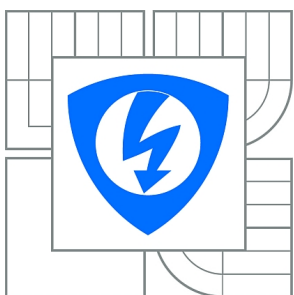


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

VERIFIKACE POZICE SERVERŮ SÍTĚ PLANETLAB

VERIFICATION OF PLANETLAB SERVERS LOCATION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. JÁN PRUŽINSKÝ

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. DAN KOMOSNÝ, Ph.D.

BRNO 2014

ABSTRAKT

Cieľom tejto práce je analyzovať uzly siete PlanetLab. Analýza spočíva predovšetkým v overovaní dostupnosti uzlov a ich fyzických polôh. Jednotlivé uzly sú testované na dostupnosť protokolu ICMP a protokolu SSH. Dostupnosť ICMP protokolu je overovaná pomocou programu ping. Hlavná časť práce je venovaná overovaniu adries uzlov. Zistené adresy sú porovnávané so zadanými adresami a výsledná zhoda je vyhodnocovaná na úroveň štátu, kraja, mesta a ulice. Na základe zadaných GPS súradníc uzlov sú počítané presnosti, s akými sú adresy uzlov zadané. Práca sa tiež zaoberá delením uzlov na základe vypočítaných indexov využiteľnosti a presnosti. V teoretickej časti sa nachádza popis experimentálnej siete PlanetLab a programátorského rozhrania pre Mapy Google.

KĽÚČOVÉ SLOVÁ

PlanetLab, Linux, zemepisná dĺžka, zemepisná šírka, SSH, ping, Geocoding API

ABSTRACT

The main objective of this thesis is to analyze the nodes of PlanetLab network. The analysis is focusing mainly on verifying the availability of nodes and on verifying the physical position of nodes. Individual nodes are tested for availability of ICMP protocol and SSH protocol. Availability of ICMP protocol is verified by using ping program. The main part of the thesis is devoted to verifying the addresses of nodes. Identified addresses are compared with entered addresses and the resulting conformity is evaluated at the level of the state, county, city and streets. Precision of specified address estimation is calculated based on given GPS coordinates. This thesis also deals with the dividing of nodes based on calculated usability index and accuracy index. The theoretical part contains a description of an experimental PlanetLab network and Google Geocoding API.

KEYWORDS

PlanetLab, Linux, latitude, longitude, SSH, ping, Geocoding API

PRUŽINSKÝ, Ján *Verifikace pozice serverů sítě PlanetLab*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2014. 47 s. Vedúci práce bol doc. Ing. Dan Komosný, Ph.D.

PREHLÁSENIE

Prehlasujem, že som svoju diplomovú prácu na tému „Verifikace pozice serverů sítě PlanetLab“ vypracoval samostatne pod vedením vedúceho diplomovej práce, využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej diplomovej práce ďalej prehlasujem, že v súvislosti s vytvorením tejto diplomovej práce som neporušil autorské práva tretích osôb, najmä som nezasiahol nedovoleným spôsobom do cudzích autorských práv osobnostných a/nebo majetkových a som si plne vedomý následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona č. 121/2000 Sb., o právu autorskom, o právach súvisejúcich s právom autorským a o zmeně niektorých zákonov (autorský zákon), vo znení neskorších predpisov, vrátane možných trestnoprávných dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákoníka č. 40/2009 Sb.

Brno

.....

(podpis autora)

POĎAKOVANIE

Ďakujem vedúcemu mojej práce doc. Ing. Danovi Komosnému, Ph.D. za užitočnú metodickú pomoc a cenné rady pri spracovávaní tejto práce.

Brno

.....

(podpis autora)

Výzkum popísaný v tejto diplomovej práci bol realizovaný v laboratóriách podporených z projektu SIX; registračné číslo CZ.1.05/2.1.00/03.0072, operačný program Výzkum a vývoj pro inovace.

OBSAH

Úvod	9
1 Experimentálna sieť PlanetLab	11
1.1 Obecné informácie o sieti PlanetLab	11
1.2 Technické riešenie siete PlanetLab	11
2 Určovanie polohy	14
2.1 Zemepisné súradnice	14
2.2 Vzdialenosť medzi zemepisnými súradnicami	15
3 Mapové služby spoločnosti Google	17
4 Príprava zdrojových dát	20
5 Programovanie skriptov	21
5.1 Overovanie dostupnosti uzlov	22
5.2 Overovanie zhody zadaných adries uzlov	25
5.3 Výpočet presnosti polohy	28
5.4 Určovanie indexu presnosti	29
5.5 Určovanie indexu využiteľnosti	30
5.6 Vytváranie výstupného dokumentu	31
6 Dosiahnuté výsledky	32
6.1 Dostupnosť uzlov	32
6.2 Úrovne zhody zadaných adries uzlov	35
6.3 Presnosť zadaných súradníc	37
6.4 Využiteľnosť uzlov	41
7 Záver	43
Literatúra	45
Zoznam symbolov, veličín a skratiek	47

ZOZNAM OBRÁZKOV

1.1	PlanetLab architektúra.	12
2.1	Rozdelenie povrchu Zeme podľa zemepisných súradníc [7].	14
2.2	Elipsoidný model Zeme.	15
3.1	Skrátený výpis formátovanej odpovede na geocoding.	18
3.2	Skrátený výpis odpovede na reverzný geocoding.	19
4.1	Časť súboru <i>web.csv</i>	20
5.1	Vývojový diagram nadväzností jednotlivých skriptov z tabuľky 5.1.	22
5.2	Vývojový diagram popisujúci princíp testovania dostupnosti uzlov.	23
5.3	Vývojový diagram hľadania výslednej úrovne zhody adresy.	28
5.4	Časť obsahu výsledného súboru <i>landmarks.csv</i>	31
6.1	Graf dostupnosti uzlov siete PlanetLab na program ping.	32
6.2	Graf dostupnosti uzlov siete PlanetLab na protokol SSH.	33
6.3	Graf dostupnosti uzlov na protokol SSH + program ping.	33
6.4	Závislosť počtu uzlov dostupných na protokol SSH a program ping v čase.	34
6.5	Histogram počtu uzlov pre každú úroveň zhody adresy.	36
6.6	Graf zastúpenia uzlov so správne zadanou adresou pre každú časť Európy.	36
6.7	Graf percentuálneho zastúpenia uzlov s indexom presnosti 1 a 0.	37
6.8	Percentuálne zastúpenie uzlov s indexom presnosti 1 pre jednotlivé časti Európy.	37
6.9	Graf distribučnej funkcie metódy založenej na porovnávaní GPS sú- radníc.	39
6.10	Graf distribučnej funkcie metódy založenej na porovnávaní medzi GPS súradnicami a údajmi zo statickej lokalizačnej databázy Ma- xMind.	39
6.11	Graf distribučnej funkcie finálnych presností zadaných polôh uzlov.	40
6.12	Graf percentuálneho zastúpenia využiteľných a nevyužitelných uzlov.	41
6.13	Percentuálne zastúpenie využiteľných uzlov pre jednotlivé časti Európy.	41

ZOZNAM TABULIEK

1.1	Prehľad vývoja siete PlanetLab.	13
2.1	Tabuľka vzdialenostných rozdielov medzi jednotkami stupňov zemepisných šírok a dĺžok pre vybrané hodnoty zemepisnej šírky.	16
5.1	Zoznam naprogramovaných skriptov spolu s ich vstupmi a výstupmi.	21
5.2	Vzorový záznam o uzle siete PlanetLab.	25
5.3	Presnosti polohy podľa počtu desatinných miest pre zemepisnú šírku 50°.	29
6.1	Tabuľka krajín spolu s percentuálnym zastúpením dostupných uzlov.	35
6.2	Tabuľka počtu uzlov s indexom presnosti 1 pre jednotlivé časti Európy.	38
6.3	Tabuľka minimálnych a maximálnych presností zadaných súradníc pre jednotlivé použité výpočetné metódy.	40
6.4	Tabuľka počtu uzlov s indexom využiteľnosti 1 pre jednotlivé časti Európy.	42
6.5	Tabuľka krajín spolu s percentuálnym zastúpením využiteľných uzlov.	42

ÚVOD

Už pred nejakým časom ľudia začali zisťovať, že základy, na ktorých stojí dnešný internet, nie sú úplne postačujúce. S nástupom nových a nových služieb sa zvyšujú nároky na prenosové kapacity, veľkosť odozvy v internete či na rýchlosť vyhľadávania zdrojov. Aby bolo možné vyvíjať nové technológie, je potrebné ich aj niekde testovať. Najlepšie v reálnom prostredí a v čo možno najväčšom merítku [8].

V roku 2002 Larry Peterson a David Culler zorganizovali stretnutie výskumníkov so záujmom o sieťové služby. Na tomto stretnutí predstavili plán o vytvorení testovacej siete o celkovom počte 100 počítačov. Od roku 2002 sa ich plán začal uskutočňovať a začala vznikať celosvetová experimentálna sieť PlanetLab [10].

Úlohou tejto práce bolo analyzovať uzly siete PlanetLab, ich údaje o polohe a dostupnosti na jednotlivé služby. V kapitole 1 je podrobnejšie popísaná sieť PlanetLab, jej architektúra a spôsob fungovania.

Predtým, než bolo možné začať testovať jednotlivé uzly, bolo potrebné vykonať prvotnú analýzu. Tá bola rozdelená medzi viacerých študentov. Spoločnou úlohou bolo vytvoriť zoznam uzlov siete PlanetLab a doplniť informácie o ich fyzickom umiestnení. Tejto úlohe je venovaná kapitola 4.

Ďalšia úloha už bola samostatná a pozostávala z viacerých častí. Z každej časti boli vyprodukované výstupné dáta, ktoré boli nakoniec spracované a pridané do výsledného dokumentu obsahujúceho zoznam všetkých uzlov. V kapitole 5 sú postupne popísané všetky samostatné úlohy spolu s popisom programovania potrebných linuxových skriptov.

Ako prvá časť 5.1 samostatnej úlohy bolo spracované overovanie dostupnosti uzlov. To bolo zamerané na program ping¹ a protokol SSH².

Ďalšou úlohou 5.2 bolo stanovenie úrovne zhody zadaných polôh jednotlivých uzlov. Tieto zhody boli vyhodnocované na úrovni ulice, mesta, kraja a štátu. Na vyhodnocovanie úrovni bola použitá služba Google Geocoding API. Popis tejto služby spolu s ukážkou jej fungovania je v kapitole 3.

¹Program ping umožňuje preveriť funkčnosť spojenia medzi dvoma sieťovými uzlami.

²SSH (Secure Shell) je zabezpečený komunikačný protokol v počítačových sieťach.

Kapitola 5.3 sa sústreďuje na výpočet presnosti zadanej polohy na základe počtu desatinných miest zadaných GPS súradníc. Teoretický úvod k tejto časti spolu s použitými vzorcami je v kapitole 2.

Posledné úlohy sú popísané v kapitolách 5.4 a 5.5. Sú zamerané na delenie uzlov podľa indexov presnosti a využiteľnosti.

Dosiahnuté výsledky pre jednotlivé úlohy sú spracované v kapitole 6. Kapitola je rozdelená podľa jednotlivých úloh na sekcie, pričom v každej z nich sú prezentované dosiahnuté výsledky.

1 EXPERIMENTÁLNA SIET' PLANETLAB

1.1 Obecné informácie o sieti PlanetLab

PlanetLab [10] je globálna výskumná sieť, ktorá podporuje vývoj nových sieťových služieb. Vznikla v priebehu roku 2002 ako konzorcium niekoľkých univerzít. Postupom času sa k nej pripojovali ďalšie a ďalšie univerzity a firmy zamerané na informačné technológie. Medzi najznámejších členov patria firmy ako Google, Intel, Hewlett Packard či AT&T. Jej členom sa môže stať každá akademická organizácia, ktorá do nej vloží svoj uzol [8].

Sieť PlanetLab si ako svoj cieľ vytýčila zmenu internetu. Vďaka tomu, že je to celosvetové distribuované laboratórium, disponuje dostatkom prostriedkov na testovanie a overovanie nových typov sieťových aplikácií. Jej prednosťou je hlavne to, že umožňuje užívateľom vytvárať nezávislé aplikácie, ktoré môžu bežať v celej sieti vedľa seba. Vrstvy používajú spoločné uzly bez toho, aby sa pri ich užívaní akokoľvek vzájomne ovplyvňovali.

Uzly siete PlanetLab sú rozdelené do troch podskupín, a to PlanetLab USA, PlanetLab Japan a PlanetLab Europe (PLE) [11]. Práve Európskej časti (PLE) je venovaná praktická časť tejto práce.

1.2 Technické riešenie siete PlanetLab

Na obrázku 1.1 je znázornená architektúra siete PlanetLab. Architektúra pozostáva z nasledujúcich hlavných častí:

- **Organizácia**

Organizácia predstavuje fyzickú polohu, kde sú umiestnené uzly, napr. organizácia CESNET. Skrátený názov organizácie je použitý ako prefix pre všetky slice patriace do danej organizácie, napr. „ple1.cesnet.cz“, „ple2.cesnet.cz“.

- **Uzol**

Uzol je server, na ktorom bežia služby siete PlanetLab.

- **Slice**

Slice je sada alokovaných zdrojov distribuovaných v sieti PlanetLab. Pre väčšinu užívateľov znamená slice konzolový prístup k PlanetLab uzlom. Aby bolo možné sa na jednotlivý slice prihlásiť, je nutné k nemu priradiť potrebné uzly. Po priradení sa na každom uzle vytvorí virtuálny server pre daný slice. Slice musí byť pravidelne obnovovaný, inak je po určitom čase vymazaný.

- **Sliver**

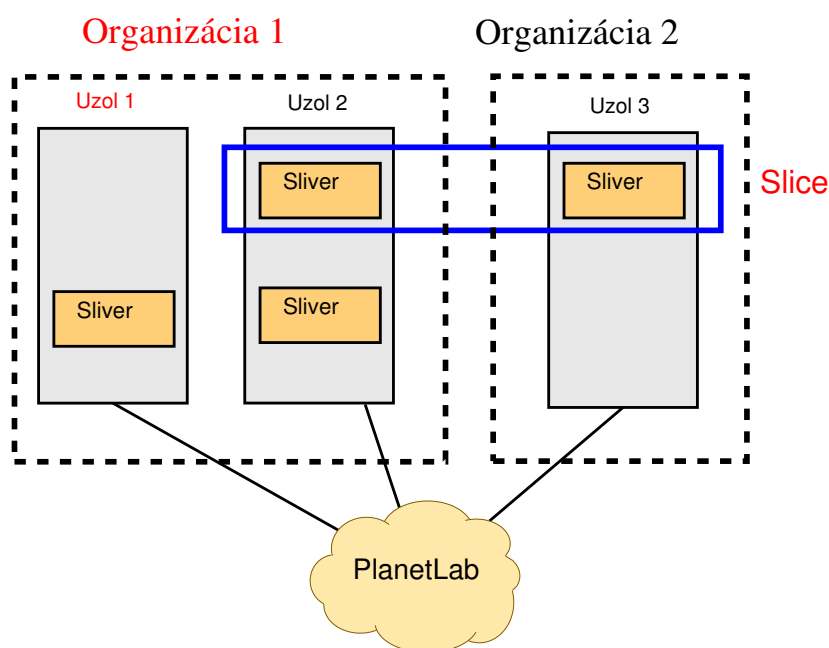
Sliver predstavuje slice bežiaci na konkrétnom serveri. Je to sada alokovaných prostriedkov na jednom konkrétnom uzle.

- **Virtuálny server**

Každý sliver je implementovaný ako linuxový virtuálny server (VServer), ktorý zabezpečuje oddelenie výkonu medzi jednotlivými slivermi na jednom uzle.

- **Užívateľ**

Užívateľom siete PlanetLab je každý, kto vyvíja alebo využíva aplikácie na nej.



Obr. 1.1: PlanetLab architektúra.

Počet jednotlivých organizácií pripojených do siete PlanetLab sa každým rokom zvyšuje. V tabuľke 1.1 je možné sledovať ako sa menil ich počet od roku 2007 do

roku 2014. Podľa štatistík zo siete Onelab [11] je aktuálne ku dňu 9. 5. 2014 do siete PLE pripojených 171 organizácií, 418 uzlov a využíva ju 2789 užívateľov.

Tab. 1.1: Prehľad vývoja siete PlanetLab.

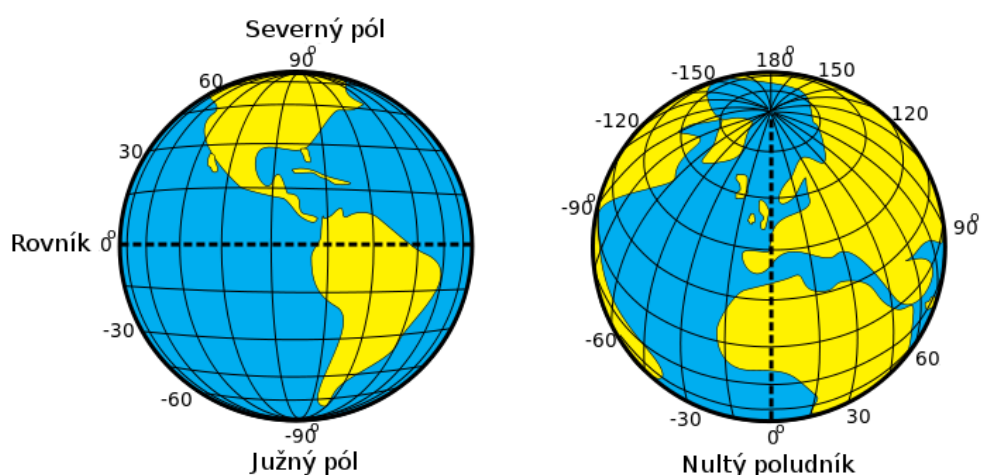
Rok	PLE			PLC	
	Organizácie	Uzly	Užívatelia	Organizácie	Uzly
2007	9	6	21	336	538
2008	41	52	129	363	695
2009	68	127	363	413	801
2010	144	336	1390	374	744
2011	151	327	1744	372	697
2012	157	354	2156	384	697
2013	158	346	2590	401	695
2014	171	418	2789	418	702

2 URČOVANIE POLOHY

Existuje viacero možností ako určiť polohu na Zemi. Najstaršou a dodnes používanou metódou je určovanie polohy pomocou zemepisných súradníc. Táto metóda bola navrhnutá skoro pred 2000 rokmi egyptanom Claudiusom Ptolemyom. Vytvorenie zemepisných súradníc bol spôsob riešenia problému ako presne reprezentovať Zem na plochý kus papiera [7].

2.1 Zemepisné súradnice

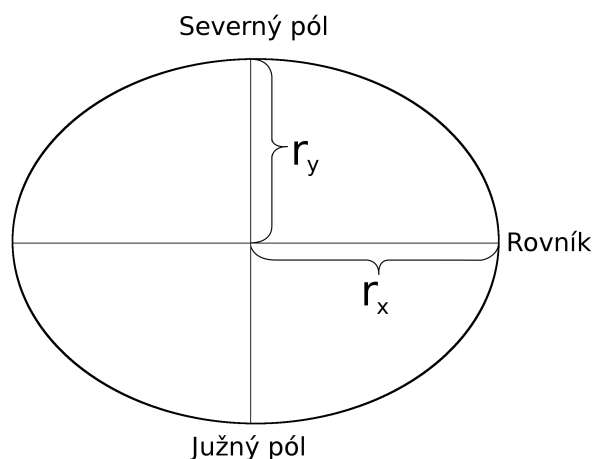
Zemepisné súradnice sa udávajú ako dve hodnoty, a to **zemepisná šírka** a **zemepisná dĺžka**. Na základe toho, na ktorej pologuli bod leží, také má zemepisné súradnice. Najčastejšie sa používa delenie severná a južná šírka a východná a západná dĺžka. Niekedy sa ale namiesto svetových strán používa rozlíšenie na základe kladných a záporných hodnôt 2.1. Body ležiace na severnej pologuli majú kladnú zemepisnú šírku a body na južnej pologuli majú zápornú zemepisnú šírku. Obdobne to platí aj pre zemepisnú dĺžku. Body na východnej pologuli majú kladnú zemepisnú dĺžku a body na západnej pologuli majú zápornú zemepisnú dĺžku. Tento spôsob zobrazovania zemepisných súradníc využíva aj mapová služba spoločnosti Google 3, ktorá je využívaná v tejto práci.



Obr. 2.1: Rozdelenie povrchu Zeme podľa zemepisných súradníc [7].

2.2 Vzdialenosť medzi zemepisnými súradnicami

Tvar Zeme je často reprezentovaný ako guľa. V skutočnosti je Zem na póloch sploštená a matematicky je možné ju reprezentovať ako elipsoid 2.2. Pre elipsoid WGS 84, ktorý je považovaný za najlepší matematický model Zeme je polomer $r_x = 6378,136$ km a polomer $r_y = 6356,751$ km [2].



Obr. 2.2: Elipsoidný model Zeme.

Zemepisná šírka môže nadobúdať hodnoty od 0° na rovníku až po 90° na póloch. Rozdiel medzi každým stupňom zemepisnej šírky je možné previesť na vzdialenosť. Pokiaľ by mala Zem tvar gule, boli by tieto vzdialenostné rozdiely medzi stupňami konštantné. Tým, že má tvar elipsoidu, sa ale tieto rozdiely smerom k pólom zvyšujú. Podľa odvodenia v literatúre [12] je možné vzdialenostné rozdiely hľadaného uhla ϕ vypočítať pomocou rovnice 2.1. Rozdiel je braný medzi zemepisnou šírkou $\phi - 0,5^\circ$ a $\phi + 0,5^\circ$.

$$\Delta_{\text{SIR}}^1 = \frac{\pi \cdot r_x \cdot (1 - e^2)}{180 \cdot (1 - e^2 \cdot \sin^2 \phi)^{3/2}} \quad [\text{km}] \quad (2.1)$$

Konštanta e predstavuje excentricitu, čo je vzdialenosť stredu elipsy od jej ohniska. Táto konštanta je počítaná pomocou vzorca 2.2 odvodeného v literatúre [12] [2].

$$e^2 = \frac{r_x^2 - r_y^2}{r_x^2} \quad (2.2)$$

Rozdiely v zemepisnej dĺžke je taktiež možné previesť na vzdialenosť ako pri zemepisnej šírke. Pre tento účel je využívaný vzorec 2.3 odvodený v literatúre [12]. Konštanta e predstavuje excentricitu ako pri zemepisnej šírke a pre jej výpočet je použitý rovnaký vzorec 2.2.

$$\Delta_{DLZ}^1 = \frac{\pi}{180 \cdot (1 - e^2 \cdot \sin^2 \phi)^{1/2}} \quad [\text{km}] \quad (2.3)$$

Zemepisná šírka a dĺžka môžu byť udávané s presnosťou na niekoľko desiatiných miest. Najčastejšie sa tieto hodnoty udávajú s presnosťou na 6–8 desiatiných miest, čo predstavuje presnosť zhruba na desiatky až jednotky milimetrov. Zväčšeniu presnosti o jedno desatinné miesto odpovedá zníženie vzdialenostného rozdielu na desaťnásobok. V nasledujúcej tabuľke 2.1 sú zobrazené vzdialenostné rozdiely medzi jednotkami stupňov zemepisných šíriek a dĺžok pre vybrané hodnoty zemepisnej šírky ϕ .

Tab. 2.1: Tabuľka vzdialenostných rozdielov medzi jednotkami stupňov zemepisných šírok a dĺžok pre vybrané hodnoty zemepisnej šírky.

ϕ	Δ_{SIR}^1 [km]	Δ_{DLZ}^1 [km]
0°	110,574	111,320
15°	110,649	107,551
30°	110,852	96,486
45°	111,132	78,847
50°	111,229	71,696
60°	111,412	55,800
75°	111,618	28,902
90°	111,694	0,000

3 MAPOVÉ SLUŽBY SPOLOČNOSTI GOOGLE

Spoločnosť Google, Inc. [13] vznikla v roku 1998. Jej zakladatelia Larry Page a Sergej Brin v roku 1996 vybudovali vyhľadávací nástroj (najprv nazývaný ako BackRub), ktorý používal odkazy na stanovenie dôležitosti jednotlivých webových stránok. Tento vyhľadávací nástroj roku 1998 pomenovali Google. Išlo o odkaz na slovo „googol“, matematický termín pre číslo 1 nasledované 100 nulami. Od tohoto roku sa začalo pôsobisko tejto firmy značne rozširovať. Firma v roku 2004 spustila e-mailovú službu Gmail, ktorá sa v krátkom čase stala veľmi rozšírenou po celom svete. V nasledujúcom roku spustila nemenej úspešnú službu Mapy Google a Google Earth.

Služba Mapy Google poskytuje digitálne mapovanie a obsahuje rôzne funkcie ako napr. aktuálne informácie o doprave, trasy verejnej dopravy, snímky na úrovni ulíc či plánovač ciest pre cestovanie. Podľa jedného z tvorcov (Lars Rasmussen) sú Mapy Google spôsobilé ako zorganizovať svetové geografické informácie. Používajú Mercatorovo zobrazenie, takže nemôžu zobrazovať oblasti okolo pólů [13].

Služba Mapy Google je hlavným nástrojom na získavanie a overovanie dát v tejto práci. Pre prístup k tejto službe je možné použiť buď internetovú aplikáciu alebo programátorské rozhranie (Google Geocoding API). Práve spomínaná druhá možnosť je využitá v tejto práci. V nasledujúcej časti budú podrobnejšie popísané spôsoby použitia tohto programátorského rozhrania.

Google Geocoding API

Google Geocoding API ponúka nasledujúce dva typy služieb.

- **Geocoding**
- **Reverzný Geocoding**

Geocoding je proces premeny adresy (napr. 2 Rabenu Yeruham St, Tel-Aviv Jaffa, Israel) na GPS súradnice (32,04283346, 34,7617836).

Reverzný geocoding je presne opačný postup, kde sa GPS súradnice prevádzajú na adresu. API rozhranie ponúka priamy prístup k týmto službám pomocou HTTP žiadostí. Ako odpoveď na túto žiadosť vracia formátovaný text s presne označenými poľami popisujúcimi napríklad názov ulice, mesta alebo kraja.

Pri vytváraní žiadostí je možné stanoviť rôzne dodatočné parametre. Pre potreby tejto práce bude ale stačiť použitie povinných parametrov. Kompletný zoznam parametrov spolu s formátom žiadostí je možné nájsť na manuálových stránkach Google Geocoding API [5].

Na obrázku 3.1 je zobrazená časť odpovede na žiadosť na Geocoding API. V tomto prípade sa jedná o službu geocoding, kde bola ako hľadaná adresa zadaná hodnota „2 Rabenu Yeruham St, Tel-Aviv Jaffa, Israel“. Zo získanej odpovede je následne možné filtrovať ulicu (route), mesto (locality), štát (country) alebo iné potrebné údaje.

```
{
  "results" : [
    {
      "address_components" : [
        {
          "long_name" : "Rabenu Yeruham",
          "short_name" : "Rabenu Yeruham",
          "types" : [ "route" ]
        },
        {
          "long_name" : "Tel Aviv",
          "short_name" : "Tel Aviv",
          "types" : [ "locality", "political" ]
        },
        {
          "long_name" : "Israel",
          "short_name" : "IL",
          "types" : [ "country", "political" ]
        }
      ],
      "formatted_address" : "Rabenu Yeruham, Tel Aviv, Israel",
      "location" : {
        "lat" : 32.0483346,
        "lng" : 34.7617836
      },
      "location_type" : "GEOMETRIC_CENTER",
      "viewport" : {

```

Obr. 3.1: Skrátený výpis formátovanej odpovede na geocoding.

Reverzný geocoding vracia taktiež formátovaný text s presne označenými poľami, ale obsahuje podrobnejšie informácie o získanej adrese. Je v ňom možné nájsť aj údaj o kraji, ktorý v odpovedi na obrázku 3.1 chýba. Skrátený výpis takejto odpovedi je

na obrázku 3.2. Ako hľadané súradnice boli zadané hodnoty „lat“ a „lng“ zistené z odpovede z výpisu 3.1.

```
{
  "results" : [
    {
      "address_components" : [
        {
          "long_name" : "Rabenu Yeruham",
          "short_name" : "Rabenu Yeruham",
          "types" : [ "route" ]
        },
        {
          "long_name" : "Tel Aviv",
          "short_name" : "Tel Aviv",
          "types" : [ "locality", "political" ]
        },
        {
          "long_name" : "Israel",
          "short_name" : "IL",
          "types" : [ "country", "political" ]
        }
      ],
      {
        "address_components" : [
          {
            "long_name" : "Tel Aviv",
            "short_name" : "Tel Aviv",
            "types" : [ "administrative_area_level_1", "political" ]
          },
          {
            "long_name" : "Israel",
            "short_name" : "IL",
            "types" : [ "country", "political" ]
          }
        ]
      }
    }
  ]
}
```

Obr. 3.2: Skrátený výpis odpovede na reverzný geocoding.

4 PRÍPRAVA ZDROJOVÝCH DÁT

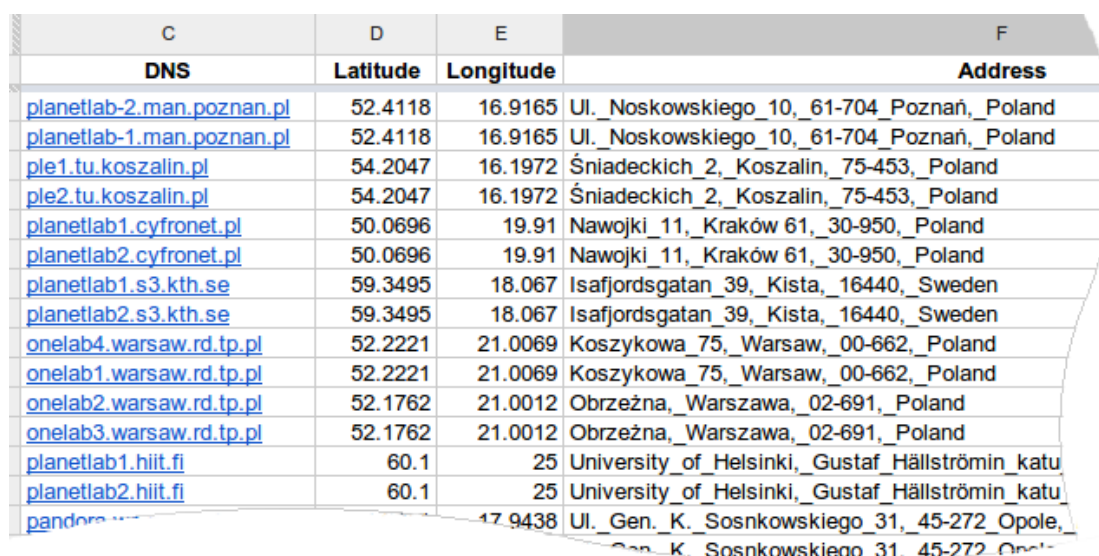
Aby bolo možné testovať a vyhodnocovať jednotlivé údaje o uzloch, bolo potrebné najprv vytvoriť počiatočný zdroj dát.

Každý pripojený uzol má na internetových stránkach siete PlanetLab okrem svojho doménového mena zverejnené aj informácie o svojom fyzickom umiestnení v podobe GPS súradníc a odkaz na internetovú adresu organizácie pod ktorú patrí.

Pre potreby práce boli potrebné GPS súradnice vyčítané zo stránok PlanetLabu a adresa organizácie, ktorá sa o daný uzol stará. K získaniu tejto adresy poslúžila internetová adresa organizácie. Z nej bola vyčítaná adresa ústavu, poprípade oddelenia zodpovedného za chod daných uzlov.

Zo získaných GPS súradníc (zemepisná šírka a dĺžka), adres inštitúcií a dns názvov uzlov bol vytvorený súbor *web.csv*, ktorý slúžil ako počiatočný zdroj dát pre vytvorené skripty. Časť vytvoreného súboru je zobrazená na obrázku 4.1.

Táto úloha bola rozdelená na tretiny medzi Pavla Ilka, Jakuba Poláška a Jána Pružinského (autora tejto práce).



C	D	E	F
DNS	Latitude	Longitude	Address
planetlab-2.man.poznan.pl	52.4118	16.9165	Ul. Noskowskiego 10, 61-704 Poznań, Poland
planetlab-1.man.poznan.pl	52.4118	16.9165	Ul. Noskowskiego 10, 61-704 Poznań, Poland
ple1.tu.koszalin.pl	54.2047	16.1972	Śniadeckich 2, Koszalin, 75-453, Poland
ple2.tu.koszalin.pl	54.2047	16.1972	Śniadeckich 2, Koszalin, 75-453, Poland
planetlab1.cyfronet.pl	50.0696	19.91	Nawojki 11, Kraków 61, 30-950, Poland
planetlab2.cyfronet.pl	50.0696	19.91	Nawojki 11, Kraków 61, 30-950, Poland
planetlab1.s3.kth.se	59.3495	18.067	Isafjordsgatan 39, Kista, 16440, Sweden
planetlab2.s3.kth.se	59.3495	18.067	Isafjordsgatan 39, Kista, 16440, Sweden
onelab4.warsaw.rd.tp.pl	52.2221	21.0069	Koszykowa 75, Warsaw, 00-662, Poland
onelab1.warsaw.rd.tp.pl	52.2221	21.0069	Koszykowa 75, Warsaw, 00-662, Poland
onelab2.warsaw.rd.tp.pl	52.1762	21.0012	Obrzeżna, Warszawa, 02-691, Poland
onelab3.warsaw.rd.tp.pl	52.1762	21.0012	Obrzeżna, Warszawa, 02-691, Poland
planetlab1.hiit.fi	60.1	25	University of Helsinki, Gustaf Hällströmin katu
planetlab2.hiit.fi	60.1	25	University of Helsinki, Gustaf Hällströmin katu
pandora		17.9438	Ul. Gen. K. Sosnkowskiego 31, 45-272 Opole,

Obr. 4.1: Časť súboru *web.csv*.

5 PROGRAMOVANIE SKRIPTOV

Tabuľka 5.1 obsahuje zoznam skriptov použitých pri vytváraní výsledného zoznamu uzlov siete PlanetLab. Pre každý skript sú definované vstupné a výstupné dáta. Vstupné dáta pochádzajú zo súboru *web.csv*, ktorý je podrobnejšie popísaný v kapitole 4. Výstupné hodnoty predstavujú všetky hodnoty, aké môže výsledok skriptu nadobúdať.

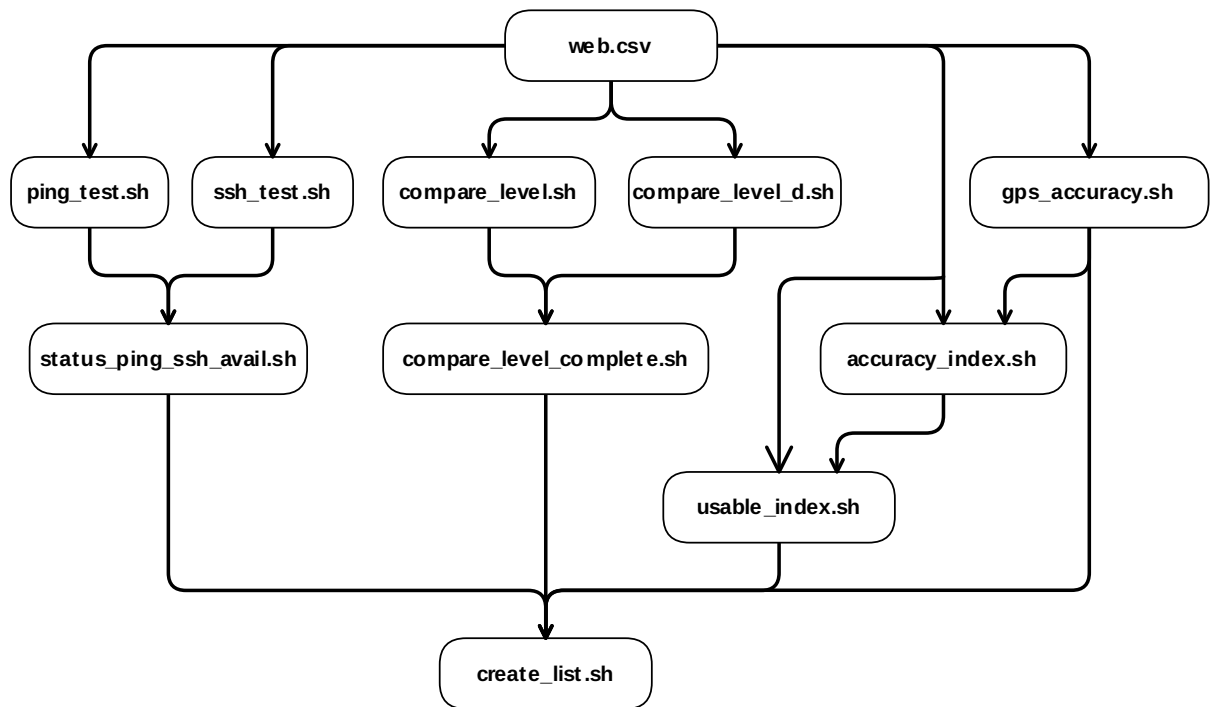
Tab. 5.1: Zoznam naprogramovaných skriptov spolu s ich vstupmi a výstupmi.

	Názov skriptu	Vstupné dáta	Výstupné hodnoty
1.	ping_test.sh	DNS názov uzlu	0 / 1 / noPLE / noIP
2.	ssh_test.sh	DNS názov uzlu	
3.	status_ping_ssh_avail.sh	Výstupy skriptov 1 a 2	
4.	compare_level.sh	Adresa, GPS súradnice	1 / 2 / 3 / 4
5.	compare_level_d.sh	Adresa, GPS súradnice	
6.	compare_level_complete.sh	Výstupy skriptov 4 a 5	
7.	gps_accuracy.sh	GPS súradnice	r_{sirka} a r_{dlzka} z tab. 5.3
8.	accuracy_index.sh	Difference_Accuracy GPS_Difference_MaxMind Výstup skriptu 7	0 / 1
9.	usable_index.sh	Výstup skriptov 3 a 8	0 / 1
10.	create_list.sh	Výstupy skriptov 3, 6, 7, 8 a 9	súbor landmarks.csv

Pre potreby testovania bol ešte vytvorený skript **address_down.sh**, ktorý v tabuľke 5.1 zahrnutý nie je. Jeho úlohou bolo vytvorenie lokálnych kópií odpovedí získaných zo žiadostí na geocoding, poprípade reverzný geocoding. Dôvodom vytvárania týchto kópií bol denný limit žiadostí, ktorý spoločnosť Google na používanie služieb geocodingu zaviedla. Použitím lokálnych kópií sa zrýchlil proces testovania a odladovania jednotlivých skriptov.

Na obrázku 5.1 je vykreslený vývojový diagram zobrazujúci logickú nadväznosť jednotlivých skriptov. Ako už bolo spomenuté, hlavným zdrojom dát je súbor

web.csv, z ktorého jednotlivé skripty využívajú dáta (DNS názov uzlu, GPS súradnice či adresu).



Obr. 5.1: Vývojový diagram nadväzností jednotlivých skriptov z tabuľky 5.1.

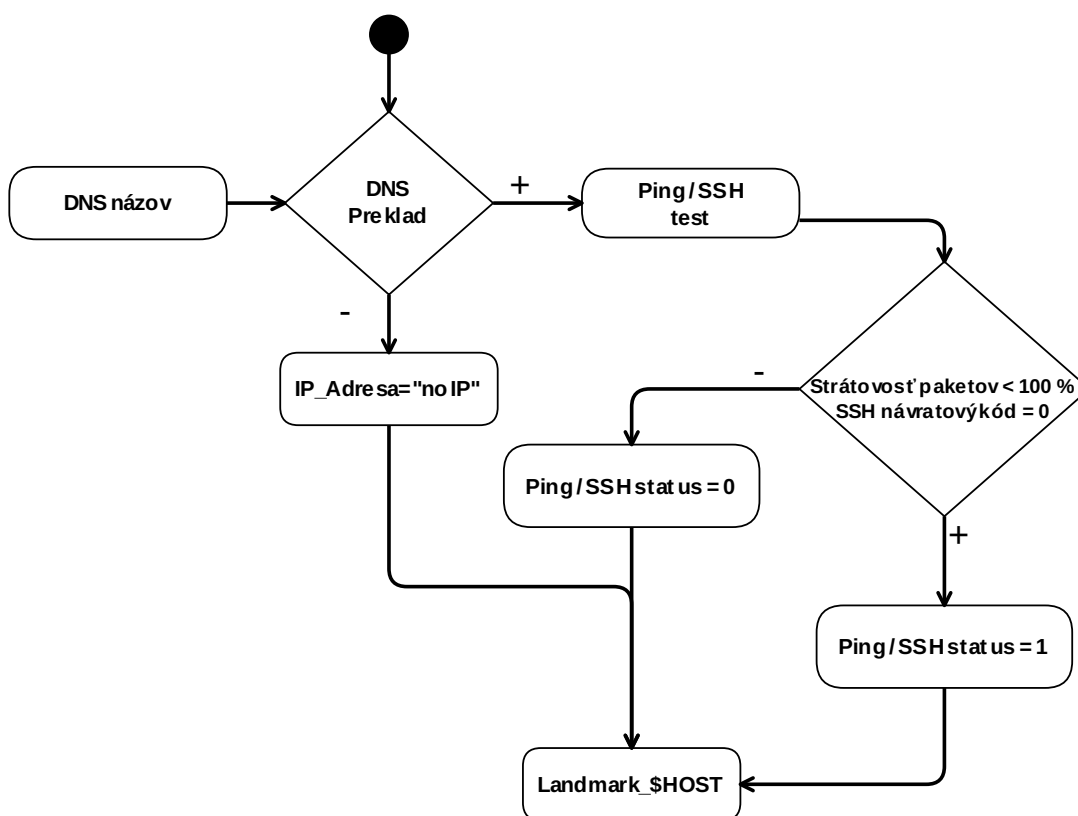
5.1 Overovanie dostupnosti uzlov

Overovanie dostupnosti uzlov bolo zamerané na program ping a protokol SSH.

K tomuto účelu boli použité skripty 1–3 z tabuľky 5.1.

Pomocou skriptov **ping_test.sh** a **ssh_test.sh** bola testovaná dostupnosť. Ako vstupný zdroj dát slúžili doménové názvy jednotlivých uzlov zo súboru *web.csv*. Na každý doménový názov uzlu boli vykonané testy dostupnosti. Tieto testy boli spúšťané automaticky v pravidelných intervaloch po dobu 180 dní. Na automatické spúšťanie bol použitý linuxový plánovač úloh „cron“ [3]. Výstupné dáta získané z testov boli ukladané do separátnych súborov pre každý uzol. Názvy súborov boli tvorené podľa doménového mena, a to *landmark_\$HOST*, kde *\$HOST* predstavuje doménový názov uzla (napr. *landmark_ple1.cesnet.cz*). Výsledky ďalších vykonaných testov boli už následne pridávané do prvotného súboru *landmark_\$HOST*.

Na obrázku 5.2 je znázornený diagram popisujúci princíp fungovania oboch vyššie spomínaných skriptov. Z dôvodu nefunkčnosti niektorých doménových názvov uzlov bolo nutné vytvoriť podmienku, aby sa v prípade nefunkčnosti daný test ukončil a uložil chybové hlásenie „noIP“. Ďalšia možná chyba „noPLE“ označuje uzol, ktorý už nepatrí do zoznamu uzlov PLE.



Obr. 5.2: Vývojový diagram popisujúci princíp testovania dostupnosti uzlov.

Na výpise 5.1 je časť skriptu `ssh_test.sh` pomocou ktorej bolo overované SSH pripojenie k uzlom. Skript na základe návratového hodu vykonaného `ssh` príkazu určuje dostupnosť uzlu. Návratový kód 0 značí úspešne vykonaný príkaz. Význam jednotlivých použitých prepínačov programu `ssh` je možné nájsť v jeho manuálových stránkach[14]. Ako časová hranica pre pripojenie sa k uzlu bola zvolená hodnota 10 sekúnd. Pokiaľ sa nepodarilo pripojiť k uzlu do tohto času, bol uzol označený ako nedostupný pre protokol SSH.

Výpis 5.1: Časť skriptu na overovanie dostupnosti SSH pripojenia.

```
# Pripojí sa k uzlu a vykoná príkaz exit
status_ssh_ou=$(ssh -o ConnectTimeout=10 -o BatchMode=yes\
-o StrictHostKeyChecking=no -q -l cesnet_feec \
-i ~/.ssh/id_planetlab \ $hosts exit )
# Uloží návratovú hodnotu vykonaného príkazu
status_ssh_out=$(echo $?)

# Návratový kód 0 znamená úspešné vykonanie príkazu
if [ $status_ssh_out -eq 0 ]
then
    status_ssh=1
else
    status_ssh=0
fi
```

Skript `ping_test.sh` používa na testovanie dostupnosti program `ping`. Časť tohto skriptu spolu s ukážkou použitia programu je na výpise 5.2. Skript na základe straty paketov určuje dostupnosť uzlov. Aby bol uzol pre daný test označený ako dostupný, musí byť úspešná aspoň 1 z 50 požiadaviek.

Výpis 5.2: Časť skriptu na overovanie dostupnosti spojenia programom ping.

```
# Spustí 50x ping na uzol
ping_output=$(ping -c50 $hosts)
# Vyfiltruje počet úspešných paketov
packet_lost_percent=$(echo $ping_output | \
grep -oP '\d+(?=% packet loss)')

# Overí úspešnosť pingu
if [ $packet_lost_percent -lt 100 ]
then
    status_ping=1
else
    status_ping=0
fi
```

Za pomoci skriptu `status_ping_ssh_avail.sh` sú získané výsledky z priebežných testov vyhodnocované. Ako vstup pre skript slúžia `landmark_$HOST` súbory.

Aby mohol byť uzol označený ako dostupný, musí byť dostupný aspoň v 95 % prípadov testovania. Výstupné dáta tohto skriptu sú ukladané do súboru *ping_ssh_avail.csv*. Tento súbor následne slúži ako vstup pre ďalšie skripty.

5.2 Overovanie zhody zadaných adries uzlov

Overovanie správnosti polôh jednotlivých uzlov spočívalo v porovnávaní adresy získanej zo stránok organizácie zodpovednej za prevádzkovanie daného uzlu a GPS súradníc zozbieraných zo stránok siete PlanetLab. Na tento účel boli vytvorené skripty 4–6 z tabuľky 5.1.

Aby bolo možné vzájomne porovnávať adresu a GPS súradnice, museli sa GPS súradnice najprv previesť na adresu. Pre tento účel boli použité mapové služby od spoločnosti Google, konkrétne ich programátorské rozhranie popísané v kapitole 3.

Princíp porovnávania adries

Na určovanie úrovne zhody jednotlivých adries boli vytvorené dva skripty, a to **compare_level.sh** a **compare_level_d.sh** z tabuľky 5.1. Ich funkcia je podobná, avšak jeden používa na overovanie údaje z geocodingu a druhý z reverzného geocodingu. Vstupom pre oba skripty je vytvorený súbor *web.csv* 4.1. Pre potreby týchto skriptov sú dôležité obsahy polí DNS, latitude, longitude a Address. Výsledná úroveň zhody jednotlivých adries je vyjadrená číselne od 1–4. Číslo 1 znamená najnižšiu presnosť, a to presnosť na úroveň štátu. Nasleduje úroveň kraja, mesta a najvyššia presnosť na úrovni ulice vyjadrená číslom 4.

Princíp fungovania skriptov bude následne vysvetlený na príklade. Ako vstupné dáta budú použité údaje z tabuľky 5.2.

Tab. 5.2: Vzorový záznam o uzle siete PlanetLab.

DNS	Latitude	Longitude	Adresa
planet01.hhi.fraunhofer.de	52.52	13.32	Einsteinufer_37,_Berlin,_Germany

compare_level.sh

Princíp fungovania skriptu **compare_level.sh** je nasledujúci:

1. Vykoná sa žiadosť na geocoding API. Ako hľadaná adresa sa zadá obsah poľa „Adresa“ z tabuľky 5.2.
„Einsteinufer_37,_Berlin,_Germany“.
2. Z odpovede sa vyfiltrujú a uložia údaje o ulici, meste a štáte.
route=„Einsteinufer“, locality=„Berlin“, country=„Germany“
3. Vykoná sa žiadosť na reverzný geocoding. Ako zdrojové údaje sa zadajú obsahy polí „Latitude“ a „Longitude“ z tabuľky 5.2.
lat=„52.52“, lng=„13.32“
4. Z odpovede sa vyfiltrujú a uložia údaje o ulici, meste, kraji a štáte.
route=„Dovestraße“, locality=„Berlin“, administrative_area_level_1=„Berlin“,
country=„Germany“
5. Údaje získané v bode 2 sa porovnávajú s údajmi z bodu 4 a výsledná úroveň sa zapíše do súboru *tmp_cmp.csv*. V tomto prípade je úroveň zhody 3, čo je presnosť na úrovni mesta.

compare_level_d.sh

Služba geocoding v niektorých prípadoch nevráti informáciu ohľadom kraja. Pre tento účel bolo potrebné použiť reverzný geocoding, ktorý túto informáciu zobrazuje. K tomuto účelu bol vytvorený skript **compare_level_d.sh**. Jeho fungovanie je nasledujúce:

1. Vykoná sa žiadosť na geocoding. Ako hľadaná adresa sa zadá obsah poľa „Adresa“ z tabuľky 5.2.
„Einsteinufer_37,_Berlin,_Germany“.
2. Z odpovede sa vyfiltrujú údaje o GPS súradniciach „lat“ a „lng“.
lat=„52.51629“, lng=„13.325“
3. Vykoná sa reverzný geocoding a ako hľadaná adresa sa zadajú GPS súradnice zistené v bode 2.
lat=„52.51629“, lng=„13.325“

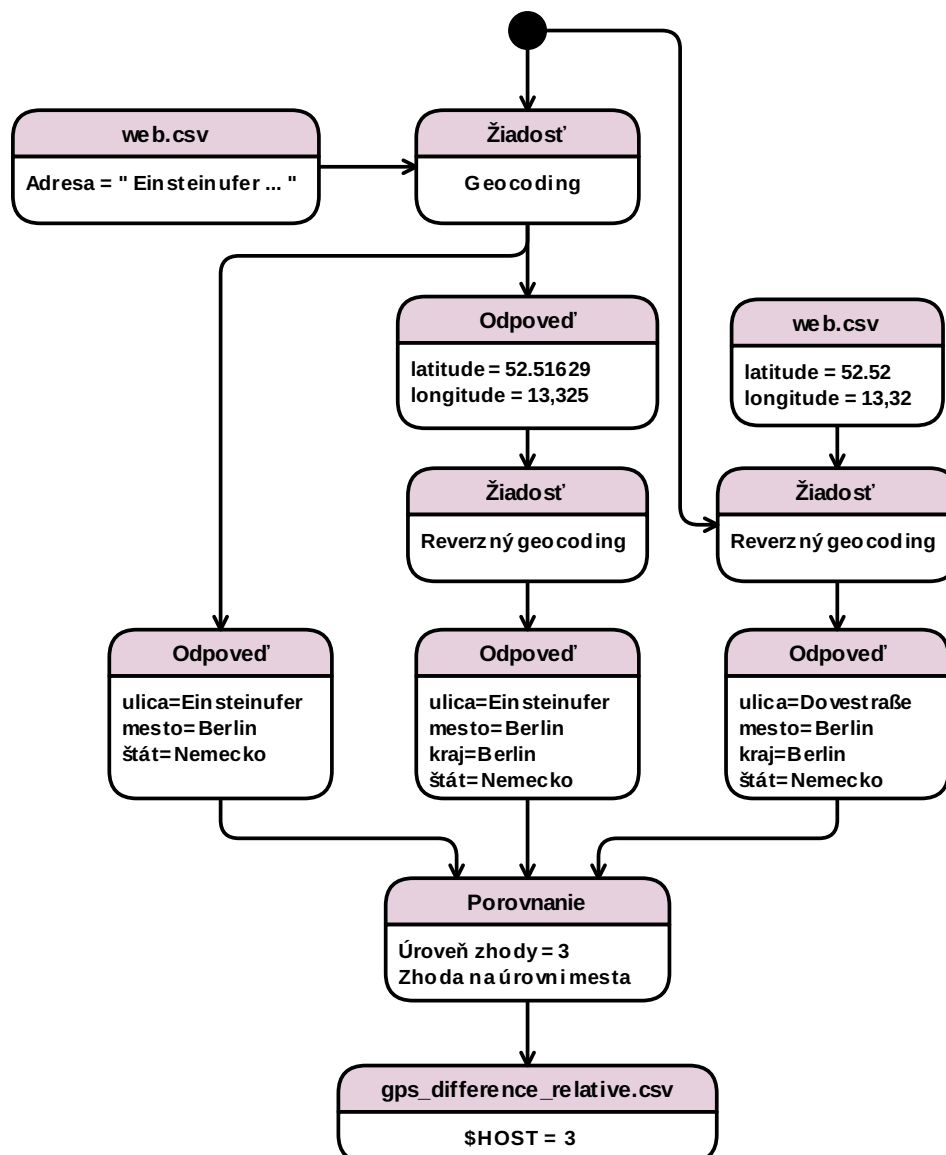
4. Zo získanej odpovede sa vyfiltrujú a uložia údaje o ulici, kraji, meste a štáte.
`route= „Einsteinufer“, locality= „Berlin“, administrative_area_level_1= „Berlin“,
country= „Germany“`
5. Vykoná sa reverzný geocoding. Ako zdrojové údaje sa zadajú obsahy polí
„Latitude“ a „Longitude“ z tabuľky 5.2.
`lat= „52.52“, lng= „13.32“`
6. Z odpovede sa vyfiltrujú a uložia údaje o ulici, kraji, meste a štáte.
`route= „Dovestraße“, locality= „Berlin“, administrative_area_level_1= „Berlin“,
country= „Germany“`
7. Údaje získané z bodu 4 sa porovnajú s údajmi z bodu 6 a výsledná úroveň sa
zapiše do súboru *tmp_cmp_d.csv*. V tomto prípade je úroveň zhody 3, čo je
presnosť na úrovni mesta.

Skript **compare_level_d.sh** vykonáva vlastne tú istú funkciu ako vyššie popísaný skript **compare_level.sh**, akurát tam pridáva ešte informáciu o kraji. Pre zisťovanie celkovej úrovne zhody adres bol ale použitý aj skript **compare_level.sh**. Dôvodom boli chybné preklady niektorých adres. Informácie o ulici (route) vrátené v odpovedi v bode 2 a odpovedi v bode 4 by mali byť rovnaké. Pri niektorých vyhľadávaniach sa všach tento údaj líši. Kvôli tomu je na zisťovanie úrovne použitý okrem **compare_level_d.sh** aj skript **compare_level.sh**. Kombináciou oboch je tak možné nájsť čo najpresnejší výsledok.

Na obrázku 5.3 je vykreslený vývojový diagram zobrazujúci princíp vyhľadávania výslednej úrovne zhody.

compare_level_complete.sh

Skript **compare_level_complete.sh** slúži na porovnanie výsledkov z oboch vyššie popísaných skriptov. Vyššia úroveň presnosti je následne spolu s doménovým názvom príslušného uzla uložená do súboru *gps_diff_relative.csv*, ktorý neskôr slúži ako vstup pre ďalší skript.



Obr. 5.3: Vývojový diagram hľadania výslednej úrovne zhody adresy.

5.3 Výpočet presnosti polohy

Výpočet presnosti polohy spočíval v určení veľkosti polomeru kružnice, do ktorej môže daný uzol patriť. K tomuto účelu bol použitý skript 7 z tabuľky 5.1.

Uzly siete PlanetLab majú zadané zemepisné súradnice s presnosťou na rôzny počet desatinných miest. Podľa počtu desatinných miest bola podľa kapitoly 2 vypočítavaná presnosť polohy, ktorá predstavovala priemer požadovanej kružnice. Z priemeru boli následne vypočítané potrebné polomery r_{sirka} a r_{dlzka} z tabuľky 5.3.

Hlavným cieľom tejto práce je analýza uzlov siete nachádzajúcich sa na území Európy. Preto bola ako referenčná hodnota na určovanie presnosti polôh uzlov stanovená zemepisná šírka 50° , čo predstavuje územie strednej Európy. V tabuľke 5.3 sú zobrazené polomery kružníc predstavujúce presnosti polohy pre zemepisnú šírku 50° s presnosťou súradníc až na 6 desatinných miest.

Tab. 5.3: Presnosti polohy podľa počtu desatinných miest pre zemepisnú šírku 50° .

Počet desatinných miest	Rozdiel v stupňoch	$r_{\text{šírka}}$ [km]	$r_{\text{dĺžka}}$ [km]
0	1,0	55,61	35,85
1	0,1	5,561	3,585
2	0,01	0,5561	0,3585
3	0,001	0,05561	0,03585
4	0,0001	0,005561	0,003585
5	0,00001	0,0005561	0,0003585
6	0,000001	0,00005561	0,00003585

Východisková hodnota vzdialenostných rozdielov pre zemepisnú šírku 50° bola prebraná z kapitoly 2 z tabuľky 2.1.

Nakoľko zadané súradnice niektorých uzlov majú rozdielny počet desatinných miest pre zemepisnú šírku než pre dĺžku, bola pre výpočet výslednej presnosti braná súradnica s menším počtom desatinných miest. V prípade zhodného počtu desatinných miest pre obe zadané zemepisné súradnice bola pre výpočet vybraná zemepisná šírka, nakoľko predstavuje väčšiu mieru nepresnosti.

5.4 Určovanie indexu presnosti

Presnosť zadanej polohy uzlu bola okrem postupu popísaného v bode 5.3 zisťovaná ešte ďalšími dvomi metódami. Tieto metódy boli spracované Pavlom Ilkom a Jakubom Poláškom. Celková hodnota presnosti sa teda určovala pomocou troch rôznych metód. Jednalo sa o tieto metódy:

- metóda založená na porovnávaní GPS súradníc uzla s GPS súradnicami zodpovednej organizácie (Pavol Ilko),
- metóda založená na porovnávaní medzi GPS súradnicami a údajmi zo statickej lokalizačnej databázy MaxMind[6] (Jakub Polášek),
- metóda založená na počítaní presnosti na základe počtu desatinných miest zadaných GPS súradníc.

Celková hodnota presnosti zadaných súradníc pre daný uzol bola počítaná podľa vzorca 5.1.

$$l = MAX(o, r, d) \quad (5.1)$$

Premenná o predstavuje rozdiel medzi GPS úradnicami uzla a GPS súradnicami zodpovednej organizácie, r je hodnota presnosti počítaná na základe počtu desatinných miest zadaných GPS súradníc 5.3 a d predstavuje rozdiel medzi GPS súradnicami a statickou lokalizačnou databázou MaxMind. Maximálna hodnota z týchto troch databáz predstavuje finálnu presnosť l .

Pomocou skriptu **accuracy_index.sh** bol vzorec 5.1 aplikovaný na všetky uzly v zozname *web.csv*. Referenčná hodnota, podľa ktorej sa určoval index presnosti, bola 5 km. Uzly s presnosťou menšou než je referenčná hodnota boli označené indexom 1, čo predstavuje presne zadanú polohu. Uzly s presnosťou prevyšujúcou referenčnú hodnotu boli označené ako nepresné s indexom 0. Výstup skriptu bol smerovaný do súboru *accuracy_index.csv*, ktorý slúži ako vstup pre ďalšie skripty.

5.5 Určovanie indexu využiteľnosti

Index využiteľnosti slúži ako parameter, pomocou ktorého je možné jednoducho oddeliť uzly vhodné na testovacie účely od nevhodných uzlov. Aby mohol byť uzol označený ako využiteľný, musí spĺňať nasledujúce tri parametre:

- dostupnosť na program ping v aspoň 95 % prípadoch testovania,
- dostupnosť na protokol SSH v aspoň 95 % prípadoch testovania,

- zadaná poloha s presnosťou menšou než 5 km.

Pomocou skriptu **usable_index.sh** sú všetky tieto podmienky overené pre každý uzol a výsledné indexy využiteľnosti sú exportované do súboru *usability_index.csv*.

5.6 Vytváranie výstupného dokumentu

Všetky doposiaľ popísané skripty vytvárali jednotlivé dielčie dáta. Ako je možné vidieť na diagrame 5.1, všetky spracované a zistené údaje sú nakoniec spracované skriptom **create_list.sh**. Jedná sa o finálny skript, ktorý má za úlohu vytvoriť dokument s dosiahnutými výsledkami. Ako zdrojové dáta sú použité výstupy zo skriptov 3, 6, 7, 8 a 9 z tabuľky 5.1. Na obrázku 5.4 je zobrazená časť obsahu výsledného súboru *landmarks.csv*.

A	B	C	D	E	F	G	H
DNS	GPS_Accuracy	GPS_Diff_Relative	Status_Ping	Status_SSH	Status_Availability	Accuracy_index	Usable_index
planet1.inf.tu-dresden.de	0.05561	3	0	0	0	1	0
planetlab-2.ida.liu.se	0.05561	4	1	1	1	1	1
planetlab-1.ida.liu.se	0.05561	4	0	0	0	1	1
ple1.cesnet.cz	0.05561	3	1	1	1	0	0
ple2.cesnet.cz	0.05561	3	1	1	1	0	0
planetlab2.cesnet.cz	0.05561	3	noPLE	noPLE	noPLE	0	0
planetlab1.cesnet.cz	0.05561	3	noPLE	noPLE	noPLE	0	0
planetlab-4.dis.uniroma1.it	0.05561	3	0	0	0	0	0
planetlab-3.dis.uniroma1.it	0.05561	3	1	0	0	0	0
plab-1.diegm.uniud.it	0.05561	3	1	1	1	1	0
plab-2.diegm.uniud.it	0.05561	3	1	1	1	1	1
planetlab1.s3.kth.se	0.03585	3	1	1	1	0	0
planetlab2.s3.kth.se	0.03585	3	1	1	1	0	0
ple1.ait.ac.th	0.03585	1	0	0	0	0	0
ple2.ait.ac.th	0.03585	1	0	0	0	0	0
plab1.create-net.org	0.03585	3	1	0	0	0	0
plab2.create-net.org	0.03585	3	1	0	0	0	0
planetlab1.cs.uit.no	0.03585	4	1	1	1	1	1
planetlab2.cs.uit.no	0.03585	4	1	1	1	1	1
planetlab2.mta.ac.il	0.03585	4	1	1	1	0	0
planetlab1.mta.ac.il	0.03585	4	1	1	1	0	0
				0	0	0	0
				0	0	0	0

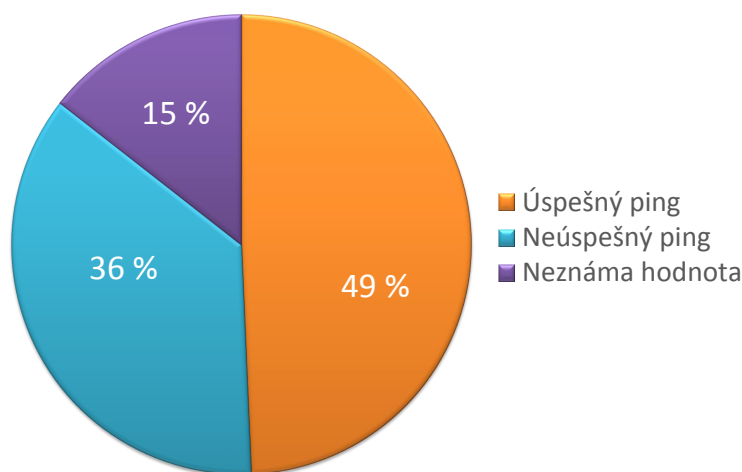
Obr. 5.4: Časť obsahu výsledného súboru *landmarks.csv*.

6 DOSIAHNUTÉ VÝSLEDKY

Na základe výsledkov získaných pomocou skriptov popísaných v sekcii 5 boli spracované nasledujúce štatistické a grafické výstupy.

6.1 Dostupnosť uzlov

Ako bolo popísané v kapitole 5.1 overovanie dostupnosti uzlov sa sústredilo na program ping a protokol SSH. Na obrázku 6.1 je vyobrazený koláčový graf zobrazujúci celkový stav dostupnosti uzlov na program ping.

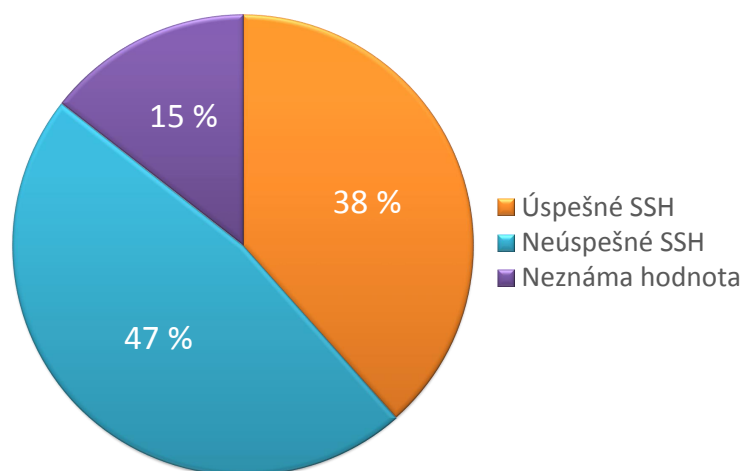


Obr. 6.1: Graf dostupnosti uzlov siete PlanetLab na program ping.

Uzlov s úspešnou odozvou na program ping je iba necelá polovica, a to 49%. Uzly, pri ktorých nebolo možné preložiť doménový názov alebo nepatria do Planetlab Europe, boli zaradené do kategórie „Neznáma hodnota“. Tie tvorili necelú 1/7 z celkového počtu uzlov. Zvyšných 36% tvorili uzly, ktoré neodpovedali na program ping. To mohlo byť zapríčinené buď blokovaním tohto programu na firewalle oddeľujúcom tento uzol od internetu alebo celkovou nedostupnosťou uzla, napr. z dôvodu fyzického vypnutia servera.

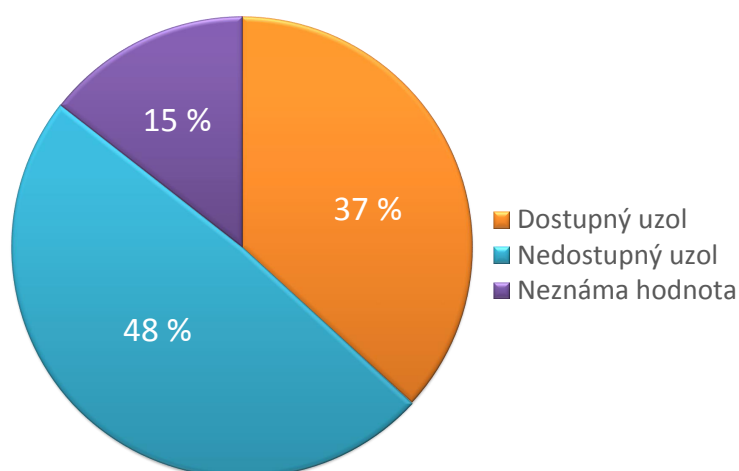
Výsledky dostupnosti uzlov na protokol SSH sú oproti pingu horšie. Z celkového počtu 367 testovaných uzlov je iba 38% z nich dostupných cez protokol SSH. Oproti dostupnosti na program ping je to zhoršenie o 11%. Uzlov zaradených do kategórie

„Neznáma hodnota“ je stále rovnaký počet, a tak zaberajú rovnakých 14 %. Počet nedostupných uzlov tvorí skoro polovicu všetkých uzlov (47 %).



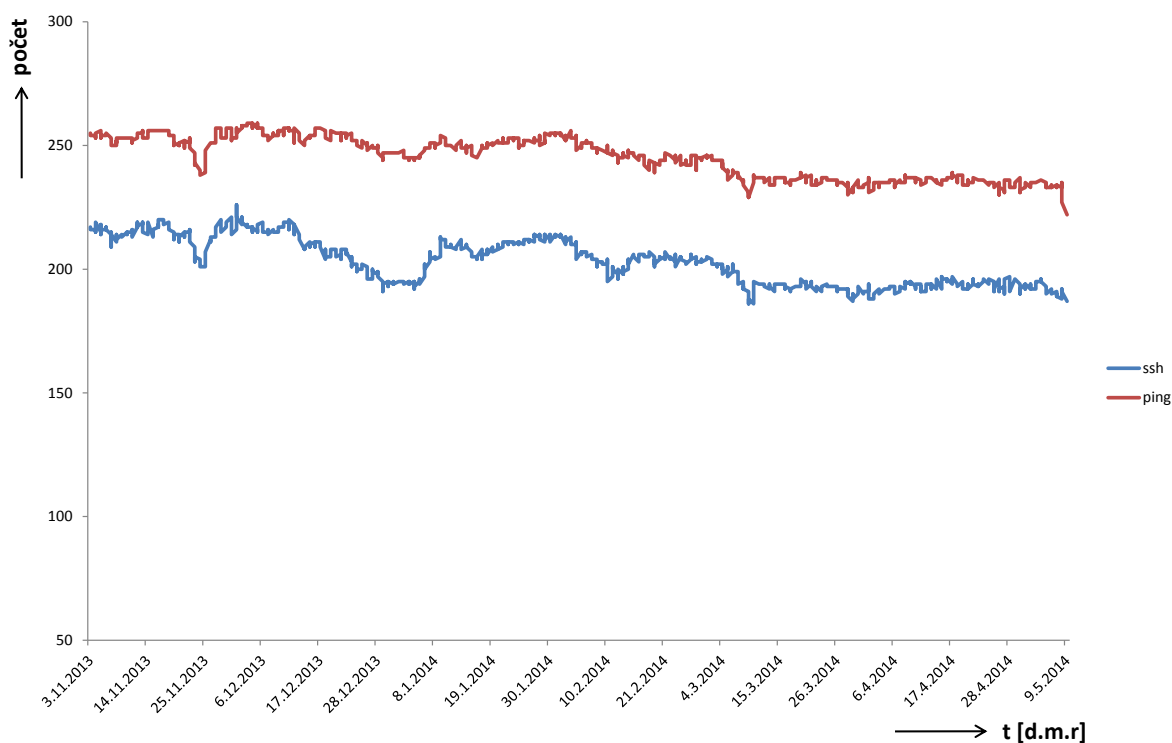
Obr. 6.2: Graf dostupnosti uzlov siete PlanetLab na protokol SSH.

Celková dostupnosť uzlu je počítaná ako súčin dostupnosti na program ping a na protokol SSH, takže obe dostupnosti musia byť rovné 1, aby bol uzol označený ako dostupný. Na obrázku 6.3 sú v grafe spracované výsledky o dostupnosti uzlov. Získaný graf je skoro totožný s grafom 6.2 vykresľujúcim dostupnosť uzlov na protokol SSH. Z dosiahnutých výsledkov vyplýva, že iba 6 uzlov, ktoré sú dostupné na protokol SSH, nie je dostupných na program ping. V grafe je to možné pozorovať ako zmenu o 1 %.



Obr. 6.3: Graf dostupnosti uzlov na protokol SSH + program ping.

Počas celej doby testovania niektoré uzly menili v čase svoju dostupnosť. Pri niektorých uzloch sa zmenila dostupnosť len na určitý okamih a pri niektorých nastala trvalá zmena. Tieto zmeny sa týkali ako dostupnosti na protokol SSH, tak aj na program ping. Na obrázku 6.4 je grafická závislosť počtu dostupných uzlov v závislosti na čase.



Obr. 6.4: Závislosť počtu uzlov dostupných na protokol SSH a program ping v čase.

Z grafu je možné pozorovať mierne klesajúcu tendenciu v počte dostupných uzlov. Na začiatku merania bolo na program ping dostupných 255 uzlov a na protokol SSH 216. V priebehu testovania tieto hodnoty nepravidelne klesali a stúpali pričom na konci testovania to bolo už iba 223 uzlov dostupných na ping a 190 na protokol SSH.

Planetlab Europe zahrňuje uzly z približne 25 štátov Európy. Tabuľka 6.1 obsahuje prvých 10 krajín s najväčším počtom dostupných uzlov. Na prvom mieste je Nemecko s najväčším počtom dostupných uzlov v Európe, a to 23. Avšak, z celkového počtu uzlov, ktoré má Nemecko v Európe, tvoria tie dostupné iba 41,82%. Obdobne

je to aj s ostatnými štátmi. Jedinou výnimkou je Nórsko, ktoré má síce celkovo iba 4 uzly ale status dostupnosti dosahujú všetky, takže má 100 % dostupných uzlov.

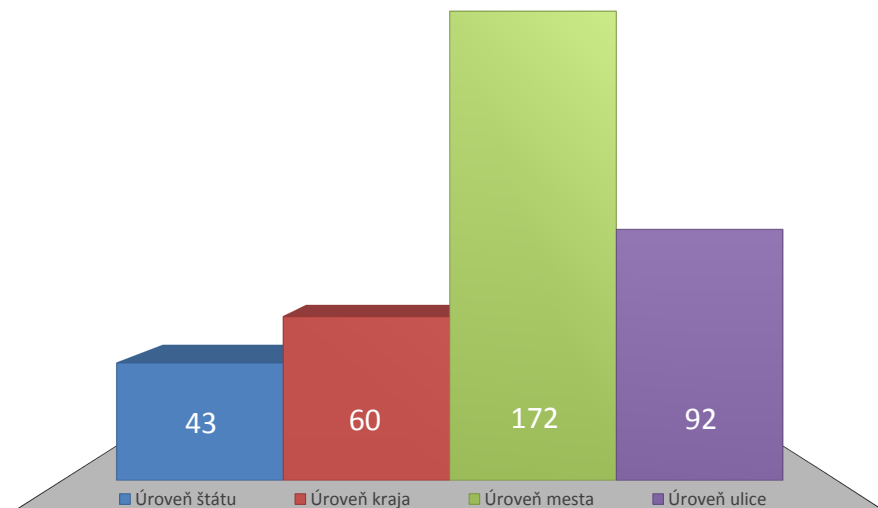
Tab. 6.1: Tabuľka krajín spolu s percentuálnym zastúpením dostupných uzlov.

Štát	Status dostupnosti uzlov			
	Dostupné	Nedostupné	Neznáma hodnota	%
Nemecko	23	20	12	41,82 %
Španielsko	13	12	0	52 %
Poľsko	13	13	6	40,63 %
Anglicko	10	5	2	58,82 %
Francúzsko	10	33	3	21,74 %
Taliansko	9	21	0	30,00 %
Fínsko	8	4	0	66,67 %
Belgicko	8	6	0	57,14 %
Portugalsko	7	9	0	43,75 %
Grécko	6	14	0	30,00 %

6.2 Úrovně zhody zadaných adres uzlov

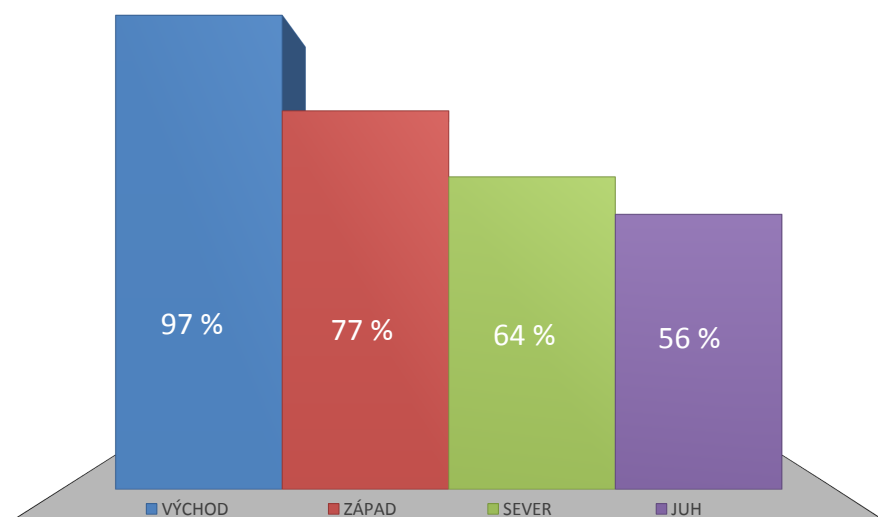
V sekcii 5.2 boli porovnávané údaje o fyzickej polohe uzlov. Výstupom z tejto časti boli číselné hodnoty predstavujúce úroveň zhody zistenej polohy pre každý uzol siete PlanetLab. Na obrázku 6.5 je histogram zobrazujúci jednotlivé úrovne zhody adres spolu s množstvom uzlov, ktoré do jednotlivých úrovní náležia. Najväčšie zastúpenie, konkrétne 172 uzlov, má zhoda na úrovni mesta. Druhú najpočetnejšiu skupinu, 92 uzlov, tvoria uzly, ktoré majú zhodu adres na úrovni ulice.

Jednotlivé uzly boli na základe správne zadanej adresy rozdelené na dve skupiny, a to uzly so správne zadanou adresou a uzly s nesprávne zadanou adresou. Za správne zadanú adresu bola považovaná zhoda adres maximálne na úrovni mesta, čiže úroveň ulice a úroveň mesta.



Obr. 6.5: Histogram počtu uzlov pre každú úroveň zhody adresy.

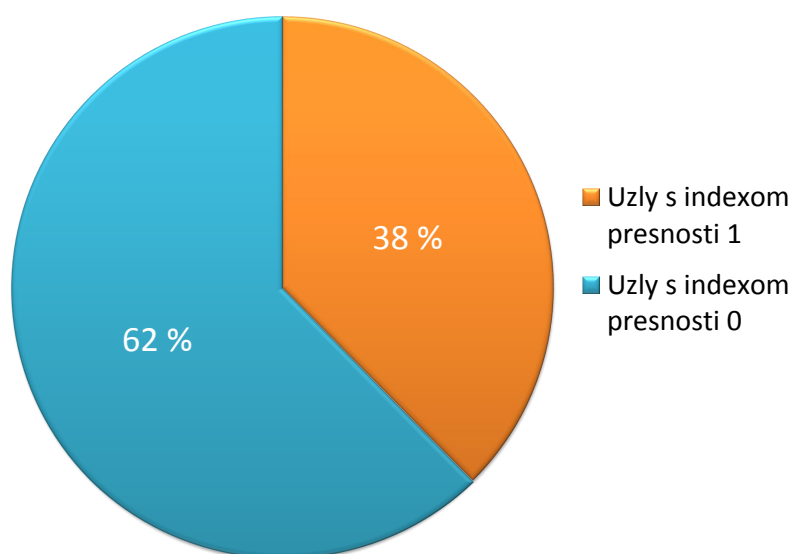
Podľa výsledkov z grafu na obrázku 6.6 majú uzly vo východnej Európe zadané adresy s najväčšou úrovňou zhody. Tomuto údaju odpovedá hodnota 97%. Pri celkovom počte 58 uzlov vo východnej Európe to vychádza iba na 2 uzly s presnosťou zadanej adresy nižšou než je úroveň mesta. Ostatné časti Európy majú adresy zadané s vyššou nepresnosťou. Pri južnej Európe to vychádza skoro až na polovicu všetkých uzlov. V južnej Európe sa celkovo nachádza 100 uzlov a z toho až 54 uzlov s nepresne zadanou adresou.



Obr. 6.6: Graf zastúpenia uzlov so správne zadanou adresou pre každú časť Európy.

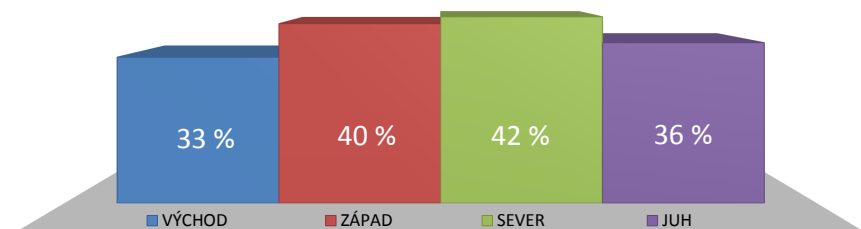
6.3 Presnosť zadaných súradníc

V časti 5.4 boli rozdelené uzly podľa indexu presnosti na dve skupiny. Prvú skupinu tvorili uzly so zadanou polohou s presnosťou do 5 km. Tieto uzly mali priradený index presnosti 1. Na vytvorenom grafe 6.7 je možné vidieť ich celkové zastúpenie. Z celkového počtu 367 testovaných uzlov tvoria tieto uzly iba 38 %. Ostatných 62 % zaberá druhá skupina uzlov. To sú uzly s indexom 0, a teda uzly s nepresne zadanou polohou.



Obr. 6.7: Graf percentuálneho zastúpenia uzlov s indexom presnosti 1 a 0.

Zo spracovaných výsledkov vyplýva, že percentuálne zastúpenie uzlov s indexom presnosti 1 je pre každú časť Európy rozdielne. Na histograme 6.8 je možné pozorovať rozdiely medzi jednotlivými časťami.



Obr. 6.8: Percentuálne zastúpenie uzlov s indexom presnosti 1 pre jednotlivé časti Európy.

Rozdiel v zastúpení je ale iba v jednotkách percent. Najlepšie zastúpenie týchto uzlov má severná Európa. Za ňou nasleduje západná Európa so zastúpením iba o 2 % horším. Následne južná a nakoniec východná Európa s najmenším percentuálnym zastúpením.

V tabuľke 6.2 sa okrem percentuálneho zastúpenia uzlov nachádzajú aj informácie ohľadom ich konkrétneho počtu. Podľa zistených výsledkov sa najviac uzlov s indexom presnosti 1 nachádza na území západnej Európy. Nasledujú ostatné časti v poradí južná, severná a východná.

Tab. 6.2: Tabuľka počtu uzlov s indexom presnosti 1 pre jednotlivé časti Európy.

Časť Európy	Celkovo uzlov	Uzlov s indexom 1	Percentuálne zastúpenie
západná	144	58	40 %
južná	100	36	36 %
severná	55	23	42 %
východná	58	19	33 %

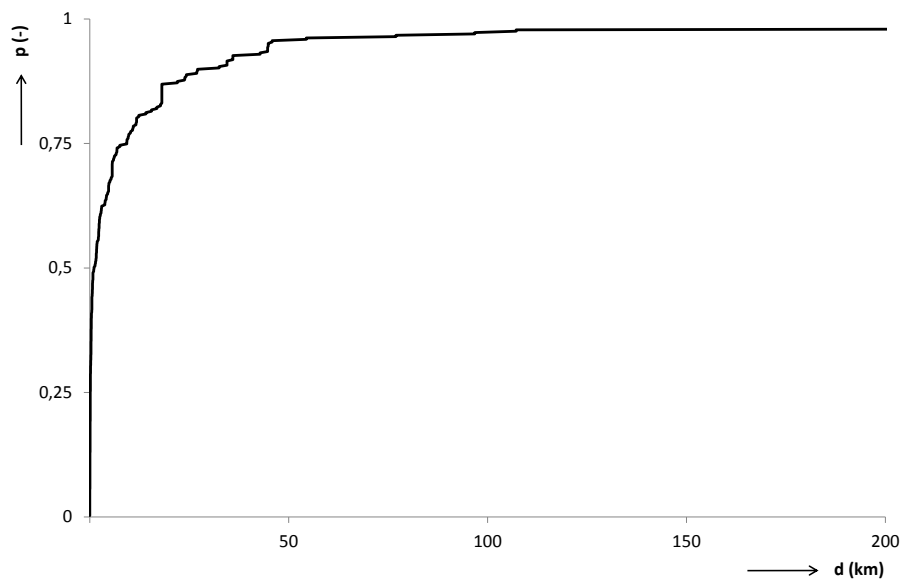
Poradie jednotlivých častí Európy podľa počtu uzlov sa od poradia podľa percentuálneho zastúpenia uzlov celkom líši. Južná a západná Európa majú síce väčší počet uzlov, no oproti ich celkovému množstvu tvoria iba relatívne malé percento.

Distribučná funkcia

Ako bolo popísané v sekcii 5.3 na určovanie výslednej presnosti zadanej polohy uzlu boli celkovo použité tri metódy.

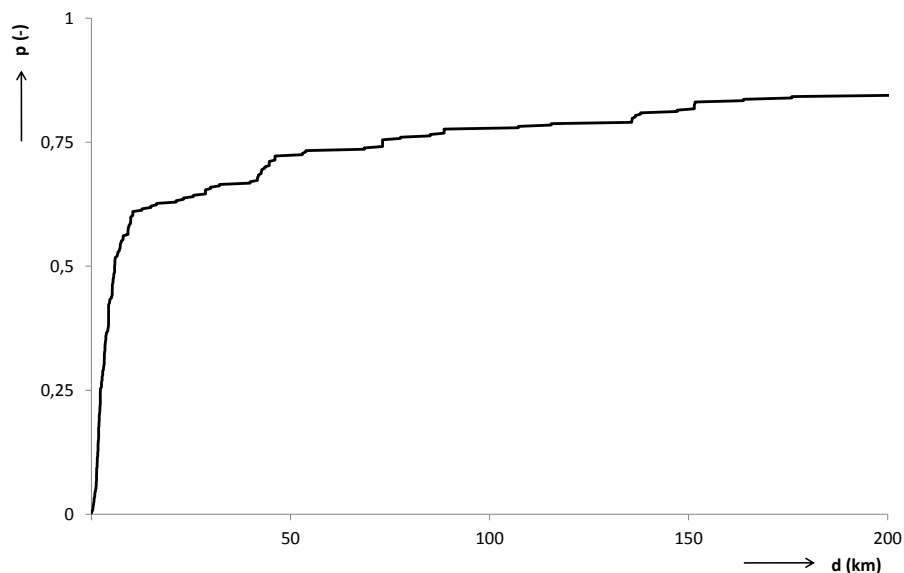
Na grafické spracovanie hodnôt získaných týmito metódami bola použitá distribučná funkcia. Distribučná funkcia je funkcia, ktorá udáva pravdepodobnosť, že je hodnota náhodnej premennej menšia ako zadaná hodnota.

Na grafickom výstupe 6.9 je distribučná funkcia metódy založenej na porovnávaní GPS súradníc uzla s GPS súradnicami zodpovednej organizácie. Získaná krivka sa blíži k pravdepodobnosti 1 zhruba pri hodnote 50 km. Zjednodušene môžeme teda tvrdiť, že podľa tejto metódy má väčšina uzlov presnosť zadanej polohy pod 50 km.



Obr. 6.9: Graf distribučnej funkcie metódy založenej na porovnávaní GPS súradníc.

Distribučná funkcia metódy založenej na porovnávaní medzi GPS súradnicami a údajmi zo statickej lokalizačnej databázy MaxMind má oproti predchádzajúcej metóde horší priebeh. Grafický výstup tejto metódy je na obrázku 6.10.

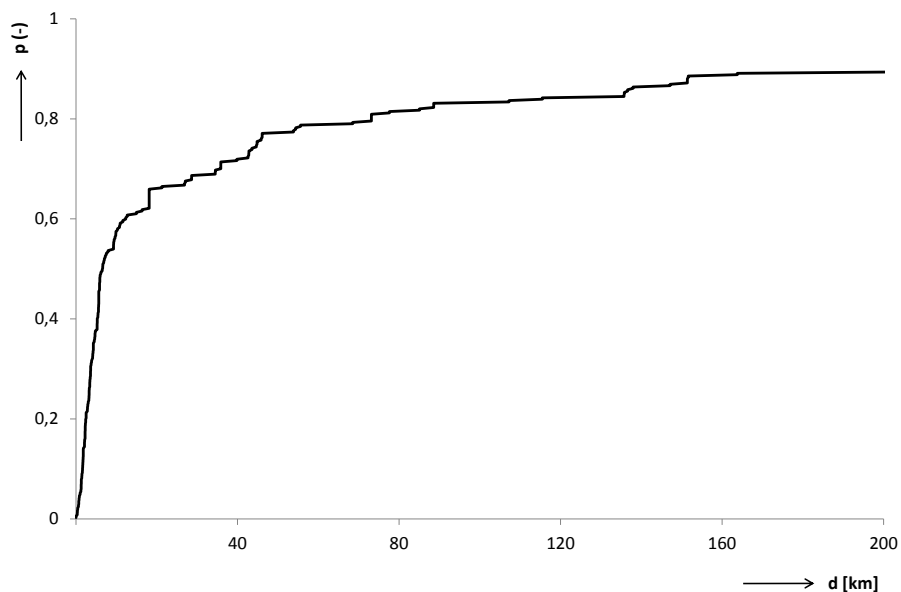


Obr. 6.10: Graf distribučnej funkcie metódy založenej na porovnávaní medzi GPS súradnicami a údajmi zo statickej lokalizačnej databázy MaxMind.

Hodnota vzdialenosti, pri ktorej sa získaná krivka blíži pravdepodobnosti 1, sa

nachádza až mimo rozsah grafu. Pri vzdialenosti 50 km dosahuje pravdepodobnosť hodnotu iba necelých 0,75 oproti pravdepodobnosti blížiaccej sa 1 v predchádzajúcej metóde.

Kombináciou týchto spomínaných dvoch metód a metódy z časti 5.3 vznikli podľa vzorca 5.1 finálne hodnoty presností polôh. Graf distribučnej funkcie týchto finálnych presností je na obrázku 6.11.



Obr. 6.11: Graf distribučnej funkcie finálnych presností zadaných polôh uzlov.

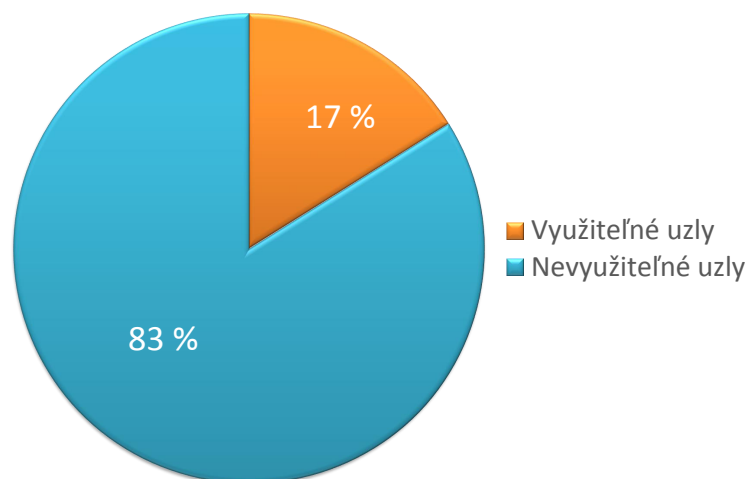
V tabuľke 6.3 sú vypísané minimálne a maximálne hodnoty presností zistené pomocou jednotlivých metód. Najpresnejšie výsledky sa podarilo dosiahnuť metódou zohľadňujúcou počet desatinných miest zadaných GPS súradníc.

Tab. 6.3: Tabuľka minimálnych a maximálnych presností zadaných súradníc pre jednotlivé použité výpočetné metódy.

Metóda	Min	Max
Porovnávanie GPS súradníc	0,00088 km	716,68 km
Porovnávanie s databázou MaxMind	0,03939 km	9380,75 km
Počet desatinných miest GPS súradníc	0,00005561 km	55,61 km

6.4 Využitelnost uzlov

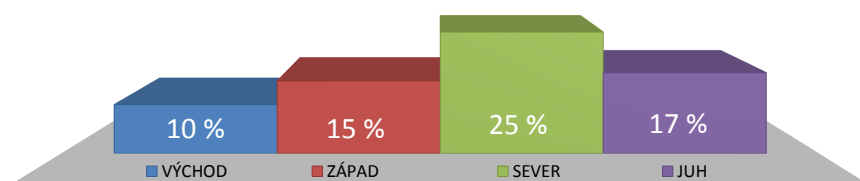
Jednoznačnému rozdeleniu uzlov na využiteľné a nevyužitelné je venovaná časť 5.5. Výsledky získané z tejto časti boli spracované do grafov. Na obrázku 6.12 je spracované percentuálne zastúpenie využiteľných a nevyužitelných uzlov.



Obr. 6.12: Graf percentuálneho zastúpenia využiteľných a nevyužitelných uzlov.

Využitelné uzly tvoria iba malú časť spomedzi všetkých uzlov. Z celkového počtu 367 testovných uzlov je iba 57 využiteľných.

Zastúpenie využiteľných uzlov pre jednotlivé časti časti Európy je vyobrazené na histograme 6.13.



Obr. 6.13: Percentuálne zastúpenie využiteľných uzlov pre jednotlivé časti Európy.

Najväčšie zastúpenie využiteľných uzlov má severná časť Európy. Podľa dosiahnutých výsledkov je to časť Európy ako s najväčším zastúpením využiteľných uzlov, tak aj s najväčším zastúpením uzlov s presne zadanou polohou (kapitola 6.8).

V tabulke 6.4 je už spomínaných 57 využitelných uzlov rozdelených medzi jednotlivé časti Európy. Podľa výsledkov má najväčší počet využitelných uzlov západná Európa a j keď percentuálne je na tom skoro najhoršie.

Tab. 6.4: Tabuľka počtu uzlov s indexom využiteľnosti 1 pre jednotlivé časti Európy.

Časť Európy	Celkovo uzlov	Uzlov s indexom 1	Percentuálne zastúpenie
západná	144	22	15 %
južná	100	16	17 %
severná	55	13	25 %
východná	58	6	10 %

Porovnanie desiatich krajín Európy s najväčším počtom využitelných uzlov je spísané v nasledujúcej tabuľke 6.5. V počte uzlov je na tom najlepšie Nemecko. Pri porovnaní percentálneho zastúpenia je na tom ale najlepšie Nórsko, ktoré má 100 % svojich uzlov označených ako využitelné.

Tab. 6.5: Tabuľka krajín spolu s percentuálnym zastúpením využitelných uzlov.

Štát	Status využiteľnosti		
	1	0	%
Nemecko	9	46	16,36 %
Taliansko	7	23	23,33 %
Belgicko	6	8	42,86 %
Poľsko	5	27	15,63 %
Francúzsko	5	41	10,87 %
Portugalsko	4	12	25,00 %
Grécko	4	16	20,00 %
Nórsko	4	0	100,00 %
Anglicko	3	14	17,65 %
Írsko	3	5	37,50 %

7 ZÁVER

Hlavným cieľom tejto práce bolo zanalyzovať uzly siete PlanetLab a overiť správnosť ich polohy. Začiatok práce bol venovaný teoretickému úvodu do siete PlanetLab, kde bola objasnená hlavná funkcia tejto experimentálnej siete spolu s popisom ako celá sieť funguje.

Prvá časť analýzy uzlov spočívala v overovaní dostupnosti jednotlivých služieb na každom uzle. Na overovanie dostupnosti bol použitý program ping. Podľa dosiahnutých výsledkov vychádza, že iba necelá polovica (49 %) všetkých testovaných uzlov je dostupná na ping. Okrem tejto dostupnosti bola testovaná aj dostupnosť protokolu SSH. Percentuálne zastúpenie uzlov dostupných na protokol SSH oproti všetkým testovaným tvorí iba 38 %. Uzlov, ktoré sú dostupné na program ping a súčasne aj na protokol SSH je ešte menej, a to 37 %.

Ďalšia časť analýzy uzlov spočívala v overovaní zadaných polôh jednotlivých uzlov siete. Overovanie spočívalo v porovnávaní zistenej adresy zo stránok zodpovednej inštitúcie a adresy zadanej na stránkach siete PlanetLab. Zhoda adries bola vyhodnocovaná relatívne na úroveň ulice, mesta, kraja a štátu. Pri porovnávaní sa využívalo programátorské rozhranie Geocoding API od spoločnosti Google. Z výsledkov získaných z tejto časti boli vytvorené grafické výstupy. Podľa týchto výstupov má najviac uzlov so správne zadanou adresou východná Európa, a to 97 %.

Okrem porovnávania zadaných adries bola vyhodnocovaná aj presnosť, s akou je adresa zadaná. Tá sa určovala pomocou troch rôznych metód. Súčasťou tejto práce bolo vyhodnotiť presnosť iba jednou z nich, a to metódou založenou na porovnávaní počtu desiatiných miest zadaných zemepisných súradníc. V práci bol vysvetlený postup pri výpočte vzdialenostných rozdielov spolu s výpisom použitých vzorcov. Práve metóda popísaná v tejto práci dosahovala najlepšie výsledky čo sa týka presnosti, s akou je adresa zadaná.

Posledná časť analýzy bola venovaná určovaniu indexu presnosti a využiteľnosti. Bol popísaný postup triedenia uzlov podľa týchto indexov a dosiahnuté výsledky boli spracované do prehľadných tabuliek a grafov. Uzlov s indexom využiteľnosti 1 bolo podľa zistených výsledkov iba 17 %. To sú uzly, ktoré sú vhodné na použitie na

rôzne testovacie účely.

Celkovým výsledkom analýzy bol zoznam všetkých testovaných uzlov spolu so zistenými hodnotami z jednotlivých testov. Tento zoznam bol doplnený ešte o ďalšie údaje, ktoré boli spracované inými študentmi a vo finále bude slúžiť na ďalšie testovacie účely.

Informácie spracované v rámci tejto práce boli použité v publikácii [1], ktorá bola prednesená na konferencii TSP¹.

¹International Conference on Telecommunications and Signal Processing (<http://tsp.vutbr.cz/>)

LITERATÚRA

- [1] 37th International Conference on Telecommunications and Signal Processing (TSP). In: KOMOSNY, D., J. PRUZINSKY, P. ILKO, J. POLASEK a K. OGUZHAN. *On Geographic Coordinates of PlanetLab Europe* [online]. 2014 [cit. 2014-05-25]. Dostupné z URL: <<https://docs.google.com/file/d/0B5Z4jYjUDxqkNTRzUW1Sb2NOQnM/>>.
- [2] CALAIS, Eric. *Elements of Geodesy* [online]. West Lafayette, 2009 [cit. 2013-12-30]. Dostupné z URL: <http://web.ics.purdue.edu/~ecalais/teaching/gps_geodesy/Elements_of_geodesy_1.pdf>. Purdue University.
- [3] CRON. *Linux Man Pages* [online]. 1996 [cit. 2013-12-27]. Dostupné z URL: <<http://www.linuxmanpages.com/man8/cron.8.php>>.
- [4] COOPER, Mendel. *Advanced Bash-Scripting Guide: An in-depth exploration of the art of shell scripting* [online]. 2010, 10 Mar 2014 [cit. 2014-05-24]. ISBN 978-1435752191. Dostupné z URL: <<https://tldp.org/LDP/abs/html/index.html>><http://tldp.org/LDP/abs/html/index.html>>.
- [5] Google Maps API Web Services: The Google Geocoding API. *Google Developers* [online]. v3. ©2007, Last updated December 26. 2013. [cit. 2013-12-27]. Dostupné z URL: <<https://developers.google.com/maps/documentation/geocoding/>>.
- [6] *MaxMind: IP Geolocation and Online Fraud Prevention* [online]. © 2014 [cit. 2014-05-11]. Dostupné z URL: <<http://www.maxmind.com>>.
- [7] MCNAMARA, Joel. *GPS For Dummies*. Indianapolis, Indiana: Wiley Publishing, Inc., © 2004. ISBN 0-7645-6933-3.
- [8] NAVRÁTIL, Jiří. *PlanetLab - model budoucího Internetu: Zpravodaj ÚVT MU*. [online]. 2006, XVI, č. 5, s. 1-5 [cit. 2014-05-20]. Dostupné z URL: <<http://www.ics.muni.cz/zpravodaj/articles/525.html>>.

- [9] NEMETH, Evi, Garth SNYDER a Trent HEIN. *Linux: Kompletní příručka administrátora*. 2. aktualizované vydání. Brno: Computer Press, 2008. ISBN 978-80-251-2410-9.
- [10] PlanetLab Consortium. *PlanetLab: An open platform for developing, deploying, and accessing planetary-scale services* [online]. © 2007 [cit. 2013-12-27]. Dostupné z URL: <<http://www.planet-lab.org>>.
- [11] PlanetLab Europe. *OneLab: Future Internet Testbeds* [online]. © 2010 [cit. 2013-12-27]. Dostupné z URL: <<http://onelab.eu/index.php/testbeds/onelab-testbeds/planetlab-europe/ple-statistics.html>>.
- [12] RAPP, Richard H. *Geometric Geodesy Part I* [online]. Columbus, Ohio, 1991-04 [cit. 2013-12-27]. Dostupné z URL: <<http://hdl.handle.net/1811/24333>>. Ohio State University Department of Geodetic Science and Surveying.
- [13] Společnost – Google. *Google* [online]. [2014] [cit. 2014-05-20]. Dostupné z URL: <<https://www.google.sk/intl/cs/about/company/>>.
- [14] SSH. *Linux Man Pages* [online]. 1996 [cit. 2014-5-10]. Dostupné z URL: <<http://www.linuxmanpages.com/man1/ssh.1.php>>.

ZOZNAM SYMBOLOV, VELIČÍN A SKRATIEK

API rozhranie pre programovanie aplikácií – Application Programming Interface

DNS systém doménových mien – Domain Name System

GPS globálny lokalizačný systém – Global Positioning System

HTTP hypertextový prenosový protokol – Hypertext Transfer Protocol

IP interne tprotokol – Internet Protocol

km kilometer – kilometre

PLC PlanetLab Central

PLE PlanetLab Europe

SSH zabezpečený prístup k príkazovému riadku – Secure Shell

TSP telekomunikácie a spracovanie signálu – Telecommunications and Signal Processing