VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

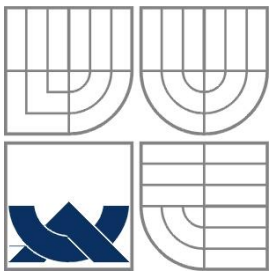
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

WEBOVÝ PORTÁL S REPORTY O SÍŤOVÉM PROVOZU

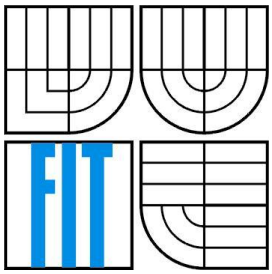
BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE PETR VÍTEK
AUTHOR

BRNO 2010



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 — PETR VÍTEK
AUTHOR

VEDOUCÍ PRÁCE — ING. JIŘÍ TOBOLA
SUPERVISOR

BRNO 2010

Abstrakt

Cílem této práce je navrhnout a realizovat systém pro přehledné zobrazení reportů o síťovém provozu. Je zde popsán proces návrhu a implementace systému. Práce také představuje dostupné technologie monitorování počítačových sítí, především technologii NetFlow.

Abstract

The aim of this work is to design and implement system for a simple presentation of reports about network traffic. There is described the design and implementation of the system. Work is also represents available technology of monitoring computer networks and especially technology NetFlow.

Klíčová slova

Webový portál, monitorování počítačových sítí, SNMP, NetFlow, NfDump, PHP, HTML, CSS, JavaScript, MySQL

Keywords

Web portal, monitoring computer networks, SNMP, NetFlow, NfDump, PHP, HTML, CSS, JavaScript, MySQL

Citace

Vítek Petr: Webový portál s reporty o síťovém provozu, bakalářská práce, Brno, FIT VUT v Brně, 2010

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 Vítek
15. 5. 2010

Poděkování

Velmi rád bych poděkovat svému vedoucímu Ing. Jiřímu Tobolovi, za jeho odbornou pomoc a konzultace, které mi poskytoval během tvorby této práce.

Obsah

1 Úvod

Informační technologie jsou jeden z nejrychleji se rozvíjejících odvětví. S každodenním rozvojem vzrůstají neustále nároky společností i jednotlivců na využití počítačových sítí. Už se nejedná o spojení několika počítačů vybraných univerzit, ale o celosvětové propojení. Společnosti i obyčejní lidé mohou komunikovat svými partnery, klienty, kamarády i příbuznými v reálném čase díky technologiím jako jsou sociální sítě, videokonference či IP telefonie.

S tímto celosvětovým propojením vznikla potřeba počítačové sítě nejen zabezpečit proti případným škůdcům ale také monitorovat aktuální dění v síti. Díky monitorování počítačových sítí mohou správci počítačových sítí získat potřebné informace pro rychlé a úspěšné vyřešení jakéhokoliv problému. Pomocí těchto informacím mohou také předcházet výpadkům, které pro mnoho firem znamenají velké finanční ztráty. Důsledná analýza těchto dat nám také usnadňuje budoucí rozvoj sítě a její infrastruktury.

V teoretické části se zpočátku věnuji oblasti monitoringu, jako takové. Je zde detailně představena architektura a možnosti využití protokolu NetFlow, z něhož se stal standard pro měření a monitorování sítí na základě IP toků.

Cílem této práce je vytvoření webového portálu pro generování a prezentaci reportů. Jeho analýza a specifikace požadavku je speciálně rozebrána ve třetí kapitole. Následující kapitola se zabývá už samotnou implementací aplikace. Zde se také dozvíte jaké technologie a knihovny byly využity.

2 Monitorování počítačových sítí

Monitorování sítě je důležité pro každou podnikovou sít jakékoliv velikosti. Systémy pro monitoring pomáhají nalézt a odstranit neduhy počítačových sítí, jako jsou ztráty emailů, podezřelé aktivity uživatelů nebo jiné problémy způsobeny pádem serveru, přetížením sítě nebo čímkoliv jiným.

Sledování sítě začíná u jejího jádra. Kontrolují se údaje o vytížení serverů, latenci a odezvy síťových zařízení a šířka přenosového pásma. Těmito údaji to nekončí, správné monitorovací nástroje doráží identifikovat, které IP adresy mají na svědomí největší traffic či neoprávněné využívání P2P sítí.

Existuje velké množství nástrojů pro monitoring. S nástroji můžeme komunikovat přes klasickou příkazovou řádku nebo využít celou řadu grafických řešení, která umožňují zobrazit detailní grafické reporty, generují grafy ale i exportovat informace do libovolného formátu.

V následujících kapitolách představím dvě základní technologie monitorování počítačových sítí: protokol SNMP a NetFlow.

2.1 SNMP

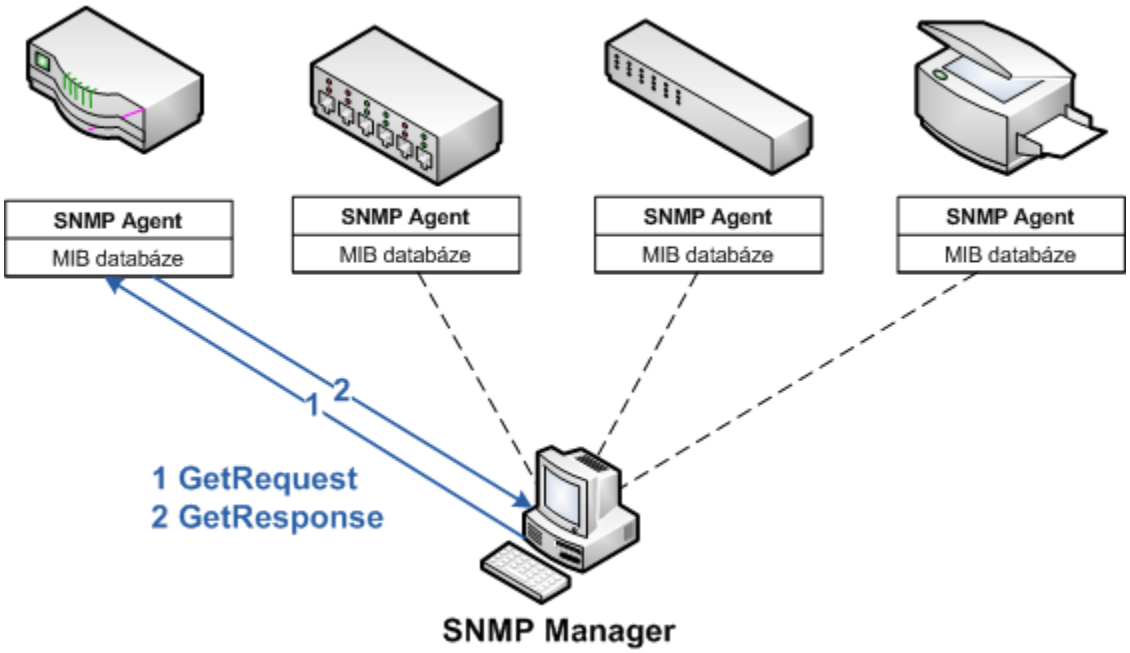
Protokol SNMP je jedním z nejrozšířenějších řídících protokolů sítí LAN a WAN. Tento protokol poskytuje prostředky ke správě a monitorování aktivních síťových prvků a jejich zařízení. Spravovaným zařízením může být obecně cokoliv, co umí protokol IP. Pomocí SNMP je možné spravovat například pracovní stanice, přepínače, směrovače, ale i tiskárny nebo záložní zdroje UPS.

2.1.1 Jak SNMP funguje

SNMP je založeno na architektuře klient-server, ve kterém jsou definovány tři základní prvky:

- SNMP Manager
- SNMP Agent
- MIB databáze

Manager a agent mohou běžet buď odděleně na různých fyzických strojích, nebo v rámci jednoho stroje. Na síťových prvcích (SNMP Agentech) je provozován software, který monitoruje stav zařízení a jejich údaje si ukládá do databáze MID. MID je stromově uspořádaná databáze, kde se k jednotlivým objektům přistupuje pomocí identifikátorů objektů. MIB obsahuje informace o objektech s přiřazeným jménem syntaxí, přístupem a stavem. SNMP Manager komunikuje s agenty v síti prostřednictvím jednoznačně definovaných zpráv. Tyto zprávy jsou přenášeny pakety protokolu UDP.



Obrázek 1 Ilustrace komunikace SNMP

2.1.2 Popis protokolu

V současné době existují tři verze protokolu SNMP:

- **SNMPv1** - První verze, největším problémem bylo slabé zabezpečení, hesla jsou uloženy a přenášeny v nezašifrované podobě.
- **SNMPv2** - Existuje několik modifikací, přidána podpora autentizace.
- **SNMPv3** - Umožňuje autentizaci a šifrování pomocí DES/AES.

Před samotnými administrativními informacemi obsahuje každý SNMP paket ještě tři údaje. Prvním je číslo verze, druhým údajem je název komunity. Komunita je textový řetězec sloužící jako primární metoda autentizace. Třetí údaj v hlavičce SNMP určuje typ odesílané operace a označuje se jako PDU.

Protokol SNMP specifikuje následujících 7 operací:

- **GetRequest** - Zjištění aktuální hodnoty proměnné.
- **GetNextRequest** - Vyžádání následující informace.
- **GetBulkRequest** - Umožňuje přenést více informací zároveň pro rychlejší komunikaci.
- **SetRequest** - Nastavení hodnoty proměnné.
- **GetResponse** - Odpověď zařízení.
- **Trap** - Informace zaslaná zařízením.
- **Inform** - Obdoba operace Trap, ale je vyžadována Odpověď.

2.1.3 Nástroje SNMP

Nagios je jedním z hlavních představitelů skupiny nástrojů pro monitoring sítě, Zaměřuje se na celkový přehled monitorované sítě a sledování všech síťových služeb a zařízení. Nagios je komplexní aktivní open source nástroj, určený pro sledování dosažitelnosti, parametrů síťových zařízení a služeb. Cílem je co nejdříve reportovat jakékoliv problémy v síti.
Mezi jeho hlavní vlastnosti a schopnosti patří:

- monitorování síťových služeb (SMTP, POP3, HTTP, NNTP, PING)
- monitorování zařízení na síti a jejich zdrojů (zatížení procesoru, využití disku a paměti)
- monitorování prostředí, kde je síť umístěna, např. teploty

2.2 NetFlow

NetFlow je otevřený protokol společnosti Cisco Systems. V současné době se jedná o nejrozšířenější průmyslový standard pro měření a monitorování počítačových sítí. Tento protokol představuje efektivní cestu pro získání komplexního přehledu aktuálního dění v síti.

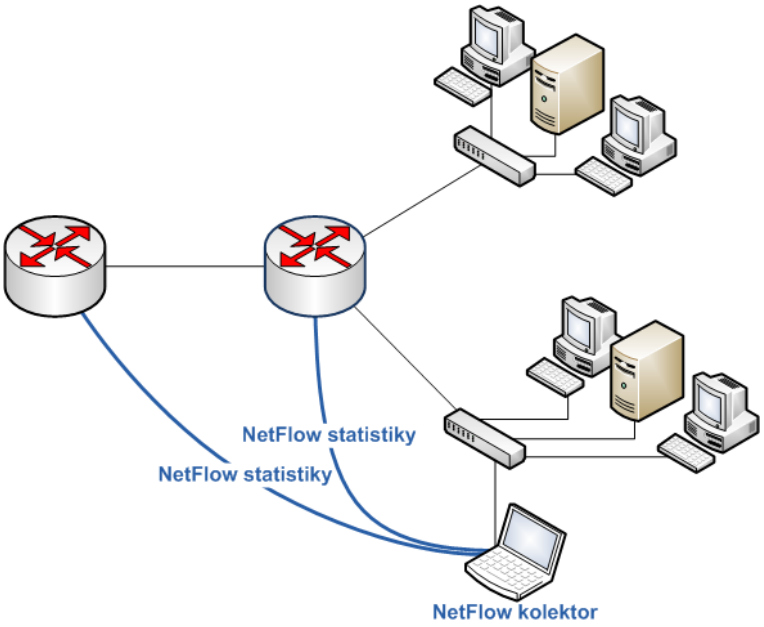
2.2.1 Jak NetFlow funguje

NetFlow je založen na principu toků síťového provozu. Tok je definován jako jedna úplná síťová konverzace. Každý tok je popsán unikátní skupinou následujících údajů:

- zdrojová IP adresa
- cílová IP adresa
- zdrojový port
- cílový port
- číslo IP protokolu
- typ služby
- vstupní rozhraní

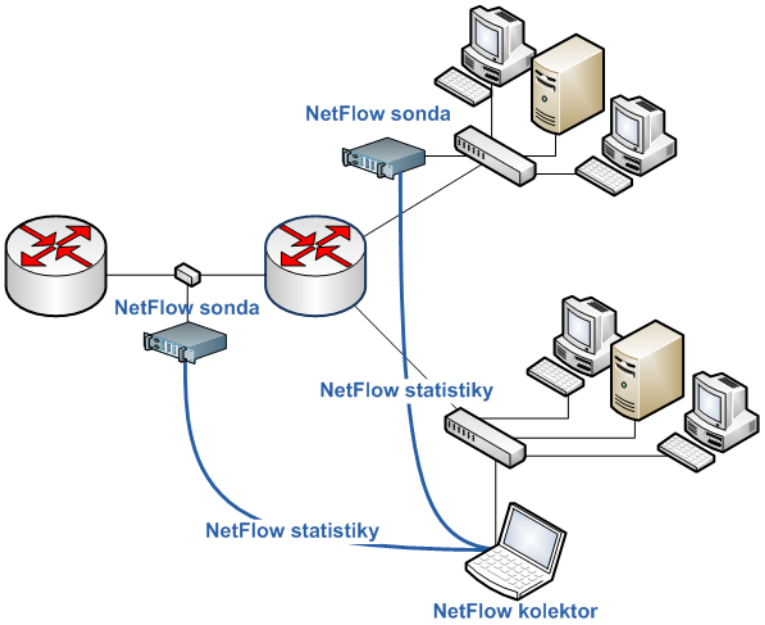
2.2.1.1 Architektura

Architektura NetFlow je tvořena NetFlow exportérem a NetFlow kolektorem. Exportér analyzuje procházející pakety, na jejímž základě generuje NetFlow statistiky a ty posílá kolektoru. NetFlow kolektor je zařízení, které sbírá statistiky z exportérů a ukládá je do souborů nebo databáze.



Obrázek 2 - Standardní architektura NetFlow

V tradiční architektuře je exportérem směrovač v síti, který provádí výpočet statistik. Toto uspořádání má neblahý vliv na výkon směrovače, proto většina směrovačů s podporou NetFlow využívá vzorkování, tzn., že se pro výpočet statistik používán každý n-tý packet.



Obrázek 3 - Moderní architektura NetFlow

V současné době se využívají pasivní NetFlow sondy, neboli specializovaná zařízení na monitorování a export NetFlow statistik. Na rozdíl od směrovače ji lze umístit do libovolného místa síti. Sondy procházející data pouze monitorují a nijak do nich nezasahují.

2.2.2 Popis protokolu

NetFlow protokol vznikl v několika verzích, první masově používanou se však stala až verze 5 (NetFlow v5), v současnosti se začíná hojně využívat i verze 9, na jehož základě vznikl IETF standard Internet Protocol Flow Information eXport (IPFIX).

Existující verze NetFlow:

- **Verze 1** - První uvedená verze.
- **Verze 2-4** Nebyli nikdy uvedeny.
- **Verze 5** - Nejrozšířenější verze, přidává podporu informací o autonomních systémech z protokolu BGP a zavádí sekvenční čísla, která slouží k detekci ztracených paketů.
- **Verze 6** - Podpora pro tunelový provoz.
- **Verze 7** - Informace z přepínačů.
- **Verze 8** - Zavádí agregaci získaných dat, díky níž se omezuje objem dat exportovaných směrovačem. Podle zvoleného agregačního schématu směrovač seskupuje toky stejných skupin do jediného agregovaného toku, jehož data se exportují jako celek.
- **Verze 9** - Flexibilní a rozšiřitelný formát, který umožňuje dále rozšiřovat a podporovat nové druhy záznamů. Je založen na šablonách, které specifikují formát přenášených dat a umožňují i nadále rozlišovat záznamy, aniž by bylo nutné vytvářet nový exportní formát. Plně podporuje IPv6 a také export multicast, Multiprotocol Label Switching, BGP nexthop a dalších informací.

NetFlow záznamy generované směrovači nebo sondami jsou exportovány na kolektor pomoci protokolu UDP nebo SCTP. Jakmile je NetFlow záznam exportován, je z důvodů větší efektivity exportérem zahozen. To má za následek ztrátu NetFlow záznamu v případě, že se paket vlivem nepříznivých okolností nepodaří doručit.

NetFlow záznam obsahuje důležité statistiky o síťovém provozu. V paketu NetFlow v5 jsou obsaženy následující informace:

- Číslo verze
- Sekvenční číslo
- SNMP index vstupního a výstupního rozhraní, který umožňuje sledovat vytížení jednotlivých síťových rozhraní
- Čas začátku a konce IP toku
- Počet bajtů a paketů v toku
- Údaje z L3 hlavičky:
    - Zdrojové a cílové IP adresy

    - Zdrojové a cílové porty
    - IP protokol
    - Type of Service (ToS)
- U TCP toků obsahuje množinu tvořenou sjednocením všem TCP flagů, které se v toku vyskytly.
- Směrovací informace:
    - IP adresa příštího hopu (důležité pro analýzu směrovacích postupů)
    - Maska cílové a zdrojové IP adresy (délky prefixů podle CIDR notace)

Některé exportéry také uvádějí hodnotu zdrojového a cílového autonomního systému (AS). Tato hodnota však nemusí být vždy přesná.

2.2.3 Možnosti využití NetFlow

Proč používat NetFlow technologii? Jaké možnosti využití poskytuje NetFlow? Zde je seznam základních oblastí, využití NetFlow.

2.2.3.1 Monitorování síťových aplikací a aktivity uživatelů

NetFlow data umožňují monitorovat a zobrazovat časové vytížení sítě, jednotlivé druhy a typy provozu v sítí. Použitím analýzy toků může být dosaženo vizualizace síťového provozu z jednotlivých síťových zařízení nebo z pohledu využití aplikacemi.

2.2.3.2 Dlouhodobější plánování architektury sítě

Pomocí NetFlow můžeme shromažďovat a analyzovat data o síťovém provozu z dlouhodobého hlediska. Na základě poskytnutých informací je možné předpovídat rostoucí nároky na kapacitu, propustnost a vlastnosti sítě, a podle těchto informací plánovat rozšiřování sítě o nová zařízeni nebo jejich inovaci. Takto je možné maximálně efektivně investovat dostupné zdroje do rozvoje a optimalizace sítě.

2.2.3.3 Bezpečnost sítě

Data získaná z NetFlow je možné téměř v reálném čase analyzovat a rozeznávat útoky na síť nebo odhalovat virová napadení jednotlivých počítačů v sítí. Jakékoliv výraznější změny v běžném chování sítě indikují nestandardní stav.

2.2.3.4 Vyúčtování provozu

NetFlow data jsou výborným prostředkem pro detailní vyúčtování služeb v síti. Je možné stanovit různé varianty placení služeb s ohledem na denní doby nebo typu stahovaných a přenesených dat.

2.2.3.5 Ukládání NetFlow dat a Data Mining

NetFlow data mohou být dlouhodoběji archivována a později použita pro různé typy analýzy. Je možné sledovat, které služby a aplikace byli používány v různých sektorech sítě nebo jednotlivými uživateli.

2.2.4 Zdroje NetFlow dat

Existuje několik způsobů jak generovat data, buď pomocí hardwarových zařízení, která automaticky generují tato data, nebo pomocí softwarových nástrojů generující NetFlow data z běžných počítačů připojených do sítě.

nProbe je softwarové řešení NetFlow sondy, podporuje NetFlow v5/v9 ale i IPFIX. nProbe je dostupný ve dvou verzích, standardní a profesionální. Standardní verze je určena pro všechny operační systémy. Profesionální verze pouze pro Linux a kromě pokročilejších funkcí nabízí možnost kompilace na embedded zařízeních.

FlowMon sonda je autonomní NetFlow sonda, která monitoruje provoz na počítačové síti a vytváří statistiky o tomto provozu ve formátech NetFlow v5/v9 či IPFIX, obsahuje vestavěný kolektor pro okamžité uložení a analýzu dat.

2.2.5 NetFlow kolektory

Pro zpracování NetFlow dat existuje velké množství nástrojů. Některé jsou pouze na obecné monitorování provozu, jiné se úzce zaměřují na bezpečnost, účtovatelnost či jinou oblast.

NFDUMP tools je sada nástrojů určených pro Unixové systémy. Všechny nástroje podporují NetFlow verze 5,7 a 9. Mezi hlavní nástroje patří:

- **nfcapd** - Démon, který pracuje jako kolektor, čte NetFlow data ze sítě a ukládá je do souborů. Pro každý tok NetFlow dat je nutné spustit jeden nfcapd proces.
- **nfdump** - Nástroj určený pro zobrazení dat uložených pomocí démona nfcapd. Umožňuje zobrazovaná data filtrovat, agregovat, dokáže generovat top N statistiky.
- **nfprofile** - Filtruje data dle zadaných kritérií a následně je ukládá do souboru pro pozdější použití.
- **nfreplay** - Umožňuje přeposílat data uložena nástrojem nfcapd jinému kolektoru.
- **nfclean.pl** - Vzorový skript pro mazání starých dat.

Práci s nástroji NfDump věnuji dále v bakalářské práci vlastní kapitolu.

NfSen je grafická webová nadstavba nad NfDump tools umožňující:

- Zobrazení NetFlow dat s využitím RRD (Round Robin Database).
- Snadnou navigaci mezi NetFlow daty.
- Zpracování dat ve zvoleném časovém intervalu.
- Vytváření upozornění dle definovaných pravidel.
- Rozšíření pomocí pluginů.

nTop je jednoduchý měřící a monitorující nástroj, který podporuje různé řídící činnosti, včetně optimalizace, plánování a detekce bezpečnosti počítačové sítě. nTop je postavený nad knihovnou libpcap a je dostupný jak pro Unixové systémy a tak i pro Windows.

3 Analýza a specifikace požadavků

Praktickou částí této práce je vytvoření webového portálu, který by uživatelům poskytl náhled na využití jejich počítačové sítě.

Obliba webových aplikací spočívá především v jejich multiplatformnosti, tzn., že jsou nezávislé na operačním systému. K jejich používání stačí pouze internetový prohlížeč, který je součástí takřka všech instalací operačního systému. Mezi další jejich nesporné výhody patří snadná instalace a následná údržba.

Jádrem systému bude technologie NetFlow a nástroj NfDump. Aplikace by měla komunikovat s nástrojem NfDump a následně zpracovávat jeho výstup do přehledných tabulek a grafů. Systém by měl být využitelný pro široké spektrum uživatelů bez ohledu na jejich znalost počítačových sítí.

V následujících částech detailně rozeberu nejdůležitější funkce a požadavky, které by měla splňovat hotová aplikace.

3.1 Uživatelské účty a profily

Aplikace bude podporovat různé typy uživatelských účtů. Uživatelský účet si může zvolit vlastní nastavení, bude si moci nastavit pravidelné odesílání reportů na zadaný email.

Administrátorský účet má stejné možnosti jako běžný uživatel, ale navíc bude mít možnost spravovat ostatní uživatelské účty, přidělovat jednotlivým uživatelům reporty o síťovém provozu a také přidělovat rozsah IP adres monitorované sítě.

3.2 Uživatelské rozhraní

Uživatelské rozhraní je jediným prostředkem pro komunikaci mezi uživatelem a aplikací. Portál by proto měl mít jednoduché, přehledné a intuitivní grafické rozhraní. Hlavní činností portálu je prezentovat data ve formě grafů a tabulek, není zde potřeba implementovat velké množství grafických efektů.

Jednotlivá rozhraní typů uživatelských účtů se budou lišit pouze rozdílnými nabídkami menu a nastavěním aplikace. Administrátorské účty budou disponovat rozsáhlejšími možnostmi, správou uživatelů a reportů.

3.3 Statistiky a reporty

Nejdůležitější částí aplikace je zobrazení reportů. Reporty jsou zpracovaná data o síťovém provozu, které by měli i méně zkušeným uživatelům umožnit získat potřebné informace.

Reporty budou tvořeny tabulkou se seznamem vybraných informací a grafem pro zobrazení procentuálního využití.

Portál musí obsahovat velké množství předdefinovaných reportů, od základních: „Seznam nejnavštěvovanějších internetových serverů?“ po konkrétní specifické zaměření „Který počítač v síti přenesl nejvíce dat?“.

Reporty by také měli zobrazovat popis, obsahující základní informace o jejich využitelnosti.

3.4 Export dat

Data získána z nástroje NfDump by měla být dostupná i ve formátech dostupných pro zpracování a analýzu v jiných softwarových nástrojích. Z tohoto důvody byly vybrány formáty XML a CSV, které jsou vhodné pro výměnu tabulkových dat.

3.5 Bezpečnost

V poslední době jsou webové aplikace často zmiňovány v souvislosti s bezpečnostními riziky. Bezpečnost systému je nutné brát vážně. Pro přístup do webového portálu bude vyžadována autentizace.

Uživatelské účty mohou vytvářet pouze administrátoři, čímž se zabrání neoprávněným a zbytečným registracím.

3.6 NfDump

Samotný nástroj NfDump není vhodný pro přímé volání a zobrazení dat ve webových aplikacích. Zpracování velkého množství NetFlow dat vyžaduje dostatek času. Z tohoto důvodu je důležité vhodně navrhnout mezivrstvu, jež bude ze strany webového portálu plně zautomatizována komunikace s aplikací NfDump.

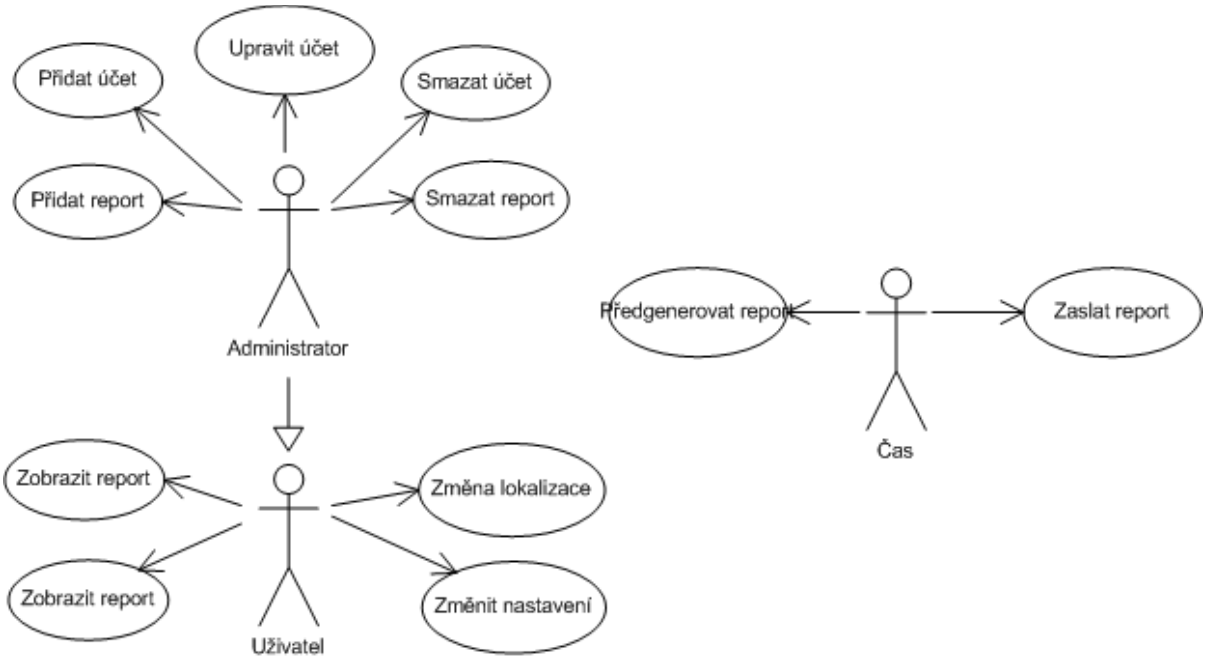
Pro přímé získávání informací z nástroje NfDump existuje nenahraditelný nástroj NfSen, ale i přes to by měli mýt administrátoři alespoň základní rozhraní pro jednorázové nadefinování filtrů a statistik.

3.7 Diagram užití

Diagram případu užití neboli „Use Case Diagram“ zobrazuje chování systému a jeho jednotlivých částí z pohledu uživatelských účtů. Zachycuje vnější pohled na modelovaný systém, tím nám pomáhá odhalit hranice systému a slouží jako podklad pro odhad rozsahu.

V aplikaci jsou tři role:

- **Administrátor** – možnost správy systému
- **Uživatel** – prohlížení reportů
- **Čas** – předgenerování reportů



Obrázek 4 - Diagram případů užití

4 Návrh a implementace systému

V této kapitole je popsán návrh a implementace systému. Pro vývoj systému byl použit skriptovací jazyk PHP a databázový systém MySQL. Detailnější informace o všech použitých technologiích jsou uvedeny na konci kapitoly.

Z důvodu velkého rozsahu aplikace jsou zde představeny pouze nejdůležitější a nejzajímavější části.

4.1 Struktura aplikace

V předcházející kapitole jsme si detailněji ujasnili požadavky a specifikaci systému. Poté přichází na řadu návrh systému a výběr dostupných technologií. Systém je potřeba vhodně navrhnout, tak aby v průběhu implementace nebylo nutné editovat spoustu věci.

Už při prvotních pokusech o získávání statistik z nástroje NfDump vznikl problém načítání dat. Zpracování delších časových úseků trvalo neúměrně dlouho a potencionální uživatele systému by tyto časové prodlevy obtěžovali. Z tohoto důvodu bylo nutné navrhnout řešení optimalizované pro rychlost.

Jako ideální řešení se jevilo rozdělit aplikaci do dvou nezávislých částí. První část se bude věnovat předzpracování statistik pro vybrané časové intervaly a uložením získaných dat do databáze. Druhá část už zajišťuje samotné zobrazení reportů.

Pro ukládání dat z nástroje NfDump do databáze jsem zvolil multiplatformní formát XML, který lze pomocí šablony XSL a dotazovacího jazyka XPath přetransformovat téměř do libovolné podoby. Tato technologie nám výrazně zjednoduší následné zpracování a export dat.

Databáze je využívána pro ukládání zpracovaných dat reportů a informace o uživatelích. V databázi si také uchovávám aktuální seznam používaných reportů a časové intervaly, pro které se mají příslušné statistiky generovat.

4.1.1 Architektura MVC

Při implementaci jsem rozhodl použít třívrstvou architekturu MVC. Tato architektura rozděluje datový model aplikace, uživatelské rozhraní a řídicí logiku do tří nezávislých vrstev tak, že modifikace některé z nich má minimální vliv na ostatní. Použitím těchto komponent se ze systému stane lehce udržitelná a rozšiřitelná aplikace.

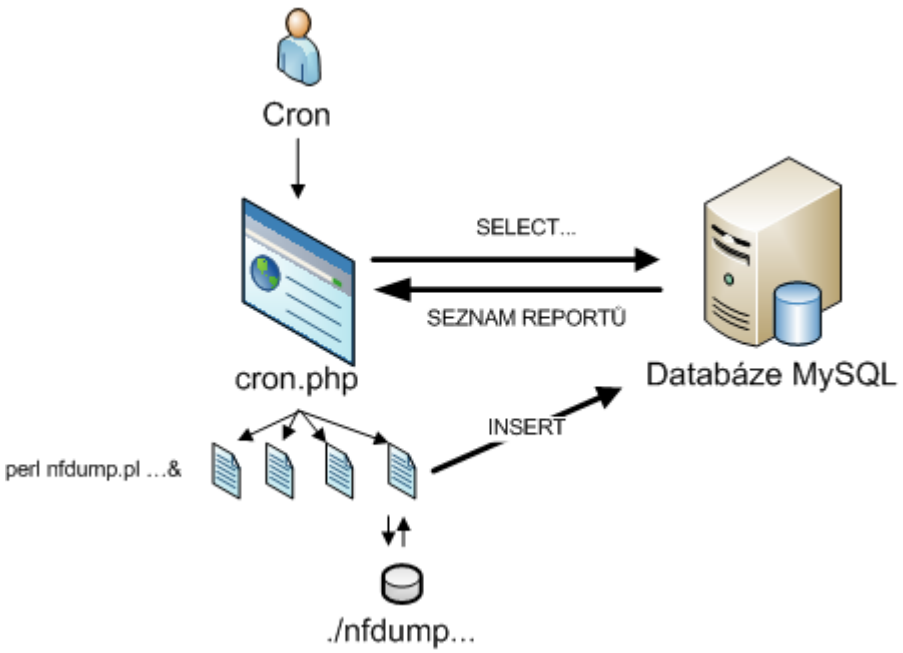
Vrstvy architektury MVC:

- *Model / Datová vrstva* - Model zajišťuje přístup k datům a manipulaci s nimi.
- *View / Prezentační vrstva* - Převádí data reprezentovaná modelem do podoby vhodné k prezentaci uživateli.
- *Controller / Aplikační vrstva* - reaguje na události pocházející od uživatele a zajišťuje změny v modelu

4.1.2 Data Mining

Předzpracování statistik probíhá v několika etapách. První etapě se získá seznam reportů, pro které se budou generovat statistiky. Pro každý předzpracovávaný report se sestavuje filtr určený nástroji NfDump. Jakmile je celý filtr sestaven, na pozadí se spustí skript napsaný v jazyce Perl. Úkolem tohoto skriptu je spustit nástroj NfDump, zpracovat jeho výsledky, které ihned uloží do databáze ve formátu XML.

Celý tento proces je vykonáván na základě automatizované činnosti, kterou si můžeme nadefinovat například v nástroji crontab. Průběh předzpracování reportů ilustruje následující obrázek, na němž je graficky znázorněn celý postup.



Obrázek 5 - Ilustrace načtení a zpracování dat z nástroje NfDump

4.1.2.1 NfDump a ukázka formátu XML

Program NfDump umožňuje vytvářet velké množství statistik. Pro správné použití je třeba sestavit sadu několika parametrů.

Prvním nezbytný parametr **–M** nebo **–R** určuje cestu k datům uložených programem nfcapd. Ihned za ním následuje parametr určující typ statistiky. Pomocí dalších parametrů si můžeme určit počet záznamů, jejich seřazení či filtrovaní podle zadaných kritérií.

Důležitým parametrem je také výstupní formát. Já jsem využil výstupní formát CSV, který poskytuje všechny dostupné informace.

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

```
nfdump -R nfdata/2010/05/16/20 -s port/flows -o csv
```

Tento příkaz vytvoří statistiku 10 nejpoužívanějších internetových portů mezi 20 a 21 hodinou dne 16. 5. 2010, seřazených podle množství toků.

Výsledek příkazu nástroje NfDump je ukládán v tomto typu formátu XML. Elementy se automaticky mění v závislosti na konkrétním výstupu dat.

```
<?xml version='1.0' standalone='yes'?>
<data>
…
  <item>
    <name>row-0</name>
    <bpp>98</bpp>
    <fl>2872</fl>
    <flP>17.7</flP>
    <ibyt>285133</ibyt>
    <ibytP>0.1</ibytP>
    <ipkt>2895</ipkt>
    <ipktP>0.7</ipktP>
    <pbs>692</pbs>
    <pps>0</pps>
    <pr>any</pr>
    <td>3296.268</td>
    <te>2010-05-17 18:59:35</te>
    <ts>2010-05-17 18:04:39</ts>
    <val>53</val>
  </item>
  <item>
…
</data>
```

Takto uložená data jsou připravena pro použití ve webovém portálu.

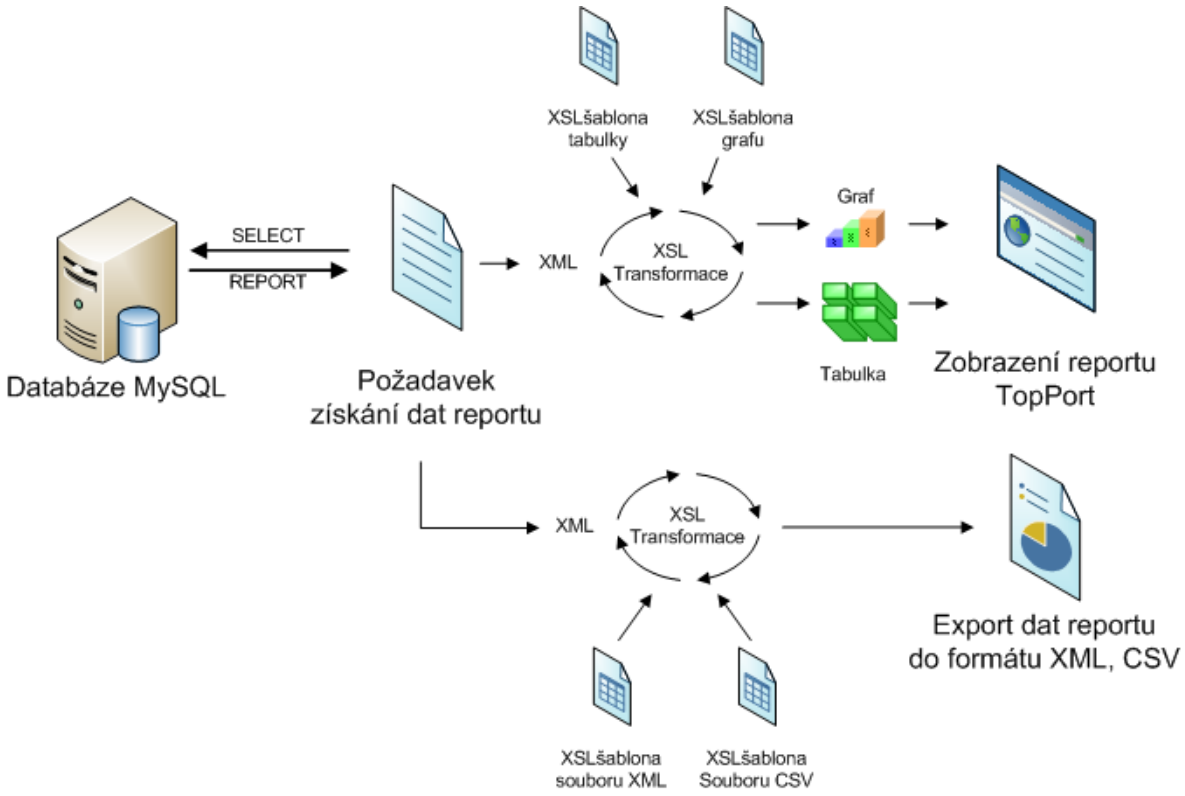
4.1.3 Princip zobrazení reportů

Máme-li data připravena v databázi, můžeme přejít k samotnému procesu tvorby reportů.
Přijde-li požadavek na zobrazení reportu učený buď jednoznačným identifikátorem dat, nebo sérií dvou informací (report a „časové razítko“) proběhne výběr XML dat z databáze.

V dalším kroku se data doplní o lokalizované hlavičky. Teď už proběhnou dvě transformace. První transformace z XML vytvoří kód tabulky HTML, druhá slouží pro úpravu formátu XML nutnou k prezentaci grafů. Na stejném principu funguje i převod do exportních formátů.

Výhodu navrženého řešení spatřuji především v jednoduché implementaci. Není nutné si uchovávat velké množství dat, a složitě zpracovávat data do použitelného formátu.Toto řešení je vhodné i pro pozdější rozšiřování statistik. Tím že není pevně definována struktura formátu XML, lze pouhou úpravou šablony XSL doplnit do potřebného formátu.

Princip zobrazení reportů je poměrně komplikovaný, proto je pro přehlednost na následující ilustraci pouze nastíněn základní princip zobrazení.



Obrázek 6 - Ilustrace zobrazení reportů

4.1.3.1 XSLT transformace

Smyslem XSLT je na základě zdrojového souboru a šablony vygenerovat jiný, třetí dokument nebo obecně soubor. Struktura tohoto výstupu XSLT není definována přímo standardem a je závislá na procesoru XSLT. Nejčastěji se používá výstup do HTML nebo XML, případně prostý textový formát.

Toto řešení se ukázalo jako jednoduché a velice efektivní pro zobrazení a úpravu dat. XSLT transformaci jsem v aplikaci použil hned v několika místech např.: pro zobrazení tabulky dat, přípravu dat určených pro načtení a zobrazení grafů ale také pro export dat v daném formátu.

Ukázka transformační funkce:

```
private function xslTransform($xmlData, $xslTemplate) {
        $xml = new DOMDocument();
        $xml->loadXML($xmlData);

        $xsl = new DOMDocument();
        $xsl->loadXML($xslTemplate, LIBXML_NOCDATA);

        $xslt = new XSLTProcessor();
        $xslt->importStylesheet($xsl);
        return $xslt->transformToXML($xml);
}
```

4.2 Databáze

V databázi jsou uchovávány veškeré informace, jak pro správu pluginů tak i uživatelských účtů. Vytvořená struktura databáze umožňuje dlouhodobé uchování dat reportů ale i rychlou navigaci v záznamech.

Pro práci s databází byla navržena struktura, které pohodlně umožní kdykoliv vyměnit databázový systém MySQL za jakýkoliv jiný. Pro použitý databázový systém je potřeba vytvořit třídu, která bude implementovat rozhraní *IDatabase (./library/IDatabase.php)*. Toto rozhraní obsahuje následující metody, které jsou nezbytné pro práci se systémem.

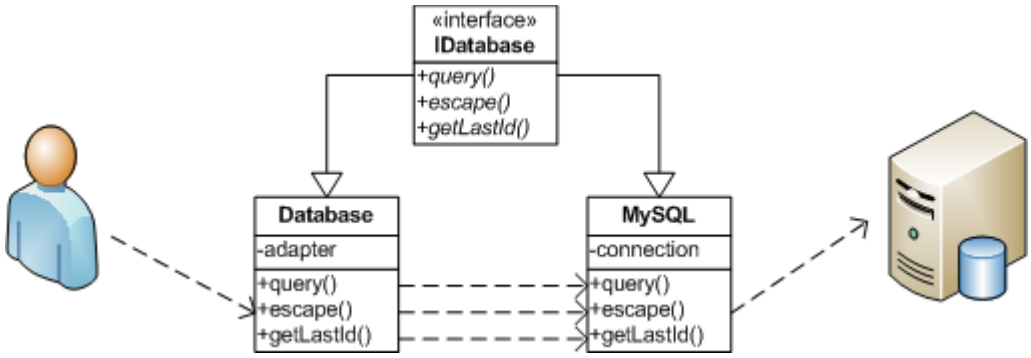
- **`public function query($sql);`**
  Nejdůležitější metoda, jejíž implementace realizuje SQL operace.

- **`public function escape($value);`**
  Escapování řetězců, používané před příkazem INSERT

- **`public function getLastId();`**
  Získání identifikátoru posledního vloženého záznamu.

V souboru *./application/loader.php* je vytvořena jediná instance objektu *Database*, kterému je v konstruktoru předáván parametr název třídy jenž bude obsluhovat databázový systém.

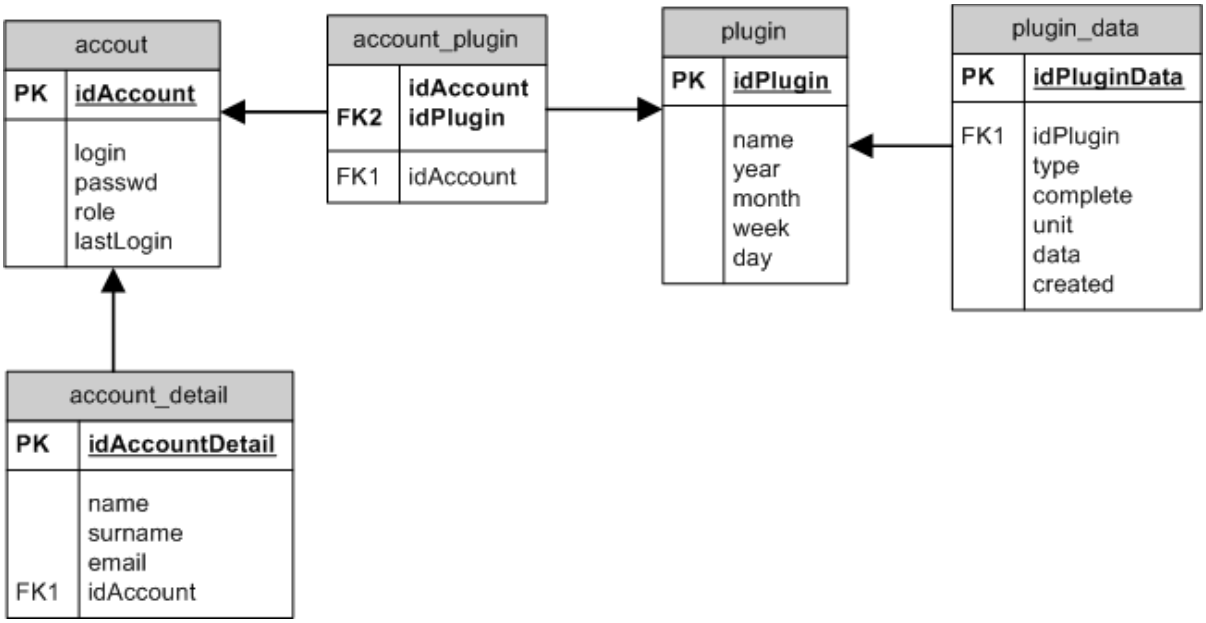
Obdobný princip je používán také pro emailové služby a načítání konfiguračního souboru.



Obrázek 7 - Delegování činnosti

4.2.1 ER diagram

ER diagram obsahuje seznam entit a vztahy mezi nimi.



Obrázek 8 - Schéma databáze

Popis entit:

account - základní informace o uživatelských účtech

account_detail - rozšiřující informace o účtech, obsahující „osobní“ informace

account_plugin - seznam pluginů přiřazených uživateli

plugin - aktuální seznam, pluginů

plugin_data - předgenerovaný seznam reportů

4.3 Plugin

Plugin je v tomto systému chápán jako předpřipravený report. Jedná se o samostatné třídy, ze kterých se sestavují nastavení filtru NfDump. K pluginu lze také přidat samotnou funkcionalitu, která bude vykonána před zobrazením reportu.

Struktura pluginu je navržena tak, aby bylo možné webový portál postupně doplňovat o všechny potřebné reporty.

4.3.1 Popis a ukázka kódu pluginu TopPort

Plugin *TopPort* zobrazuje seznam nejpoužívanějších internetových portů. Pomocí těchto informací můžeme zjistit seznam nejvyužívanější sítové služby a nástroje.

Seznam jednotlivých proměnných:

Popis pluginu:

- **name** - Pole názvů jednotlivých lokalizací.
- **description** - Pole popisů jednotlivých lokalizací.
- **type** - Typ reportu.

Sestaveni příkazu NfDump:

- **function** - Funkce NfDump, např.: statistiky (-s), agregace (-a).
- **filter** - Definice filtru NfDump a řazení výsledku.
- **other** – Doplňkové doplnění filtru NfDump.

Zobrazení reportu:

- **visible** – Seznam zobrazitelných hlaviček a dat v tabulce reportu.
- **header** – Lokalizace hlaviček souboru do potřebných jazyků.
- **special** – Pole určené pro přídavnou funkcionalitu datových sloupců.

Proměnná **special** je v pluginu *TopPort* využita pro náhradu čísel známých portů za sítové služby. Tato přídavná funkcionalita umožňuje tvořit reporty pro široké spektrum uživatelů i bez znalosti počítačových sítí. Nemusíme tedy zasahovat do kódu aplikace.

Výhody tohoto řešení jsem nalezl v pohodlném a konfigurovatelném vytváření pluginů. Přidaná funkcionalita se projeví ve všech datech reportu uložených v databázi. Je proto možné už vytvořené reporty nadále upravovat, nebo rozšiřovat o další položky.

Ukázka implementace pluginu TopPort.

```
class TopPort {

    public $name = array(
            'cz' => 'Nejpoužívanější porty',
            'en' => 'Nejpoužívanější porty'
        );

    public $description = array(
            'cz' => 'Seznam nejpoužívanějších internetových portů.'
            'en' => 'Seznam nejpoužívanějších internetových portů.'
        );

    public $type;

    public $function = '-s port';

    public $filter = '';

    public $other = '';

    public $visible = 'val,fl,ipkt';

    public $header = array(
        'cz' => array (
            'val' => ' - Služba (Port) - ',
            'fl' =>  ' - IP Toky -',
            'ipktP' => '- ipkP - ',
        ),
        'en' => array (
            'val' => ' - Service (Port) - ',
            'fl' =>  ' - Flows -',
            'ipktP' => '- ipkP - ',
        ),
    );

    public $special = array(
            'val' => 'servicePort',
    );

    public function servicePort($value) {
        $listPort = array(
            '22' => 'SSH',
            '25' => 'SMTP',
            '53' => 'DNS',
            '80' => 'HTTP',
        );
        return (isset($listPort[$value]) ? $listPort[$value].' ('.$value.')' : $value);
    }
}
```

4.4 Konfigurační soubor config.ini

Soubor config.ini je obsahuje všechny nezbytné údaje nutné pro běh aplikace: přístup k databázi, SMTP serveru, nastavení cest k NetFlow datům, a základní nastavení SEO parametrů webového portálu.

Pro konfigurační soubor byl zvolen formát „ini“ pro svoji jednoduchost a přehlednost. Tento formát má výbornou podporu v jazyku PHP, rychlost načítání souborů je tak rychlá že se nevyplatí načtené informace cachovat, ale znovu načíst.

4.5 Jazyková lokalizace

Pro správu jazykových překladů existuje velké množství existujících produktů, jako například Gettext. Tato knihovna je bezesporu velmi kvalitní. Já jsem však zvolil vlastní řešení. Důvodů je několik – nejzávažnější problém vidím v tom, že je pro úpravu nebo vložení nového překladu nutné pokaždé vytvořit jeho binární podobu.

Každá jazyková mutace obsahuje ve vlastním souboru asociativní pole, ve kterém klíče definují použitý řetězec, a hodnota obsahuje jazykový překlad. Jazykové lokalizace jsou uloženy v adresáři *./locale*

Ukázka souborů z adresáře ./locale:

Soubor: system.cz.php

```
$lang = array(
// Uživatel
'user_title_add' => 'Přidat uživatele',
'user_title_edit' => 'Upravit uživatele',
'user_title_delete' => 'Smazat uživatele',
'user_title_list' => 'Seznam uživatelů',
...
```

Soubor: system.en.php

```
$lang = array(
      // User
'user_title_add' => 'Add user',
'user_title_edit' => 'Edit user',
'user_title_delete' => 'Delete user',
'user_title_list' => User list',
...
```

4.6 Použité technologie

4.6.1 HTML

Jazyk HTML (HyperText Markup Language) je značkovací jazyk, který od své druhé verze patří do rodiny jazyku SGML (Standard Generalized Markup Language).

HTML je jedním z jazyků pro vytváření webových stránek, byl navržen společně s protokolem HTTP v roce 1990. Jeho vývoj měl být původně ukončen 4 verzí a poté přejít na XHTML (následník jazyka HTML využívající univerzální jazyk XML). Toto se však nezamlouvalo některým společnostem a proto byla vytvořena skupina, jejímž cílem bylo vytvořit novou verzi HTML, která se začala označovat jako „HTML 5“.

4.6.2 CSS

Kaskádové styly neboli CSS (Cascading Style Sheets) byly navrženy pro tvorbu vzhledu jazyků HTML, XHTML a XML. Hlavním smyslem jeho vzniku bylo oddělit vzhled dokumentů od jeho struktury a obsahu. Přidáním kaskádových stylů lze definovat vzhled jedné stránky, ale i celé webové prezentace.

4.6.3 JavaScript

JavaScript je multiplatformní, objektově orientovaný skriptovací jazyk za jehož vývojem stojí společnost Netscape. Využívá se především jako interpretovaný programovací jazyk pro webové prezentace.

4.6.4 PHP

Jazyk PHP patří mezi skriptovací programovací jazyk, určený především pro implementaci dynamických internetových stránek. PHP skripty jsou interpretovány na straně serveru a k uživateli je přenášen až výsledek operace. PHP je platformě nezávislé a podporuje velké množství knihoven k rozličným účelům např. zpracování textu, grafiky, práci se soubory a podporu řady internetových protokolů.

4.6.5 MySQL

Relační databázový systém šířený pod licencí GLP. MySQL je optimalizována pro rychlost za cenu některých zjednodušeni. Velmi oblíbená a často nasazovaná je kombinace Linux, MySQL, PHP a Apache jako základní software webového serveru.

4.6.6 XML

Extensible Markup Language je obecný značkovací jazyk, který byl vyvinut a standardizován konsorciem W3C. Je zjednodušenou podobou staršího jazyka SGML. Umožňuje snadné vytváření konkrétních značkovacích jazyků pro různé účely a různé typy dat.

Jazyk je určen především pro výměnu dat mezi aplikacemi a pro publikování dokumentů, u kterých popisuje strukturu z hlediska věcného obsahu jednotlivých částí, nezabývá se vzhledem. Prezentace dokumentu (vzhled) může být definována pomocí kaskádových stylů. Další možností zpracování je transformace do jiného typu dokumentu, nebo do jiné aplikace XML.

4.7 Použité knihovny

4.7.1 PhpMailer

Jazyk PHP ve své implementaci umožňuje zasílat email vestavěnou funkci mail(), která bohužel poskytuje omezené prostředky. Proto bylo vhodné použít knihovnu optimalizovanou pro využití emailových funkcí, např. PhpMailer.

Objektově řešená knihovna PhpMailer nabízí velké možnosti nastavení, využití služeb SMTP serveru vyžadující autentizaci, podporu pop3 protokolu atd.

4.7.2 mPDF

Formát PDF slouží pro ukládání dokumentů nezávislé na softwaru i hardwaru, na kterém byly použity. Soubor typu PDF umožňuje obsahovat text i obrázky, přičemž tento formát zajišťuje, že se libovolný dokument na všech zařízeních zobrazí stejně. Z tohoto důvodu byl tento formát zařazen mezi výstupní exportující formáty.

Třída mPDF umožňuje snadný a efektivní způsob exportu HTML kódu do PDF formátu za pomocí jazyka skriptovacího PHP. mPDF mimo jiné zvládá kaskádové styly, číslování a změnu orientace stránky, obrázky tabulky atd…

4.7.3 ExtJs

ExtJs je javascriptová knihovna určena pro vytváření interaktivních webových aplikací s využitím technik, jako je Ajax, DHTML a DOM skriptování. Vyniká výbornou podporou téměř všech současných internetových prohlížečů a také přináší velké množství grafických komponent.

Při implementaci systému byli z toho frameworku využity komponenty pro práci s formuláři, zobrazování seznamů a také při použití technologie AJAX.

4.7.4 amCharts a Libchart

Pro jednoduché a zobrazení dat v podobě grafů je využívána flashová knihovna amCharts, která poskytuje velké množství různých typů grafů. Knihovnu je možné využít ve spoustě různorodých jazyků. Grafy lze prezentovat na základě XML a CSV dat.

Jelikož animované grafy nelze ukládat pro pozdější analýzu, a exportovat ostatních formátů byla použita knihovna Libchart, která ze zadaných dat dokáže vytvořit přehledné grafy.

5 Instalace a testování

Proces vývoje software samotnou implementací aplikace nekončí. Systém je potřeba důkladně otestovat na všech možných zařízení.

5.1 Doporučená konfigurace

Server

- Apache HTTP Server 2.0
- PHP verze 5.2 a vyšší
    - Grafická knihovna PHP
- MySql verze 5
- NfDump 1.6.1
- perl verze 5 a vyšší,
    - knihovna Simple XML
    - knihovna DBD-mysql

Klient

- Internetový prohlížeč s podporou JavaScriptu
- Adobe Flash Player 10
- PDF reader

5.2 Instalace

Máme-li k dispozici webový server s odpovídající konfigurací, můžeme zahájit instalaci webového portálu. Proces instalace je velmi jednoduchý, stačí nakopírovat jednotlivé soubory do adresáře serveru, importovat soubor *install.sql* do databázového serveru a poté upravit konfigurační soubory *./cron/nfdump.pl* a *./application/config.ini*, kde nastavíme cestu k adresáři obsahující NetFlow data, přístup k databázi. Posledním krokem je nastavení automatického spouštění skriptů *cron.php* pomocí crontabu.

Nyní se můžeme do aplikace přihlásit defaultně vytvořeným administrátorským účtem „admin“ s heslem „demo“. Zde už můžeme provádět veškeré administrátorské činnosti.

Podrobnější popis instalace se nachází v souboru INSTALL, který je součástí webového portálu.

5.3 Testování

Stejně tak důležité jako návrh a implementace aplikace je testování. Úkolem testování je odchytit co nejvíce možných chyb před jejím reálným nasazením.

Vývoj systému probíhal v několika iteracích. Během každé iterace byl systém otestován různými uživateli. V první fázi byli testy zaměřeni na samotnou funkcionalitu aplikace a v dalších etapách jsme se zaměřovali na správu portálu, efektivitu a intuitivnost grafického rozhraní.

Během testování bylo objeveno několik závažných chyb ovlivňující správné zobrazení dat, Před dokončením portálu byly všechny nalezené chyby opraveny.

6 Možná rozšíření

6.1 Reporty a statistiky

Nástroj NfDump nabízí velké množství nastavení filtrování dat, ze kterých lze získat nejrůznější statistiky. Portál by mohl být rozšířen o další kategorie reportů. Reporty by také mohli být archivovány na straně serveru pro pozdější analýzu.

V současné rozvržení, aplikace umožňuje zobrazit především objemy přenášených dat. Pro kompletní monitoring by bylo vhodné systém rozšířit a reporty umožňující detekci útoků, nebo reagovat zasíláním informací při definovaných událostech.

6.2 Uživatelské skupiny a role

Aplikace umožňuje administrátorovi přidat reporty jednotlivým uživatelům. Tato práce může být administrátorovi ulehčena vytvořením skupin. Které by měli přednastavený seznam vybraných reportů. Skupiny by mohly být rozděleny podle využití např.: správci sítě, bezpečnost sítí nebo manažeři. Tím by jednotliví uživatelé mohli efektivně využít svůj dostupný seznam reportů.

6.3 Uživatelské rozhraní

V této práci byl kladen velký důraz na jednoduché a efektivní využití uživatelského rozhraní. Systém byl optimalizován především pro uživatelé počítačů, ale stále více uživatelů se pohybuje na internetu pomocí mobilního telefonu či PDA. Aplikace by se mohla automaticky detekovat rozlišení přístroje a přizpůsobit jim šablonu vzhledu.

6.4 Optimalizace

Aplikace je navržena pro použití v malých až středně velkých společností. Pro velké množství uživatelů by systém mohl zaznamenat značné zpomalení. Abychom se této nepříjemnosti vyhnuli, bylo by potřeba již jednou načtené a zpracované reporty ukládat do cache. Kdyby poté přišly další požadavky na stejný report, data by byla zobrazena z cache.

Závěr

V této práci jsem navrhnul a implementoval webový portál s reporty o síťovém provozu. Během této realizace jsem prošel všemi etapami vývoje softwaru. Od prvotní analýzy požadavku až po samotné testování hotové aplikace.

Hotový systém je připravený pro potenciální reálné nasazení. Aby se mohl systém označit za plně funkční, bylo by ho potřeba ještě detailně monitorovat a opravovat případné neduhy, které nebyly objeveny při testování.

Při tvorbě této bakalářské práce jsem také zúročil znalosti nabyté během studia na této Fakultě Informačních technologií. V teoretické části především z předmětů zaměřených na oblasti počítačových sítí. U samotného vývoje jsem čerpal z předmětů: informační systémy, uživatelská rozhraní ale i databázové systémy.

Díky této práci jsem se detailně seznámil s možnostmi a technologiemi monitoringu počítačových sítí a to především s technologií NetFlow.

Literatura

[1] Dostálek Libor, Kretchma James: Administrace a diagnostika sítí pomocí OpenSource utilit a nástrojů. Computer Press, 2005, ISBN: 80-251-0345-5

[2] Behr Alyson, Khundhur Patrik: Abeceda monitoringu sítě. In: Business World, roč. 2008, č 9, Praha, CZ, s. 46-48, ISSN 1213-1709

[3] SNMP - Simple Network Management Protocol. [online]. [cit. 2010-11-15].
URL: < http://www.samuraj-cz.com/clanek/snmp-simple-network-management-protocol/ >

[4] Wikipedia: NetFlow. [online]. [cit. 2010-11-15].
URL: < http://cs.wikipedia.org/wiki/Netflow/ >

[5] Caligare: NetFlow. [online]. [cit. 2010-11-15].
URL: < http://netflow.caligare.com/ >

[6] Cisco IOS NetFlow. [online]. [cit. 2010-11-15].
URL: < www.cisco.com/web/go/netflow >

[7] INVEA-TECH: FlowMon. [online]. [cit. 2010-11-15].
URL: < http://www.invea.cz/produkty-sluzby/flowmon/ >

[8] NfDump. [online]. [cit. 2010-11-15].
URL: < http://nfdump.sourceforge.net/ >

[9] NfSen [online]. [cit. 2010-11-15].
URL: < http://nfsen.sourceforge.net/ >

[10] nTop [online]. [cit. 2010-11-15].
URL: < http://www.ntop.org/ >

[11] Nagios [online]. [cit. 2010-11-15].
URL: < http://cs.wikipedia.org/wiki/Nagios/ >

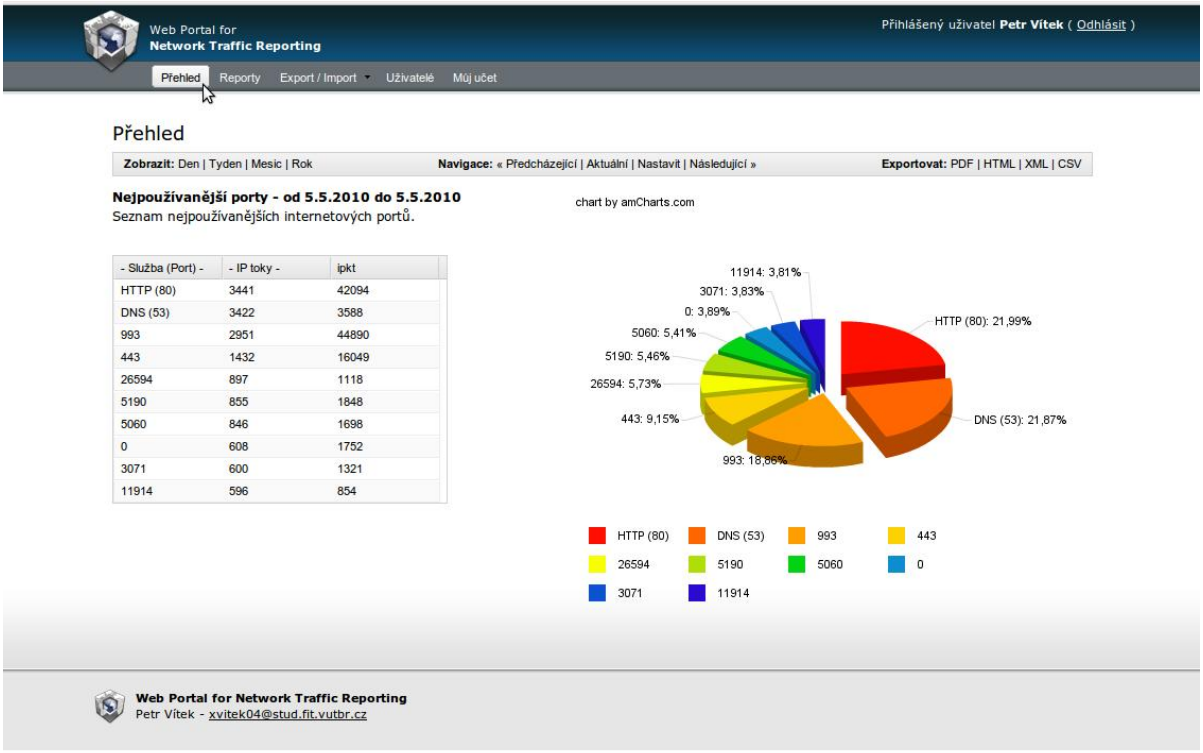
[12] Jazyk XSL [online]. [cit. 2010-11-16].
URL: < http://www.kosek.cz/clanky/swn-xml/xsl.html >

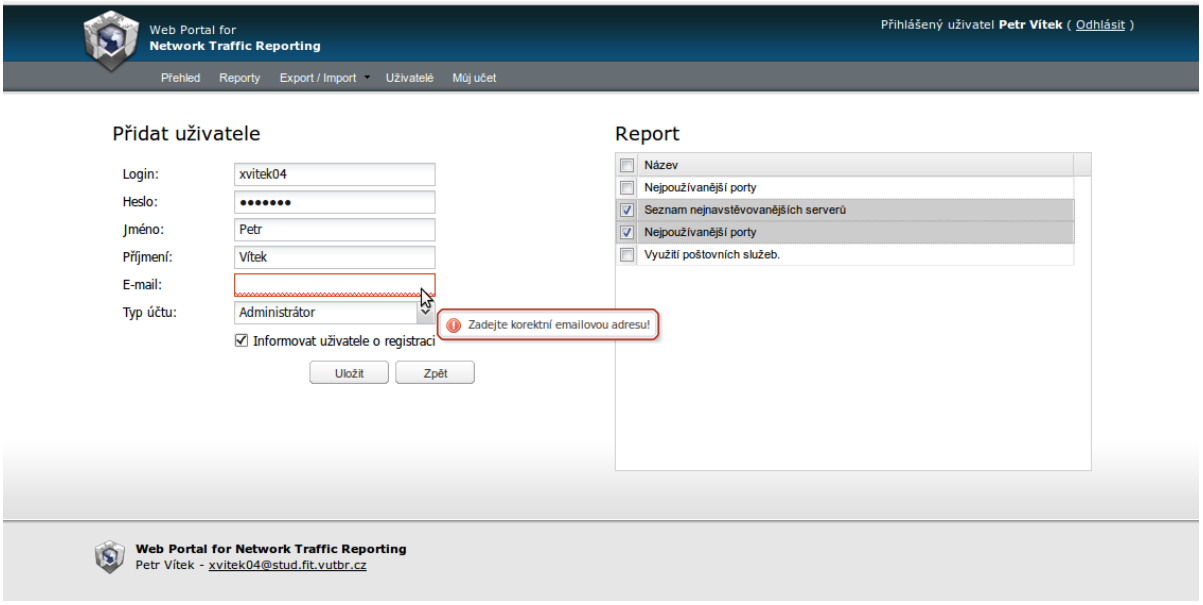
[13] Jazyk XML [online]. [cit. 2010-11-16].
URL: < http://www.kosek.cz/clanky/xml/index.html >

[14] PHP Manuál [online].
URL: < http://www.php.net/manual/en/ >

Seznam příloh

Příloha č. 1: Uživatelské rozhraní.





Příloha č. 2: Adresářová struktura

/application	zdrojové texty aplikace
/controller	aplikační logika
/model	datová logika
/template	prezentační logika
/cron	skripty automatizované činnosti
/document_root	jediný přístupny adresář z internetu
/css	kaskádové styly
/img	adresář s použitými kaskádovými styly
/js	javascriptové soubory
/swf	data ve formátu flash
/libraries	jádro aplikace a knihovny
/config	
/database	
/mail	
/locale	jazykové mutace
/plugin	dostupný seznam pluginů
/docs	programová dokumentace