveru - zatížení systému [Zdroj: vlastní].....	36
Obrázek 8: Záloha hlavního serveru – využití paměti [Zdroj: vlastní].....	36
Obrázek 9: Záloha hlavního serveru – využití procesoru [Zdroj: vlastní].....	37
Obrázek 10: Záloha klientů – zatížení systému [Zdroj: vlastní].....	38
Obrázek 11: Záloha klientů – využití paměti [Zdroj: vlastní]	38
Obrázek 12: Záloha klientů – využití procesoru [Zdroj: vlastní]	39
Obrázek 13: Záloha klientů – využití ethernetového portu [Zdroj: vlastní]	40
Obrázek 14: Graf poměru přenesených dat a režijních dat [Zdroj: vlastní]	41

SEZNAM TABULEK

Tabulka 1: Tabulka protokolů [Zdroj: vlastní]	15
---	----