VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

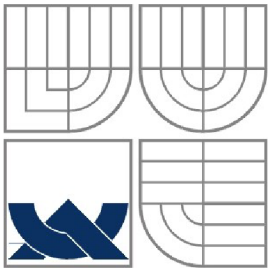
WEBOVÝ PORTÁL S REPORTY O SÍŤOVÉM PROVOZU

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

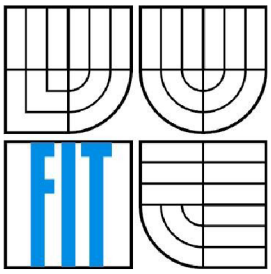
AUTOR PRÁCE
AUTHOR

PETR ZAPLETAL

BRNO 2009



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

WEBOVÝ PORTÁL S REPORTY O SÍŤOVÉM PROVOZU

WEB PORTAL FOR NETWORK TRAFFIC REPORTING

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

PETR ZAPLETAL

VEDOUCÍ PRÁCE
SUPERVISOR

ING. JIŘÍ TOBOLA

BRNO 2009

Zadání bakalářské práce/8098/2008/xzaple01

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav počítačových systémů Akademický rok

Zadání bakalářské práce

Řešitel: **Zapletal Petr**
Obor: Informační technologie
Téma: **Webový portál s reporty o síťovém provozu**
Kategorie: Web

Pokyny:

1. Seznamte se s technologiemi pro tvorbu webových informačních systémů (HTML, CSS, PHP, Javascript, MySQL apod.).
2. Stručně se seznamte s technologií NetFlow pro monitorování sítí.
3. Proveďte analýzu požadavků pro systém umožňující tvorbu reportů, grafů a tabulek na základě NetFlow dat. Systém musí poskytovat podporu široké škály statistik (top uživatelé, nejnavštěvovanější servery, doby činnosti na síti, souhrnné statistiky sítě atp.).
4. Vytvořte detailní návrh tohoto systému a vhodně jej modelujte.
5. Navržený systém realizujte a otestujte, funkčnost systému demonstrujte na vhodně zvoleném vzorku dat.
6. Zhodnoťte dosažené výsledky a diskutujte možnosti dalšího rozšíření systému.

Literatura:

- Dle pokynů vedoucího.

Při obhajobě semestrální části projektu je požadováno:

- Splnění prvních tří bodů zadání.

Podrobné závazné pokyny pro vypracování bakalářské práce naleznete na adrese http://www.fit.vutbr.cz/info/szz/

Technická zpráva bakalářské práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap (20 až 30% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Tobola Jiří, Ing.**, UPSY FIT VUT
Datum zadání: 1. listopadu 2008
Datum odevzdání: 20. května 2009

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta informačních technologií
Ústav počítačových systémů a sítí
612 66 Brno, Božetěchova 2

L.S.

doc. Ing. Zdeněk Kotásek, CSc.
vedoucí ústavu

Abstrakt

Tato práce popisuje návrh a implementaci webového portálu s reporty o síťovém provozu. Tento systém je určen pro snadné monitorování sítě pomocí přehledných grafů a tabulek. Pro získávání statistik o provozu na síti je použita technologie NetFlow, která je klíčovou součástí sytému. Portál je nezávislý na platformě a je postaven na technologiích HTML a PHP. Portál umožňuje export zpracovaných dat prostřednictvím pravidelného zasílání emailů nebo pomocí souborů ve formátu PDF.

Abstract

This thesis deals with the design and implementation of web portal with reports about network traffic. This system is designed for easy network monitoring with transparent charts a tables. To obtain statistics from network traffic, NetFlow technology, which is the key part of system, is used. Portal is platform independent and is built around HTML and PHP technologies. Portal also allows export proccessed data through periodic email service or files in PDF format.

Klíčová slova

Webový portál, reporty, síť, monitorování sítě, síťový provoz, NetFlow, nfdump, HTML, PHP

Keywords

Web portal, reports, network, network monitoring, network traffic, NetFlow, nfdump, HTML, PHP

Citace

Petr Zapletal: Webový portál s reporty o síťovém provozu, bakalářská práce, Brno, FIT VUT v Brně, 2009

Webový portál s reporty o síťovém provozu

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Jiřího Toboly. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Petr Zapletal
5. května 2009

Poděkování

Velmi rád bych poděkoval svému vedoucímu Ing. Jiřímu Tobolovi za poskytnutou pomoc, odborné vedení a čas strávený při konzultacích při tvorbě této práce.

Obsah

1 Úvod

V současné době dochází ke stále většímu rozšíření internetu a síťových technologií jak do domácností, tak i do firem. Současně s rozšiřováním sítí dochází ke zvyšování rychlostí datových přenosů po těchto sítích. Ruku v ruce s tímto rozvojem se zvyšují i možnosti zneužití sítě. Může se jednat o zneužití sítě vlastními zaměstnanci, kteří navštěvují zakázané stránky nebo používají nepovolené služby. Může se jednat o zákazníky poskytovatelů internetového připojení, kteří se snaží stahovat z internetu nelegální obsah, případně se může jednat o útočníky, kteří se snaží o proniknutí nebo o vyřazení cizí sítě.

Typickým řešením těchto problémů jsou různé blokace a omezení přístupu, ať už uživatelů nebo služeb. Abychom mohli tato opatření provádět, je třeba o nich vědět. Potřebné informace můžeme v obecné podobě získat z mnoha zdrojů zabývajících se zabezpečením sítí. Avšak specifická data pro konkrétní síť je třeba získat jejím pravidelným monitorováním. Monitorování a následná pečlivá analýza provozu umožňuje nejen pružně reagovat na pokusy o zneužití sítě, ale také usnadňuje plánování dalšího rozvoje sítě a odstraňování slabých míst její infrastruktury.

Dlouhá léta byl synonymem pro monitorování a dohled nad počítačovou sítí protokol SNMP (Simple Network Management Protocol) [1]. SNMP je jednoduchý a široce rozšířený protokol, pracující na principu dotaz-odpověď a umožňující získání odpovědi na souhrnné dotazy o provozu na síti. Protokol SNMP má několik nevýhod daných jeho jednodušší implementací. Protokol má problémy s rychlostí práce v rozsáhlých sítích. Dále má jisté bezpečnostní slabiny a nemá možnost získat konkrétní informace o dění v síti, jako jsou dotazy typu: „Která stanice síť nejvíce zatěžuje?“.

Současná doba, kdy na dostupnosti a funkcionalitě závisí velké množství společností, si však žádá modernější a efektivnější metody. Ty musí být schopny v reálném čase detailně informovat o dění na síti. Právě tyto schopnosti nabízí technologie NetFlow.

V současnosti je NetFlow nejrozšířenější průmyslový standard pro měření a monitorování sítí na základě IP toků. Tok je v terminologii NetFlow definován jako sekvence paketů mající shodné klíčové vlastnosti, jako jsou IP adresy, čísla portů a typ protokolu. Pro každý tok je zaznamenávána doba jeho trvání, čas vzniku, počet přenesených dat a další údaje. Tyto podrobné záznamy jsou použity k vytvoření vysoce přesných a detailních statistik o dění v síti.

Cílem této práce je navrhnout a implementovat portál, který pomocí technologie NetFlow umožní tvorbu různých statistik o dění v síti a jejich vizualizaci. Portál dále nabídne uživatelům možnosti pro export těchto dat.

Práce je členěna do několika kapitol. Klíčovou částí je technologie NetFlow, jejíž teoretický popis se nachází ve druhé kapitole. V následující, tedy třetí kapitole, naleznete vše potřebné

o specifikaci požadavků na výsledný portál. Čtvrtá kapitola se zabývá návrhem výsledné aplikace, který je doplněn diagramem případů užití. Nejobsáhlejší pátá kapitola se zabývá implementací systému a dozvíte se, jaké technologie a knihovny byly v aplikaci použity. Budou Vám objasněny principy chodu portálu, popsány jeho nejdůležitější funkce a klíčové skripty. Další, tedy šestá kapitola, se zabývá instalací a požadavky aplikace jak na straně serveru, tak na straně klienta. Najdete v ní i zkušenosti z testování portálu. Poslední kapitola shrnuje výsledky dosažené v této práci a jejich zhodnocení z hlediska dalšího vývoje.

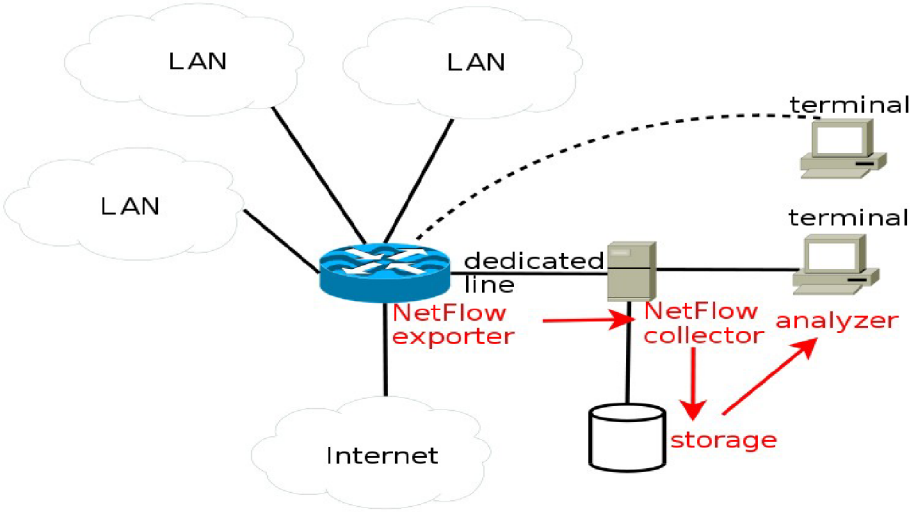
2 Technologie NetFlow

Tato práce se zabývá tvorbou portálu pro zobrazování dat o síťovém provozu. První krokem k zobrazování dat je samozřejmě jejich sběr. K tomuto bude aplikace využívat široce rozšířenou technologii NetFlow, která byla vyvinuta společností Cisco Systems, Inc..

NetFlow je otevřený, ale patentovaný síťový protokol, vyvinutý pro běh na zařízeních podporující Cisco IOS (Internetwork Operating Systems). Jedná se o operační systém používaný na směrovačích a přepínačích od společnosti Cisco Systems, Inc.. Tento protokol představuje velice efektivní cestu pro získání komplexního přehledu o dění na IP síti.

2.1 Popis protokolu

Pokud je na zařízení povolen NetFlow, zařízení začne generovat NetFlow záznamy (NetFlow records), které popisují veškerý síťový provoz. Data pro tvorbu záznamů se získávají analýzou paketů procházejících zařízením. Tyto záznamy jsou pak v pravidelných intervalech posílány na kolektor (Neflow collector), který se stará o jejich archivaci. Zařízení, odesílající NetFlow záznamy, bývá označováno jako exporter (NetFlow exporter). Na jeden kolektor připadá zpravidla více exporterů. Pro přenos dat ze zdrojového zařízení se požívá protokol UDP (User Datagram Protocol), případně protokol SCTP (Stream Control Transmission Protocol), pokud je vyžadována vyšší spolehlivost a bezztrátovost přenosu. Schéma protokolu NetFlow ukazuje ilustrace 1.



Ilustrace 1: Schéma NetFlow

Protokol NetFlow má několik verzí, z nichž alespoň některé je potřeba zmínit. Verze 1 je původní verze, která se ale příliš neujala. Masového využití v praxi dosáhla až verze 5, která mimo mnoha vylepšení přidala podporu BGP (Border Gateway Protocol), a tudíž byla použitelná i k monitorování komunikace mezi autonomními systémy. Verze 5 je stále nejpoužívanější verzí. Verze 7 byla navržena exkluzivně pro přepínače. Verze 8 přidala bohaté možnosti agregace dat podle různých kriterií, což umožnilo snížit nároky na přenosovou linku mezi exportery a kolektorem, a dále zvýšila rychlost zpracování dat. V současnosti se začíná silně prosazovat verze 9, která je založená na šablonách (templates), umožňujících budoucí rozšíření beze změn formátu záznamů. O kvalitě protokolu svědčí i fakt, že nový IETF standard – Internet Protocol Flow Information eXport (IPFX) je v podstatě totožný s plánovanou verzí 10. Očekává se, že IETF se stane průmyslovým standardem a že bude hojně podporován výrobci sítových zařízení.

2.2 NetFlow záznamy

Jak již bylo dříve nastíněno, datům, které vytváří exporter, se říká záznamy (NetFlow records). Jeden záznam může obsahovat široké množství informací pro daný tok (Flow). Tok je v terminologii NetFlow definován jako sekvence paketů se shodnou pětící údajů – zdrojová a cílová adresa, zdrojový a cílový port a číslo protokolu. Pro každý tok je zaznamenávána doba jeho vzniku, délka jeho trvání, počet přenesených bajtů a paketů a případně další údaje.

Nejpoužívanější verze 5 obsahuje tyto informace:

- číslo verze
- sekvenční číslo
- Simple Network Management Protocol (SNMP) indikátory
- časové označení začátku a konce toku
- počet paketů a bytů v toku
- hlavička síťové vrstvy (zdrojová a cílová adresa, zdrojový a cílový port, protokol, hodnota Type of Service (ToS))

2.3 Nástroje pro práci s NetFlow

Pro práci s NetFlow existuje celá řada nástrojů. V této práci se setkáme s nástroji nfcapd a především s nástrojem nfdump, jehož použití v této práci bude podrobněji popsáno v kapitole 5.5. Mezi hlavní nástroje pro práci s NetFlow patří:

- **nfcapd** – démon, který čte data ze sítě a ta následně archivuje do souborů. Výstupní soubor je pravidelně, zpravidla po pěti minutách, měněn. Součástí jména souboru je čas a datum ve tvaru RRRRMMDDHHMM. Například soubor nfcapd.2000903010000 obsahuje data o síťovém provozu od půlnoci 1. března 2009 do času určeném hodnotou střídání souboru, tj. v našem případě do pěti minut po půlnoci. Více informací o programu nfcapd je k nalezení v [2].
- **nfdump** – program, který zpracovává data uložená nfcapdem. Umožňuje záznamy filtrovat a vytvářet velké množství statistik. Více informací o programu nfdump se nachází v kapitole 2.5, v kapitole 5.5 a v [3].
- **nfprofile** – program podobný nfdumpu. Čte data ze souborů uložených nfcapdem, a tato data dále zpracovává podle uložených profilů. Je používán především programem nfsen. Více na [4].
- **nfsen** – jedná se o webovou grafickou nástavbu programu nfdump. Nfsen nabízí bohaté možnosti přehledného zobrazení NetFlow dat. Dále umožňuje vytváření událostí reagujících na předem definované síťové okolnosti (Alerts) a vytváření historie NetFlow dat. Více informací se nalézá v[5].

- **nfreplay** – čte data uložená programem nfcapd a ta pak posílá po síti na další stanici, případně skupinu stanic. Umožňuje filtrování odesílaných dat obdobnými filtry jako nfdump. Více informací o programu nfreplay se nachází na [6].
- **nfclean.pl** – jednoduchý skript pro mazaní starých dat.
- **ft2nfdump** – pomocný konvertor vstupních dat pro nfdump.
- **nfexpire** – zajišťuje vypršení platnosti NetFlow dat podle zadaných parametrů

2.4 Použití NetFlow

NetFlow je v současnosti nejrozšířenějším průmyslovým nástrojem pro monitorování a měření počítačových sítí na základě IP toků. Použití NetFlow se dá shrnout do následujících bodů:

- **monitorování sítě** (Network Monitoring) – umožňuje monitorování sítě téměř v reálném čase. Techniky jsou založené na analýze dat z NetFlow exporterů. Používají se k přehlednému zobrazování datových toků, které procházejí jednotlivými směrovači. Toky poskytují aktivní detekci problémů na síti a pomáhají s jejich odstraňováním.
- **monitorování aplikací** (Application Monitoring and Progiling) – poskytuje detailní statistiku používání aplikací v časových úsecích. Toho se využívá k plánování a navržení správné topologie sítě.
- **monitorování a analýza uživatelů** (User Monitoring and Profiling) – detailní statistika aktivit jednotlivých uživatelů na síti. Používá se k efektivnímu plánování rozložení zatížení, umístění serverů a podobně. Často se používá k detekci a řešení potencionálních bezpečnostních problémů.
- **podpora účtování a plateb** (Accounting and Billing) – součástí informací o datovém toku jsou i záznamy zdrojové a cílové IP adresy, počtu přenesených paketů a bytů, doby spojení atd.. Tyto informace je možno využít pro podrobné účtování mezi jednotlivými poskytovateli připojení. Poskytovatelé tyto statistiky používají k proplácení svých služeb, zpravidla na základě objemu přenesených dat.
- **plánování a analýza sítě** (Network Planning and Analysis) – data z NetFlow exportérů se dají použít pro optimalizaci strategického plánování sítě. Cílem je minimalizace celkové ceny síťových operací při maximalizaci výkonu, kapacity a dostupnosti sítě.
- **ukládání dat o provozu v síti** (NetFlow Data Warehousing) – data z exportérů mohou být uložena k pozdější analýze. Z této analýzy se pak dá zrekonstruovat veškerý síťový provoz.

Tyto služby jsou často využívány pro generování statistik a grafů o vytíženosti jednotlivých linek. Umožňují zjistit, které služby používají uživatelé vnitřních sítí a které uživatelé z vnějšího světa. Zvláště cenné jsou tyto informace pro poskytovatele připojení. Poskytovatelé připojení jsou často povinni ze zákona tyto informace archivovat. V České republice to upravuje vyhláška 485/2005 Sb. Podle této vyhlášky jsou poskytovatelé povinni archivovat klíčové informace o spojení po dobu šesti měsíců. Více informací naleznete v [7].

2.5 Využití NetFlow v této práci

V této práci je technologie NetFlow využita ve dvou oblastech. Nejprve k zaznamenávání provozu sledované sítě. To spočívá v nastavení a běhu programu nfcapd. Touto činností se zde ale dále zabývat nebudeme. Pokud Vás tato problematika zajímá, obraťte se na [2]. Jakmile jsou data k dispozici, je třeba je zpracovat. K tomu účelu bude aplikace používat program nfdump. Pomocí něho bude aplikace získávat data potřebná k dalšímu zpracování. Podrobným popisem použití programu nfdump v této aplikaci se zabývá kapitola 5.5.

Nfdump umožňuje tvorbu mnoha různých statistik a má relativně obsáhlou syntaxi, podobnou programu tcpdump. Další informace jsou k dispozici v [3].

3 Specifikace požadavků

Na začátku každého většího projektu je nezbytné si se zadavatelem práce ujasnit požadavky na výsledný produkt. Cílem této práce je vytvoření webového portálu, který umožní uživatelům (zpravidla správcům sítě) náhled na různé top statistiky o síti. Systém dále bude mít možnost jednoduchého exportu těchto statistik.

Jádrem systému má být technologie NetFlow, která umožní monitorování sítí a dodá data, která pak budou zpracována a zobrazena na webovém portálu. Tyto reporty se mají zabývat top statistikami síťového provozu a budou zobrazeny ve formě přehledných tabulek a grafů.

Pro tvorbu webového portálu bude použito standardního jazyka HTML. Pro efektivní vytvoření přívětivého uživatelského rozhraní bude použita technika kaskádových stylů, známých jako CSS. Pro dynamickou generaci obsahu a zpracovávání dat bude použit skriptovací jazyk PHP. Tvorba grafů bude implementována pomocí technologie Flash, která umožňuje tvorbu moderních, případně animovaných grafů. Podle potřeby je možné použít JavaScript. Dále systém bude umožňovat správu více uživatelů a jejich autentizaci.

Aplikace bude schopna sama komunikovat s nfdumpem pomocí příkazové řádky a jeho výstupy dále zpracovávat. Získávání vstupních dat pro nfdump se bude řešit externě.

Hlavní komunikační kanál s uživatelem bude pomocí webového rozhraní. Rozhraní bude navrženo funkčně a přehledně a bude optimalizováno pro nejrozšířenější internetové prohlížeče a rozlišení 1024x768. Implementace portálu bude striktně dodržovat oddělení textů a grafiky od zdrojového kódu. Díky tomu bude umožněno snadno manipulovat s texty, respektive snadno měnit grafický styl.

Systém bude podporovat dva typy uživatelů – administrátora a běžného uživatele. Administrátor bude mít na rozdíl od uživatele možnost vytvářet a rušit uživatele a bude moci upravovat důležitá nastavení aplikace. Oba si budou moci editovat svůj profil a plně využívat služeb aplikace.

Export získaných reportů z aplikace bude umožněn přímo ve webovém rozhraní ve formátu PDF nebo pomocí vyžádané emailové služby.

4 Návrh

Po specifikaci požadavků na výslednou aplikaci je třeba tyto řádně analyzovat a vytvořit model systému. Vytvoření kvalitního návrhu je základem vývoje každého většího projektu.

4.1 Analýza požadavků

4.1.1 Stručný popis návrhu

Činnost výsledné aplikace se dá rozdělit na dvě základní oblasti. První musí být schopna používat nfdump a zpracovávat jeho výsledky. Druhá musí být schopna tyto výsledky vhodně prezentovat pomocí webového rozhraní. Jako ideální řešení se jeví rozdělit aplikaci na dvě samostatné části. První část se bude starat o obsluhu nfdumpu, jeho volání se správnými parametry, ukládání jeho výsledků a případně provádění dalších činností bez interakce s uživateli. Pro příklad uveďme automatické zasílaní emailů. Druhá, webová část se bude starat o prezentaci získaných dat, nastavování offline části a o další činnosti, které vyžadují uživatelskou interakci. Toto rozdělení se jeví výhodné i z důvodu nutnosti předzpracovávat data offline. Zpracování dat nfdumpem online se ukázalo jako příliš pomalé pro potřeby webové aplikace. Další výhodou tohoto konceptu je, že obě aplikace mohou

běžet na dvou, na sobě nezávislých serverech. Pouze je nutné zařídit pravidelnou, ale nepříliš častou výměnu dat. Portál bude samozřejmě schopen činnosti i na jednom serveru.

Implementovat aplikaci, jako webový portál, je výhodné hned z několika hledisek. Hlavní výhodou je vysoká dostupnost díky rozšířenosti internetu. Další důležitou vlastností je, že veškerá případná rozšíření či opravy aplikace se budou provádět pouze u provozovatele. Jediným požadavkem na uživatele je nutnost internetového připojení a webového prohlížeče. Toto ale, vzhledem k cílové skupině uživatelů, nejsou podstatné nedostatky. Pojetí aplikace jako webového portálu též zajišťuje nezávislost na platformě.

Jak již bylo zmíněno, portál musí být schopen ukládat data. Jedná se především o data zpracovaná nfdumpem, která se použijí při vizualizaci. Dále se jedná o konfigurační data. Ta jsou potřeba pro uložení nastavení aplikace a pro přenos informací mezi webovou a offline částí. Celkový objem těchto dat je relativně nízký. Z toho důvodu nebude použita databáze, ale pouze textový XML soubor pro konfiguraci a textové soubory pro uložení zpracovaných dat. Toto řešení dále snižuje nároky na server, protože nebude nutné instalovat a udržovat v běhu databázový program.

Portál bude schopen automatického exportu pomocí emailu nebo bude umožňovat ruční stažení reportů do formátu PDF. Výhodou tohoto formátu je jeho rozšířenost a platformní nezávislost.

4.1.2 Uživatelské role a profily

Portál bude podporovat správu více uživatelů a bude umět ukládat jejich osobní nastavení. Každý uživatel bude mít své vlastní nastavení aplikace, bude mu umožněno vytvářet si vlastní reporty, bude si moci nastavit pravidelné odesílání vybraného reportu na svůj email a podobně. Druhou rolí, která se v systému bude vyskytovat je administrátor. Administrátor bude moci provádět všechny úkony jako běžný uživatel. Navíc bude mít možnost upravovat vnitřní nastavení aplikace a spravovat všechny uživatelské účty. Bude mít tedy pravomoc přidělovat práva novým uživatelům systému, případně je rušit.

4.1.3 Uživatelské rozhraní

Uživatelské rozhraní portálu bude jednoduché a intuitivní. Hlavním důvodem je efektivita práce s portálem. Portál bude navržen pro správu sítě a není tedy nutné, aby rozhraní bylo plné grafických efektů. Rozhraní bude implementováno pomocí CSS (viz kapitola 5.1.2), což umožňuje v případě potřeby jednoduše měnit grafický výstup portálu. Administrátorské a uživatelské rozhraní se budou

graficky lišit pouze minimálně. Interakce rozhraní s uživatelem bude řešena možnostmi HTML (viz kapitola 5.1.1), tedy pomocí formulářů.

Zobrazování reportů bude řešeno pomocí přehledných tabulek a grafů. Pro vykreslování grafů bude použita především technologie Flash (viz kapitola 5.1.5), která zajistí jejich přívětivý a moderní vzhled. Pokud bude uživatel potřebovat, bude možné si report stáhnout ve formátu PDF, případně si bude moci nechat poslat data na email v textové podobě.

4.1.4 Reporty

Klíčovou součástí této práce jsou reporty. V našem kontextu reporty rozumíme různě zpracovaná data o provozu na síti. Tato práce se zabývá tzv. top reporty. Příkladem dotazu na top report může být například: „Který uživatel stáhl minulý týden nejvíce dat?“ nebo „Mezi kterými počítači probíhala včera datová výměna nejčastěji?“

Portál bude obsahovat širokou paletu předdefinovaných reportů rozdělených do kategorií podle zaměření, například bezpečnostní reporty, či reporty pro přenos dat. U těchto reportů bude možnost si je zobrazit v různých časových úsecích. Předdefinované reporty jsou určeny především běžným uživatelům, kterým poskytnou základní pohled o dění v síti.

Samozřejmě není možné předdefinovat všechny možné reporty. Proto je zkušenějším uživatelům nabídnuta možnost si vytvořit své vlastní reporty. Aplikace se bude snažit vytvoření vlastního reportu uživatelům co nejvíce usnadnit. Stejně ale bude potřeba, aby uživatel znal alespoň základy syntaxe nfdumpu (viz kapitola 2.5).

4.1.5 Automatická činnost aplikace

Jak již bylo zmíněno na začátku 3. kapitoly, je nezbytné, aby některé úkony aplikace byly prováděny samostatně, bez interakce s uživatelem. Aplikace tedy musí být schopna volat nfdump v pravidelných intervalech z několika důvodů. Hlavní důvodem je doba zpravování dat nfdumpem, která vylučuje jeho použití přímo ve webových portálech. Dále snaha o co nejaktuálnější data pro reporty, jejichž zpracovávaný časový interval musí být co nejblíže aktuálnímu času. Konečně je nutné generovat nové, uživatelem definované reporty.

Další činností, kterou musí aplikace umět provádět samostatně, je práce s pravidelnými emaily pro uživatele. Aplikace musí být schopna zjistit, jestli si některý uživatel systému přeje zaslat email. Pokud ano, musí to provést.

4.1.6 Bezpečnost

Aplikace je navržena jako webová. Z toho plynou jistá bezpečnostní rizika. Je možné, že se někdo může pokusit tuto aplikaci zneužít. Možnosti zneužití nejsou příliš velké, protože aplikace pouze zobrazuje data o síťovém provozu. V lokální síti by případný útočník mohl maximálně zjistit, na jaké servery se počítače ze sítě připojují nejčastěji. Bezpečnost systému je ale nutné brát vážně. Z tohoto důvodu bude pro přístup do aplikace vyžadována autentizace. Hesla samozřejmě budou uložena v zahashovaném tvaru. Další zabezpečení bude umožněno pomocí Apache (viz kapitola 5.1.8).

4.1.7 Předpokládané nasazení

Systém je primárně navržen pro nasazení v malé nebo středně velké síti. Typickým příkladem nasazení je menší firma s několika desítkami stanic, jejíž síť odděluje od internetu směrovač, na kterém běží procesy pro sběr NetFlow dat. Tento systém pak umožní správci lepší přehled nad sítí. Umožní mu lépe monitorovat provoz na síti a lépe identifikovat možné zdroje problémů.

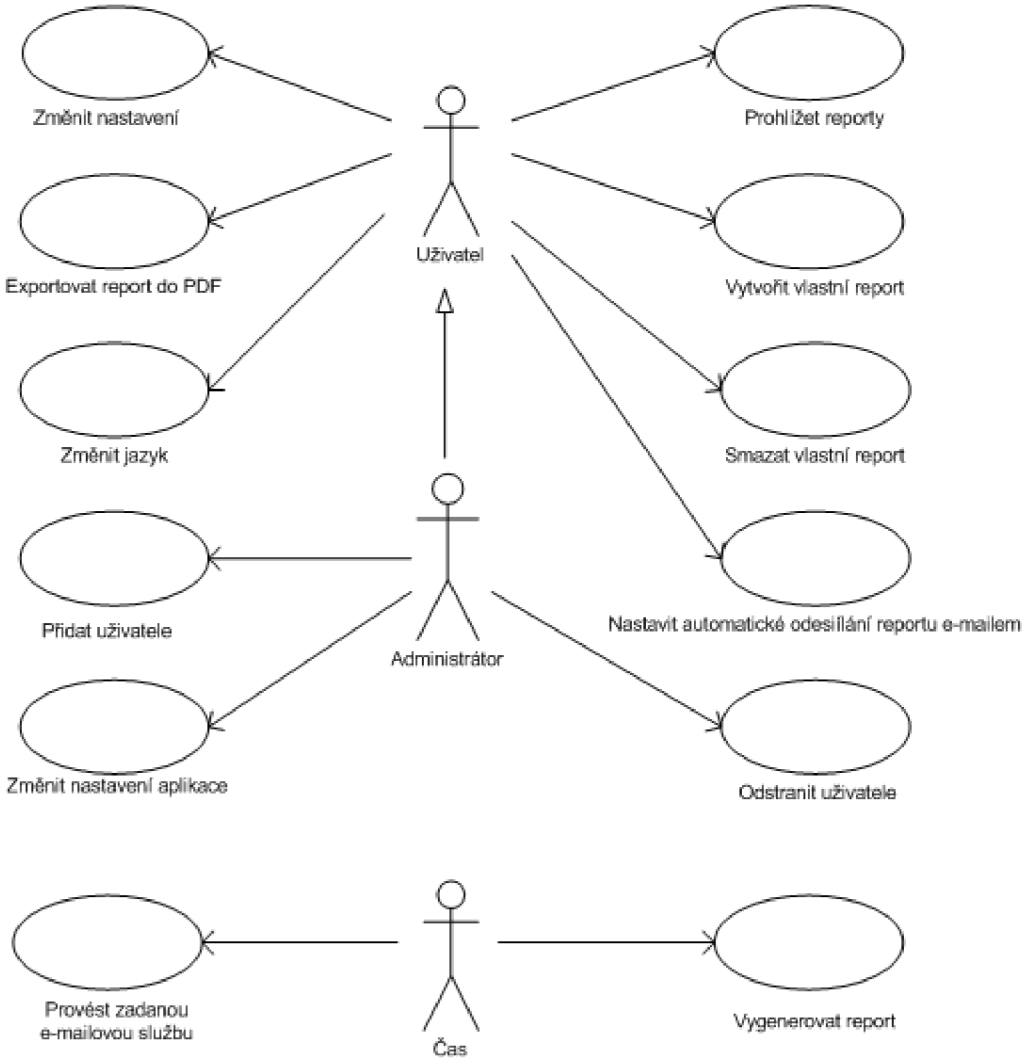
4.2 UML

UML (Unified Modeling Language) je v softwarovém inženýrství označení pro grafický jazyk, vhodný pro vizualizaci, specifikaci, navrhování a dokumentaci programových systémů. UML umožňuje standardizovat zápis návrhu systémů, jak konceptuálních, tak konkrétních prvků systému. UML popisuje tvorbu mnoha různých diagramů. Pro vývoj této aplikace je vhodné použít diagram případů užití.

4.3 Diagram případů užití

Diagram případů užití (Use Case Diagram) znázorňuje chování systému z pohledu uživatele. Umožňuje definovat systém a jeho vztahy s vnějším okolím. Jeho úkolem je podávat obraz funkčnosti systému, který je vyvolán vnějšími podněty.

Diagram případů užití zahrnuje aktéry (actors), systém a případy užití (use case). Aktér reprezentuje činnost uživatele nebo jiný prvek, ovlivňující chování systému. Případy užití reprezentují možnou interakci aktérů se systémem. Diagram případů užití aplikace ukazuje ilustrace 2.



Ilustrace 2: Diagram případů užití

5 Implementace

5.1 Použité technologie

5.1.1 HTML

Jazyk HTML (HyperText Markup Language) je značkovací jazyk, který vznikl aplikací dříve vyvinutého značkovacího jazyka SGML (Standard Generalized Markup Language).
Standard SGML vznikl v rámci projektu ODA (Open Document Architecture). Cílem ODA bylo vytvořit standardní architekturu pro vytváření, předávání, uchovávání a zpracovávání různých dokumentů v elektronické podobě. Pro potřeby ODA bylo nutno vytvořit formát, který by umožňoval uložení textů nezávisle, jak na softwarové, tak hardwarové platformě. Jako řešení se ukázalo použití značkovacího jazyka (odtud Markup Language). Jazyk je koncipován tak, že přímo v textu jsou umístěny značky určující sémantiku textu, který obklopují. SGML také zavádí DTD (Document Type Definition), který upravuje, jaké značky můžeme v konkrétním dokumentu použít a v jakých vzájemných vztazích mohou být použity. Tohoto konceptu bylo využito při tvorbě HTML.

HTML je SGML dokument, ve kterém je značkám přiřazena sémantika webového hypertextového dokumentu. Kořeny jazyka HTML sahají do roku 1989, kdy se pro tvorbu dokumentů používaly jazyky jako TeX, PostScript, případně SGML. Tim Berners-Lee si uvědomoval, že je potřeba vytvořit jednoduší jazyk a v roce 1990 navrhl jazyk HTML společně s protokolem HTTP (HyperText Transfer Protocol). Zároveň napsal první internetový prohlížeč. Tyto události odstartovaly rychlý rozvoj webu a bylo nutné pravidelně vytvářet standardy, o které se stará konsorcium W3C.

Současná HTML verze 4.01 se postupně nahrazuje novým jazykem XHTML. XHTML je jazyk, který vyhovuje požadavkům pro tvorbu XML (viz. kapitola 5.1.6), ale je zpětně kompatibilní s jazykem HTML. Aktuální verze jazyka je 1.1, ale očekává se jeho verze 2.0. Mezi rozdíly HTML a XHTML patří například vyloučení nepárových značek (tag) v XHTML, či nutnost psát je malými písmeny. Dále dokument začíná deklarací DTD.

Tímto způsobem je implementována i výsledná aplikace.

5.1.2 CSS

Cascading Style Sheet, neboli kaskádové styly jsou jazykem, určeným pro tvorbu vzhledu HTML, XHTML nebo XML dokumentů. Úkolem jazyka je oddělení formátování a vzhledu dokumentů od

obsahu. CSS soubor bývá zpravidla umístěn externě, aby bylo možné měnit vzhled stránky pouze změnou CSS dokumentu. Jazyk CSS má velice jednoduchou syntaxi, obsahující pouze selektor elementů (značek), na které se má styl aplikovat a popis výsledného vzhledu. Aktuální verze je CSS2 a jazyk je široce používán. Negativním jevem při vývoji aplikací s využitím CSS je různá podpora standardů výrobci prohlížečů a tedy různý vzhled stejných dokumentů v rozdílných prohlížečích. Jazyk je vyvíjen konsorciem W3C.

5.1.3 PHP

Skriptovací jazyk PHP (dříve Personal Home Pages, nyní PHP: Hypertext Preprocesor) vytvořil v roce 1994 Rasmus Lerdorf při vytváření počítadla přístupů na své stránky v Perlu. Neustálé spouštění interpretu Perlu zbytečně zatěžovalo server, což vedlo k přepsání principů Perlu do jazyka C. Tento systém se velmi osvědčil a rychle se rozšířil.

PHP je skriptovací jazyk, běžící na straně serveru. PHP se zpravidla začleňuje přímo do HTML souboru, který je pak zpracován skriptem. Ten projde zdrojový soubor, postupně zpracuje všechny PHP příkazy a klientovi pak pošle jen čistý HTML kód, který je výsledkem běhu skriptu. PHP je schopno běžet i samostatně jako klasický skript, což je využito i v implementaci této aplikace.

Velkou výhodou PHP je rozšiřitelnost pomocí obrovského množství doplňujících knihoven, které činí z PHP velice univerzální jazyk. Popis knihoven použitých v této práci se nalézá v kapitole 5.2.

PHP je schopno pracovat s libovolným webovým serverem a je tedy platformě nezávislé. V této aplikaci poběží na serveru Apache (viz kapitola 5.1.8). V současné době se používají verze 4 a 5, přičemž verze 5 umožňuje objektovou práci v PHP a je rovněž použita v této aplikaci.

5.1.4 JavaScript

JavaScript je multiplatformní, objektově orientovaný, skriptovací jazyk, syntaxí podobný jazykům C++ a Java. Autorem jazyku je Brendad Eich a jeho první standard pochází z roku 1997.

JavaSript se zpravidla používá jako interpretovaný jazyk, vkládaný přímo do HTML dokumentu. Obvykle jsou jím tvořeny různé interaktivní prvky webového rozhraní nebo různé grafické efekty. JavaSript je interpretován na straně klienta zpravidla webovým prohlížečem. Z toho plyne obrovská výhoda a možnosti provádět operace bez odeslání na server. Nevýhodou je, že JavaScript může být z bezpečnostních důvodů zakázán, případně omezen. JavaScript je možné použít i na straně serveru. První implementace je LiveScript od společnosti NetScape, uvedený v roce 1996.

Mimo webové prostředí se JavaScript používá k psaní různých rozšíření pro různé aplikace, případně ke skriptování pod Windows (Windows Script Host).

5.1.5 Flash

Adobe Flash je grafický vektorový program, používaný především pro tvorbu interaktivních animací, prezentací a her. Flash nalezl uplatnění ve webových aplikacích, především díky malé velikosti souborů, které bylo dosaženo uložením dat ve vektorovém formátu. Součástí Flashe je i vlastní programovací jazyk jménem ActionScript. Soubory Flashe, určené pro webové aplikace, mají koncovku .swf a k jejich běhu je nutný přehrávač. Soubory Flash lze generovat i pro použití bez přehrávače. Tento formát se ale z důvodu větší velikosti ve webovém prostředí příliš nepoužívá. V naší aplikaci je použita .swf varianta pro vykreslování grafů. Více v kapitole 5.2.1.

5.1.6 XML

XML (eXtensive Markup Language) je obecný značkovací jazyk vyvinutý konsorciem W3C. Patří do rodiny jazyků SGML. Jazyk umožňuje snadné vytváření konkrétních značkovacích jazyků. Jazyk sám o sobě nemá žádné předdefinované značky. XML byl především navržen pro výměnu dat mezi aplikacemi a pro publikování dokumentů. Syntaxe jazyka XML je podstatně přísnější a čistější, než syntaxe HTML. Součástí XML je i kontrola struktury dokumentu.

Dokument XML je vždy textový a Unicode. V Česku se obvykle používá kódování UTF-8. Efektivita XML je silně závislá na obsahu, struktuře a integritě. Aby mohl být dokument považován za správně strukturovaný (well-formed), musí splňovat následující podmínky:

- musí mít právě jeden kořenový element (root)
- neprázdné elementy musí být ohraničeny startovací a ukončovací značkou
- prázdný element musí být označen značkou prázdný element
- hodnoty atributů musí být uzavřeny v jednoduchých nebo dvojitých uvozovkách, nikoliv však kombinací obou
- elementy mohou být vnořeny, ale nesmějí se překrývat

Jména XML elementů rozlišují malá a velká písmena. V současnosti je aktuální verze 1.1. Příklad XML dokumentu použitého v aplikaci se nachází v kapitole 5.3.2.

5.1.7 PDF

PDF (Portable Document Format) je hardwarově a softwarově nezávislý formát pro ukládání dokumentů, vyvinutý společností Adobe Systems. Jedná se o otevřený standard, díky němuž je velice

oblíbený a rozšířený. PDF umožňuje ukládat do souboru text, obrázky a případně i hypertextové odkazy. Je zaručeno, že se dokument zobrazí na libovolné platformě stejným způsobem.

5.1.8 Apache

Apache HTTP Server (označovaný jako Apache) je softwarový multiplatformní webový server s otevřeným kódem. Apache je vyvíjen otevřenou komunitou, známou jako Apache Software Foundation, a v současnosti je nejpoužívanějším webovým serverem na světě. Apache je velmi ceněn z mnoha různých důvodů. Je třeba zdůraznit fakt, že je zcela zdarma. Dále pak jeho nezávislost na platformě operačního systému. V neposlední řadě i jeho rozšiřitelnost pomocí modulů, které nabízejí velké množství dalších funkcí (například autentizaci, podporu tvorby dynamického obsahu nebo podporu proxy serveru). Hlavním důvodem použití Apache v této práci je jeho vhodnost pro tvorbu dynamicky generovaných stránek. V současnosti je aktuální verze 2.0.

5.2 Použité knihovny

5.2.1 Vykreslování grafů

Pro jednoduché a rychle pochopitelné zobrazení dat se v této aplikaci používají grafy. Jsou zde použity moderní a uživatelsky příjemné grafy. Tyto grafy jsou vytvářeny pomocí technologií JavaScript a Flash, což umožňuje jejich animaci a interaktivitu.

Pro práci s grafy je použita velmi dobře navržená knihovna FusionCharts. Výhodou této knihovny je, že je v základní verzi distribuována zdarma. Velice ceněnou vlastností knihovny je možnost změnit typ vytvářeného grafu pouhým změněním zdrojového flashového souboru, bez nutnosti měnit implementaci aplikace nebo formát vstupních dat. Více informací o této knihovně naleznete v [8].

5.2.2 Tvorba PDF

Často je potřeba si uložit zobrazené informace k pozdějšímu použití. Jako formát exportu dat byl pro tuto aplikaci vybrán velice rozšířený a univerzální formát PDF. Pro tvorbu PDF souborů je použito knihovny FPDF, která je opět volně šiřitelná. Jedná se o velice dobře navrženou a implementovanou knihovnu s mnoha možnostmi tvorby výsledného dokumentu. Více se můžete dozvědět na domovských stránkách knihovny v [9].

5.2.3 Práce s emaily

Aplikace počítá s možností, že bude třeba zaslat výsledná data o nějakém reportu emailem. Přímo v PHP je možné emaily zasílat vestavěnou funkcí ***mail()***. Bohužel použití této funkce je celkem omezené, a tak bylo rozhodnuto použít specializovanou knihovnu. Volba padla na objektově řešenou knihovnu PHP Mailer. Tato knihovna umožňuje bohaté možnosti nastavení a hlavně je distribuována jako volně šiřitelná. K přenosu emailů je použit protokol SMTP, využívající zpravidla port 25. Pro více informací se obraťte na [10]

5.2.4 Automatické spouštění skriptů

Automatické spouštění skriptů je nezbytné pro běh aplikace. Je nutné především pro pravidelnou offline práci s nfdumpem. Dále je toto spouštění použito při odesílání vybraných reportů na email. Je tedy nutné zařídit, aby byly tyto skripty spouštěny v pravidelných intervalech.
Řešení tohoto problému nabízí softwarový démon cron. Cron umožňuje v operačních systémech na bázi Unix automatizované a pravidelné spouštění nastavených příkazů nebo procesů. Nastavení cronu se provádí konzolovým příkazem crontab. ***Crontab -l*** zobrazí aktuální nastavení a ***crontab -e*** umožňuje toto nastavení měnit. Syntaxe v crontab je následující:

*** * * * * příkaz_k_provedení**

příkaz_k_provedení je příkaz, který se cron pokusí spustit. Hvězdičky postupně znamenají minuty, hodiny, dny v měsíci, měsíce a dny v týdnu, kdy se má příkaz spouštět. Například pro spuštění skriptu **run.php** v adresáři **/usr/local** každý den, minutu po půlnoci provedeme takto:

1 0 * * * /usr/bin/php /usr/local/run.php ,

kde v adresáři **/usr/local/php** se nachází binární soubor jazyka PHP.

Pro podrobnější informace o crontab můžete použít manuálové stránky, které jsou dostupné na [11].

5.3 Popis klíčových skriptů a souborů

V této části práce budou stručně popsány nejdůležitější soubory aplikace. Všechny soubory se dají rozdělit na PHP skripty a ostatní soubory.

5.3.1 Popis klíčových PHP skriptů

- **index.php** – hlavní skript celé aplikace. Zajišťuje generování HTML hlavičky, připojení CSS souboru a správu zobrazeného obsahu aplikace. Dále zajišťuje inicializaci a správu Sessions, kontroluje autentizaci a dobu nečinnosti uživatele, případně zajišťuje jeho odhlášení.
- **func.php** – soubor s mnoha pomocnými funkcemi, především pro generování obsahu, kontrolu autentizace a další všeobecně použitelné funkce. Do tohoto souboru jsou rovněž inkludovány všechny konstanty.
- **const_arrays.php** – obsahuje používané konstanty a pole používaná při běhu aplikace. V něm lze nastavit dobu pro automatické odhlášení uživatele, spravuje cesty k souborům nebo obsahuje tabulku umožňující převod čísla portu na odpovídající službu.
- **texts.php** – obsahuje pole pro generování všech vypisovaných textů. Texty jsou odděleny od zbytku aplikace z důvodu snadné modifikace či přidání dalšího komunikačního jazyka.
- **loader.php** – obsahuje všechny potřebné funkce pro zpracovaní dat, vykreslení tabulek a grafů, dále se zajišťuje volání skriptu pro export dat.
- **parser_long.php** – zajišťuje zpracování výstupních dat z programu nfdump a jejich správné uložení.
- **login.php** – zajišťuje kontrolu přihlašovacích údajů. Je volán skriptem index.php, pokud se o práci s aplikací pokouší neautentizovaný uživatel.
- **nfcall.php** – offline skript zajišťující práci s nfdumpem. Spouštěn je pravidelně cronem. Po svém spuštění načte z konfiguračního XML souboru požadavky na volání programu nfdump a provede je. Dále se stará o pravidelné generování předdefinovaných reportů, kontroluje dostupnost zdrojových souborů pro nfcapd a mnoho dalších věcí. Ke komunikaci s webovou aplikací používá soubor config.xml.
- **mailer.php** – automaticky spouštěný offline skript. Stará se o zasílaní vyžádaných emailů s reporty na požadované adresy. Při svém spuštění zkontroluje, zda je potřeba nějaký email zaslat. Pokud ano a potřebná data jsou k dispozici, mail odešle. Pokud data k dispozici nejsou, požádá o jejich vygenerování. K zaslání emailů používá knihovnu PHP Mailer. Dále obsahuje konfigurační data o používaném SMTP serveru.
- **offlinelib.php** – knihovna funkcí offline spouštěných souborů. Obsahuje především funkce pro práci s XML a pomocné funkce pro správu a kontrolu zdrojových souborů nfdumpu.

- **xmlhandler.php** – knihovna všech potřebných funkcí pro práci s XML soubory, často využívaná různými skripty. Pro svou činnost používá pouze standardní funkce jazyka PHP.
- **fpdfhandler.php** – zajišťuje vytvoření PDF souboru z dodaných dat. Pro svou činnost používá knihovnu FPDF.

Další skripty aplikace se starají o generování konkrétního obsahu stránek, voláním výše uvedených skriptů. Jejich popis zde není nutný.

5.3.2 Popis dalších významných souborů a adresářů

V této kapitole budou uvedeny soubory, které sice nejsou skripty, ale podílejí se na běhu aplikace.

- **config.xml** – konfigurační soubor a zároveň komunikační prostředek mezi webovou a offline částí aplikace. Obsahuje všechna důležitá konfigurační data, nezbytná pro běh aplikace. Nalézají se v něm klíčové cesty, informace o souborech vygenerovaných nfdumpem. Dále pak informace o uživatelích, jejich profilech a o objednaných emailových exportech. Následuje krátká ukázka struktury dat:

<users>
<user>
<nick>Petr</nick>
<name>Petr</name>
<surname>Zapletal</surname>
<commlang>cz</commlang>
<animate>false</animate>
<mail>xzaple01@stud.fit.vutbr.cz</mail>
<dns>true</dns>
. ..
</user>
. ..
</users>

- **style** – adresář obsahující soubory s kaskádovými styly aplikace. Tyto soubory zajišťují grafické rozhraní aplikace. V aplikaci jsou použity dva soubory s kaskádovými styly. Jeden je určen pro prohlížeče Microsoft Internet Explorer a druhý pro ostatní prohlížeče. Tyto soubory jsou volány skriptem index.php.
- **.htaccess** – soubor nastavující zabezpečení Apache. Znemožňuje přístup k vybraným souborům. Více v kapitole 5.8.
- **auth.txt** – skrytý soubor obsahující zahashované autentizační údaje. Více v kapitole 5.8.
- **FusionCharts** – adresář obsahující Flashové.swf soubory potřebné pro tvorbu grafů. Jeho umístění je uloženo v konfiguračním souboru.
- **lib** – adresář obsahující soubory knihoven pro obsluhu elektronické pošty a pro tvorbu souboru ve formátu PDF. Více informací naleznete v kapitole 5.2.

V adresáři aplikace se zpravidla nacházejí i další adresáře. Ty mohou být použity pro uložení zdrojových dat programu nfdump. Jedná se tedy o soubory vytvořené programem nfcapd. Dále je potřeba ukládat výstupní data z nfdumpu. Výstup nfdumpu bývá zpravidla uložen v dalším adresáři a data jsou ukládána beze změn v textové podobě. Umístění těchto adresářů může být libovolné, protože cesta k nim je uložena v konfiguračním souboru.

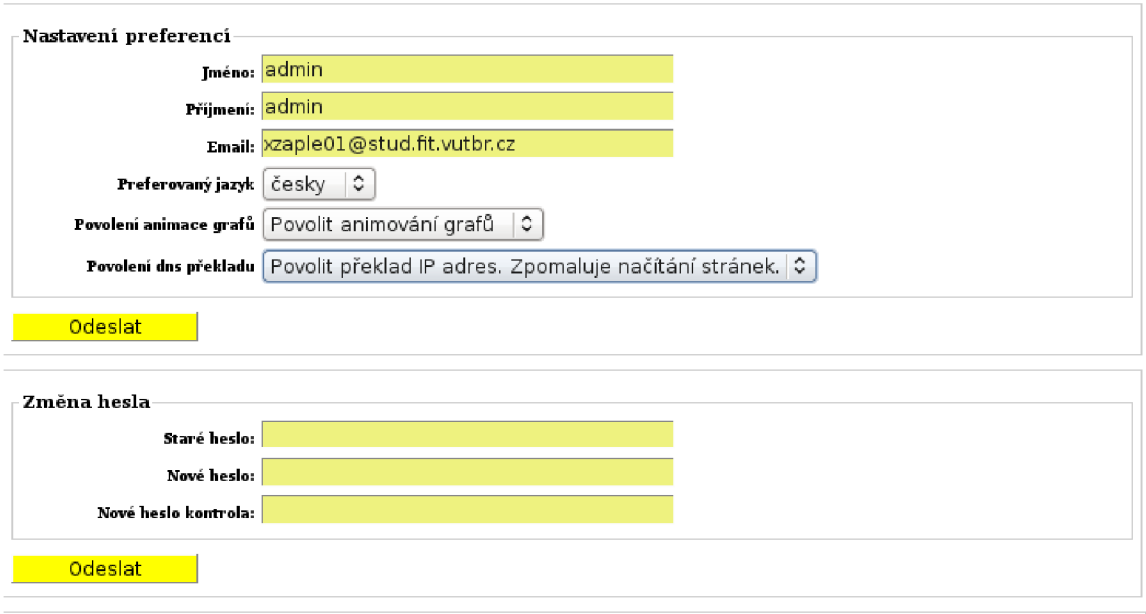
5.4 Uživatelské prostředí

5.4.1 Uživatel

Uživatelské prostředí bylo navrženo jako jednoduché a intuitivní, ale zároveň umožňující všechny požadavky na funkčnost. Do grafického rozhraní nebyly implementovány bohaté grafické prvky. Vše bylo podřízeno přehlednosti a funkčnosti.

Grafické rozhraní umožňuje uživateli prohlížet uložené reporty, případně je ukládat do PDF souboru. Předdefinované reporty mají možnost zobrazení v několika časových intervalech. Uživatel si může v rozhraní nastavit pravidelné posílaní vybraných reportů na emailovou adresu, případně si nadefinovat reporty vlastní. Přestože se aplikace snaží definici vlastních reportů co nejvíce ulehčit, je nutné, aby měl uživatel alespoň minimální znalosti syntaxe nfdumpu. Proto se vytváření vlastních reportů doporučuje spíše zkušenějším uživatelům. Dále má uživatel možnost spravovat svůj profil. Z možností profilu stojí za zmínku především možnost zakázat překlad IP adres na doménová jména,

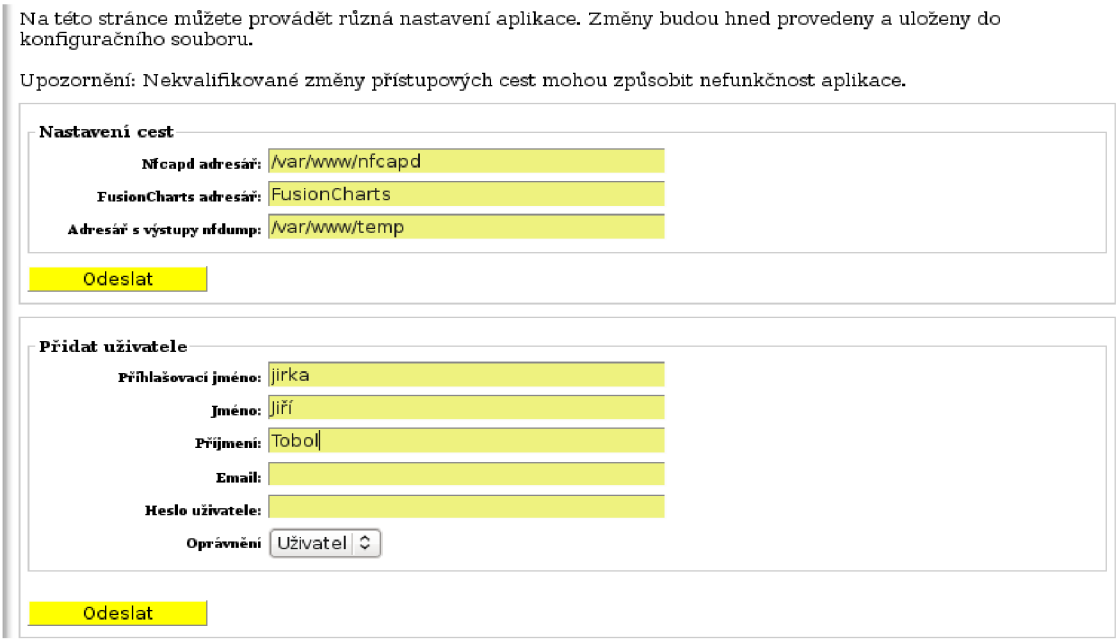
což znatelně urychlí generování reportů. Dále je možno zakázat animaci grafů. Samozřejmě si uživatel může editovat údaje o své osobě. Nastavení uživatelské profilu ukazuje ilustrace 3.



Ilustrace 3: Nastavování profilu uživatele

5.4.2 Administrátor

Vzhled administrátorského prostředí se nijak neliší od prostředí uživatelů. Rozdíly jsou však ve funkčnosti. Funkčnost uživatelského prostředí je plně přenesena do administrátorského. Administrátor má navíc možnosti pro správu uživatelů. Je mu umožněno vytvářet nové uživatele, případně rušit stávající. Jeho další důležitou pravomocí je možnost správy klíčových vlastností aplikace. Jedná se především o nastavení cest k datovým a zdrojovým adresářům. Toto nastavení je přístupné pouze administrátorům, protože nekvalifikované změny by mohly způsobit nefunkčnost aplikace. Administraci portálu ukazuje ilustrace 4.



Ilustrace 4: Administrace portálu

5.5 Nfdump a zpracování reportů

Klíčovou součástí aplikace je způsob zpracování dat o síťovém provozu. Úkolem aplikace je zpracovat data dodaná programem nfcapd. K tomuto účelu je použit program nfdump, který byl pro tuto činnost navržen. V této kapitole je podrobný popis jeho použití v aplikaci.

Program nfdump umožňuje vytvářet velké množství statistik. Pro správné použití je třeba předat programu několik parametrů. V prvé řadě je potřeba určit, ze kterých zdrojových souborů má nfdump čerpat, a tím i časový rozsah výsledného reportu. To se provádí pomocí parametru ***-M***, pro určení hlavního adresáře s nfcapd soubory a parametrem ***-R***, pro určení potřebných podadresářů a jednotlivých souborů. Dále je třeba specifikovat, jaké statistiky nás zajímají. V tomto případě se jedná o top statistiky. Parametrem pro vytváření top statistik a případně určení klíče pro jejich seřazení je parametr ***-s***. Další důležitou volbou je formát výstupních dat. V tomto případě byl použit implicitní formát line, který poskytuje dostatečné množství informací a navíc nfdump umí pracovat s některými top statistikami pouze v tomto formátu.

Nfdump umožňuje vytvářet top statistiky o těchto datech:

- zdrojová IP adresa
- cílová IP adresa
- jakákoliv IP adresa (zdrojová nebo cílová)
- zdrojový port
- cílový port
- jakýkoliv port
- zdrojové AS
- cílové AS
- jakékoliv AS
- vstupní rozhraní
- výstupní rozhraní
- jakékoliv rozhraní
- číslo protokolu
- záznam (record) obsahuje informace o relaci mezi zdrojovou a cílovou IP adresou

Výsledná data je pak možno seřadit podle:

- toků
- paketů
- bytů
- paketů za sekundu
- bitů za sekundu
- bytů na paket

Dále je možno zadat klíč pro filtrování těchto statistik. Možnosti filtrování dat jsou velice bohaté. Nfdump umožňuje filtrovat data podle IP adres, podle použitých portů, podle protokolu, podle síťové masky, podle počtu přenesených bytů a podobně. Je možno použít logických a matematických operátorů.

Nfdump implicitně používá standardní výstup, a proto je výhodné převést tento výstup do souboru.

Výsledný příkaz, používaný pro volaní nfdumpu, pak vypadá například takto:

nfdump -M /var/www/nfcapd -R 02/22/nfcapd.200902221000: 03/01/nfcapd.200903011000 -s ip/bytes 'port 80 or port 443' >top_www_servers.txt

Tento příkaz vytvoří top statistiku o IP adresách, které jsou seřazeny podle počtu bytů. Nfdump bude zpracovávat toky od desíti hodin dne 22.2.2009 do jedenácti hodin dne 1.3.2009, jejichž číslo portu bylo 80 anebo 443. Výstup bude uložen do souboru top_www_servers.txt.

Ukázka formátu uložených dat:

Top 10 Dst IP Addr ordered by bytes:

Date first seen Duration Proto IP Addr Flows Packets Bytes pps bps bpp

2009-02-27 23:59:31.544 2675057.880 any 192.168.3.110 125914 164663 32.0 M 0 100 204

2009-03-02 10:10: 52.778 2435608.405 any 192.168.3.106 31775 61858 19.7 M 0 67 333

2009-03-06 09:50:32.932 522.018 any 192.168.3.103 30 16084 16.1 M 30 258 10

...

Summary: total flows: 372398, total bytes: 114.8 M, total packets: 533568, avg bps: 359, avg pps: 0, avg bpp: 225

Time window: 2009-02-27 23: 59: 30 – 2009-03-31 00: 04: 34

Total flows processed: 992068, Records skipped: 0, Bytes read: 51694540

Sys: 0.360s flows/second: 2755576.0 Wall: 0.333s flows/second: 2975680.3

Takto uložená data pak mohou být použita webovým portálem pro prezentaci. Syntaxe programu nfdump je velmi rozsáhlá a propracovaná. Účelem této práce není její podrobný popis. Více informací o programu nfdump, jeho syntaxi a možnostech použití lze nalézt v [3].

5.6 Předdefinované reporty

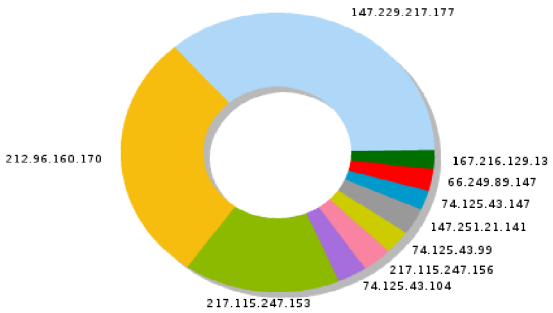
Součástí aplikace jsou dvě desítky předdefinovaných reportů. Tyto reporty jsou pevnou součástí aplikace. Byly vytvořeny s ohledem na okamžité praktické použití správci sítí. Jedná se o top reporty, které jsou v praxi nejčastěji používané.

Předdefinované reporty jsou rozděleny do tří kategorií. První se zabývá obecně využitelnými reporty, použitelnými pro standardní sledování sítě. Jsou zde reporty zobrazující nejčastější relace v síti, nejvíce a nejčastěji využívané služby, nejčastěji navštěvované internetové servery, či stanice s nejrychlejším downloadem a uploadem.

Druhá kategorie se zabývá reporty sledujícími datový přenos. Nachází se zde reporty ukazující stanice s nejvíce využívanými datovými službami, stanice s největším downloadem nebo uploadem, ať už v lokální síti nebo do internetu.

Poslední kategorie je poměrně specifická. Zabývá se bezpečnostními reporty. Tyto reporty jsou použitelné v monitorování různých pokusů o narušení bezpečnosti sledované sítě. Pokud sledované statistiky těchto reportů budou výrazně přesahovat obvyklé hodnoty, je pravděpodobné, že na určitou stanici byl veden útok. Reporty v této kategorii zobrazují počet různých relací stanice nebo umožňují detekci možných útoků, jako SYN útok, smurf útok nebo UDP flood útok.

Součástí aplikace je vždy i popis, co přesně zkoumaný report zobrazuje, podle jakého klíče jsou data seřazena a co by měl uživatel provést v případě, že data neodpovídají očekávaným standardům. Používání předdefinovaných reportů je vhodné pro všechny, poněvadž nekladou na uživatele žádné další nároky. Vykreslení reportu ukazuje ilustrace 5.



	Doménové jméno	Cílová IP adresa	Bajty	Pakety	Bity/s
1	a04-0917b.kn.vutbr.cz	147.229.217.177	1.3 M	1500	322
2	invea.invea.cz	212.96.160.170	75823	706	11
3	217-115-247-153.customer.optimax.cz	217.115.247.153	460547	615	10405
4	bw-in-f104.google.com	74.125.43.104	29953	87	6
5	217-115-247-156.customer.optimax.cz	217.115.247.156	15491	118	73
6	bw-in-f99.google.com	74.125.43.99	20095	79	5
7	svn.liberouter.org	147.251.21.141	8656	76	8
8	bw-in-f147.google.com	74.125.43.147	22761	77	5
9	jp-in-f147.google.com	66.249.89.147	16774	75	5
10	shopping.netsuite.com	167.216.129.13	15017	130	55

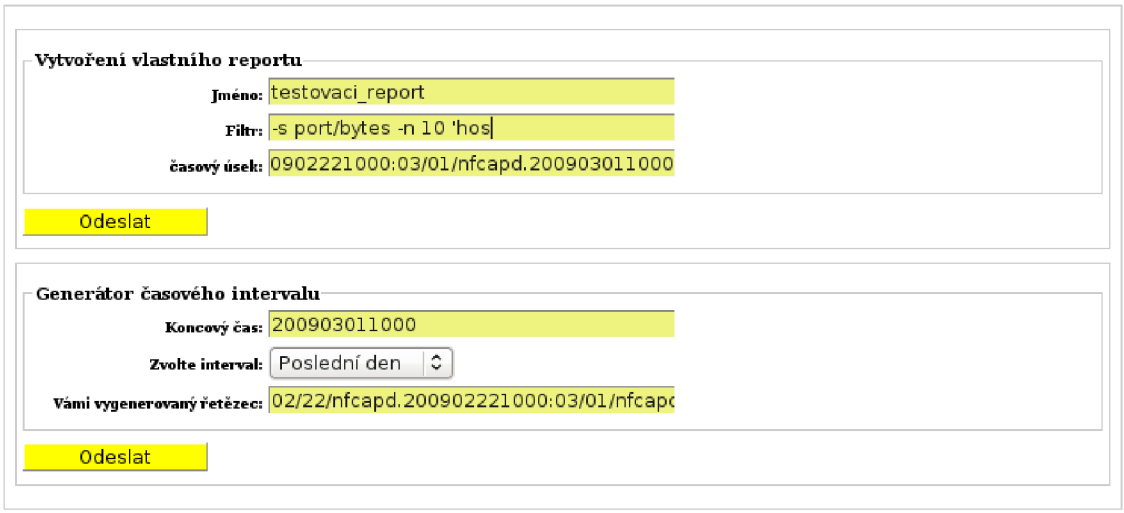
Ilustrace 5: Ukázka zobrazení reportu

5.7 Tvorba vlastních reportů

Předdefinovaných reportů může být velký počet a mohou být i velmi dobře navrženy. Je téměř jisté, že někdy uživatel bude potřebovat reporty, které připraveny nebyly. Z tohoto důvodu je v aplikaci implementována možnost tvorby reportů, podle vlastních parametrů. Tato možnost je určena především pro zkušenější uživatele.

Pro vytvoření vlastního reportu je nutné správně vyplnit připravený formulář. Je třeba vyplnit identifikační jméno reportu, dále pak časový rozsah vytvářeného reportu a nakonec popis vyžadovaného filtru pro nfdump. Rozsah reportu se vloží včetně správných zdrojových souborů pro nfdump. Aplikace však umožňuje tento řetězec vygenerovat automaticky podle zadaných údajů. Dále je nutné vyplnit filtr, tedy parametry, se kterými bude nfdump volán. Doporučuje se používat především parametr -s, který slouží k vytváření top reportů, pro které je aplikace určená. Další nastavení filtru jsou omezena pouze možnostmi programu nfdump.

Po vytvoření bude report uložen a při příštím spuštění se ho skript pro obsluhu nfdumpu pokusí vygenerovat. Pokud bude report správně nadefinován, zobrazí se v kategorii obsahující uživatelem definované reporty. Pokud ne, bude vygenerováno chybové hlášení. S takto vytvořeným reportem je pak možno pracovat stejně jako s předdefinovanými reporty. Tvorbu vlastního reportu ukazuje ilustrace 6.



Ilustrace 6: Vytvoření vlastního reportu

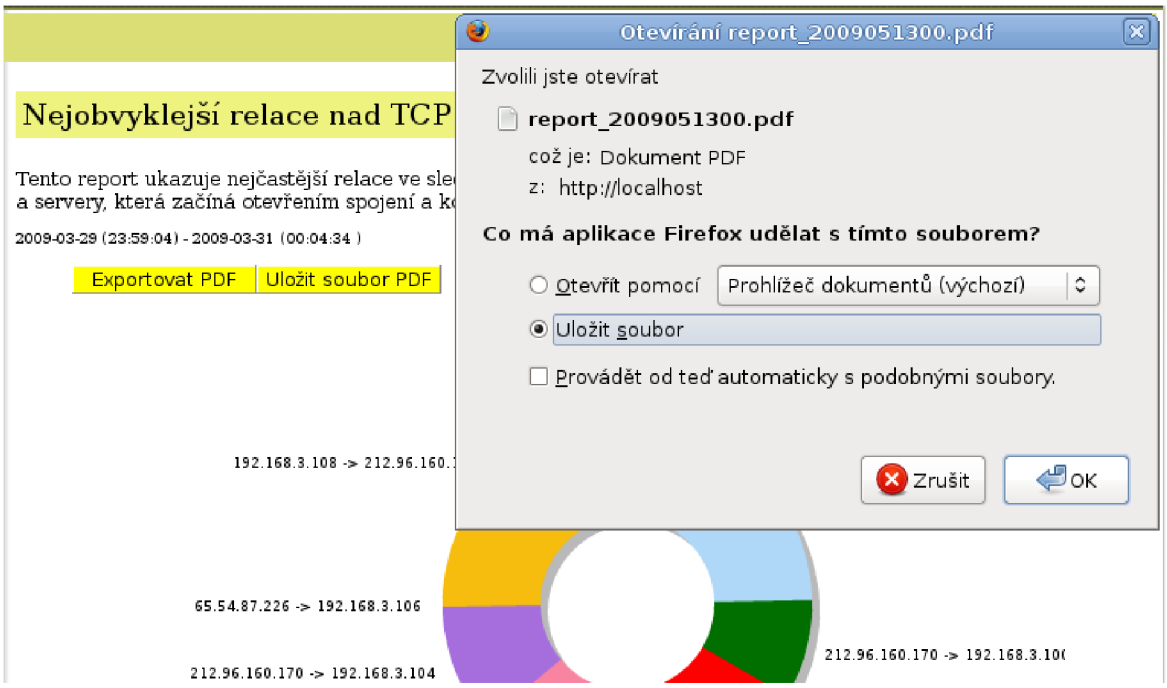
5.8 Export dat

V aplikacích tohoto typu je důležité umožnit uživatelům snadný export dat. Tato aplikace dovoluje uživatelům ukládat zobrazené reporty a dále umožňuje pravidelné zasílání reportů na email.

Export zobrazených reportů je navržen pro uložení, archivaci a případně další použití reportů. Data exportovaného reportu jsou uložena do formátu PDF. V souboru je uložen název exportovaného

reportu a jeho popis, dále datum exportu a samozřejmě data exportovaného reportu. Ukázku exportu reportu do PDF souboru ukazuje ilustrace 7.

Aplikace umožňuje nastavit pravidelné odesílání emailů s vybraným reportem na zvolenou emailovou adresu. Při vytváření emailového exportu musí uživatel vybrat typ reportu, jak často a na jakou emailovou adresu chce report posílat. Tato data budou uložena do konfiguračního souboru. Později budou předána automaticky spouštěnému skriptu, který zajistí provoz této služby.



Ilustrace 7: Export reportu do PDF souboru

5.9 Zabezpečení a šifrování

V současnosti je nezbytné, aby byla každá webová aplikace dobře zabezpečena. Více než kdy jindy platí, že bezpečnost není radno podceňovat. Tato aplikace však nepotřebuje špičkové zabezpečení jako systémy, které pracují s citlivými a snadno zneužitelnými údaji. Přesto je nutné aplikaci standardně zabezpečit. Zabezpečení stojí na dvou základních pilířích. První z nich je vlastní autentizační systém, druhý je zabezpečení poskytované serverem Apache.

Při přihlášení do systému je uživatel vyzván k zadání přihlašovacího jména a hesla. Zadané údaje se porovnají s daty uloženými v autentizačním souboru auth.txt. Pokud dojde ke shodě, jsou pro uživatele vygenerovány Sessions, umožňující přístup do portálu. Systém si dále kontroluje dobu

nečinnosti uživatele. Pokud je doba nečinnosti vyšší než nastavená hodnota, dojde k jeho automatickému odhlášení. Při manipulaci s hesly jsou použity šifrovací algoritmy MD5 a SHA-256 (označovaný i jako SHA-2). Pro práci s uvedenými algoritmy se používají vestavěné funkce PHP. Oba šifrovací algoritmy pracují na hashovacím principu.

Pro zvýšení bezpečnosti je využito služeb serveru Apache. K nastavení Apache se používá soubor .htaccess, umístěný v kořenovém adresáři portálu. Tento soubor umožňuje bohaté nastavení serveru Apache nejen v oblasti bezpečnosti. V našem případě bude použit pro omezení přístupu k privátním stránkám serveru. Další informace o souboru .htaccess jsou nalezení v [12].

5.10 Ovládání a nároky na uživatele

Ovládání portálu je intuitivní a odpovídající webovým standardům. Nemělo by činit problémy nikomu, kdo již má alespoň minimální zkušenosti s prostředím internetu a webu. O něco horší může být situace s pochopením obsahu a funkce portálu. Zde je potřeba mít základní povědomí o síťových technologiích. Aplikace je primárně určena pro správce sítí a ti nebudou mít s ovládáním aplikace problémy.

5.11 Nápověda a chybová hlášení

Přes veškerou snahu je možné, že bude mít uživatel problémy s ovládáním nebo funkčností aplikace. Součástí práce je i HTML soubor s nápovědou. Tento soubor obsahuje základní informace pro práci s portálem, případně pro jeho instalaci a nastavení. Dále obsahuje často kladené otázky (FAQ) a odpovědi na ně i kontakt na autora a vedoucího práce. Tento soubor je přístupný, jak samostatně, tak i přímo z menu aplikace.

Portál může při své činnosti vypisovat chybová hlášení. Hlášení jsou generována v PHP a vypisována pomocí HTML. Chybové zprávy rovněž obsahují popis konkrétní chyby a pravděpodobný důvod, proč k chybě došlo. Jednotlivé chyby jsou rovněž popsány v nápovědě.

6 Instalace a testování

Vlastní implementací systému práce na projektu zdaleka nekončí. Je třeba systém zprovoznit a důkladně otestovat.

6.1 Doporučená konfigurace

6.1.1 Server

- Apache HTTP Server 2.0
- PHP verze 5 a vyšší
- Nfdump 1.5.7
- Unixový operační systém

6.1.2 Klient

- Moderní internetový prohlížeč (Firefox 2.0, Internet Explorer 7.0, Google Chrome 2.0, Opera 9.64)
- Adobe Flash Player 10
- JavaScript 1.5 a vyšší

6.2 Instalace

Pokud je k dispozici server s odpovídající konfigurací, je možno zahájit instalaci aplikace. Instalace je velmi jednoduchá. Stačí překopírovat adresář se zdrojovými soubory aplikace do webového adresáře serveru. Oprávnění a cesty mezi soubory jsou správně připraveny a nastaveny na zdrojovém nosiči. Posledním krokem je nastavení automatického spouštění skriptů ***nfcall.php*** a ***mailer.php*** pomocí crontabu. To je vše a aplikace bude správně fungovat. Pokud tomu tak není, jsou pravděpodobně špatně nastaveny cesty k jednotlivým souborům a adresářům. Podrobný popis instalace se nachází v souborech ***napoveda.html*** a ***help_eng.html***.

6.3 Testování

Účelem testovaní je odhalit co nejvíce možných chyb v aplikaci a postarat se o jejich opravu. Dá se očekávat, že zejména v prvních týdnech provozu aplikace se mohou nějaké objevit.

Tato aplikace byla testována různými uživateli, jak na místní stanici, tak na veřejně přístupném serveru. Pro testování byla použita skutečná data od provozovatele NetFlow. Testování bylo zaměřeno na kompletní uživatelskou funkčnost aplikace. Dále bylo testováno, zda aplikací generované reporty odpovídají očekávaným výsledkům. Testováním se podařilo odhalit několik drobných chyb aplikace. Všechny byly odstraněny.

7 Závěr

Na základě analýzy požadavků byla vytvořena aplikace, určená pro snadné monitorování počítačových sítí. Pro získávání dat o síťovém provozu byla použita technologie NetFlow, především program nfdump. Aplikace umožňuje generování různorodých statistik a jejich následné zobrazení na webovém rozhraní. Dále nabízí export zpracovaných dat pomocí souborů ve formátu PDF. Důležitou komunikační vlastností portálu je možnost exportovat data formou pravidelného zasílání emailů.

Aplikace je implementačně rozdělena na dvě části: webový portál a pravidelně spouštěné offline skripty. Celý systém je postaven na jazycích HTML a PHP. Uživatelské rozhraní aplikace je jednoduché, intuitivní, ale zcela splňující funkční požadavky. Při jeho implementaci bylo použito kaskádových stylů. Vykreslování grafů, což je nejdůležitější část uživatelského rozhraní, je implementováno pomocí technologie Flash. Aplikace byla testována různě zkušenými uživateli na skutečných datech. Systém se ukázal ve všech ohledech funkční a v praxi efektivně použitelný.

Jako každý systém je i tento možné dále zdokonalovat. První budoucí rozšíření aplikace by mělo být v oblasti bezpečnosti, především použitím prostředků protokolu HTTPS. Dále by bylo dobré více využít možností tvorby reportů v nfdumpu. Tato práce je zaměřena na tvorbu top reportů, ale možnosti programu nfdump jsou mnohem širší. Jako ve všech aplikacích spolupracujících s uživateli přes grafické rozhraní je možné, že v průběhu používání se objeví efektivnější možnosti ovládání, než které nabízí původně navržené rozhraní. Vylepšení uživatelského rozhraní je tedy další možný směr budoucího vývoje aplikace.

Literatura

[1] Cisco Systems Inc.: Simple Network Management Protocol (SNMP) [online]. 2009 [cit. 10.5.2009]. Dostupný z WWW: <http://www.cisco.com/en/US/docs/internetworking/technology/handbook/SNMP.html>

[2] Benjamin Poulain: Nfcapd [online]. 2006 [cit. 10.5.2009]. Dostupný z WWW: <http://www.linuxcertif.com/man/1/nfcapd/>

[3] Benjamin Poulain: Nfdump [online]. 2006 [cit. 10.5.2009]. Dostupný z WWW: <http://www.linuxcertif.com/man/1/nfdump/>

[4] Benjamin Poulain: Nfprofile [online]. 2006 [cit. 10.5.2009]. Dostupný z WWW: <http://www.linuxcertif.com/man/1/nfprofile/>

[5] WWW stránka projektu nfsen [online]. [cit. 10.5.2009]. Dostupný z WWW: <http://nfsen.sourceforge.net/>

[6] Benjamin Poulain: Nfreplay [online]. 2006 [cit. 10.5.2009]. Dostupný z WWW: <http://www.linuxcertif.com/man/1/nfreplay/>

[7] Ministersto informatiky ČR, Ministerstvo vnitra ČR: Vyhláška 485/2005 Sb. [online]. 2005 [cit. 15.5.2009]. Dostupný z WWW: <http://www.sagit.cz/pages/sbirkatxt.asp? zdroj=sb05485&cd=76&typ=r>

[8] InfoSoft Global Linited: FusionCharts [online]. 2008 [cit. 1.5.2009]. Dostupný z WWW: <http://www.fusioncharts.com/>

[9] WWW stránka projektu FPDF [online]. [cit. 3.5.2009]. Dostupný z WWW: <http://www.fpdf.org/>

[10] Codeworx Technologies: PHP mailer [online]. 2006 [cit. 1.5.2009]. Dostupný z WWW: <http://phpmailer.codeworxtech.com/>

[11] A Little Technology Shoppe LLC: Event Scheduler (cron) [online].21.1.2000 [cit. 1.5.2009]. Dostupný z WWW: <http://www.alts.net/servers/extensions/cron/>

[12] Portál AskApache.com:.htaccess Tutorial – The Ultimate Htacces Guide [online]. 10.1.2009 [cit. 10.5.2009]. Dostupný z WWW: <http://www.askapache.com/htaccess/apache-htaccess.html>

[13] Castagnetto J., Rawat H., Schumann S., Scollo Ch., Veliath D.: Programujeme PHP profesionálně. Brno, Computer Press, 2001, ISBN: 8072263102

[14] Croft J., Lloyd I., Rubin D.: Mistrovství v CSS. Brno, Computer Press, ISBN: 978-80-251-1705-7

[15] Pošmura V.: Apache, Příručka správce WWW serveru. Praha, Comtuter Press, 2002, ISBN: 80-7229-696-9

[16] Cisco Systems Inc.: Cisco IOS NetFlow [online]. 2009 [cit. 4.4.2009]. Dostupný z WWW: <http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html

[17] Portál Wikipedia: NetFlow [online] 26.3.2009 [cit. 4.4.2009]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/Netflow>

[18] Portál Wikipedia: PHP [online] 25.5.2009 [cit. 14.5.2009]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/Php>

[19] Portál Wikipedia: HTML [online] 27.4.2009 [cit. 14.5.2009]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/HTML>

[20] Portál Wikipedia: Cascading Style Sheets [online] 2.4.2009 [cit. 14.5.2009]. Dostupný z WWW: <http://cs.wikipedia.org/wiki/Cascading_Style_Sheets>

[21] Janovský D.: Jak psát web [online]. 2001. [cit. 3.3.2009]. Dostupný z WWW: <http://www.jakpsatweb.cz/>

[22] Zajíc P.: Seriál o PHP [online]. 2004 [cit. 2.3.2009]. Dostupný z WWW: <http://www.linuxsoft.cz/php>

[23] Portál php.net: PHP manual [online]. 2009 [cit. 2.3.2009]. Dstupný z WWW: <www.php.net>

Seznam příloh

Příloha č. 1. Struktura souborů a adresářů aplikace
Příloha č. 2. CD/DVD – zdrojové soubory včetně knihoven a testovacích dat jsou umístěny v adresáři www. Součásti toho adresáře jsou i soubory s nápovědou. V kořenovém adresáři disku je dále uložen PDF soubor s touto zprávou a soubor README.txt obsahující základní informace o aplikaci a o možnostech jejího použití a testování.

Příloha č. 1: Struktura souborů a adresářů aplikace

- index.php - hlavní soubor
- *.php - zdrojové PHP skripty
- *.js - zdrojové soubory JavaScriptu
- ./css - adresář obsahující soubory stylů
- ./lib - adresář obsahují soubory použitých knihoven
- ./FusionCharts - adresář se soubory knihovny FusionCharts
- ./images - adresář s použitými obrázky
- ./nfcapd - adresář s výstupy programu nfcapd, určený pro testování
- ./temp - adresář pro ukládání výstupů programu nfdump
- config.xml - konfigurační soubor
- auth.txt - autentizační soubor
- htaccess - implicitní konfigurační soubor pro Apache
- napoveda.html - soubor s českou nápovědou
- help_eng.htmp - soubor s anglickou nápovědou